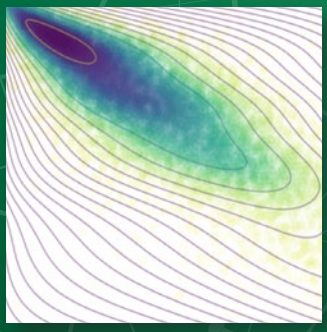
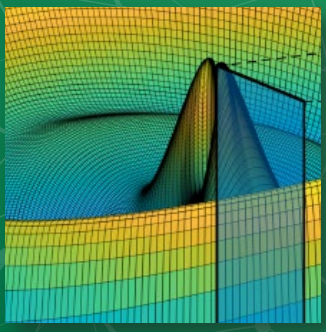
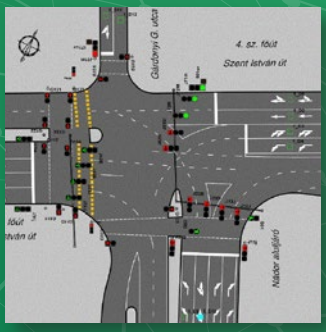
HU–MATHS–IN

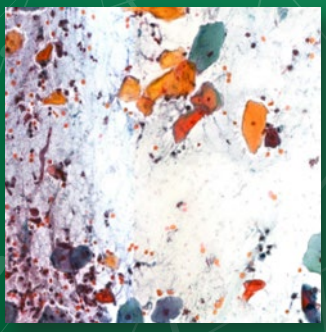
Hungarian Service Network of Mathematics for Industry and Innovation

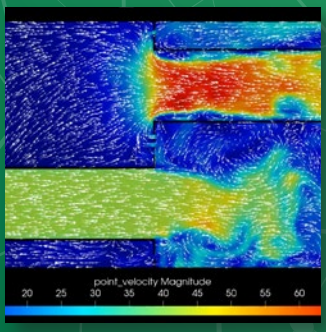
HU-MATHS-IN SUCCESS STORIES OF MATHEMATICAL SHORT-TERM PROJECTS FOR INDUSTRY IN 2017-2021











SZÉCHENYI 2020

SZÉCHENYI EGYETEM UNIVERSITY OF GYŐR

SZTE SZEGEDI TUDOMÁNYEGYETEM UNIVERSITY OF SZEGED

DEBRECENI EGYETEM

HUNGARIAN GOVERNMENT

European Union
European Social Fund

INVESTING IN YOUR FUTURE

HU-MATHS-IN

The Hungarian Service Network of Mathematics for Industry and Innovation is a cooperation of 22 Hungarian interdisciplinary mathematical research groups with the aim of providing services with mathematical solutions to companies and organizations of all sectors. The research groups have joined the Network from seven major universities and two research institutes. HU-MATHS-IN welcomes the research and innovation challenges of industry in its One-Stop-Shop.

HU-MATHS-IN is strongly integrated in EU-MATHS-IN, the European network with the same purpose.

For further information on organization, concluded projects and the One-Stop-Shop, please visit http://hu-maths-in.hu.

EFOP-3.6.2-16-2017-00015 project

The consortium of three universities, Széchenyi István University, University of Szeged and University of Debrecen participated in the project EFOP-3.6.2-16-2017-00015 with the aim of deepening and developing the activities of HU-MATHS-IN. The leader of the consortium and the beneficiary of the project was Széchenyi István University. The project helped to build the institutional organisation and operation of HU-MATHS-IN and contributed to the targeted and efficient development of mathematical methods and to the exploitation of the results in industrial, in particular international R+D+I projects.

The contracted amount of funding was HUF 1 389 996 580 HUF, with a 100% funding rate. The project ended on 31. 03. 2021.

The project has been supported by the European Union.

HU-MATHS-IN success stories of mathematical short-term projects for industry in 2017–2021

Editor: Zoltán Horváth

Publisher: Széchenyi István Egyetem

Publishing managed by: Zoltán Horváth

Printed by: Innovariant Nyomdaipari Kft., Algyő, 2021

ISBN 978-615-6443-00-7

TABLE OF CONTENT

EDITORIAL

The core of the activities of a service network are the projects that are delivered by the members for the clients. HU-MATHS-IN, the service network of industrial mathematics operating in Hungary, has delivered dozens of short-term projects in the past years, thanks to the EFOP-project. The short reports on the projects are compiled into this book.

The HU-MATHS-IN projects are based on an open procurement procedure and quality management. The HU-MATHS-IN office announced calls for projects and researchers belonging to the network could submit proposals to support their aimed basic research with significant mathematical content on a problem arising from industry. The industrial problems were collected by the researchers or, in some cases, the HU-MATHS-IN office. In the latter case the office searched for suitable proposer. Each proposal contained a support letter from the industrial partner whose problem drove the research. Propositions were presented to the whole HU-MATHS-IN community during Center Days and were subject for discussion. Two researchers from the Scientific Board of HU-MATHS-IN evaluated the proposals in short reviews and then this board decided on the rankings. The members of the HU-MATHS-IN Scientific Board are Tibor Csendes, Sándor Fridli, Balázs Harangi, Zoltán Horváth, Tibor Illés, Tamás Kis, László Márkus, Gergely Röst, László Szeidl. Then the EFOP-project, according to the available budget, supported the proposals in the order of ranking. During the projects' lifetime, researchers regularly presented short reports to the community at Center Days. Since these projects were of aimed basic research, intellectual property related topics were discussed and contracted between the institution of the project and industrial partners directly.

The length of the projects varies from half to three years, depending on the arising of the industrial problem and the specialities of the EFOP-project. The achievements of the research started in the HU-MATHS-IN projects, in particular basic research related aspects are often to be discussed in forthcoming papers, e.g. in the planned special issue of a journal.

Zoltán Horváth

president of HU-MATHS-IN,
EFOP-project leader, editor

The effect of epidemics on mortality

Miklós Arató[1], **Patrik Hegel**[2] and **László Márkus**[3]

[1]Department of Probability Theory and Statistics, Loránd Eötvös University, and Department of Mathematics and Computational Sciences, Széchenyi István University,
[2]NN Hungary, and Department of Mathematics and Computational Sciences, Széchenyi István University,
[3]Department of Probability Theory and Statistics, Loránd Eötvös University, and Department of Mathematics and Computational Sciences, Széchenyi István University,

[1]E-mail: miklos.arato@ttk.elte.hu
[2]E-mail: hegel.patrik@gmail.com
[3]E-mail: laszlo.markus@ttk.elte.hu

1 Executive Summary

As a result of the research, we proposed a simple but new measure of annual mortality changes, which can clearly identify significant external factors (epidemics, weather). Calculations for more than 30 countries and more than 60 years show spatial correlations. In our view, the actual impact of epidemics is better measured by the total change in mortality than by the number of deaths caused by the disease. We used this method to measure the impact of COVID-19 on the few countries for which data are already available. As is well known, the mortality of the life insurance portfolios differs significantly from the population mortality. However, the above procedures can also be applied to their own data.

In the next step of the research, we developed and expanded Skiadas and Janssen's dynamic mortality model to model consecutive years and decades' mortality and health state, while cohort effects are taken into account as well. Likelihood-based parameter estimations are shown for the models, and in the case of the multi-year model, we describe a stochastic optimization algorithm to attain the maximum-likelihood estimation. The generalized Skiadas model can already incorporate external factors' influence, so we hope that more accurate models can be developed to model insurers' longevity risk.

2 Challenge overview

Over the last two centuries, the mortality of the Earth's population has been gradually declining, and as a result, life expectancy has roughly doubled. Examining the economic and social impacts due to increasing life expectancy has special importance in many sectors. For example, concerning the pension system, pension insurance, and life insurance, much of the calculation is based on mortality rates. Since an increasing proportion of the population has an exceptionally long lifetime, these institutions need to assess long-term and longevity risks properly.

However, changes in mortality did not show a steady improvement, and relapses also occurred frequently. These mortality increases may also play an important role in the growth rate of future life expectancy, so it would be important to model them as accurately as possible. The COVID-19 epidemic has also highlighted the importance of epidemics, as sudden spikes in mortality can also cause an unexpected increase in payments.

3 Implementation of the project

The industrial partner formulated the research problem, analyzed and evaluated the results in regular consultations. It committed itself to the project by letting one of its employees work on it, ensuring permanent attention and guidance. The academic party collected and processed the data and created the models. Miklós Arató and Patrik Hegel searched the data in different databases. László Márkus and Miklós Arató modeled the mortality changes and made suggestions for the measures of the changes. Patrik Hegel determined the maximum likelihood estimates in the modified Skiadas model and performed the predictions.

4 The research problem, methods, and results

Determining the factors influencing mortality has always been one of the most interesting and important tasks for demographers, actuaries, and statisticians. With the help of these factors, we attempted to make as accurate forecasts as possible for the development of life expectancy in the future. However, to date, there is no generally accepted methodology for measuring past changes in mortality and determining epidemics' impact. In our research, we proposed a measure of change and a new method of predicting mortality.

4.1 The research problem

The objective of the research is

- to determine the inference on yearly mortalities in association with historic epidemics/pandemics,
- modeling the difference of years with/without pandemics,
- to estimate the impact of COVID-19 on mortality in Hungary and other counties,
- to model the evolution of mortality in the observed countries over decades,
- to predict the development of mortality.

During the research, we encountered several problems

- Data from previous epidemics are incomplete.

- Epidemics **cause unexpected**, prevention measures **avert expected** deaths.
- Several countries have not yet published their 2019 mortality data and January 1, 2020 population.

The following data sources were used.

- Human Mortality Database (mortality.org)
- KSH Central Statistical Bureau of Hungary
- WHO (inaccurate databases)

4.2 The applied methods

In the first step, mortality changes were defined.

- Predict in a given year the number of deaths in gender and age cohorts, based on previous year's data; summarize for population
- Relative change: Registered minus predicted death number over predicted
- Treat and display it as a time series

Relative change for the i^{th} year in formulas:

- $t_{i,a,j}$: exposure time for the gender a and age j in year i
- $d_{i,a,j}$: number of deaths for the gender a and age j in year i
- $md_{i,a,j} = t_{i,a,j} \cdot \frac{d_{i-1,a,j}}{t_{i-1,a,j}}$: number of "expected" deaths for the gender a and age j in year i
- $d_i = \sum_{a,j} d_{i,a,j}$: number of deaths in year i
- $md_i = \sum_{a,j} md_{i,a,j}$: number of "expected" deaths in year i
- $rc_i = \frac{d_i - md_i}{md_i}$: relative change for the i^{th} year

For 2020, however, it was not possible to determine the relative change using this method. As the population numbers on 1 January 2020 were already available in several countries, the expected mortality was also determined according to the populations on 1 January and the mortality by year of birth. Although this calculation is more inaccurate than using the risk exposure, it still gave similar results to previous years as the more accurate method, so we thought it was applicable for 2020.

Our new mortality prediction model was based on the model of Janssen and Skiadas ([1]). In this model, let S be a stochastic process that characterizes the state of health at the level of the individual, which is defined by the following formula:

$$S(t) = S(0) + \int_0^t \mu(s)ds + \int_0^t \sigma(s)dW(s)$$

where $S(0)$ is the initial value of the stochastic process, $\mu(t)$ is the drift and $W(t)$ is a standard Wiener process. In order to obtain a quantity representative of the whole population, we need an aggregation of these. Denote this process by H and define it as the expected value of the S process at every t time:

$$H(t) = E(S(t)) = \int_{t_0}^{t} \mu(s)ds$$

This function $H(t)$ is called the "health state function".
A possible parametrization of the health state function used in many cases:

$$H(t) = c - (lt)^b$$

In order to model mortality, they define the time of death when the level of health state function decreases below a certain level (in the case of the previous parametrization of $H(t)$, this level can be set to 0). For estimating the parameters, Janssen and Skiadas use a squared error minimizing the iterative algorithm. Alternatively, if we have observations regarding the l_j number of years alive at age j and the d_j number of deaths at age j, we can use a maximum-likelihood estimate for the parameters:

$$L = \prod_{j=j_0}^{j_{max}} \binom{l_j}{d_j} p_j^{d_j} (1-p_j)^{l_j - d_j},$$

where p_j is the probability of death in year j.

Suppose that data are available for a set of consecutive calendar years I and ages J. In order to model the change in mortality and health state over time, the parameters are assigned to calendar years and cohorts. The examined calendar years I must have a different starting level, i.e., a parameter c_i $(i \in I)$. For cohorts, l_{i-j} should be the relevant parameter $(j \in J)$. That is, in the i -th calendar year, the health state of a j years old person is described by the term $c_i - (l_{i-j}j)^b$.
Thus, the parameters c are intended to characterize the effects specific to a given year; for example, if there were a war or a serious epidemic, we would expect a lower value for the parameter belonging to the year. A salient change in cohort parameters may indicate a generational effect.
The parameters are estimated using the maximum-likelihood method as described earlier. The likelihood to maximize is

$$L^* = \prod_{k \in I} N_k! \prod_{i,j:i-j=k} \left(\frac{g_k(j)^{d_{ij}}}{d_{ij}!}\right) \times \frac{(1 - g_k(j_0) - \ldots - g_k(j_{max}))^{N_k - d_{ij_0 - \ldots - d_{ij_{max}}}}}{(N_k - d_{ij_0 - \ldots - d_{ij_{max}}})!}.$$

where k is the examined cohort, N_k is the population of the k -th cohort at j_0 years of age, j_0 and j_{max} are the examined minimum and maximum ages, $g_k(j)$ is the probability that one person in the k-th cohort dies at the age of j, and d_{ij} is the number of deceased in the i -th year at the age of j.

The basic concept of our new prediction method is to predict the evolution of the parameters over time and then use them to calculate the evolution of mortality in different cohorts over the years.

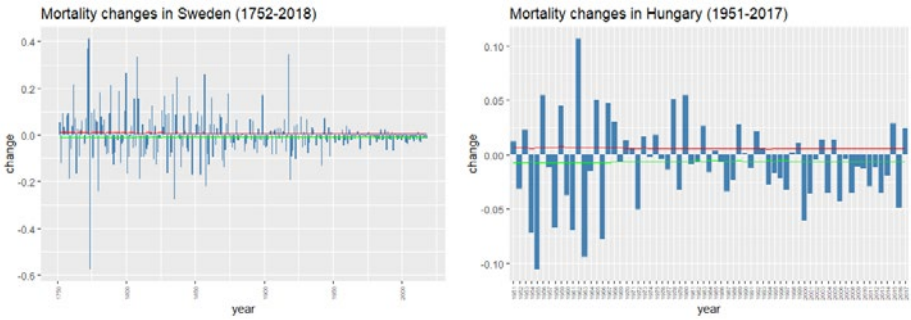


Figure 1: Mortality changes in Sweden and Hungary up to 2018

4.3 The achieved research results

Mortality relative changes were determined for more than 30 countries. The results are shown in the Figures 1, and 2 for Sweden and Hungary. It is possible to notice the enormous impact that the Spanish flu had in Sweden in 1918. In the much shorter data for Hungary the year 1962 stands out with its awful increase in deaths as compared to expectation based on the previous year's data. Most likely, it is a result of a larger than usual flu epidemic, reported for that year. With the results for 2020

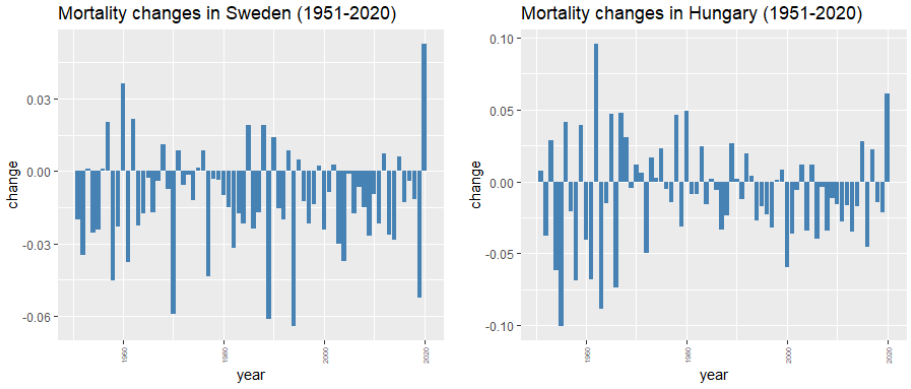


Figure 2: Mortality changes in Sweden and Hungary, including the year 2020

included, the Figure 2 is obtained for the Swedish and Hungarian relative changes of mortality, starting from 1949. The clearly visible negative trend in the last several decades corresponds well to the already mentioned growing life expectancy. The year 2020 emerges in both plots with strikingly high changes in mortality.

Restricting data to the period 1991-2016, we compute the correlation of the relative changes as a similarity measure between several European countries. Using multidimensional scaling (MDS finds the best planar representation of a d-dimensional point set), Figure 3 displays the similarity of these countries. Despite very different expected life spans, the variability is similar for countries geographically close to each

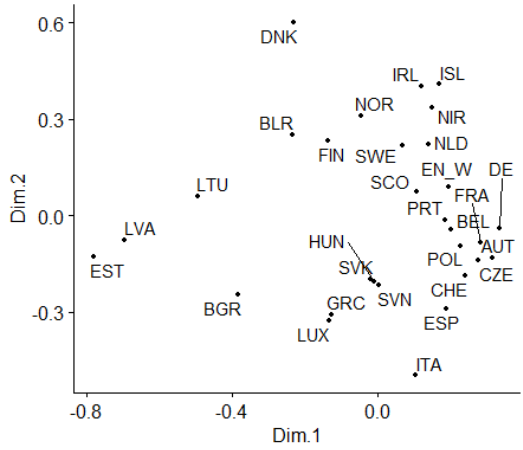


Figure 3: Similarity of countries by the correlation of mortality changes

other! Therefore, a spatial effect can be hypothesized, with extreme weather similarities possibly in the background. However, a more definite claim of this statement needs further study. Our new estimate based on the Skiadas model can be used for

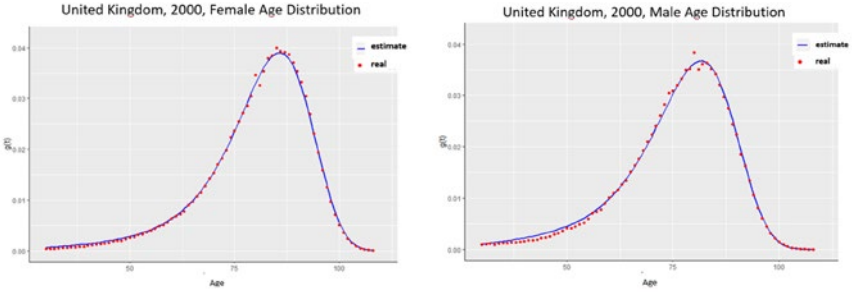


Figure 4: Mortality distribution estimation for males and females in the United Kingdom in 2000.

distribution estimation in mortality tables. As an illustration, in Figure 4 we present the excellent fit of a density function to UK male and female mortality data in the year 2000.

Supposing available mortality data for a set of consecutive calendar years and ages, the model parameters b and l do not change with the years, unlike c. Thus, the parameters c_i characterize the effects specific to the i-th calendar year. For example, if there was a war or a serious epidemic, we would expect a lower value for c_i. The drop of the c parameter signifying spectacularly the Spanish flu's effect in Sweden just demonstrates that.

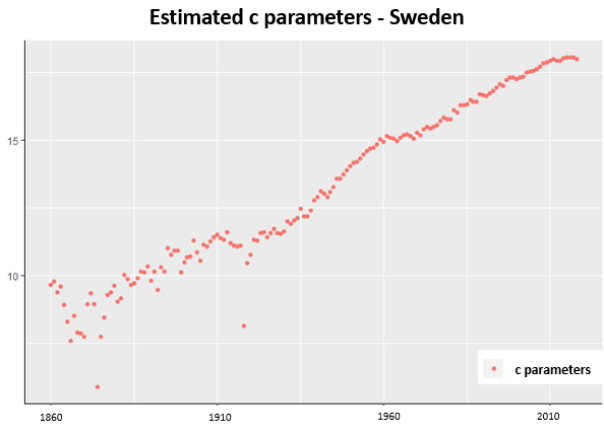


Figure 5: The estimated c_i parameter values for the Swedish mortality data. A great drop in the year 1918 signifies the outbreak of the Spanish flu epidemic.

5 Solution of the industrial problem and its benefits

The project lasted only for several months; nevertheless, the industrial partner followed it closely by having one of its employees committed to it. The primary goal set for this short period of time has been achieved, and the industrial partner has acknowledged that. The research topic fits the company's broader research objectives, and in their opinion, the current results can serve as a starting point for a new mortality forecasting methodology. We have serious indications for a continuing collaboration in the field. One of its scopes can be the comparison of the new forecasting method with existing ones. Beyond this, the publication of new mortality data is expected in the coming months in several other countries, and a much more detailed European scale comparison of the COVID effect on mortality in various countries is of paramount interest to the company.

6 Conclusions

We proposed a simple but new measure of annual mortality changes which can clearly identify significant external factors (epidemics, weather). In our view, the actual impact of epidemics is better measured by the total change in mortality than by the number of deaths caused by the disease. We used this method to measure the impact of COVID-19 on the case of a few countries for which data are already available.

By developing the original model of Skiadas ([2],[3]), we showed an expansion to model the evolution of mortality over several consecutive calendar years, taking into account the specific effects of each cohort. Since the number of parameters increased in the multi-year model, an MCMC-like stochastic algorithm was developed and used for parameter estimation. The effects of wars and pandemics during the 20th century are also perceived as changing the related parameters. In the project, a possible

forecasting method using ARIMA models is elaborated, and the future mortality is predicted for Swedish, American, and British populations. However, this forecast is a draft first cast version, and its plausibility is yet to be tested. Hence we do not present it in this report, and its refinement and testing is a topic for future research.

References

[1] J. Janssen, C. H. Skiadas, *Dynamic modelling of life table data*, Applied Stochastic Models and Data Analysis,Volume 11,pp.35-49, 1995.

[2] C.H. Skiadas, C. Skiadas, *Exploring the health state of a population by dynamic modeling methods*, The Springer Series on Demographic Methods and Population Analysis, 2018.

[3] C.H. Skiadas, C. Skiadas, *Exploring the state of a stochastic system via stochastic Simu- lations: An interesting inversion problem and the health state function*, Springer Science, 2014.

[4] P. Hegel, *Estimation of mortality by using the health state function* (in Hungarian), submitted

Network epidemic modelling of COVID-19 in Hungary

Ágnes Backhausz[1], **Edit Bognár**[2] and **Péter Simon**[3]

[123] ELTE Eötvös Loránd University, Budapest, Hungary, Institute of Mathematics,

[1]E-mail: agnes.backhausz@ttk.elte.hu
[2]E-mail: bognaredith@student.elte.hu
[3]E-mail: peter.simon@ttk.elte.hu

1 Executive Summary

The two central questions of epidemic modelling is to create a good mathematical model and then tune its parameters in order to recover available infection data and to predict the course of the epidemic. This can help decision makers to introduce the necessary measures. The most frequently used propagation models are the so-called compartment models that divide the society into groups according to their phase of infection and age. The estimation of parameters in these models has been widely studied. More accurate mathematical models can be developed by using network theory. In this project, network epidemic models are considered and their main parameters, the infection and recovery rates are determined from simulation data. It is shown that the parameters can be accurately estimated, and hence valuable predictions can be made, when detailed information is available about the infectiousness of the households. It is also studied how the quality of the prediction declines as the available data is becoming weaker. The research is carried out in a simple network situation that is to be continued for more complex and realistic networks. Nevertheless, it sheds light to the importance of detailed infection data in accurate prediction of future scenarios.

2 Challenge overview

Considerable amount of infection data has been collected during COVID-19 epidemic, challenging researchers to explain the data and to predict possible outcomes. Several mathematical models have been developed since the pioneering works of the first modelers [1, 5] and applied in the present pandemic. Network science has also significant impact on understanding disease propagation by introducing several network epidemic models, see e.g. [6]. Network models have different parameter sets, on one hand network parameters, like the average degree and on the other hand epidemic parameters, like the infection rates along different edges and recovery rates at different nodes. Most studies known in the network epidemic literature focus on the dynamical properties of the propagation process on networks with given parameters. Relatively few results are available about estimating the parameters from data. Hence even a simple network structure is interesting from the parameter estimation point of view. This project aims at computing the infection and recovery rates in a two-level random network, in which there are households with stronger inner connections and weaker

ones between them. The model is motivated by the deterministic model studied in [4]. The details of a mathematical method for finding the parameter values are worked out and the corresponding algorithm is implemented in MATLAB. It is shown that the more is known about the number of susceptible, infected and recovered persons in households the better is the estimation of the parameters.

3 Implementation of the project

The study presented in this document is motivated by the Hungarian governmental partner, the Ministry of Innovation and Technology. The Ministry has an epidemiological modelling research group helping the government in decision making concerning epidemiological measures. Our research is related to the work of this modelling group by extending their efforts towards the direction of network epidemic modelling. The task of our group was to develop a mathematical model that is simple enough to derive mathematical methods to determine its parameters from data and is more realistic than a simple compartmental model. The tasks carried out by our group were as follows. We chose a network model forming a compromise between complexity and reality. Implemented an individual-based stochastic simulation algorithm that has been developed before to handle simulations on large networks. Developed a mathematical method for estimating the infection and recovery rates in network propagation models. Finally, we carried out extensive numerical studies on these networks to answer the questions relevant to our governmental partner.

We thank Balázs Boros for giving consent to the use of his MATLAB scripts for SIR processes on certain random graphs, which became the starting point of the simulations.

4 The research problem, methods, and results

Our general goal is the estimation of the paramaters describing epidemic spread in multilayer random graph models. These models separate different layers of the society, and consists of groups of people living in the same household, working in the same place, going to the same school or live close to each other and have some change to meet and transfer contagion. During the current project, we studied an SIR (susceptible, infected, recovered) process on a two-layer model (household model) with an underlying random network, by using computer simulations. The goal was to understand what kind of epidemic statistics (e.g. the number of infected people, the number of household with exactly two infected people) are sufficient for the estimation of the parameters, and hence for the forecasting of the epidemic curve.

4.1 The research problem: estimation of contagion parameters in a random-graph based household model

Our model has two main components: an underlying weighted graph, and a random process modelling epidemic spread on this network. For the latter, we implemented the commonly used SIR model. In the basic setup, the graph consists of N vertices

(individuals), and everyone lives in a household of size N_{hh}. We also studied cases when the household sizes are chosen according to the corresponding Hungarian data [3], but for the results presented here, we used a fixed number. Individuals living in the same household are connected to each other with an edge of weight 1. In addition, every pair of individuals living in different households are connected independently with an edge of weight w with a fixed probability, such that the average number of neighbors of a vertex is d. That is, households are modelled with smaller complete graphs, while other connections are modelled with an Erdős–Rényi graph of average degree d (in another version, we used a preferential attachment random graph). As for the epidemic spread, let τ be the infection rate and γ be the recovery rate in the SIR model. Since we have a weighted graph, this means that the infection rate along an edge within a household is τ, while it is $w \cdot \tau$ along an edge going between different households.

In this setup, we focused on the following question. Suppose that we know N, N_{hh}, and also w. What kind of epidemic statistics are sufficient to estimate the unknown values of τ and γ? In addition, we also made some experiments for examining other cases, for example, estimating w when the infection rate within a household is known, or regarding the size of households as a variable [3].

4.2 The applied methods

Our starting point was the paper [7], which provides maximum likelihood and bayesian estimation of the SIR parameters in the case of a complete graph with edge weights c/N, where N is the number of vertices. This can be considered as a deterministic version of the Erdős–Rényi graph, where the average degree is fixed (e.g. $d = 6$), and every pair is connected with probability $d/(N - N_{hh})$. Thus, the first task was to find an analogue of the estimate (we used the simpler maximum likelihood version), which gives good results for the random graph-based household model. We found that the number of SI edges (the edges having an endpoint in state S and another in state I) has a key role here, but this is not always easy to estimate in reality, and mean-field approximation is not very good, due to the inhomogeneous structure of the random graph. Hence we needed new ideas to estimate this quantity.

To describe our methods more precisely, we need some more notation. Let S_t, I_t, R_t be the number of susceptible, infected and recovered individuals at time t. We have $N = S_t + I_t + R_t$ for every $t \geq 0$. We will use time span $[0, T]$, with an appropriately chosen T. Let z_I and z_R be the total number of infection and recovery events until T. Let E_t^{SI} be number of SI edges at time t.

Notice that in our model the average number of people infected by a single individual is roughly

$$R_0 = \frac{(wd + (N_{hh} - 1))\tau}{\gamma}.$$

We have chosen the parameters such that the value of R_0 is between 1.15 and 2.4, which is a realistic choice for periods of growth of the SARS-COV-2 virus (for the outbreak in the spring of 2020 in Italy, the estimations are even higher, stating that R_0 was between 2.43 and 3.1 [2]).

For the modelling of the epidemics spread we implemented the Gillespie and the event-driven fast SIR Algorithm for weighted garphs according to [6]. Both of them simulate the same stochastic process in continuous time, however computationally Gillespie is more time consuming. Gillespie algorithm firstly calculates the time to the next event based on the number of infected and susceptible individuals, then calculates wether the event will be an infection or a recovery. Since it is important to know the exact node changing its status, a random node is drawn uniformly from the infectious ones in case of recovery, while a susceptible-infectious pair is chosen proportionally to the weight of the edge connecting them in case of an infection. The process jumps to the next time, updates the statuses then repeats until there is no infectious node. The event-driven algorithm is based on the idea that when a node gets infected, nothing can affect when it recovers, who it transmits to or the timing of those transmissions. Therefore, when a node becomes infected, we calculate its possible transmission(s) and recovery times, and put them into a priority queue ordered by the time of the events.

4.2.1 Estimation of the recovery rate

In an SIR model, the recovery rate is easy to estimate, as the graph structure does not have any effect here. The processes describing the number of infected and recovered individuals provide sufficient information. We simulated the continuous-time process till the end of the epidemics (no infectious node in I) on an interval $[0, T]$, where T is different for each trajectory. Gillespie and fast SIR algorithm as well provide data on a non-equidistant mesh $0 = t_0 < t_1 < \ldots < t_K = T$, where t_js are the times when an event occours (e.g. a node gets infected of recovered). Therefore S_t, I_t and R_t are step functions. Hence, we could use the common estimate for γ which was derived in [7] for example, and the integral of the step function became a sum, which was easy to calculate from the trajectories:

$$\hat{\gamma} = \frac{z_R}{\int_0^{\overline{T}} I_t dt} = \frac{z_R}{\sum_{t_i \leq \overline{T}} I_{t_i} \cdot (t_i - t_{i-1})}.$$

Here z_R denotes the total number of events when an individual recovered in $[0, \overline{T}]$. Time $\overline{T}$ is our choice for the length of the period which was used for the estimate. Usually this was a function of T, the total time until the epidemic dies out, e.g. $\overline{T} = (5/6)T$.

4.2.2 Estimation of the infection rate using the SI statistics

In [7], the authors proposed the following estimate for the infection rate between two individuals, with z_I denoting the total number of events when an individual gets infected:

$$\frac{z_I}{\int_0^T S(t)I(t)\, dt}.$$

In their deterministic, homogeneous model the product $S(t)I(t)$ in the formula corresponds to the number of possible contacts between infectious and susceptible individuals. Therefore, by using the weighted sum of $SI(t)$ edges (the edges having an

endpoint in state S and another in state I) instead of $S(t)I(t)$, we get the generalization of the formula to graphs. (In a complete graph equation $S(t)I(t) = SI(t)$ holds exactly.) Since the model in [7] can be implemented as a complete graph with edge weights of $\frac{1}{N}$, similarly as before we used the following estimation to the infection rate:

$$\hat{\tau} = \frac{z_I}{\int_0^{\overline{T}} E_t^{SI} dt} = \frac{z_I}{\sum_{t_i \leq \overline{T}} E_{t_i}^{SI} \cdot (t_i - t_{i-1})}.$$

However, measuring and collecting data for SI in a real-life network is an impossible challenge. In addition, due to random structure of the underlying graph, mean-field approximation $E_t^{SI} \approx S(t)I(t)$ did not work well. Therefore in the following we try to find and utilize more accessible data in terms of the virus spread for the estimation of $SI(t)$ values. We examined the problem broken down into two subparts: the estimaton of E_t^{SI} edges between and within households ($\hat{E}_t^{SI,o}$ and $\hat{E}_t^{SI,h}$).

4.2.3 Estimation of the number of SI edges between households

The next step is to estimate the number of edges connecting a susceptible and an infected individual who live in different households. The simplest mean-field model would be the following:

$$I_t \cdot d \cdot \frac{S_t}{N},$$

as every infected individual has d neighbors from other households on average, and a randomly chosen individual is susceptible with probability S_t/N. However, this estimate was not accurate, due to the fact that some people get infected through a connection outside their households, and they only have $d-1$ neighbors on average who can be in state S. We approximated the probability of getting infected from outside by simply the proportion of neighbors from different households among all the neighbors (although our simulations presented in [3] show that this quantity significanlty varies over time). Hence we used the following estimate:

$$\hat{E}_t^{SI,o} = I_t \cdot \left(d - \frac{wd}{wd + N_{\text{hh}} - 1} \right) \cdot \frac{S_t}{N}.$$

4.2.4 Estimation of the number of SI edges within households

The estimation of the number of SI edges within households turned out to be harder to handle. Our idea was to look at households as they were separate units, and can be in state S^* (if every member is in state S), or in state I^* (if at least one member is infected), or in state R^* (otherwise). Since typically infection reaches a household only once, this can be approximated with an SIR process. We looked for the estimation in the following form

$$\hat{E}_t^{SI,h} = \frac{I_t}{I_t^*} \cdot \left(N_{hh} - (1 + f(S_t^*, I_t^*)) \cdot \frac{I_t}{I_t^*} \right) \cdot \frac{N}{N_{\text{hh}}}.$$

Here S_t^* and I_t^* denote the number of households in state S^* and I^*, and f is an appropriately chosen function. The first factor is the average number of infected

individuals in a household, the second factor is an estimate for the average number of susceptible members in a household. Since using R^* was not successful in our simulations, instead we looked for a function f, which, based on the current intensity of the epidemic, estimates the average proportion of household member in state R and I. With this formula, for certain values of parameters, we could achieve some partially good estimates for the growth period, but we did not succeed in finding a general f which works in a more general setup and throughout the whole process.

Notice that if we exactly know the number of households with a given S, I, R configuration (i.e. one S, two I, one R), then the number of SI edges within a household can be calculated easily. That is, until we do not have a better estimate, these data are necessary for parameter estimation in our model.

4.3 The achieved research results

For the figures presented in this section we run the Gillespie or the event-driven fast SIR process on underlying random graphs of $N = 10000$ nodes. The size of every houshold is fixed $N_{\text{hh}} = 4$, while connections between different household are modelled with Erdős–Rényi graph of average degree $d = 6$ with weight $w = 0.4$ on the edges. At the beginning of the process 1% of the individuals are infected, which is chosen random and uniformly. The scatter plots in Figure 1 and 2 proves that the use of formulas in Section 4.2.1 and 4.2.2 for estimating γ and τ with the weighted sum of SI edges are appropiate in case of our weighted graphs (modelling non-household connections with Erdős–Rényi and Preferential attachment graph as well). In the figures below we can see the convergence of the estimation: each blue line at t shows the estimation of γ and τ on time interval $[0, t]$ based on a trajectory. By increasing the length of the interval, the esimation becomes more stable and converges to the real parameter. In case of τ we overestimate the parameter in the beginning of the process. However, this is not a model or estimation related characteristic, and can only be derived from the unrealistic start of the process (i.e. 1% of the individuals get infected at $t = 0$).

In Figure 3 scatter plot of real and estimated parameters is displayed when we use mean-field assumptions presented in Section 4.2.3 for the estimation of SI edges outside households. Since the mean-field approximation the estimated values does not converge to the real value, we used a different $\overline{T}$ as before ($\overline{T} = \frac{T}{4}$). The quality of the estimation strongly depends on the length of the interval we use. Furthermore, with parameter settings resulting in R_0 close to the critical value of 1, the estimations are less stable since outliers can occur.

5 Solution of the industrial problem and its benefits

Based on our analysis of the random-graph based household model, we could determine the epidemic statistics which are sufficient to provide good estimations for the unknown parameters. These are (1) the number of susceptible, infected, recovered people; (2) the estimated proportion of the infection rate within and between households; (3) the average number of connections of an individual to people in other

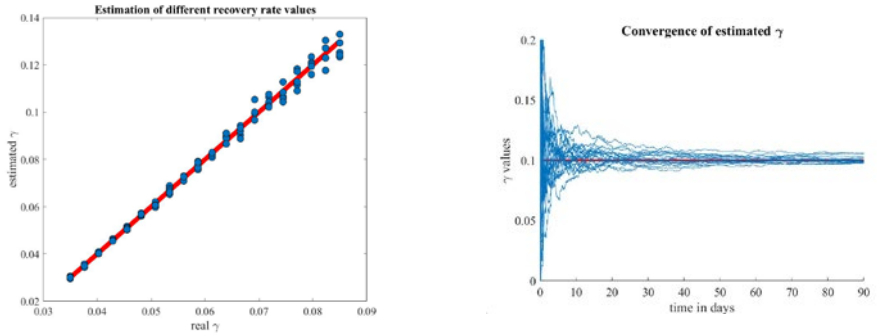


Figure 1. Scatter plot of real and estimated γ values (left) and the convergence of the estimation as the length of the observed period increases (right)

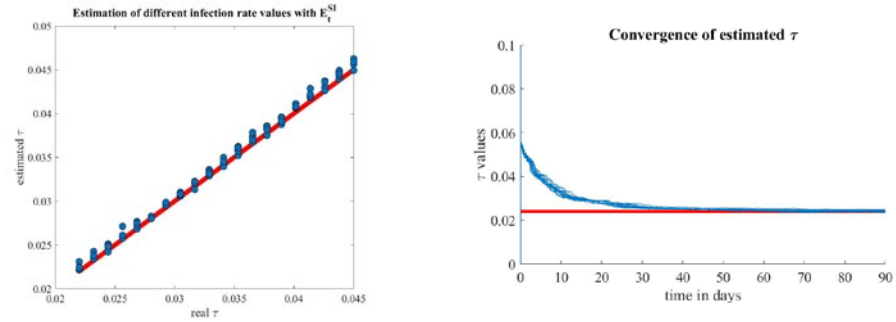


Figure 2. Scatter plot of real and estimated τ values by using the exact number of SI edges (left) and the convergence of the estimation as the length of the observed period increases (right)

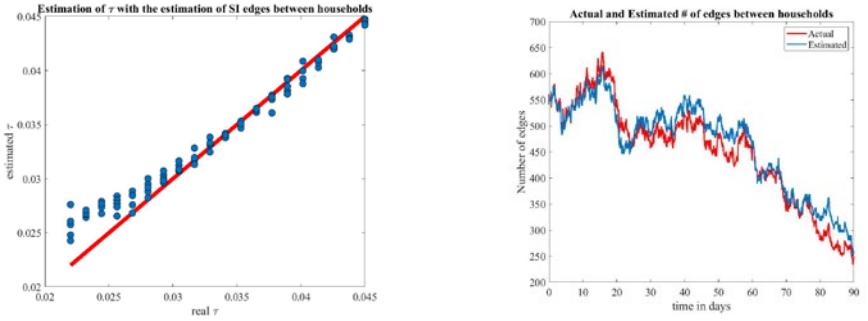


Figure 3. Scatter plot of real and estimated τ values using only the number of SI edges within households, estimating the number of SI edges between households (left); here R_0 goes from 1.15 to 2.4. On the right we compare the actual number of SI edges between households, and the estimated value on a typical trajectory.

households; (4) the number of households with a given S, I, R configuration (i.e. one S, two I, one R).

This does not mean that these numbers have to be exactly known, some kind of sampling can be good enough. For example, a study like H-UNCOVER [8] from a sample with approximately 10000 participants gives good estimates for (1). If a similar study can be performed such that complete households are chosen and tested, not only individuals, then (4) can also be estimated. However, these have to be monitored regularly, as we needed the whole process to provide good estimates. As for (2), the number of connections can be estimated from surveys like MASZK (`https://covid.sed.hu/tabs/response`). The average proportion of infection rate between and within households might be estimated if we keep track of the proportion of people who get infected at home, which also assumes that not only the total number of infected people is available, but we also have information on the number of people who got ill when there was at least one infected person in their household.

As for the currently available data in Hungary on COVID-19 (`https://koronavirus.gov.hu/`), our calculations showed that it is already not easy to estimate the recovery rate. The estimate presented in Section 4.2.1 lead to $\hat{\gamma} \approx 0,01$ (from September to November 2020), which would mean that the average time of recovery is 100 days – this is far from being realistic. Hence we conclude that in order to make our parameter estimation scheme work, it would be necessary to keep track of the state of individuals more precisely, and to access more data on the status of the households, regarding the number of people in state S, I or R.

6 Conclusions

Our study shows that in a two-level, random graph-based network model the key step for parameter estimation is understanding the number of SI edges. As for applications, this means that it is not only the number of individuals in states S, I, R is necessary, but we also need information of the number of households with a given infection configuration, e.g. consisting of one susceptible, two infected and one recovered members.

Further steps in this directions would be (1) to introduce the latent period (2-10 days for COVID-19), and use an SEIR model; (2) to use currently available data to replace one or more of the above statistics; (3) to make the model more realistic, and include groups at workplaces, schools etc., and understand how these ideas can be applied in the more general case. These extensions of our study could lead to better methods for parameter estimations and forecasting.

References

[1] Anderson R.M. & May, R.M, 1991, Infectious diseases of humans: dynamics and control, Oxford University Press.

[2] M. D'Arienzo, A. Coniglio, *Assessment of the SARS-CoV-2 basic reproduction number, R_0, based on the early phase of COVID-19 outbreak in Italy*, Biosafety and Health, Volume 2 (2020), Issue 2, Pages 57-59.

[3] Á. Backhausz, E. Bognár, *Járványterjedés paramétereinek becslése többrétegű véletlengráf-modellekben*. Kézirat. Tervezett publikáció: Alkalmazott Matematikai Lapok.

[4] F. Ball, D. Mollison, G. Scalia-Tomba, *Epidemics with two levels of mixing*. Ann. Appl. Probab. 7 (1997) no. 1., 46-89.

[5] Diekmann, O. & Heesterbeek, J.A.P, 2000, Mathematical epidemiology of infectious diseases: model building, analysis and interpretation. John Wiley & Sons Ltd, Chichester, UK.

[6] Kiss, I. Z, Miller, J. C., Simon, P. L., *Mathematics of Epidemics on Networks; From Exact to Approximate Models*, Springer, 2017.

[7] W. R. KhudaBukhsh, B. Choi, E. Kenah, G. A. Rempała, Survival dynamical systems: individual-level survival analysis from population-level epidemic models. Interface Focus. Published online: December 2019. https://doi.org/10.1098/rsfs.2019.0048

[8] B. Merkely, A. J. Szabó, A, Kosztin et al., Novel coronavirus epidemic in the Hungarian population, a cross-sectional nationwide survey to support the exit policy in Hungary. GeroScience 42, 1063-1074 (2020).

A reliable algorithm to determine adversary example free zones for artificial neural networks

Nándor Balogh[1], **Balázs Bánhelyi**[2], **Tibor Csendes**[3], **István Megyeri**[4], **Richárd Tóth**[5], and **Dániel Zombori**[6]

[1]Redink Ltd., Szeged,
[23456]University of Szeged, Institute of Informatics

[1]E-mail: bn@redink.hu
[2]E-mail: banhelyi@inf.u-szeged.hu
[3]E-mail: csendes@inf.u-szeged.hu
[4]E-mail: imegyeri@inf.u-szeged.hu
[5]E-mail:
[6]E-mail: zomborid@inf.u-szeged.hu

1 Executive Summary

Automatic licence plate recognition systems are based on artificial neural networks. Surprisingly small data perturbations, the so-called adversarial examples can result in false classification. This research gives a reliable verification algorithm for realistic size image recognition problems.

One of the hottest topics in present artificial intelligence research is to understand the phenomenon of adversarial examples for machine learning techniques applying artificial neural networks [8, 16]. The typical such image classification problem is the following. After the proper training of the network, there exist pictures surprisingly similar to the positive sample images that result in a wrong denial decision. As an illustration of the problem, see some real life images of car licence plates on Figure 1, that could not be detected correctly.

Figure 1: Real life examples with car licence plate number classification problems.

2 Challenge overview

We have more and more artificial neural network based automated recognition systems. Important application fields both in transport and health care In recent years more and more problems have been reported: adversarial examples can be given and the well trained NN systems give bad classification We aimed to give an interval arithmetic based reliable algorithm that is capable to determine large adversary example free Zones.

Our aim was to check the limitation of interval arithmetic based reliable verification algorithm. By applying interval arithmetic, we can determine the level of noise that will not change a correct classification.

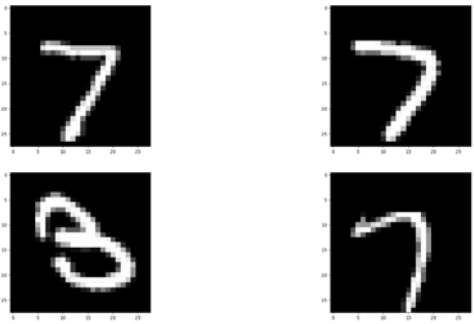


Figure 2: Handwritten number classification test problem instances from the MNIST test set.

3 Implementation of the project

The project team was led by the Redink Ltd. of Szeged (Nándor Balogh). The main collaborator in this project was the Department of Computational Optimization at the University of Szeged (Balázs Bánhelyi, Tibor Csendes, István Megyeri, Richárd Tóth, and Dániel Zombori). The industrial partner provided its long-term experience in image processing, picture recognition and learning algorithms.

4 The research problem, methods, and results

There are already many available shocking results regarding adversarial examples for artificial neural networks (see e.g. those in [9, 14, 19]). Also, many approximate procedures are suggested for e.g. locating the nearest adversarial example to a given correctly accepted image. On the other hand, we do not know about existing verified implemented techniques being capable of providing adversarial example free zones. Obviously there are approaches in this direction [7, 11, 17, 18]. This latter feature is important for mathematically correct statements, especially on a field, where the expected behavior of a computational method differs sometimes from the anticipated one. Interval arithmetic based verified numerical calculations are the proper tool for handling both rounding errors and their consequences, and also for proving statements on positive measure sets of high dimension. We applied interval methods to prove that the damped forced pendulum is chaotic [2], we proved most of the Wright conjecture on a delayed differential equation [3], and verified new optimal circle packing instances [15].

4.1 The research problem

Using adversarial examples generated by existing attack algorithms like those in [12, 4, 14], an offending driver can easily prevent the system to identify him. Even in black box cases, i.e. when the attacker has no access to model parameters, the attack can be successful [13]. This makes it difficult to apply these state-of-the-art techniques in

any safety critical settings. One might naively use attack algorithms for evaluating robustness. However, as show in [4], robustness against some attacks does not mean that the network is robust. Later, stronger attacks may be developed, which will be able to fool the network. A certified evaluation may end this arms race. Further, it motivates to develop reliable methods for evaluating neural networks.

4.2 The applied methods and some results

There are already many available shocking results regarding adversarial examples for artificial neural networks (see e.g. those in [9, 14, 19]). Also, many approximate procedures are suggested for e.g. locating the nearest adversarial example to a given correctly accepted image. On the other hand, we do not know about existing verified implemented techniques being capable of providing adversarial example free zones. Obviously there are approaches in this direction [7, 11, 17, 18].

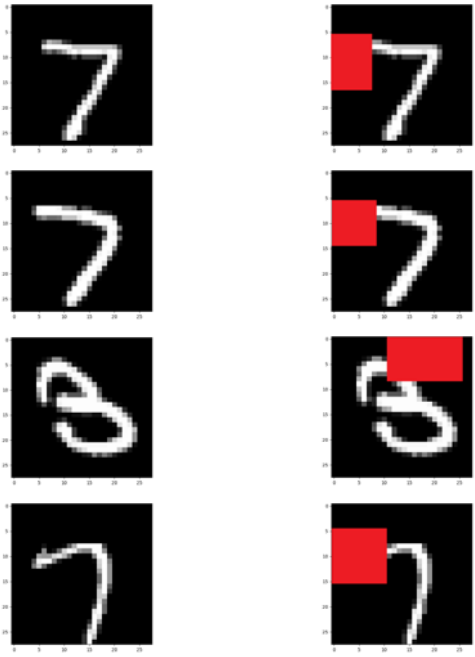


Figure 3: Original pictures and proven rectangles where we can change *everything* without having an adversarial example.

This latter feature is important for mathematically correct statements, especially on a field, where the expected behavior of a computational method differs sometimes from the anticipated one. Interval arithmetic based verified numerical calculations are the proper tool for handling both rounding errors and their consequences, and also for proving statements on positive measure sets of high dimension. We applied interval methods to prove that the damped forced pendulum is chaotic [2], we proved

most of the Wright conjecture on a delayed differential equation [3], and verified new optimal circle packing instances [15].

The simplified pseudo code of the algorithm is:

0. If $F(p0) > 0.5$ then $greater = true$, otherwise $greater = false$
1. Iterate until $percent <= 100$
2. Let P be an n dimensional interval
3. For $i = 1$ to n do
   (a) If $p_i = 0$, then $P_i = [0, 2 * percent/100]$
   (b) Otherwise, if $p_i = 1$, then $P_i = [1 - 2 * percent/100, 1]$
   (c) Otherwise $P_i = [p_i - percent/100, p_i + percent/100]$, and check the end points: if the lower one is negative, then set it to zero, if the upper one is larger than 1, then set it to 1.
4. If $greater = true$ and $F(P) \geq 0.5$, or $greater = false$ and $F(P) < 0.5$ then do:
   (a) If $percent < 1$, then $maxpercent = percent$, and break the main cycle, Stop.
   (b) Otherwise $maxpercent = percent$, and $percent = percent + 1$
5. Otherwise if $percent < 1$, then set $percent = percent - 0.1$
   (a) If now $percent = 0$, then set $maxpercent = 0$ and STOP
   (b) Otherwise break the outer loop
6. End of the cycle started in the first step

5 Main benefits

We have designed, implemented and tested and interval arithmetic based reliable verification algorithm that is capable to check simple artificial neural networks and realistic size images. The involved firm, Redink Ltd. can improve its licence plate recognition application, and provide more reliable solutions for the costumers.

After the completion of the project, we summarized our results in the paper of [6]. Then focused on the adversarial example problem, and could achieve nice some nice results fooling some complete neural network verifier algorithms [20].

6 Conclusions

We are still in the phase when we explore the capabilities of interval arithmetic based algorithms for describing the sensitivity of trained natural neural networks to changes in object to be classified, but we find our present results encouraging enough to continue our research project. The next issue can be the frightening case of small

patches changing the recognized meaning of traffic signs. On the other hand, we could also be capable of proving where the next adversarial example is relative to a given image.

Acknowledgements. This research was supported by the project "Extending the activities of the HU-MATHS-IN Hungarian Industrial and Innovation Mathematical Service Network" EFOP3.6.2-16-2017-00015, and that of 2018-1.3.1-VKE-2018-00033.

References

[1] Ba, J. and Kingma, D., Adam, A Method for Stochastic Optimization. 3rd Intl. Conf. on Learning Representations (ICLR), 2015, http://arxiv.org/abs/1412.6980

[2] Bánhelyi, B., Csendes, T., Garay, B.M., and Hatvani, L., A computer-assisted proof for Sigma_3-chaos in the forced damped pendulum equation. *SIAM J. on Applied Dynamical Systems* Vol. 7. (2008), 843-867.

[3] Bánhelyi, B., Csendes, T., Krisztin, T., and Neumaier, A., Global attractivity of the zero solution for Wright's equation. *SIAM J. on Applied Dynamical Systems* Vol. 13 (2014), 537-563.

[4] Carlini, N. and Wagner, D.A., Towards Evaluating the Robustness of Neural Networks. *IEEE Symposium on Security and Privacy*, SP 2017, San Jose

[5] Csendes, T., An interval method for bounding level sets of parameter estimation problems, *Computing* Vol. 41 (1989), 75-86.

[6] Csendes, T., Balogh, N., Bánhelyi, B., Zombori, D., Tóth, R., and Megyeri, I., Adversarial Example Free Zones for Specific Inputs and Neural Networks. ICAI Proceedings, 2020.

[7] Fazlyab, M., Morari, M., and Pappas, G.J., Safety Verification and Robustness Analysis of Neural Networks via Quadratic Constraints and Semidefinite Programming. *arXiv:*1903.01287v1.

[8] Goodfellow, I., Shlens, J., and Christian Szegedy, Explaining and Harnessing Adversarial Examples. *International Conference on Learning Representations*, 2015

[9] Ilyas, A., Santurkar, S., Tsipras, D., Engstrom, L., Tran, B., and Madry, A., Adversarial Examples Are Not Bugs, They Are Features. *arXiv:*1905.02175.

[10] Lecun, Y., Bottou, L., Bengio, Y., and Haffner, P., Gradient-Based Learning Applied to Document Recognition, *Proc. of the IEEE*, 86 (1998) 2278-2324.

[11] Lin, W., Yang, Z., Chen, X., Zhao, Q., Li, X., Liu, Z., and He, J., Robustness Verification of Classification Deep Neural Networks via Linear Programming. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), IEEE Explore, 2019, 11418-11427, DOI: 10.1109/CVPR.2019.01168.

[12] Moosavi-Dezfooli, S.M., Fawzi, A., and Frossard, P., DeepFool: A Simple and Accurate Method to Fool Deep Neural Networks. *The IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2016, 2574-2582.

[13] N. Narodytska and S. Kasiviswanathan, Simple Black-Box Adversarial Attacks on Deep Neural Networks. *IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2017

[14] Su, J., Vargas, D.V., and Kouichi, S., One pixel attack for fooling deep neural networks. *arXiv:*1710.08864.

[15] Szabó, P.G., Markót, M.Cs., Csendes, T., Specht, E., Casado, L.G., and García, I., New Approaches to Circle Packing in a Square - With Program Codes, Springer, Berlin, 2007.

[16] Szegedy, C., Zaremba, W., Sutskever, I., Bruna, J., Erhan, D., Goodfellow, I.J., and Fergus, R., Intriguing properties of neural networks. *International Conference on Learning Representations*, 2014

[17] Vincent Tjeng, V., Xiao, K., and Tedrake, R., Evaluating Robustness of Neural Networks with Mixed Integer Programming. *arXiv:*1711.07356v3.

[18] Xiang, W., and Johnson, T.T., Reachability Analysis and Safety Verification for Neural Network Control Systems. *arXiv:*1805.09944v1.

[19] Zaj, M., Zolna, K., Rostamzadeh, N., and Pinheiro, P.O., Adversarial Framing for Image and Video Classification. *The Thirty-Third AAAI Conference on Artificial Intelligence (AAAI-19).*

[20] Zombori, D., Bánhelyi, B., Csendes, T., Megyeri, I., Jelasity, M., Fooling a Complete Neural Network Verifier. *International Conference on Learning Representations (ICLR 2021*, https://openreview.net/forum?id=4IwieFS44l.

Modelling and control of COVID-19 in a bank environment

Ferenc Á. Bartha[1], **Tamás Ferenci**[2], **Gergely Röst**[3], **Péter Wiandt**[4], **Zsuzsanna T. Vőneki**[5], and **Gábor Horváth**[6]

[134] University of Szeged,
[2] Óbuda University,
[5] Corvinus University of Budapest; OTP Bank,
[6] OTP Bank

[1]E-mail: barfer@math.u-szeged.hu
[2]E-mail: ferenci.tamas@nik.uni-obuda.hu
[3]E-mail: rost@math.u-szeged.hu
[4]E-mail: wpeti88@gmail.com
[5]E-mail: zsuzsanna.voneki@uni-corvinus.hu

1 Executive Summary

The COVID-19 pandemic reached Hungary in March 2020. The industrial partner (OTP Bank) has introduced several measures early on, in order to prevent or, if that's not possible, mitigate the spread within the organization. The primary purpose of the collaboration was to analyze the relative effect of such measures, assess their necessity and usefulness from a risk-management perspective and, additionally, identify possible improvements. We carried out the investigation via utilizing Gillespie's Stochastic Simulation Algorithm (SSA) on an empirical network model of the company. In that network, each node represents an employee of the organization, and two employees are linked if they work in the same office or if they meet during their regular work. Constructing such networks required data analytics efforts from the industrial partner that were driven by consultations with the partnering research team. The simulations had to account for the progress of the epidemic in the "environment" so that the risk of infections of external origin are realistically modeled. The dynamics of the environmental effects and the temporal nature of the intervening measures were making it a necessity to carry out the simulations using a customized, temporal version of the Gillespie SSA. The results are justifying the use and demonstrating the benefits of the implementation of intra-company restrictive measures and provide some ideas for further improvements.

2 Challenge overview

The first confirmed COVID-19 cases were reported early March 2020 in Hungary. At that time, the epidemic has readily reached high levels in Italy [1]. The industrial partner (OTP Bank) put in place numerous restrictive measures to mitigate intra-company transmission of the disease. Such measures carry an inevitable cost for any business and, thus, have to be well-justified from a risk-management standpoint so

that they are seen as a reasonable compromise from other management perspectives with conflicting interests.

3 Implementation of the project

The joint efforts of the industrial and academic partners were mainly coordinated by biweekly online meetings. From the industrial side, the project required data analysis efforts to construct a model network of the employee structure of the company and to design the effect and timing of measures modeled. The industrial partner compiled a database capturing the intricacies of the inter-company connection network *e.g.* headcounts per office, geographical locations, inter-office connections. Next, a tentative list of restrictive measures were created with their respective timeline of application. The research group, first, compiled daily incidence statistics of COVID-19 spread in Hungary for the modeling time frame. Then, a temporal Gillespie algorithm was designed with dynamic environmental conditions and intra-network measures. Finally, a statistical analysis was conveyed to create the final results.

4 The research problem, methods, and results

In the following, we describe the research problem in detail in Section 4.1. Then, we discuss what methods were utilized and how they were applied in the problem context in Section 4.2. Finally, we present the results in Section 4.3.

4.1 The research problem: Assessing the efficacy of restrictive measures aimed at mitigating an epidemic on real connectivity networks

The Gillespie contagion algorithm generates statistically correct trajectories for stochastic systems with known reaction rates [2]. Albeit originally designed to model chemical reactions, it is often applied for modeling epidemics on connection networks [3]. A vital aspect of the method is that both the rates describing transitional behaviours and the underlying network structure are assumed to be constant. However, in practice, all of these features may change. For example, introducing preventive measures *e.g.* mask wearing, increased hygiene, separator walls in office environment decrease the likeliness of personal contacts resulting in a new infection. Also, connections within the network may be removed (and reinstated) based on *e.g.* home office policies or testing campaigns aimed at identifying and confining infectious individuals. Such dynamic changes need to be accommodated by the modeling framework so that reliable answers may be obtained.

In our case, the system governing the epidemics under simulation is a stochastic network model tailored to COVID-19, following SEIR-type dynamics. The states of the nodes are from $\{S, E, I, R\}$, representing the *susceptible*, *exposed (latent)*, *infected (infectious)*, and *recovered (removed)* individuals (nodes) in the network. Upon infection, an S individual's state becomes E, then, after the so-called incubation time, turns into I. While being in the infectious state, it may generate new infections

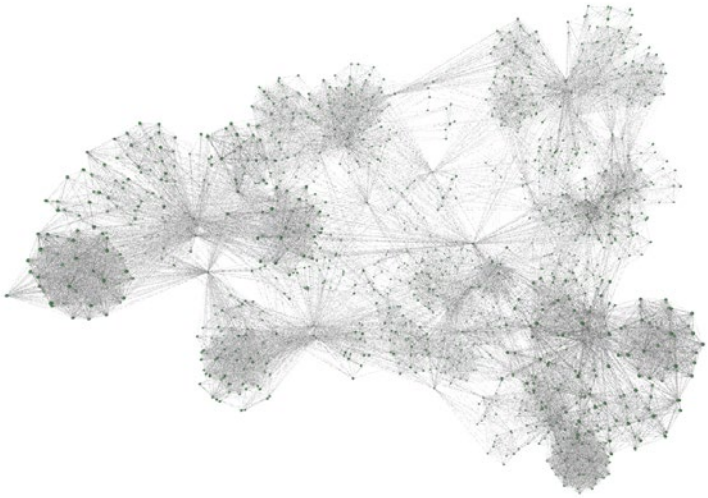


Figure 1: Practical example of a real life connection network.

through its connections to other, still susceptible individuals. The adjacent edges of the connectivity network originating from the node representing the infected individual describe all such potential connections. Finally, after its infectious period, the individual node's state is changing to R, and it will not transmit the disease any more and can not be reinfected. We note that this is a simplistic description of the entire system, most of the aforementioned states are realized in multiple stages and transitions between them, representing disease progression, resulting in a total of 29 possible states of nodes.

Real life networks carry some of the characteristics of the so-called small world graphs [4, 5] on which such epidemics methods are applied in research contexts. Yet, it is both an interesting and challenging task to model an epidemic stochastic process on actual networks describing contact structures; both the magnitude and structure of the graph describing the contacts between nodes carries new problems to tackle and questions to ask. Figure 1 demonstrates an example for a real life connection network. It was derived based on our actual model network with some alterations to protect any proprietary information. In fact, the networks we had to work with are considerably larger and more complex. The example depicted on Figure 1, generated using Graphia [6], is merely similar to some sub-graph of the whole structure.

As we have discussed, even running contagion simulations on large scale networks may be challenging. For our case the primary goal was to assess how preventive measures affect the spread in such simulations. A list of candidate interventions needed to be compiled based on the actions of the industrial partner. In what follows, we present the four measures investigated within our cooperation. Their naming is but tentative, as it also serves as a short, approximate description of their actions.

1. Transmission Reduction: this measure comprises *e.g.* mask wearing policies,

increased hygiene, regular sanitizing of public spaces and equipment, personal distancing, etc. As a result, the probability that an in-person contact results in a new infection is decreased.

2. Home Office: this intervention aims to break edges in the contact network either permanently (w.r.t. the modeling time frame) or temporarily *e.g.* creating multi-week shifts for employees being on-site or off-site.

3. Testing Campaign: regular, periodic testing is used to localize and isolate infectious individuals in the network. Simply put, these nodes may be removed from the network until they recover.

4. Centrality Protection: isolating or protecting central nodes facilitates the breaking of the connection network into smaller components, hence, may be a great asset in mitigating an ongoing epidemic.

Implementing the last measure has an interesting aspect for research. There exist several known centrality metrics that could be applied to a connection network in order to filter the "most central" nodes. Several authors investigated the relationship between centrality and characteristics of an epidemic. Bucur and Holmes related the expected outbreak size to the centralities of the network using 7 different centrality measure [7]. Instead of ranking the nodes, they used an overall measure, and calculated the results for all possible networks with less than 11 nodes. Their results, however, are not directly applicable here as they used a regression, i.e., they used several centrality measures at the same time to predict the final size of the outbreak. Goering et al investigated the relationship between epidemic hitting time (expected time for an epidemic starting in a given to reach another specified node) and centrality using real-life networks instead of simulated ones [8]. They found that – with a single exception – the effective resistance centrality was best correlated with epidemic hitting time. Perhaps the most relevant to our investigation is that of Doostmohammadian et al, who used an SIS epidemic on simulated – scale-free, clustered scale-free and small world – networks using 6 different centrality scores. Outcome metric was the number of nodes needed to be 'cured' so that the reproduction number is reduced below 1 [9]. They found that for scale-free networks betweenness centrality outperforms other measures, but this is not necessarily the case for other network structures.

4.2 The applied methods

As an initial step, we had to reconstruct the daily COVID-19 incidence in Hungary. The officially reported confirmed cases serve as a guideline, but due to the considerable portion of asymptomatic and undetected infections, we had to take into account the unavoidable underreporting. Figure 2 depicts our model incidence data that was used to create a discrete step function like approximation of the probability that an individual in Hungary gets infected during the time frame of our modeling.

As we have discussed in Section 4.1, our framework requires dynamic modeling of changes both in transmission rates and network structure. We have constructed a temporal modification of the Gillespie algorithm described in [10] for our purposes.

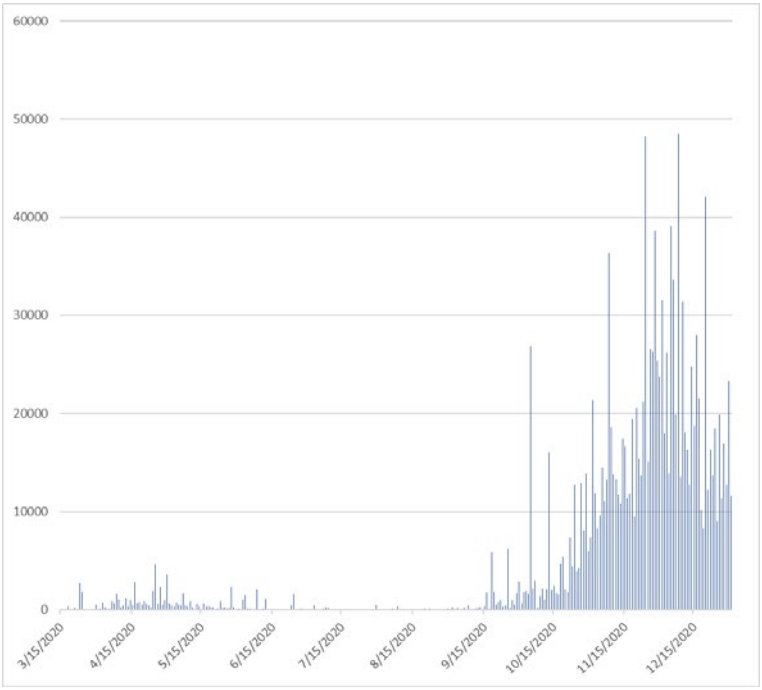


Figure 2: Reconstructed daily COVID-19 incidence for the modeling time frame in Hungary.

In practice, we have custom tailored the Python package, EoN: Epidemics on Networks [11] for our needs as it does not contain a temporal SSA algorithm directly. A considerable effort was spent on creating an efficient temporal implementation that can be adapted to the measures we considered.

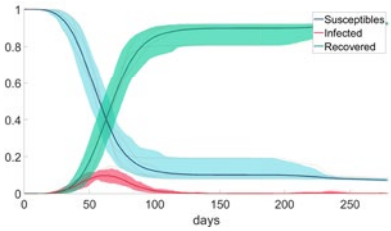
Several centrality measures were calculated using the package `CINNA` version 1.1.54 [12]. Calculations were carried out using the `R` statistical environment version 4.0.4 [13] using the package `igraph` version 1.2.6 for basic graph manipulation [14].

List of nodes to be removed was generated as the the top $\{1, 2, \ldots, 20\}$ percentile from the list of ordered centralities. In case of ties in the centrality values being larger than 1 percent of the number of nodes, some percentiles will be missing. For non-connected networks, the lists were generated independently for all components.

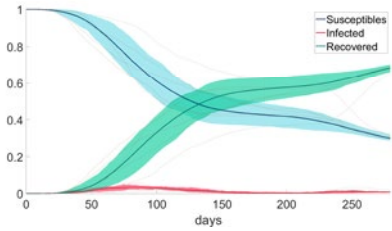
4.3 The achieved research results

We have successfully implemented our modeling framework and environment. The computations were carried out on the large-scale network and enabled us to evaluate and compare some of the measures in various settings. We have carried out the evaluation for Transmission Reduction, Home Office, and Centrality Protection. These measures were investigated both individually and in a combined setting. Figure 3 depicts the dynamics of an envisioned COVID-19 epidemic on the contact network. The

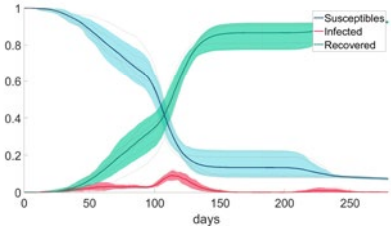
modeling timeline was roughly 290 days from mid March 2020 until mid December 2020. The scenarios considered are: unmitigated epidemic, individual application of measures, combined application of all measures.



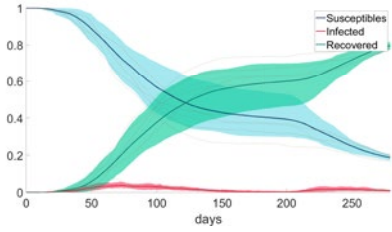
(a) No measures.



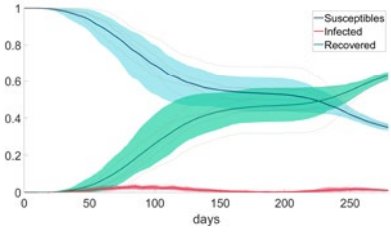
(b) Transmission Reduction.



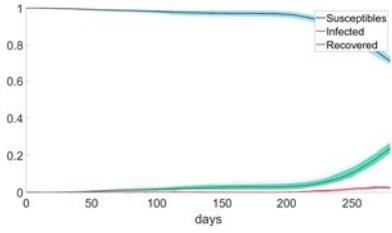
(c) Centrality Protection (short).



(d) Centrality Protection (full).



(e) Home Office.



(f) Combined Measures.

Figure 3: Dynamics of the epidemics in various scenarios with 90% confidence intervals.

The Centrality Protection measure was applied in two variants. First, for a limited time frame in accordance with some aspects of its actual counterpart installed by the industrial partner. Second, for the full length of the modeling. It is remarkable that, even though it removes solely about 2% of nodes from the connection network, even the first variant achieves a definite delay of the outbreak.

Some of the most critical indicators of an epidemic are the peak size and the final size. In our context, they correlate with the peak of employees on sick leave and with how large fraction of the employee network of was affected during the modeling time frame. Our results are summarized by Figure 4.

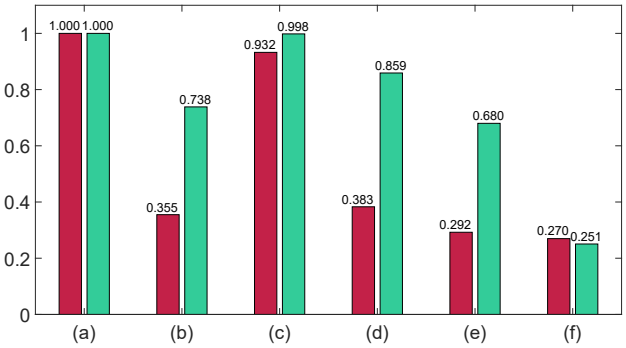


Figure 4: Relative peak size and final size of the epidemic. (a) No measures. (b) Transmission Reduction. (c) Centrality Protection (short). (d) Centrality Protection (full). (e) Home Office. (f) Combined Measures.

5 Solution of the industrial problem and its benefits

The results presented in Section 4.3 serve as a basis for justifying the usefulness and timeliness of the preventive measures put in use by the industrial partner. Our work displays that all measures combined are effective against the ongoing COVID-19 epidemic. Independent application analysis shows that their combination is a must in order to achieve a desired level of mitigation. Comparing our centrality analysis of the industrial network and the actual Centrality Protection carried out by the industrial partner revealed that the company readily identified about 70% of their central employees. Our analysis may serve as a rigorous method for such selections and, hence, as an enhancement of this measure.

6 Conclusions

As a result of this collaboration, a large-scale model of the employee network of the industrial partner was built. On this network, the efficacies of various preventive measures were analysed. Our results serve as a justification for the installment of the interventions from a risk-management perspective. In addition, the analysis results in a potential improvement of epidemic mitigation.

In the near future, we are planning to carry out an in-depth study of centrality protection and its effect on the course of on epidemic on the network. We believe that custom tailored approaches might carry great benefits for mitigation efforts. Another exciting future goal of ours is to analyse testing campaigns, with particular emphasis on identifying an optimal "trigger event" (*e.g.* the number of concurrently hospitalized cases) that would imply an amplification of the measure, ie. increase its frequency.

References

[1] G. Sebastiani, M. Massa, E. Riboli, *Covid-19 epidemic in Italy: evolution, projections and impact of government measures*. Eur. J. Epidemiol., 35, 341-345, 2020.

[2] D.T. Gillespie, *Exact Stochastic Simulation of Coupled Chemical Reactions*. J. Phys. Chem., 81, 25, 2340-2361, 1977.

[3] I.Z. Kiss, J.C. Miller, P.L. Simon, *Mathematics of Epidemics on Network: from exact to approximate models*. Springer, 2017.

[4] R. Albert, A.L. Barabási, *Statistical mechanics of complex networks*. Reviews of Modern Physics, 74 (1): 47-97, 2002.

[5] P. Erdős, A. Rényi, *On Random Graphs. I.* Publicationes Mathematicae, 6: 290-297, 1959.

[6] T.C. Freeman et al., *Graphia: A platform for the graph-based visualisation and analysis of complex data*.

[7] D. Bucur, P. Holme P, *Beyond ranking nodes: Predicting epidemic outbreak sizes by network centralities*. PLoS Comput Biol., 16(7):e1008052, 2020.

[8] M. Goering, N. Albin, P. Poggi-Corradini, C. Scoglio, F.D. Sahneh, *Numerical investigation of metrics for epidemic processes on graphs*. 49th Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, USA, pp. 1317-1322, 2015.

[9] M. Doostmohammadian, H.R. Rabiee, U.A. Khan, *Centrality-based epidemic control in complex social networks*. Soc. Netw. Anal. Min. 10, 32, 2020.

[10] C.L. Vestergaard, M. Génois, *Temporal Gillespie Algorithm: Fast Simulation of Contagion Processes on Time-Varying Networks*. PLOS Computational Biology 11(10): e1004579, 2015.

[11] J.C. Miller et al., *EoN (Epidemics on Networks): a fast, flexible Python package for simulation, analytic approximation, and analysis of epidemics on networks*. Journal of Open Source Software, 4(44), 1731, 2020.

[12] M. Ashtiani, M. Mirzaie, M. Jafari, *CINNA: an R/CRAN package to decipher Central Informative Nodes in Network Analysis*. Bioinformatics, Volume 35, Issue 8, Pages 1436-1437, 2019.

[13] R Core Team,*R: A language and environment for statistical computing*.R Foundation for Statistical Computing: Vienna, Austria, 2020.

[14] G. Csardi, T. Nepusz, *The igraph software package for complex network research*. InterJournal, Complex Systems, 1695, 2006.

Combined vehicle and driver scheduling with fuel consumption and parking constraints: a case study

József Békési[1] and **Albert Nagy**[2]

[1]University of Szeged, Faculty of Science and Informatics, Department of Computer Algorithms and Artificial Intelligence,
[2] Obuda University, Applied Informatics,

[1]E-mail: bekesi@inf.u-szeged.hu
[2]E-mail: albert.nagy@me.com

1 Executive Summary

Efficient operation is an important question for transport companies, and that can be most easily achieved by reducing their operational costs. These costs represent a large part of the expenses. Their most important parts are vehicle fleet costs, fueling and maintenance costs and driver salaries. Consequently, the budget can be improved significantly by decreasing these costs. The most commonly used technique to reduce these costs is the usage of a powerful, computer-aided information system. Due to the ICT (Information and Communication Technology) development today almost every public transportation company has its own information system. We can make such system capable to automatically generate optimized scheduling of vehicles and the work of the drivers. Such an optimization task can be very complex. In this project we studied how to solve the combined vehicle and driver scheduling problem. We defined the applied mathematical model and presented the calculation results for practical examples. Our experiences show that these kind of methods can help the planning process of transport companies.

2 Challenge overview

The problem is to automatically calculate optimal or approximately optimal vehicle and driver schedules for a given list of trips based on the master data and the company specific requirements and parameters in compliance with labor regulations. The optimality is measured by a given objective function and the aim of the optimization is to increase economy and efficiency.

3 Implementation of the project

There were 3 participants of the project. The main industrial partner, who specified the problem was the Budapest Transport Privately Held Corporation (BKV Zrt.). They gave the test data and helped the test process. BKV is owned by the Municipality of Budapest runs most of the vehicles (bus, tram, metro and trolleybus) of the extensive network of public transportation in Budapest. The second industrial partner

was Optin Ltd. from Szeged. Optin Ltd. is an information technology research and development company, which develops innovative technical solutions primarily in the fields of logistics, transportation, medical technology, environmental analysis and IoT, keeping in mind the individual needs of the partners. Their task was to provide the technical background and take part in the software development part of the project. The academic partners were the researchers of Discrete Optimization and Data Mining Laboratory (DOMLAB) at the University of Szeged. Their task was to develop the mathematical model and the necessary algorithms.

4 The research problem, methods, and results

The scheduling problems of public transport are very complex. When looking at the problem from an operational research perspective, we expect a global optimum that minimizes the cost of both vehicle-related tasks and driver scheduling. In the most widely used models today, the vehicle scheduling problem is formulated as an integer multicommodity network flow problem ([2, 6, 7]). In this model, optimal scheduling can be calculated as a solution to an integer linear programming problem. Other models exist to solve this problem. The problem can also be formulated as a set partition problem (see for example [8, 4]). Dávid and Krész proposed a heuristic method[3].

In our experience, the disadvantage of the models discussed and used in the literature in a specific, practical situation is that it only takes into account the rules relating to timetabled and overhead trips, bus types and capacities required for them. However, it is not possible to include specific conditions that come from a real application environment. When planning the operational tasks of public transport, such typical vehicle-specific conditions are fuel consumption rules, various maintenance requirements (weekly, monthly, etc.), and parking rules. Parking rules may apply to both daytime and nighttime parking: where to park, what capacity the parking places have, etc. The length of the parking period may influence where, in which geographical location it may be performed.

4.1 The research problem

The research problem is to develope a combined vehicle and driver scheduling model that can automatically solve real problems and takes into consideration the following characteristics of the schedule planning for city buses:

- the problems are given by packages,
- the trips of a package can belong to a single line or a group of lines,
- the trips of a package can overlap 2 days,
- different day types are possible (e.g. working weekday, feast day, school day, etc.),
- deadhead trips from or to the depots can be possible from each end station,
- the maximum number of depots in a package is 5,

- the maximum number of a vehicles that can be used in the solution of a package is 30,
- the maximum number of vehicle types in a package is 4,
- the type of the vehicle can determine the fuel consumption, which should be taken into consideration in the schedule,
- breaks with stand-off or driver change are possible,
- stand-off or parking is possible on more end stations or depots,
- parking place capacities are given for end stations in 5 minute intervals,
- driver change can be possible on given end stations,
- the labor regulations can be defined by several parameters.

4.2 The applied methods

To solve the problem described in the previous chapter we use a combined vehicle and driver scheduling optimization model. Several versions of this are known in the literature. The mathematical formalization we use is a modified version of the model described in the paper published by Huisman et al.[5].

The most important components of the model are the set of timetabled trips and the available vehicles. Timetabled trips are all trips where vehicles carry passengers. Each such trip is determined by the departure and arrival times, the departure and arrival stations, and the distance between them. The vehicles are divided into (physical) depots based on their location. This may be the garage, parking space or site where the vehicle is parked. Vehicles can also have various important features that allow us to group them further. Based on physical (geographical) locations and features, vehicles are classified into disjoint subsets. The subsets thus formed are called depots. In addition to timetabled trips, vehicles must also carry out other types of drivings. These are called overhead trips. For example, for the first trip of the day the vehicle must leave the night parking lot and return after the last trip of the day. Such deadhead trips can occur during the day, e.g. for longer breaks. Typical additional overhead trips are when a vehicle goes to another station after completing a trip to make another trip from there. We also need to allow these in order to get an effective schedule. For each timetabled trip, the user can specify the depots from which the trip can be served. In practice, this might mean, for example, that certain lines may be served by given type of buses or from given locations. These requirements may be determined by the location of the station and the characteristics of the traffic. We can define certain relationships between timetabled trips. Two trips are said to be compatible if after finishing the first trip the vehicle is able to arrive at the place of departure of the second trip in time. If the first trip arrives at the destination location of the second trip, the only condition is that the first trip arrives earlier as the second trip. If the arrival station of the first trip is not the same as the departure station of the second trip, we must take into account the overhead time between the stations. There may be rules that include mandatory technological time between

two trips, which should be taken into account when examining compatibility. The network is described by a directed graph. The nodes represents the trips, to which the departure and arrival depot nodes are added. Because multiple depots problem is handled, more depot nodes are possible. This corresponds to the usual technique. Two nodes of the graph are connected by a directed edge if the trips representing them are compatible. The edge length always represents the net working time of the driver, which corresponds to the objective function. We also connect the depot nodes to the appropriate trip nodes, which can be accessed from that depot. Huisman et al. classifies the edges into two main groups, namely short and long ones. The short edges always represent the shorter events when the driver remains with the vehicle, while the long edges represent the events when the driver stops the vehicle in the parking space and it remains unattended. We tried to build such a graph, which better represents the real situation we have. So we introduced the following node types:

- source_depot: nodes representing the source depots,
- sink_depot: nodes representing the sink depots,
- trip: nodes representing timetabled trips.

There are 11 types of edges in our model:

- start_of_schedule: edge representing the sign-on event with the first deadhead trip to the location of the first timetabled trip of a schedule,
- end_of_schedule: edge representing the last deadhead trip of a schedule to the location of the depot with the sign-off event,
- short_wait: edge representing a short wait after completing a timetabled trip, when the driver remains with the vehicle,
- short_break_endstation: edge representing a break of the driver spent at an end station after completing a timetabled trip,
- short_break_depot: edge representing a break of the driver spent in a depot after completing a timetabled trip,
- short_driverchange_endstation: edge representing a driver change event at an end station after completing a timetabled trip,
- short_driverchange_depot: edge representing a driver change event in a depot after completing a timetabled trip,
- long_stop_endstation: edge representing a driver change with a long parking of the vehicle at an end station without attendance,
- long_stop_depot: edge representing a driver change with a long parking of the vehicle in a depot without attendance,
- long_dividedstop_endstation: edge representing a long parking of the vehicle at the end station while the driver works in a divided schedule,

- long_dividedstop_depot: edge representing a long parking of the vehicle in the depot while the driver works in a divided schedule,

The main steps of the calculation process are the followings:

1. In the first phase of the calculation, the appropriate input data and parameters are read and the graph is created. Here we take into account the technological and compensatory times, as well as certain rules regarding the duration of the waitings or breaks. These can also be controlled by parameters. This means, for example, that if two trips are too close in time, they will not be connected, which means that they will not be executed one after the other by the same vehicle or driver. The first step in generating a graph is to create nodes based on the timetabled trips, location information, and vehicle types received. Then we add the above listed edges, also taking into account the parameters and labor rules. There may be different types of parallel edges that have different weight values. For example, the driver may take a break at the end station or in a depot, depending on parking possibilities and station capabilities. However, these mean different working times, so if we want a model that reflects the real situation fairly well, we need to include both options. This increases the number of edges, which itself would not necessarily be a problem in the size of the optimization model, but practical experience has shown that the number of possible driver schedules can be critically large in some cases and this must be handled some way.

2. In Phase 2, if possible, we generate all regular driver schedules. There is already a check that takes into account all the rules specified in the specification. Generation is done systematically by traversing the base graph with depth first search. However, every vertex is checked on the fly, and if the subschedule already formed does not conform to any rule, then this branch is cut off. When we get a complete schedule during the generation, a final check runs, which only accepts the schedule if it is found regular. The generated driver schedules are stored. If too many schedules are generated, the process is stopped and a heuristic method is executed, to decrease the complexity of the problem. This method will be described later. Parameters can also be used to limit the number of schedules.

3. In Phase 3, we construct the mathematical model, which is essentially an extended version of the already mentioned VCSP model described in detail in [5]. The model includes constraints on fuel consumption, parking places and manages different vehicle types and locations. More detailed description of the mathematical model can be found in [1].

4.3 The achieved research results

Several real problems were solved by the method. They were part of the tasks that arose during the daily work of the company. The problems were selected by experts to represent the various cases that can happen in practice. The mathematical models were generated by the system and they were solved by an optimization solver.

Problem	Solution status	Running time (sec)	Trip grouper
1	Optimal	8	No
2	Optimal	998	Yes (2)
3	Optimal	138	No
4	Optimal	11364	Yes (3)
5	Optimal	543	No
6	Optimal	5199	No
7	Optimal	49	No
8	Optimal	134	No
9	Optimal	4467	No
10	Optimal	42	No
11	Optimal	27	No
12	Optimal	35	No
13	Optimal	679	No
14	Optimal	1212	No
15	Stopped	6432	No
16	Optimal	17570	Yes (4)
17	Stopped	1506	No
18	Stopped	25170	Yes (4)
19	Stopped	1512	No
20	Stopped	1531	No

Table 1: Properties of the solutions

5 Solution of the industrial problem and its benefits

We can say that this method can automate manual vehicle and driver scheduling in large part. The complexity of the problem highly depends on the characteristics of the input, such as the number of trips, vehicle types and depots, and the average length of trips. In some cases the number of valid driver schedules was too large to solve the original problem. The trip grouper heuristic was able to handle this situation in most of the cases. The running times are very diverse, but 14 of the 20 problems were solved in 30 minutes and 5 of them in 1 minute. These times include all phases, i.e. the graph and the driver schedule generation and the solution of the mathematical program. 12 problems were solved to optimality, without using heuristic method or optimization process stopping.

6 Conclusions

In this project we studied the use of a combined vehicle and driver scheduling model for practical problems. We presented the real problem, the mathematical model and some calculation results. Our experiences show that these kind of methods can help the planning process of transport companies.

References

[1] J. Békési and A. Nagy, Combined Vehicle and Driver Scheduling with Fuel Consumption and Parking Constraints: a Case Study, *Acta Polytechnica Hungarica*, 17, 45-65, 2020.

[2] L. Bodin, B. Golden, A. Assad and M. Ball, Routing and Scheduling of Vehicles and Crews: The State of the Art, *Computers and Operations Research*, 10, 63-211, 1983.

[3] B. Dávid and M. Krész, Application Oriented Variable Fixing Methods for the Multiple Depot Vehicle Scheduling Problem, *Acta Cybernetica*, 21(1), 53-73, 2013.

[4] A. Hadjar, O. Marcotte, and F. Soumis, A Branch-and-Cut Algorithm for the Multiple Depot Vehicle Scheduling Problem, *Tech. Rept.* G-2001-25, Les Cahiers du Gerad, Montreal, 2001.

[5] D. Huisman, R. Freling and A.P.M. Wagelmans, Multiple-depot integrated vehicle and crew scheduling, *Transportation Science*, 39, 491-502, 2005.

[6] N. Kliewer, T. Mellouli and L. Suhl, A time-space network based exact optimization model for multi-depot bus scheduling, *European Journal of Operational Research*, 175, 1616-1627, 2006.

[7] A. Löbel, Optimal Vehicle Scheduling in Public Transit, *Ph.D. thesis*, Technische Universitaet at Berlin, 1997.

[8] C.C. Ribeiro and F. Soumis, A Column Generation Approach to the Multiple-Depot Vehicle Scheduling Problem, *Operations Research*, 42(1), 41-52, 1994.

Application of artificial intelligence techniques for situation analysis based on multimodal spatial data

Csaba Benedek[1,2] and **Balázs Nagy**[1,2] and **Lóránt Kovács**[1,2] and **Örkény Zováthi**[1,2]

[1] University of Debrecen, Faculty of Informatics, 4028 Debrecen, Hungary
[2] Institute of Computer Science and Control (SZTAKI), Kende u. 13-17, Budapest, Hungary, email: {lastname.firstname}@sztaki.hu

1. Executive Summary

This project proposed novel algorithmic solutions for the automated analysis of 3D spatial measurements collected by up-to-date sensors and heterogeneous sensor arrays composed of high-resolution cameras and different types of Lidar laser scanners. Various critical problems of environment perception have been addressed by the new developed methods including object detection, tracking, map-based localization, motion and change detection, which can be directly utilized in self-driving and ADAS applications. In addition, a new method has been introduced for target-less automatic camera-Lidar registration, which enables efficient data fusion of these two very different sensors, without needing a cumbersome manual calibration process. Finally new techniques have been developed for image based urban area analysis and recognition, including brick segmentation, and artifact inpainting in masonry wall images, and hierarchical function and style based classification of buildings from freely taken photos. The practical usability of the project developments has been demonstrated at proof-of-concept level, and methodological novelties were presented in scientific publications, including three top journal papers.

2. Challenge overview

Automated situation analysis based on multiple sensorial measurements is a key issue in several applications, including autonomous driving, surveillance, smart city management, state assessment of public premises and archeological geo-data processing for cultural heritage preservation.

Real time dynamic object detection (RTDOD) in 3D sparse point clouds is a key challenge in autonomous driving. During the past few years, several geometric and deep learning based approaches appeared in the literature, which operate on raw Rotating Multi-Beam (RMB) LiDAR frames, and provide as output sets of oriented bounding boxes for various object categories such as vehicles, pedestrians or bicycles. However, due to the low spatial density of the RMB LiDAR sensor data, which quickly decreases as a function of the objects' distance form the sensor, the typical ring patterns of the point clouds, and various occlusions effects in dense urban environment, there are a number of limitations of these approaches. On one hand, *false positive* hits

may be detected for point cloud regions containing static scene objects with similar appearance and location parameters to the focused dynamic scene objects. On the other hand, the point cloud blobs of several dynamic objects can be heavily merged or occluded by static street furniture elements, yielding many *false negative* detections. Mobile Laser Scanning (MLS) technologies may be used to obtain High Density Localization (HDL) maps the cities, with providing dense and accurate point clouds from the static environment with homogeneous scanning of the surfaces and a nearly linear reduction of points as a function of the distance. In our project work, we presented a new approach which utilizes these HDL maps, to decrease in parallel both the *false negative* and *false positive* hits of RTDOD algorithms.

Nowadays, state-of-the-art autonomous systems rely on wide range of sensors for environment perception such as *optical cameras*, *radars* and *Lidars*, therefore efficient sensor fusion is a highly focused research topic in the fields of self driving vehicles and robotics. While real time Lidars, such as Velodye's rotating multi-beam (RMB) sensors provide accurate 3D geometric information with relatively low vertical resolution, optical cameras capture high resolution and high quality image sequences enabling to perceive low level details from the scene. In the project we proposed a new fully automatic and target-less extrinsic calibration approach between a camera and a rotating multi-beam (RMB) Lidar mounted on a moving car.

Urban area analysis and in particular building classification are key issues for numerous applications such as urban planning and management, urban visualization and population estimation. Although extensive research work has already been conducted on the detection and localization of buildings in remote sensing images, only a few existing methods deal with the automated analysis and characterization of the extracted buildings. In the project we introduced a new approach on building classification in urban environments based on photographs of their facades. The proposed method can be added to the top of a remote sensing data processing pipeline for virtual city modeling, providing specific information about the functionality and the architectural styles of the buildings. Our work focused on two main aspects of urban building categorization: functional purpose and architectural style. These two characteristics are useful in a number of applications such as: real-estate valuation, declaration as a monument, virtual city tours, etc.

Over the past years, we have observed a significant shift to digital technologies in Cultural Heritage (CH) preservation applications. By investigating masonry walls of buildings, bricks are considered as the vital components of the masonry structures. Accurate detection and separation of bricks is a crucial initial step in various applications, such as stability analysis for civil engineering, managing the damage in architectural buildings, and in the fields of heritage restoration and maintenance. Various ancient heritage sites have experienced several conversions over time, and many of their archaeologically relevant parts have been ruined. Reassembling these damaged areas is an essential task for archaeologists studying and preserving these ancient monuments. In particular, by examining masonry wall structures, it is often necessary to estimate the original layout of brick and mortar regions of wall components which are damaged, or occluded by various objects such as ornamenting elements or vegetation. In the project, we proposed an end-to-end deep learning-based algorithm for masonry wall image analysis and virtual structure recovery.

3. Implementation of the project

In the project, the research team led by Csaba Benedek (PI) fulfilled various research a development tasks in the fields of image processing, geometry, machine learning, deep learning, and geoinformatics. All team members, Balázs Nagy, Örkény Zováthi and Lóránt Kovács have been Ph.D. students supervised by the PI. The supporting company, *IP Camp* is a German/Hungarian company working on automotive software development projects and product development, performance cloud computing applications, also covering automated driving and AI topics. Their expertize helped the project in defining the research goals, and evaluating the results.

4. The research problem, methods, and results

The research team proposed contributions and published international research papers regarding the following main research problems:

- Enhancing AI based object recognition from Lidar measurements used in self-driving and ADAS applications
- Exploitation of high density point cloud maps for navigation and situational analysis
- Optimizing camera-Lidar fusion algorithms
- Extending the methods to further applications including urban area management and archeology

In the following, the main research results of the project will be briefly summarized, with highlighting the main publications.

4.1 Multi-object Detection in Urban Scenes Utilizing 3D Background Maps and Tracking

We have proposed a workflow of several new algorithms (Fig. 1) for the analysis of the dynamic 3D environment using up-do-date sensor configurations, including moving car-mounted Rotating Multi-beam (RMB) Lidar sensors, and very dense 3D point clouds of the environment obtained by mobile laser scanning by Mobile Laser Scanning (MLS) technology [1]. First real time methods have been presented for moving object detection and recognition in RMB Lidar sequences. Thereafter we introduced a new 3D convolutional neural network (CNN) based approach to segment dense MLS point clouds into nine different semantic classes, which can be used for high definition city map generation.

Next, we proposed a RMB Lidar based real time and accurate self-localization approach for self-driving vehicles in high resolution MLS point cloud maps of the environment [7]. We have also demonstrated that by the exploitation of High Density Localization (HDL) maps obtained by Mobile Laser Scanning we can increase the performance of state-of-the-art real time dynamic object detection (RTDOD) methods

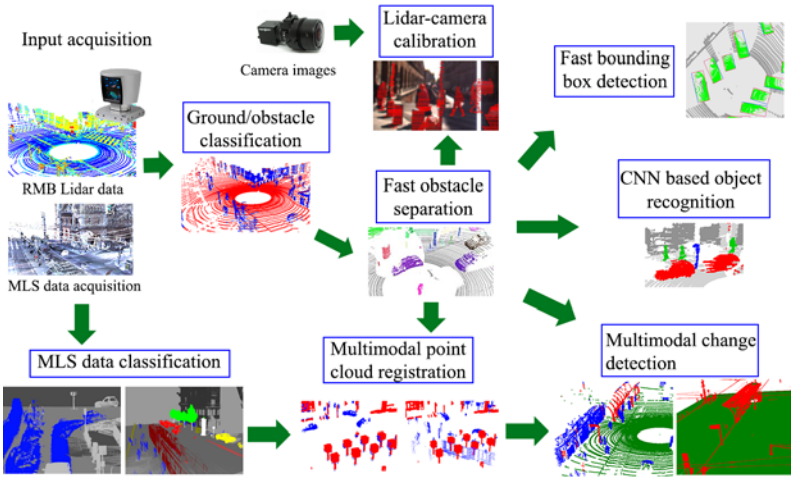


Figure 1. Workflow of instant environment perception, composed of the proposed algoritms.

utilizing RMB Lidar measurements. Here, we align the onboard measurements to the 3D HDL map with a multimodal point cloud registration algorithm operating in the Hough space. Next we apply a grid based probabilistic step to filter out the object regions on the RMB Lidar data which were falsely predicted as dynamic objects by RTDOD, although they are part of the static background scene. On the other hand, to find objects erroneously missed by the RTDOD predictions, we implement a Markov Random Field based point level change detection approach between the map and the current onboard measurement frame. Finally, to analyse the changed but previously unclassified segments of the RMB Lidar clouds, we apply a geometric blob separation and a Support Vector Machine based classification to distinguish the different object types. Comparative tests are provided in high traffic road sections of Budapest, Hungary, and we show an improvement of $5,96\%$ in precision, $9,21\%$ in recall and $7,93\%$ in F-score metrics against the state-of-the-art RTDOD algorithm.

To find additional objects missed by the object detector on the current measurement, we apply a Kalman-filter based object tracking [6]. Hereby we first predict the current state of the previously detected and tracked objects. Next, we adopt a Hungarian matcher based assignment between the tracked and the current objects and update the object list according to the result. For better accuracy, we keep all predictions through a couple of frames. Results have been published in [1, 6, 7].

4.2 On-the-Fly Camera and Lidar Calibration

Sensor fusion is one of the main challenges in self driving and robotics applications. We have proposed an automatic, online and target-less camera-Lidar extrinsic calibration approach (see Fig. 1 top-middle image). We adopt a structure from motion (SfM) method to generate 3D point clouds from the camera data which can be matched to the Lidar point clouds; thus, we address the extrinsic calibration problem as a registration

task in the 3D domain. The core step of the approach is a two-stage transformation estimation: First, we introduce an object level coarse alignment algorithm operating in the Hough space to transform the SfM-based and the Lidar point clouds into a common coordinate system. Thereafter, we apply a control point based nonrigid transformation refinement step to register the point clouds more precisely. Finally, we calculate the correspondences between the 3D Lidar points and the pixels in the 2D camera domain.

To evaluate our proposed *target-less* and *fully automatic* self-calibration method we compared it to two state-of-the-art target-less techniques, and an *offline target-based* method. We quantitatively evaluated the considered methods by measuring the magnitude and standard deviation of the pixel level projection error values both in the x and y directions along the image axes. We collected *10* km long test data from different dense urban scenarios such as *boulevards*, *main roads*, *narrow streets* and *large crossroads*. To take into account the daily traffic changes, we recorded measurements both in rush hours in heavy traffic, and outside the peak hours as well.

We manually measured the projection error of the proposed and the reference methods by projecting the 3D Lidar points onto the 2D image pixels, which was followed by visual verification. Before starting the test, we carefully calibrated the Lidar and camera sensors using the offline target-based reference method, which served as a baseline calibration for the whole experiment. During our *10* km long test drive, we chose *200* different key frames of the recorded measurement sequence, and by each key frame - based on the actual a Lidar point cloud and $N = 8$ consecutive camera images - we executed new calibration processes in parallel by our proposed approach and by the remaining automatic reference techniques. We paid attention to collect these key frames from diverse scenarios (*main roads*, *narrow streets*, etc.), and we also noticed the corresponding vehicle speed values to assign each key frame either to the *Slow* or to the *Fast* test set. We evaluated the performance of all considered methods by each key frame (i.e. after each re-calibration). To calculate the pixel projection error we selected *10* well defined feature points (e.g. corners) from different regions of the 3D point cloud, and using the calibration results of the different methods we projected these 3D feature points to the image domain. Finally we measured the pixel errors versus the true positions of the corresponding 2D feature points positions detected (and manually verified) on the images.

The results confirmed that our proposed extrinsic calibration approach is able to provide accurate and robust parameter settings on-the-fly, and it significantly surpasses the state-of-the-art reference techniques. The new approach was published in [4].

4.3 HierarchyNet: Hierarchical CNN-Based Urban Building Classification

Automatic building categorization and analysis are particularly relevant for smart city applications and cultural heritage programs. Taking a picture of the facade of a building and instantly obtaining information about it can enable the automation of processes in urban planning, virtual city tours, and digital archiving of cultural artifacts. In this work, we have gone beyond traditional convolutional neural networks

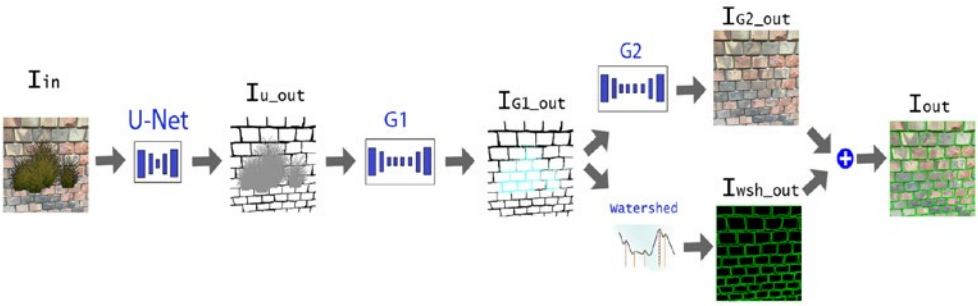


Figure 2. Workflow of the proposed masonry wall image analysis approach.

(CNNs) for image classification and proposed the HierarchyNet: a new hierarchical network for the classification of urban buildings from all across the globe into different main and subcategories from images of their facades. We introduced a coarse-to-fine hierarchy on the dataset and the model learns to simultaneously extract features and classify across both levels of hierarchy. We proposed a new multiplicative layer, which is able to improve the accuracy of the finer prediction by considering the feedback signal of the coarse layers. We have quantitatively evaluated the proposed approach both on our proposed building datasets, as well as on various benchmark databases to demonstrate that the model is able to efficiently learn hierarchical information. The HierarchyNet model is able to outperform the state-of-the-art convolutional neural networks in urban building classification as well as in other multi-label classification tasks while using significantly fewer parameters. The new approach was published in [5].

4.4 Deep Learning-Based Masonry Wall Image Analysis

We introduced a novel machine learning-based fully automatic approach for the semantic analysis and documentation of masonry wall images, performing in parallel automatic detection and virtual completion of occluded or damaged wall regions, and brick segmentation leading to an accurate model of the wall structure. For this purpose, we proposed a four-stage algorithm which comprises three interacting deep neural networks and a watershed transform-based brick outline extraction step (Fig. 2). At the beginning, a U-Net-based sub-network performs initial wall segmentation into brick, mortar and occluded regions, which is followed by a two-stage adversarial inpainting model. The first adversarial network predicts the schematic mortar-brick pattern of the occluded areas based on the observed wall structure, providing in itself valuable structural information for archeological and architectural applications. The second adversarial network predicts the pixels' color values yielding a realistic visual experience for the observer. Finally, using the neural network outputs as markers in a watershed-based segmentation process, we generate the accurate contours of the individual bricks, both in the originally visible and in the artificially inpainted wall regions. Note that while the first three stages implement a sequential pipeline, they interact through dependencies of their loss functions admitting the consideration of

hidden feature dependencies between the different network components. For training and testing the network a new dataset has been created, and an extensive qualitative and quantitative evaluation versus the state-of-the-art is given. The experiments confirmed that the proposed method outperforms the reference techniques both in terms of wall structure estimation and regarding the visual quality of the inpainting step, moreover it can be robustly used for various different masonry wall types. The new approach was published in [2, 3].

4.5 The applied methods

To reach the project goals, various mathematical and computational methods and techniques have been applied. First of all, we conducted research on Bayesian classification and geometry based modeling techniques, especially Markov Random Fields, where to formally define an image segmentation task, we assign to each s pixel of the image pixel lattice (R) a $l_s \in \{F, S, B\}$ class label so that we aim to minimize the following energy function:

$$E = \sum_{s \in R} -\log(\mathcal{F}_{l_s}(s)) + \sum_{s \in R} \sum_{s^* \in N_s} \beta \cdot 1\{l_s \neq l_{s^*}\}.$$

In addition, we applied and improved recent deep learning techniques, through the development of new neural network architectures. Fusion of classical and AI based modeling approaches and decision techniques for fusing machine learning and probabilistic approaches were also key issues in our project.

4.6 The achieved research results

Using the support of the project, the project team published 7 referred and international scientific paper [1, 2, 3, 4, 5, 6, 7], including 3 articles in top (Scimago's D1 category) journals [3, 4, 5], and a paper of an invited *keynote lecture* by Csaba Benedek in the ROBOVIS 2020 international scientific conference [1].

During the project period, Csaba Benedek obtained the D.Sc. degree from the Hungarian Academy of Sciences, and he also received the Michelberger Master Prize from the Hungarian Academy of Engineering. Balázs Nagy obtained the Ph.D. degree from the Pázmány Péter Catholic University, and Örkény Zováthi won a scholarship of the Cooperative Doctoral Program supporting industrial-academic cooperation.

5 Solution of the industrial problem and its benefits

During the project new multi-sensorial solutions have been developed, which provide higher level of automation and reliability of environment perception for automotive, surveillance and cultural heritage applications. The cooperating company, *IP-Camp*, can also benefit from various results of the project, which can be used for implementing new innovative technology demonstrators for their future advanced AI projects. In particular, the following project contributions can affect the application: (i) New AI based object detection approach for camera images and Lidar point cloud streams (ii)

New building classification method using images (based on function and style) (iii) New targetless fully automatic camera-Lidar calibration technique.

6 Conclusions

During the project, the working team has achieved progress in multiple, application oriented research tasks on spatial data analysis with machine learning, and the results have been demonstrated in real data samples under realistic conditions. Further steps will focus on bringing the results closer to products at an industrial level.

References

[1] Csaba Benedek, Multi-sensorial Environment Perception in Urban Environment. In *International Conference on Robotics, Computer Vision and Intelligent Systems*, Communications in Computer and Information Science. Springer, Budapest, Hungary, 2020. KEYNOTE TALK article, to appear.

[2] Yahya Ibrahim, Balázs Nagy, and Csaba Benedek, A GAN-Based Blind Inpainting Method for Masonry Wall Images. In *International Conference on Pattern Recognition (ICPR)*, Milan, Italy, 2020.

[3] Yahya Ibrahim, Balázs Nagy, and Csaba Benedek, Deep learning-based masonry wall image analysis. *Remote Sensing*, 12(23), 2020.

[4] Balázs Nagy and Csaba Benedek, On-the-Fly Camera and Lidar Calibration. *Remote Sensing*, 12(7), 2020.

[5] Salma Taoufiq, Balázs Nagy, and Csaba Benedek, HierarchyNet: Hierarchical CNN-Based Urban Building Classification. *Remote Sensing*, 12(22), 2020.

[6] Örkény Zováthi, Lóránt Kovács, Balázs Nagy, and Csaba Benedek, Multi-Object Detection in Urban Scenes Utilizing 3D Background Maps and Tracking. In *International Conference on Control, Artificial Intelligence, Robotics Optimization (ICCAIRO)*, pages 231-236, 2019.

[7] Örkény Zováthi, Balázs Nagy, and Csaba Benedek, Exploitation of Dense MLS City Maps for 3D Object Detection. In *International Conference on Image Analysis and Recognition (ICIAR)*, volume 12131 of *Lecture Notes in Computer Science*, pages 393-403. Springer, Póvoa de Varzim, Portugal, 2020.

Lane Modelling Algorithm for Video-Based Driver Assistance System

Norbert Bogya[1], **Róbert Fazekas**[2], **Judit Nagy-György**[3] and **Zsolt Vizi**[4]

[1]Bolyai Institute, University of Szeged,
[2]Robert Bosch Kft., Budapest,
[34]Bolyai Institute, University of Szeged,

[1]E-mail: nbogya@math.u-szeged.hu
[2]E-mail: Robert.Fazekas@hu.bosch.com
[3]E-mail: Nagy-Gyorgy@math.u-szeged.hu
[4]E-mail: zsvizi@math.u-szeged.hu

1 Executive Summary

An autonomous car is a vehicle that is able to sense environmental data and to navigate based on them without any human action. Advanced driver assistance systems (ADAS) are electronic devices which help the drivers during the process of driving. One of the most important task of an ADAS is managing the data produced by a sensor.

There is currently no reliable algorithm for road reconstruction because the situation is much more complicated than in a highway environment. The main goal is to develop a dynamic process, which computes the lanes continuously and directly from the data of a stereo-camera and give a lane model.

We developed two clustering algorithms. The comparision of the two algorithms on several test cases does not suggest that one of the algorithms should be preferred. With the popular and faster threshold method, a lot of information can be lost, and the theoretical investigation of spectral clustering suggests, that it should be better. These algorithms can be combined.

2 Challenge overview

In this research we considered a stereo video camera as a sensor of perception. The main goal is to provide a clustering of the available lane segments by a dynamic process, which computes the lanes continuously and directly from the data of a stereo-camera and give a lane model. Most of the solutions in the industry assume simple environments, e.g. highway, for a functionality like adaptive cruise control (ACC), but these concepts fail, if the scene has a complexity. In this work we provided algorithms, which are robust when the environment is quite chaotic, e.g. in construction zones. Another difficulty is that we have to care about the complexity of the provided algorithms, because these are implemented in an embedded system with a small computing capacity and the decision has to be made quickly.

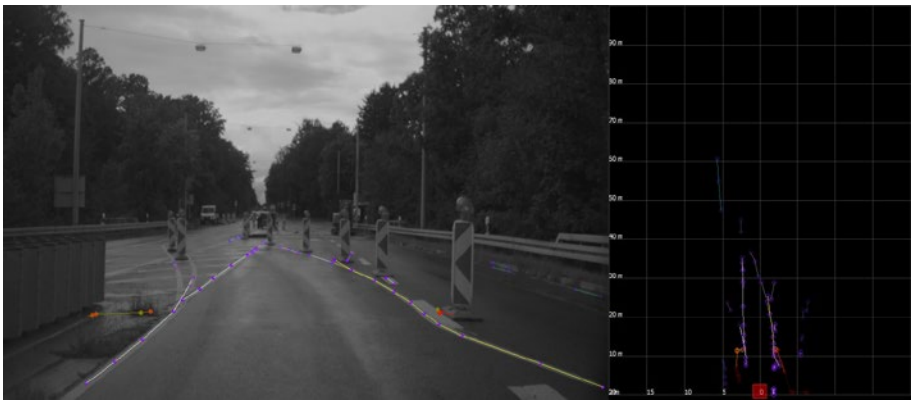


Figure 1: A complicated situation

3 Implementation of the project

The industry partner, Bosch, provided the task and data, and the academic partner, the University of Szeged, provided the researchers and the methods leading to the solution. The leader of the team was the senior researcher Judit Nagy-György. Other members of the team are a young researcher, Norbert Bogya, another young researcher, Zsolt Vizi, who was also employed by Bosch at the time of the research and is currently a researcher at University of Szeged, and Róbert Fazekas, who was an MSc student at University of Szeged at the time of the research, is currently employed by Bosch. The methods were developed by the two young researchers and implemented by the MSc student.

4 The research problem, methods, and results

An autonomous car is a vehicle that is able to sense environmental data and to navigate based on them without any human action. Nowadays, the development of self-driving cars and driver assistance systems belongs to the most dynamic industrial projects. Advanced driver assistance systems (ADAS) are electronic devices which help the drivers during the processing of driving. One of the most important task is to manage the data produced by sensors. In this research we considered a stereo video camera as a sensor of perception and we implemented two algorithms to provide a clustering of the available lane segments. In this work we provided algorithms, which are robust in more complicated situations. As we know, there are no stateof-the-art solution for these situations. The algorithms in use are very sensitive for the quality and quantity of input data, and most of the well-know algorithms are hybrid algorithms for multiple sensors.

4.1 The research problem:

A stereo-camera includes two separate cameras, taking pictures of the same focus used to model the environment in 3-dimension. At first, the pictures go through a process that recognises the road markings and it makes a segmentation of them. This method is not considered in this paper and we use only a little part of the data given by the segmentation. We use only on the geometrical representation of the lane segment and ignore e.g. colour of the lane marking, which would improve the consistency of the lane model. Additionally, we have to care about the complexity of the provided algorithms. In this environment, it is complicated to apply neural networks, mostly model-based algorithms are considered.

Our input is a set of polygonal paths, whose vertices are the points of the road markings given by the pre-segmentation method. Two points are connected with a straight segment if the pre-segmentation procedure marks them as points of a same road marking. In this work, we considered a polygonal path as one object and investigate a relationship between these objects. Our goal is to classify which polygonal path belongs to which road marking, thus this can be interpreted as a clustering problem.

4.2 The applied methods

A key concept in all clustering algorithm to define a similarity function, which measures the "conceptual distance" of the sample data. If H is the set of data points, similarity function f is an $H \times H \to \mathbb{R}^+$ symmetric map. We use a normalised similarity function in order to interpret $f(h_1, h_2)$ as the probability of h_1 and h_2 belong to the same cluster.

In [7], the authors investigate a very similar question raised from image processing, namely a finite set of objects is given, and what can be said about the global connection of the objects, if we have information about all connection of each pairs of objects. Their global concept is a clustering into two clusters (foreground and background). They use the concept of Gestalt psychology, that have several principles about what kind of segments can be thought as a continuation of each other. Following their idea, we used similarity function

$$f(i,j) = \exp\left(-\frac{d_{ij}^2}{d_0^2} - \frac{2 - \cos(2\alpha_i) - \cos(2\alpha_j)}{\alpha_0} - \frac{g_{i,j}}{g_0}\right)$$

if i and j denotes different segments, otherwise $f(i,j)$ is zero. The distance $d_{i,j}$ is measured between the midpoints of the segments i and j; and $g_{i,j}$ denotes the distance between these segments. The angles α_i and α_j can be understood from Figure 2. The constants $d_0,\ \alpha_0,\ g_0$ are real parameters.

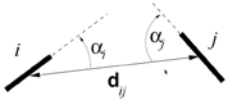


Figure 2: Parameters for similarity of segments

Instead of similarity of individual line segments we define similarity of provided polygonal paths. Investigating the inputs, we found that the connected polygonal paths can be contained in a very narrow but much longer rectangle. Hence, polygonal paths can be substituted by a segment between their first and last point.

From the similarity function, we can build a similarity matrix $\mathbf{W} \in \mathbb{R}^{n\times n}$, where the rows and columns represent the polygonal paths and $W_{i,j} = f(i,j)$. If we consider this matrix as an edge-weight matrix of a graph G, the vertices of the defined graph are the polygonal paths and there is an edge between them if their similarity is positive.

In this approach, the original problem is converted to the clustering of the vertices of G and our goal now is to construct an efficient algorithm for edge-weighted graph clustering. We developed two algoritmh to solve this problem.

Algorithm #1 uses a thresholding method considered in [2]. In the first step the weight matrix $\mathbf{W}$ is transformed to a 0-1 matrix $\mathbf{A}$: $A_{i,j} = I(W_{i,j} > \varepsilon)$, where I is the indicator function and $\varepsilon = \mu - \sigma + h$ is an adaptive threshold, where μ, σ and h are the average, standard deviation and entropy of the nonzero elements of $\mathbf{W}$, respectively. Instead of it we used a larger threshold (thus stricter truncation of the elements), namely $\varepsilon = \mu + \sigma - h$, and this was implemented in our algorithm.

Algorithm #2 is based on the so-called spectral clustering presented in [6]. Applying this method, the following question can be answered: how can we make vectors from the vertices of a graph, and embed them into $\mathbb{R}^k$ saving as much information as possible about the structure of the original graph [2, 4]. Let $S = \{s_1, \ldots, s_n\}$ be a set of d-dimensional points in the Euclidean space, but it can be used in arbitrary feature space with an appropriate measure of distance. The input is the edge-weight matrix $\mathbf{W}$ (which is calculated with the similarity function f) and the steps are the followings.

1. Compute $\mathbf{D} \in \diagdown^{n\times n}$ diagonal matrix, such that $D_{i,i} = \sum_{j=1}^{n} W_{i,j}$.
2. Compute $x_1, \ldots, x_n$ diagonal eigenvectors of matrix $\mathbf{D}^{-1/2}\mathbf{W}\mathbf{D}^{-1/2}$ corresponding to the k largest eigenvalues and let $\mathbf{X} = [x_1, \ldots, x_n] \in \mathbb{R}^{n\times k}$.
3. Normalise the rows of $\mathbf{X}$ to obtain $\mathbf{Y} \in \mathbb{R}^{n\times k}$.
4. We cluster the row vectors of $\mathbf{Y}$, that are unit vectors.
5. The points $s_i, s_j \in S$ are declared to belong to the same cluster, if the corresponding row vectors of $\mathbf{Y}$ are declared into the same cluster.

More details about the mathematics behind the algorithm can be found in [2] or [8].

4.3 The achieved research results

The results are illustrated in the example shown in Figure 3. On the left side of the figure, the results of Algorith #1 can be observed. The two most important lanes can be recognised easily: the green one is the original road marking, the red one is a detour. On the right side of this row, the result of the spectral clustering is presented, and the main lanes are almost the same as on the left.

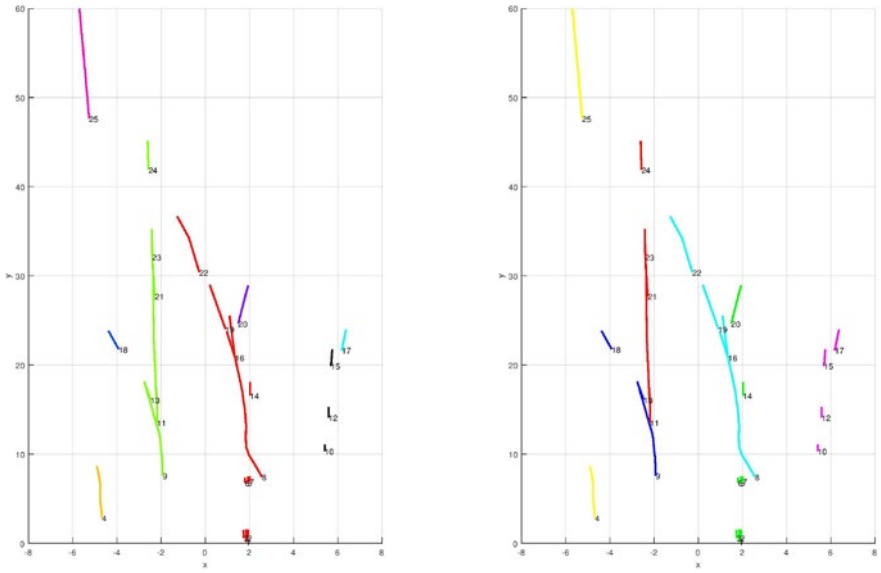


Figure 3: Result of Algorith #1 and Algorith #2, respectively

The run time of Algorithm #1 is less in the most cases. In Table 1 we summarise a manual validation of our precesses.

Algo #1 won	Tie	Algo #2 won
27%	40%	33%

Table 1: Comparison of algorithms

Tie means that the outputs are almost the same (99%) and they show the lanes correctly or they produce bad result, and it is hard to decide which is the worse. Algorithm #2 lost several times because it made connection between segments that are clearly not related to each other, just there was a pressure due to the predefined number of clusters. These numbers were chosen separately, manually by the pre-checking of images.

The comparision of the two algorithms on several test cases does not suggest that one of the algorithms should be preferred. For more details see [1].

5 Solution of the industrial problem and its benefits

We have given two solutions to the task given by Bosch. The benefits of Algorith #1 are easy implementation, efficient runtime and the number of clusters are computed automatically and dynamically. With the popular and faster thresholding method, a lot of information can be lost, and the theoretical investigation of spectral clustering suggests that it should be better, because more information can be saved and used. However, Algorithm #2 works with a predefined number of clusters, it may happen, that such polygonal paths can be clustered into the same group, that are probably not the parts of the same lane. Because of the predefined number of clusters, each path is put into one of the clusters, so cluster validation can be reasonable.

Bosch is satisfied with the work of our team. Algorithms was benchmarked in a concept phase of a series project and it will likely be used. This work is a base of further cooperation projects.

6 Conclusions

As a conclusion, running the algorithms on several test cases, we cannot state, that one of the algorithms should be preferred over the other.

Another idea is to use machine learning methods. Both of the algorithms works with changeable parameters that can be tuned via training using good collections of input data.

Algorithm #1 can be implemented easier, and the clusters can be computed with depth-first search in contrast to the numeric eigenvalue problem of spectral method. And, of course, its running time is also less. The number of clusters are computed automatically and dynamically, we do not need any prediction about it.

Algorithm #2 uses k-means method, which needs an initialisation of centres (typically random) and pre-defined number of clusters. One idea to resolve both needs is combining the two algorithms: Algorithm #1 is run and centres of provided clusters initialises the k-mean method and we choose the number of clusters given by Algo #1.

More sophisticated evaluation of the algorithms and the latter idea will be investigated in the future.

References

[1] N. Bogya, R. Fazekas, J. Nagy-György, Zs. Vizi, *Clustering Algorithm Exploring Road Geometry in a Video-Based Driver Assistant System*, In: Progress in Industrial Mathematics at ECMI 2018, Springer International Publishing, pp 589-95 (2019)

[2] F.R. Chung, *Spectral graph theory.* CBMS Regional Conference Series in Mathematics, vol. 92 (1997)

[3] K.M. Hall, *An r-dimensional quadratic placement algorithm.* Manag. Sci. Theory Ser. 17(3), 219-229 (1970)

[4] A.R. Kelly, E.R. Hancock: *Grouping-line segments using eigenclustering.* In: Proceedings of the British Machine Vision Conference 2000, pp. 1-10 (2000)

[5] Y. Koren, *On spectral graph drawing.* In: Proceedings of the 9th Annual International Conference on Computing and Combinatorics, COCOON-03, pp. 496-508 (2003)

[6] A.Y. Ng, M.I. Jordan, Y. Weiss, *On spectral clustering: analysis and an algorithm.* In: Proceedings of the 14th International Conference on Neural Information Processing Systems: Natural and Synthetic, NIPS-01, pp. 849-?856 (2002)

[7] P. Perona, W. Freeman, *A factorization approach to grouping.* In: Computer Vision ECCV?98. Lecture Notes in Computer Science, vol. 1406 (1998)

[8] J. Shi, J., J. Malik, *Normalized cuts and image segmentation.* IEEE Trans. Pattern Anal. Mach. Intel. 22(8), 888-905 (2000)

Dimension Reduction of High Frequency and High Dimensional Data in Time and Space

Marianna Bolla[1], **Máté Baranyi**[2] and **Fatma Abdelkhalek**[3]

[123]Budapest University of Technology and Economics, Department of Stochastics

[1]E-mail: marib@math.bme.hu
[2]E-mail: baranyim@math.bme.hu
[3]E-mail: fatma@math.bme.hu

1 Executive Summary

Our industrial partner was the I-QRS International Kft. It is a R&D company studying ECG and its relation to other physiological parameters. With their unique in-house created hardware, they monitor wirelessly and precisely vital and motion parameters of subjects. Our objective was to find the underlying main signals or driving latent sources detected by many motion sensors worn by different subjects, mainly athletes.

This is a topic of multidimensional time series analysis. Our studies raised some theoretical questions too, some of which we were able to answer. We proved asymptotic relation between the eigenvalues of the block Toeplitz matrix of autocovariances in the time domain and the union of spectra of the spectral density matrices at the Fourier frequencies in the frequency domain. It has nice consequences related to dynamic Principal Component Analysis (d-PCA). The theoretical results of our research was summarized in [1]. We implemented an algorithm in Python to calculate/estimate the d-PCs of the multidimensional motion sensor data of the athletes. The company is planning to further investigate the delivered algorithm/implementation in the future.

2 Challenge overview

The objective of this research was to find the underlying main signals or driving latent sources detected by many motion sensors worn by different subjects, mainly athletes. Main challenge was the handling of the high and mixed frequency multidimensional time series data that came from noisy sensors. In Fig. 1 the log-eigenvalues of the estimated spectral density process from a preprocessed sample of the I-QRS data is illustrated. This also shows the complexity of their signals. The nature of the data together with its size is not entirely apt for simple algorithms. Our goal was to investigate the d-PCs behind the underlying signals. These have many nice properties, and advantage over static techniques. They provide explainable dimension reduction by their loadings.

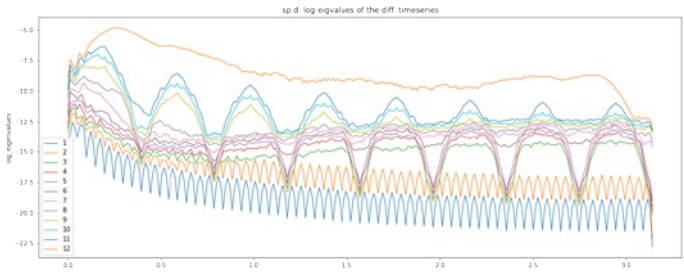


Figure 1: Log-eigenvalues of the spectral density estimates of the mixed frequency data after its preprocession

3 Implementation of the project

After a series of meetings we could find a common platform of the company's needs and our expertise in time series analysis. We have agreed on to investigate their sensor data based on PCA techniques. The industrial partner provided the data and their expert knowledge about the motion sensors. Our role was the theoretical understanding of the topic, and later on the implementation of some novel algorithm (that may help them better understand their data) as deliverable.

4 The research problem, methods, and results

Given a weakly stationary, multivariate time series with absolutely summable autocovariances, asymptotic relation is proved between the eigenvalues of the block Toeplitz matrix of the first n autocovariances and the union of spectra of the spectral density matrices at the n Fourier frequencies, as $n \to \infty$. The theorem, proved in [1], has important consequences as for the analogies between the time and frequency domain calculations. In particular, the complex PCs are used for low-rank approximation of the process. The results are applicable to high frequency time series, in our case, to dimension reduction of medical and motion sensor data. In this way, we can find the important signals behind many sensors with our d-PCA method.

The book [4] to be published soon discusses the background spectral theory of multivariate, weakly stationary processes in details. In our case, the word "spectra" has two meanings: the spectral distribution, and the collection of eigenvalues of the spectral density matrices at the Fourier frequencies. Papers [2,5] are rather related to longitudinal data, where the structural equations are solved with block matrix techniques, but mainly applied to demographic and sociological data. Article [5] extends the seminal Kálmán's filtering, discussed in [4] in details, to the case of multiple state and observation equations. In contrast to considering the time horizon, paper [3] rather focuses on prediction based on causal relations between multidimensional, mixed type data, represented by chain graphs in space.

4.1 The research problem: spectra of spectra

Let $\{\mathbf{X}_t\}$ be a weakly stationary d-dimensional real time series, $t \in \mathbb{Z}$, $\mathbb{E}(\mathbf{X}_t) = \mathbf{0}$ and the sequence of the $d \times d$ autocovariance matrices $\boldsymbol{C}(h)$ $(h \in \mathbb{Z})$ is absolutely summable (entrywise). The $nd \times nd$ covariance matrix $\mathfrak{C}_n$ of the compounded random vector $(\mathbf{X}_1^T, \ldots, \mathbf{X}_n^T)^T$ is a symmetric, positive semidefinite block Toeplitz matrix $\mathfrak{C}_n$, the (i, j) block of which is $\boldsymbol{C}(j - i)$. The symmetry comes from the fact that $\boldsymbol{C}(i - j) = \boldsymbol{C}^T(j - i)$.

At the same time, we consider the (under our conditions, existing) spectral density matrix $\boldsymbol{f}$ of the process. This is a self-adjoint, positive semidefinite, $d \times d$ matrix, which is, in fact, the Fourier transform of the autocovariance matrices:

$$\boldsymbol{f}(\omega) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \boldsymbol{C}(h) e^{-ih\omega}, \quad \omega \in [0, 2\pi].$$

In view of $\boldsymbol{f}(-\omega) = \overline{\boldsymbol{f}(\omega)}$, it suffices to confine ourselves to the $[0, \pi]$ interval (here $\bar{\cdot}$ denotes the complex, entrywise conjugation).

In our main theorem, we state that for "large" n, the eigenvalues of $\mathfrak{C}_n$ asymptotically comprise the union of the spectra of the spectral density matrices $\boldsymbol{f}(\omega_j)$, $\omega_j = \frac{2\pi j}{n}$ $(j = 0, 1, \ldots, n - 1)$. The theorem also has the computational benefit that instead of finding the eigenvalues of an $nd \times nd$ matrix (that needs $\mathcal{O}(n^3 d^3)$ operations) we can find the eigenvalues of number n of $d \times d$ matrices (that means only $\mathcal{O}(nd^3)$ operations).

To characterize the eigenvalues of the block Toeplitz matrix $\mathfrak{C}_n$, we need the symmetric block circulant matrix $\mathfrak{C}_n^{(s)}$ that we consider for odd n, say $n = 2k+1$ here (for even n, the calculations are similar). $\mathfrak{C}_n^{(s)}$ is a symmetric block Toeplitz matrix, and it is the same as $\mathfrak{C}_n$ within the blocks (i, j)s for which $|j - i| \leq k$ holds. The point is that $\mathfrak{C}_n^{(s)}$ has k double eigenvalues, to which both a real and a complex system of eigenvectors belong, the columns of $\boldsymbol{U}$ and $\boldsymbol{W}$, respectively. With the real ones,

$$\mathfrak{C}_n^{(s)} = \boldsymbol{U} \boldsymbol{\Lambda}^{(s)} \boldsymbol{U}^T \tag{1}$$

is the corresponding spectral decomposition.

Theorem 1. *In the above setup, let*

$$\boldsymbol{D}_n := \operatorname{diag}(\operatorname{spec} \boldsymbol{f}(0), \operatorname{spec} \boldsymbol{f}(\omega_1), \ldots, \operatorname{spec} \boldsymbol{f}(\omega_k), \operatorname{spec} \boldsymbol{f}(\omega_k), \ldots, \operatorname{spec} \boldsymbol{f}(\omega_1)).$$

Then, with the spectral decomposition (1),

$$\boldsymbol{U}^T \mathfrak{C}_n \boldsymbol{U} - 2\pi \boldsymbol{D}_n \to \boldsymbol{O}, \quad n \to \infty,$$

i.e. the entries of the matrix $\boldsymbol{U}^T \mathfrak{C}_n \boldsymbol{U} - 2\pi \boldsymbol{D}_n$ tend to 0 uniformly as $n \to \infty$.

4.2 The applied methods

The complex PC transform of the collection of random vectors $\mathbf{X} = (\mathbf{X}_1^T, \ldots \mathbf{X}_n^T)^T$ of real coordinates is the random vector $\mathbf{Z} = (\mathbf{Z}_1^T, \ldots, \mathbf{Z}_n^T)^T$ of complex coordinates

obtained by $\mathbf{Z} = \boldsymbol{W}^*\mathbf{X}$. To relate the PC transformation to a discrete Fourier transformation, we also make PC transformations within the blocks. By Theorem 1, $\mathbb{E}\mathbf{Z}\mathbf{Z}^* \sim 2\pi \boldsymbol{D}_n$, so the coordinates of $\mathbf{Z}$ are asymptotically uncorrelated, for "large" n. Instead, we consider the blocks $\mathbf{Z}_j$s of it, and perform a "partial PC transformation" (in d-dimension) of them. Let $\mathbf{w}_{1j}, \dots, \mathbf{w}_{dj}$ be the columns of $\mathbf{W}$ corresponding to the coordinates of

$$\mathbf{Z}_j = \frac{1}{\sqrt{n}}(\boldsymbol{V}_j^* \otimes \mathbf{r}^*)\mathbf{X},$$

where $\mathbf{r}^* = (1, \rho_j^{-1}, \rho_j^{-2}, \dots, \rho_j^{-(n-1)})$ (ρ_j is the jth primitive nth root of 1) and $\boldsymbol{V}_j$ is the $d \times d$ unitary matrix in the spectral decomposition $\boldsymbol{M}_j = \boldsymbol{V}_j \boldsymbol{\Lambda}_j \boldsymbol{V}_j^*$.

It is also important that the spectrum of $\mathfrak{C}_n^{(s)}$ is the union of spectra of the matrices

$$\boldsymbol{M}_j = \boldsymbol{C}(0) + \sum_{h=1}^{k} [\boldsymbol{C}(h)\rho_j^h + \boldsymbol{C}^T(h)\rho_j^{-h}] = \boldsymbol{C}(0) + \sum_{h=1}^{k} [\boldsymbol{C}(h)e^{i\omega_j h} + \boldsymbol{C}^T(h)e^{-i\omega_j h}]$$

for $j = 0, 1, \dots, n-1$. Because of $\mathbb{E}\mathbf{Z}_j\mathbf{Z}_j^* = \boldsymbol{\Lambda}_j$, we have that

$$\mathbb{E}(\boldsymbol{V}_j\mathbf{Z}_j)(\boldsymbol{V}_j\mathbf{Z}_j)^* = \boldsymbol{V}_j\boldsymbol{\Lambda}_j\boldsymbol{V}_j^* = \boldsymbol{M}_j.$$

At the same time,

$$\mathbf{T}_j = \boldsymbol{V}_j\mathbf{Z}_j = \frac{1}{\sqrt{n}}\boldsymbol{V}_j(\boldsymbol{V}_j^* \otimes \mathbf{r}^*)\mathbf{X} = \frac{1}{\sqrt{n}}(\boldsymbol{I}_d \otimes \mathbf{r}^*)\mathbf{X} = \frac{1}{\sqrt{n}}\sum_{t=1}^{n} \mathbf{X}_t e^{-it\omega_j}, \quad j = 1, \dots, n.$$

This is the discrete Fourier transform of $\mathbf{X}_1, \dots, \mathbf{X}_n$.

To find the best m-rank approximation ($1 < m \le r$) of the process, we project the d-dimensional vector $\mathbf{T}_j$ onto the subspace spanned by the m leading eigenvectors of $\mathbf{V}_j$. Important that the eigenvalues in $\boldsymbol{\Lambda}_j$ are in non-increasing order. Let us denote the eigenvectors corresponding to the m largest eigenvalues by $\mathbf{v}_{j1}, \dots, \mathbf{v}_{jm}$. Then

$$\widehat{\mathbf{T}}_j := \mathrm{Proj}_{\mathrm{Span}\{\mathbf{v}_{j1},\dots,\mathbf{v}_{jm}\}}\mathbf{T}_j = \sum_{\ell=1}^{m}(\mathbf{v}_{j\ell}^*\mathbf{V}_j\mathbf{Z}_j)\mathbf{v}_{j\ell} = \sum_{\ell=1}^{m} Z_{j\ell}\mathbf{v}_{j\ell},$$

and $\widehat{\mathbf{T}}_{n-j} = \overline{\widehat{\mathbf{T}}}_j$, for for $j = 1, \dots, k$ (by the previous considerations), were $n = 2k+1$. Further, $\widehat{\mathbf{T}}_0 := \sum_{\ell=1}^{m} Z_{0\ell}\mathbf{v}_{0\ell}$. So, for each j, the resulting vector is the linear combination of the vectors $\mathbf{v}_{j\ell}$s with the corresponding coordinates $Z_{j\ell}$s of $\mathbf{Z}_j$, $\ell = 1, \dots, m$.

Eventually, we find the m-rank approximation of $\mathbf{X}_t$ by inverse Fourier transformation:

$$\widehat{\mathbf{X}}_t = \frac{1}{\sqrt{n}}\sum_{j=0}^{n-1} \widehat{\mathbf{T}}_j e^{it\omega_j},$$

where the vectors $\widehat{\mathbf{X}}_t$ ($t = 1, \dots, n$) all have real coordinates ($n = 2k+1$).

In this way, we have a lower rank process with spectral density of rank $m \le r$. Note that if the process is regular (e.g., it has a rational spectral density), then so

is its low-rank approximation. The "larger" the gap between the mth and $(m+1)$th eigenvalues (in non-increasing order) of the spectral density matrix, the "smaller" the approximation error is.

To back-transform the PC process into the time domain, note that $Z_{j\ell} = \mathbf{v}_{j\ell}^* \mathbf{T}_j, \quad \ell = 1, \dots m$ defines the coordinates of an m-dimensional approximation of $\mathbf{T}_j$, $m \leq r \leq d$. This is the m-dimensional vector $\tilde{\mathbf{T}}_j = (Z_{j1}, \dots, Z_{jm})^T$. The other $d-m$ coordinates of $\mathbf{Z}_j$ are disregarded (they are taken zeros in the new coordinate system $\mathbf{v}_{j1}, \dots, \mathbf{v}_{jd}$). The proportion of the total variance explained by the first m principal components at the jth Fourier frequency is $\sum_{\ell=1}^{m} \lambda_{j\ell} / \sum_{\ell=1}^{d} \lambda_{j\ell}$.

Then the m-dimensional approximation of $\mathbf{X}_t$ by the PC process is as follows:

$$\tilde{\mathbf{X}}_t = \frac{1}{\sqrt{n}} \sum_{j=0}^{n-1} \tilde{\mathbf{T}}_j e^{it\omega_j} = \frac{1}{\sqrt{n}} \left(\sum_{j=0}^{n-1} Z_{j1} e^{it\omega_j}, \dots, \sum_{j=0}^{n-1} Z_{jm} e^{it\omega_j} \right)^T$$

that again results in real coordinates.

4.3 The achieved research results

2beginitemize

We applied d-PCA for the I-QRS data: football players, medical sensor data (ECG, 3-dimensional accelerometer, gyroscope, and magnetometer data with position on the playground). In Fig. 2, a small segment of the first 6 dynamic PCs is visualized.

In Fig. 3, the dynamic factor loadings inside the first two PCs are illustrated. The horizontal axis corresponds to the frequency domain, whereas the vertical one to the weights on the given frequency. The loadings (corresponding to the peak frequencies) showed that two coordinates of the accelerometer and gyroscope had the highest loadings, while magnetometer had low weights. This is in accord with the company's observation, that the measurements by the sensors are highly redundant, and d-PCA is able to find near linear relations between them.

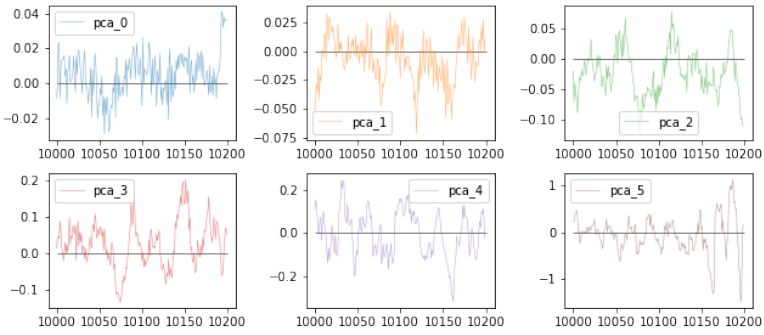


Figure 2: A small segment ($\sim 0.4s$) of the dynamic PCs estimated from the motion sensor data, reduction from 12 to 6 dimensions

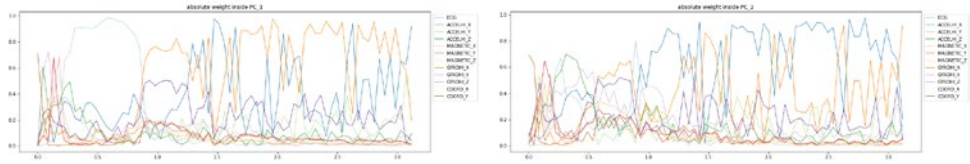


Figure 3: Absolute values of the complex weights (loadings of each dimension) inside the first two PCs on small segment of the data ($\sim 40s$)

5 Solution of the industrial problem and its benefits

We developed and implemented an algorithm in Python to calculate/estimate the dynamic principal components of the multidimensional motion sensor data of the athletes. A schematic workflow of the algorithm is illustrated in Fig. 4. It may help the company understand their data even better in the future. The algorithm can be utilized as a lossy data compression technique too. An illustration for this is Fig. 5 where the low rank approximation (equivalent to the inverse transformation from the PCs) of one sensor is shown. The company is planning to further investigate the delivered algorithm/implementation in the future.

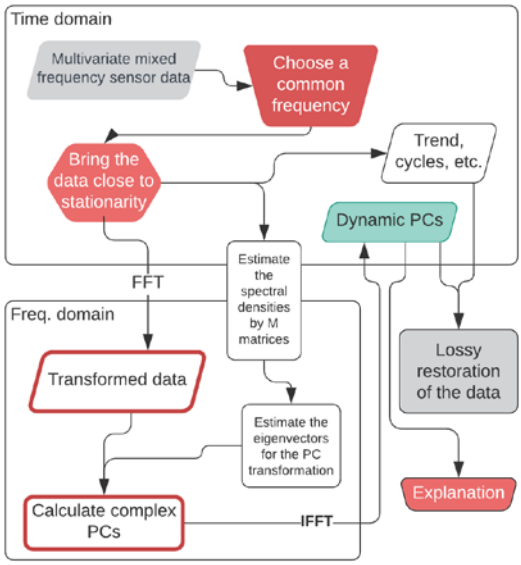


Figure 4: Schematic workflow of the algorithm.

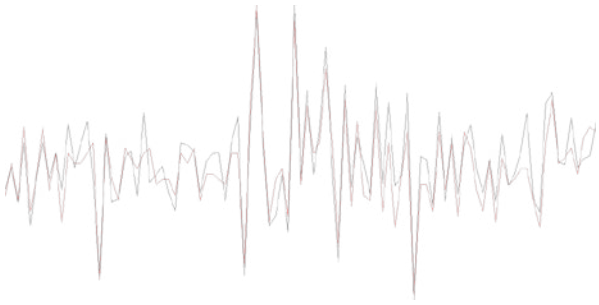


Figure 5: A small segment ($\sim 0.2s$) of one sensor (red) with its low-rank approximation (black). A result of the inverse transformation back from the PCs.

6 Conclusions

Given a weakly stationary, multivariate time series with absolutely summable entries, we proved asymptotic relation between the eigenvalues of the block Toeplitz matrix of the first n autocovariances and the union of spectra of the spectral density matrices at the n Fourier frequencies, as $n \to \infty$. The theorem unveils important analogies between the time and frequency domain calculations. Above touching upon the computational benefits of our algorithm, the calculations are illustrated on medical and motion sensor data.

We plan to continue our investigations on the I-QRS signals. The aim of our further research is also to connect the space and time horizons.

References

[1] Bolla, M., Szabados, M., Baranyi, M., Abdelkhalek, F., *Block circulant matrices and the spectra of multivariate stationary sequences*, Special Matrices 9, pp. 36-51, De Gruyter Open, 2021, doi:10.1515/spma-2020-0121.

[2] Ray Langsten, Fatma Abdelkhalek & Tahra Hassan *Arabic Language Skills: A Comparative Study of Community and Government Schools in Rural Upper-Egypt*, Compare: A Journal of Comparative and International Education, pp. 1-16, Routledge, Taylor&Francis Group, 2020, doi:10.1080/03057925.2020.1843003.

[3] Baranyi, M., Bolla, M., *Iterated Conditional Expectation Algorithm on DAGs and Regression Graphs*, Econometrics and Statistics (Elsevier), published online, 2020, doi:10.1016/j.ecosta.2020.05.003.

[4] Bolla, M., Szabados, T., *Multidimensional stationary time series: dimension reduction and prediction*, to be published by Taylor and Francis (CRC Press).

[5] Bolla, M., Abdelkhalek F., *Kálmán's Filtering Technique in Structural Equation Modeling*, Studia Universitatis Babeş-Bolyai Mathematica 66, No. 1, pp. 179-196, 2021, doi:10.24193/subbmath.2021.1.15.

3D mapping and localization of mobile robots, an application to UV-C disinfection robots

Barnabás Bugár-Mészáros[1], **Iván Eichhardt**[2], **Sándor Gazdag**[3] and **András Majdik**[4]

[1234]University of Debrecen, Faculty of Informatics, Informatics, 4028 Debrecen, Hungary and Institute for Computer Science and Control (SZTAKI), Kende u. 13-17, Budapest, Hungary,

[1]E-mail: bugar-meszaros@unideb.hu, barnabas.bugar-meszaros@sztaki.hu
[2]E-mail: eichhardt.ivan@unideb.hu, ivan.eichhardt@sztaki.hu
[3]E-mail: gazdag.sandor@unideb.hu, gazdag.sandor@sztaki.hu
[4]E-mail: majdik@unideb.hu, majdik.andras@sztaki.hu

1 Executive Summary

The global COVID-19 pandemic highlighted once again the importance of disinfection robotics in general and the availability at large scale of the connected technologies. The industrial problem tackled in this project is the 3D mapping of a UV-C disinfection robot that is able to quantise the irradiance on the surrounding surfaces. This project proposed a novel algorithmic solution for the automated irradiance estimation. Therefore providing guarantees, that the required amount of UV-C light to disinfect the surrounding surfaces was successfully applied. Furthermore, the goal of this project is to perform fundamental research in the area of computer vision and mobile robot localization and mapping.

The major results, outcomes and benefits are: (i) a voxel based algorithm to reconstruct the environment and to estimate the 3D irradiance; (ii) a new method to solve the relative pose between two robot positions from deep learned depth and a single affine correspondence, (iii) an accurate appearance-based mono-camera localization system in 3D maps.

The practical usability of the project developments has been demonstrated at proof-of-concept level, and methodological novelties were presented in scientific publications, including a premier, top tier computer vision conference (ECCV 2020).

2 Challenge overview

The industrial problem tackled in this project is the 3D mapping of a UV-C disinfection robot, that can quantise the irradiance on the surrounding surfaces. Furthermore, its goal is to perform fundamental research in the area of computer vision to compute and track the displacement between consecutive robot positions, and to use 3D maps to accurately localize robots.

To summarize the specific challenges and goals of this projects are as follows:

- To develop a 3D mapping algorithm using RGB-D sensors and to quantise the UV-C irradiance on the surrounding surfaces in near real time operation

- To perform fundamental research to speed up the algorithm for relative position estimation between consecutive robot positions
- To investigate and develop a system to accurately localize a mono-camera equipped robot in 3D mesh models and point-clouds

3 Implementation of the project

In the project, the research team led by András L. Majdik (PI) and all team members, Barnabás Bugár-Mészáros, Iván Eichhardt, and Sándor Gazdag fulfilled various research and development tasks in the fields of image processing, geometry, machine learning, deep learning, and spatial artificial intelligence. The supporting company, working on healthcare cleaning and disinfection robots had the role and expertise to help the project in defining the research goals and designing the experimental setup for the irradiance calculation algorithm.

4 The research problem, methods, and results

In the previous years the usage of autonomous robots for cleaning has become very popular as these machines can take care of various chores such as mopping a floor or vacuum-cleaning [1, 2]. The advantage of these robots is that they can robustly operate in challenging unknown environments without any supervision. Most of the robots which were designed for the aforementioned activities are limited to ground operation however in some cases a whole indoor environment has to be disinfected. The distribution of ozone based gas in the air is an effective way to tackle this problem especially if there are many hard-to-reach nooks in the room and the highest level of disinfection is required [3]. Another solution is the usage of UV-C light which is proven to be successful in reducing the number of bacteria and viruses on surfaces and in the air [4]. In this project we focused on the latter and developed a software which can simulate the process of irradiation and compute the UV-C irradiance in the local environment of the robot. See Fig. 1 for a sketch of a UV-C disinfection robot.

Figure 1. Left: sketch of a UV-C disinfection robot. Right: the sensors used in the experiments (Xbox 360 Kinect and Intel RealSense d435i depth camera).

We detailed our solution to this industrial problem in the next Section. Further below, we present the outcomes of the two connected scientific challenges addressed in this research project.

4.1 Relative Pose from Deep Learned Depth and a Single Affine Correspondence

We investigated [5] the challenges and viability of combining deep-learned monocular depth with affine correspondences (ACs) to create minimal pose solvers. Estimating pose is a fundamental problem in computer vision [6] (SfM). In this project is fundamental to estimate the consecutive robot positions with great accuracy and fastest runtime. TO achieve these, the proposed approach uses monocular depth predictions together with ACs to estimate the relative camera pose from a single correspondence. We used an off-the-shelf deep monocular depth estimator [7] to provide a 'relative' (non-metric) depth for each pixel (examples are in Fig. 2) and use the depth together with ACs for estimating the relative pose from a single correspondence.

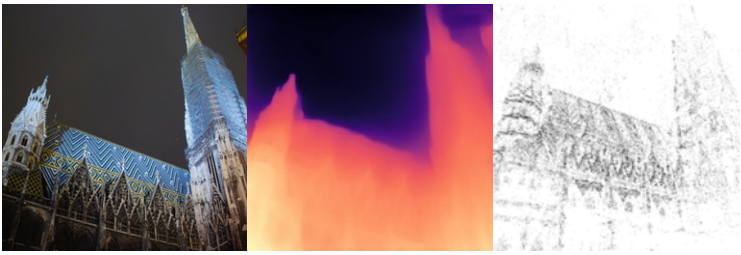


Figure 2. Vienna Cathedral scene of the 1DSfM dataset [8]. From left to right: color image from the dataset, predicted monocular depth [7], and 3D reconstruction using global SfM [9]. The pose graph was initialized by the proposed solver, 1AC+D.

We showed that it is possible to use the predicted imperfect depth signal to robustly estimate the pose parameters. The depth and the AC together constrain the camera geometry so that the relative motion and scale can be determined as the closed-form solution of the implied least-squares optimization problem. The proposed new constraints were derived from general central projection and, therefore, they valid for arbitrary pairs of central-projective views. It was shown that the proposed solver has significantly lower computational cost, compared to state-of-the-art pose solvers. The imperfections of the depth signal and the AC were alleviated through using the solver with a modern robust estimator [10], providing state-of-the-art accuracy at exceptionally low run-time. The reduced number of data points required for the model estimation lead to linear time complexity when combined with robust estimators, *e.g.*, RANSAC [11]. The resulting 1-point RANSAC has to check only n model hypotheses (n = nr. of points) instead of, *e.g.*, $\binom{n}{5}$, which the five-point solver (5PT) [12] implies.

Experimental results. It was demonstrated, on 110 395 image pairs from the 1DSFM dataset [8], that the proposed methods are similarly robust to image noise while being up to 2 orders of magnitude faster, when applied within Graph-Cut RANSAC [10], than the traditional 5PT algorithm. See Fig. 3 for our experimental results on relative pose estimation.

Also, we demonstrated, that when using the resulting pose-graph in the global SfM pipeline [13, 8] as implemented in the Theia library [9], the accuracy of the reconstruction is similar or better compared to when the 5PT algorithm is used.

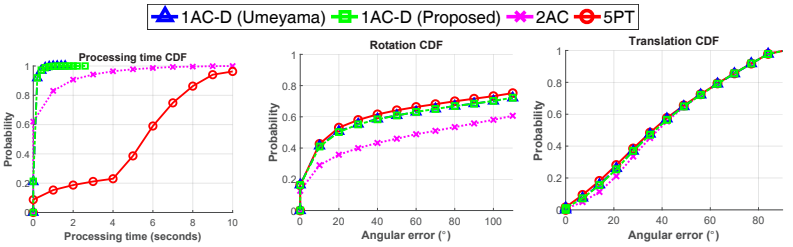


Figure 3. Relative pose estimation on a total of 110 395 image pairs from the 1DSfM dataset, using 1AC+D. The cumulative distribution functions are shown for the processing time (in seconds), angular error of the estimated rotations and translations (in degrees). Being accurate or fast is interpreted as a curve close to the top-left corner.

4.2 Accurate appearance-based mono camera localization in 3D maps

A research challenge for the autonomous cleaning robot is the accurate localization inside maps. Although the localization problem for such robot is constrained to 3 degrees of freedom (DoF), we widened the objective to 6 DoFs. The proposed method can accurately localize a monocular camera in a 3D point cloud or mesh map of the environment. Recently, different image-matching based algorithms were proposed, that render virtual images from 3D models to compute the position of the online aerial view with respect to them. These Simultaneous Tracking and Rendering algorithms were designed and demonstrated in rather small, indoor, and confined settings [14, 15], or assume well structured environments where edges can be robustly detected and aligned [16]. Our approach was inspired by [17], where a Normalized Mutual Information (NMI) based registration method was proposed to localize a ground vehicles on planar roads by matching the online images containing mainly road painted sings to an external map obtained from LiDAR measurements, where everything was removed except the road surface. We advance [17] by extending the algorithm to localize flying MAVs in 3D environments, therefore solving the full 6 DoF pose estimation problem.

4.2.1 The applied methods: Normalized Mutual Information (NMI)

The mutual information of two images is defined using the entropy of the histogram of the images, and the joint entropy of the joint histogram of the two images. This is formulated usually as in (1), where H_A and H_B are the Shannon entropies of the images, and H_{AB} is the joint entropy of the two image.

$$MI(A,B) = H_A + H_B - H_{AB} \tag{1}$$

$NMI_{SUC}(A,B)$ (2)is a normalized form of $MI(A,B)$ which is more robust against differences in overlap [17, 18, 19].

$$NMI_{SUC}(A,B) = 2\left(1 - \frac{H_{AB}}{H_A + H_B}\right) \tag{2}$$

4.2.2 Proposed algorithm

The outline of the proposed algorithm can be seen on Fig. 4. We use an iterative pose refinement algorithm explicated on the right of the flowchart where an initial pose estimation is refined. Virtual images are rendered around the pose estimation along the 3 translational DoFs and the camera image is warped along the rotational DoFs. Then the refined pose is calculated using the image pairing with the highest NMI value. This method is iterated until the end condition is met.

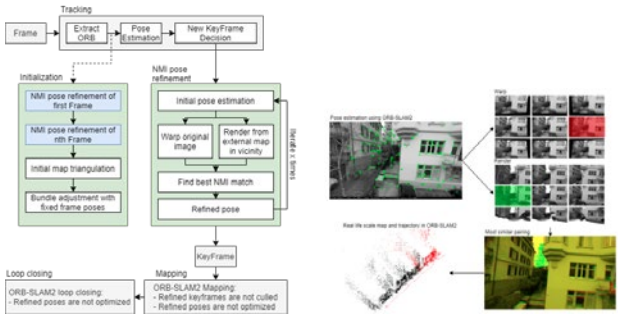


Figure 4. Left: flowchart of the algorithm. Right: outline of the steps of the algorithm.

During the real world scale initialization method the pose of the first and the nth frame is refined with this method and then using the matched ORB feature points an initial map is triangulated and optimized with bundle adjustment. To correct the drift of the ORB-SLAM2 algorithm the key frame poses are iteratively refined and then the local map is optimized. This results in a trajectory and online map, that is consistent with the offline map of the environment, thus the drift of the underlying SLAM algorithm is eliminated.

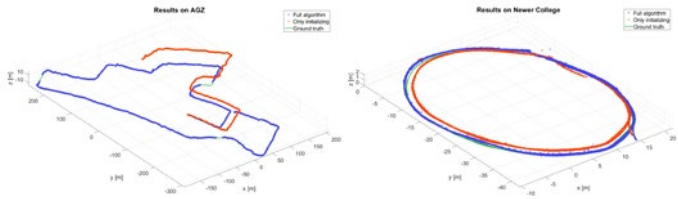


Figure 5. Left: Results on ZU-MAV mesh model, Right: Results on Newer College point cloud model. For both datasets the ground truth is visualized as a green line, the calculated camera positions are blue and with red we visualized the positions calculated by ORB-SLAM2 without pose refinement, after it was initialized with correct scale.

4.2.3 Results

We tested the algorithm on the Zurich urban micro aerial vehicle dataset (ZU-MAV) [20] and the Newer College dataset [21]. See Fig. 5 for the resulting trajectories.

5 Solution of the industrial problem and its benefits

The visualization of UV-C light irradiation created by an autonomous disinfection robot is a challenging task. Firstly, the environment of the machine has to be reconstructed. While the low computational cost of this process is essential, the 3D representation should preserve as many details of the surrounding objects as possible. During its movement the robot builds a global 3D map along the whole path, however in the simulation those parts of the model have to be selected which are currently affected by the UV-C light. To create a realistic simulation of the process we need to investigate those factors which have a significant impact on the efficiency of irradiation. Although the precise computation of a UV-C light tube's characteristic is out of the scope of this work, the most important factors in the calculation should be taken into consideration. The last part of the problem is the visualization of the computed irradiance on the reconstructed surfaces. On the one hand a time independent model has to be created in order to demonstrate the impact of the previously mentioned factors. On the other hand a dynamic simulation is also necessary to illustrate how the absorbed UV-C irradiation increases on the surrounding surfaces as the time goes by.

5.1 Irradiance calculation and simulation software for autonomous UV-C disinfection robots

The development of the simulation software of UV-C light irradiation requires access to the sensors of the autonomous robot. This task can most generally be accomplished using the Robot Operating System (ROS) [22]. In addition to robot communication ROS provides a basis for 3D reconstruction and visualization. To efficiently store the model of the environment we used octomap [23] which is an octree-based voxel representation. The mapping of the surrounding objects into the octomap was first accomplished by an Xbox 360 Kinect and later by an Intel RealSense d435i depth camera. The devices can be seen on the right side of Fig. 1, while the RGB and depth images of the RealSense camera are shown on Fig. 6.

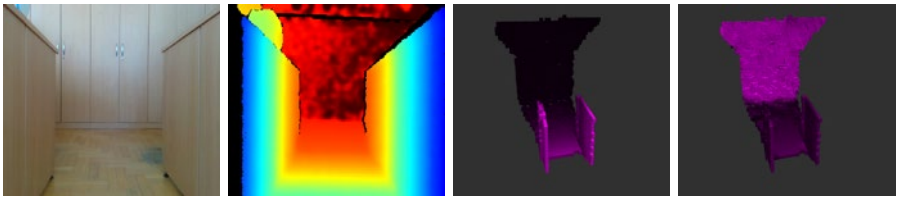


Figure 6. From left to right: RGB image of the RealSense camera, depth image of the RealSense camera, impact of distance on irradiance, impact of the angle of incidence on irradiance.

As we assumed in the simulation that the UV-C light source coincides with the camera center, the currently radiated parts of the model can be selected using either the depth point cloud provided by the device or the field of view of the camera. The former solution is time-consuming, however opposite to the latter it does not take those voxels into account which are covered by closer objects which results in a more precise simulation. To visualize the estimated irradiance on the surfaces we

considered some fundamental factors which affects the calculation. While the impact of the distance r from the light source and the angle of incidence Θ are described by

$$I(r) = \frac{k}{r^2} \tag{3}$$

$$I(\Theta) = I_{dir} \cdot cos(\Theta) \tag{4}$$

where k is a constant and I_{dir} denotes a beam of irradiance perpendicular to the surface. The affect of these factors on UV-C light irradiance is illustrated on Fig. 6.

5.2 Results of the proposed system

The proposed algorithm can not only compute the irradiance on surrounding surfaces based on distance and the angle of incidence but it is capable of visualising the process over time which is important as the efficiency of disinfection depends on the time of irradiation. If the dynamic simulation is selected, the software regularly updates the color of the currently affected voxels according to the computed irradiance multiplied by the elapsed time. The result of the dynamic simulation is illustrated on Fig. 7. The structure of the software allows offline calculation of irradiance if the 3D model of the environment, the trajectory of the motion and the time of irradiation are provided. However, a more expressive near real time simulation can also be achieved if the number of voxels in the 3D map is limited. Finally, as the algorithm was developed using ROS, it can easily be adapted to different mobile robots and depth cameras.

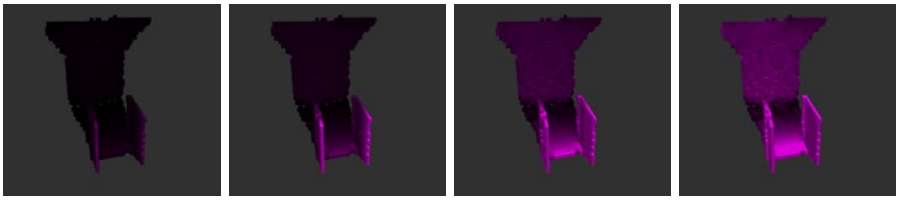


Figure 7. Dynamic simulation of irradiation computed from both distance and angle of incidence over time.

6 Conclusions

To conclude, in this project we created a simulation software for autonomous UV-C disinfection robots, while providing novel solutions for relative pose estimation and accurate mono camera localization. In the future we aim to continue our research and to develop new methods for computer vision and other robotic applications.

References

[1] Y. Hong, R. Sun, R. Lin, S. Yu, and L. Sun, "Mopping module design and experiments of a multifunction floor cleaning robot," in *Proceeding of the 11th World Congress on Intelligent Control and Automation*. IEEE, 2014, pp. 5097-5102.

[2] M. Kukde, S. Nagpurkar, A. Dhakulkar, and A. Amdare, "Automatic & manual vacuum cleaning robot," *International Research Journal of Engineering and Technology (IRJET)*, Vol. 5, No. 02, pp. 2196-2198, 2018.

[3] D. Zoutman, M. Shannon, and A. Mandel, "Effectiveness of a novel ozone-based system for the rapid high-level disinfection of health care spaces and surfaces," *American journal of infection control*, Vol. 39, No. 10, pp. 873-879, 2011.

[4] C.-C. Tseng and C.-S. Li, "Inactivation of viruses on surfaces by ultraviolet germicidal irradiation," *Journal of occupational and environmental hygiene*, Vol. 4, No. 6, pp. 400-405, 2007.

[5] I. Eichhardt and D. Barath, "Relative pose from deep learned depth and a single affine correspondence," in *Proc. European Conf. on Computer Vision.* Springer, 2020, pp. 627-644.

[6] D. Nistér, "An efficient solution to the five-point relative pose problem," *IEEE Trans. Pattern Analysis and Machine Intelligence*, pp. 756-770, 2004.

[7] Z. Li and N. Snavely, "MegaDepth: Learning single-view depth prediction from internet photos," in *Computer Vision and Pattern Recognition*, June 2018.

[8] K. Wilson and N. Snavely, "Robust Global Translations with 1DSfM," in *Proc. European Conf. on Computer Vision*, 2014, pp. 61-75.

[9] C. Sweeney, "Theia multiview geometry library," http://theia-sfm.org.

[10] D. Barath and J. Matas, "Graph-cut RANSAC," in *Computer Vision and Pattern Recognition*, 2018, pp. 6733-6741.

[11] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, 1981.

[12] H. Stewenius, C. Engels, and D. Nistér, "Recent developments on direct relative orientation," *Journal of Photogrammetry and Remote Sensing*, Vol. 60, No. 4, pp. 284-294, 2006.

[13] A. Chatterjee and V. Madhav Govindu, "Efficient and robust large-scale rotation averaging," in *Proc. International Conf. on Computer Vision*, 2013, pp. 521-528.

[14] S. Oishi, Y. Kawamata, M. Yokozuka, K. Koide, A. Banno, and J. Miura, "C*: Cross-modal simultaneous tracking and rendering for 6-dof monocular camera localization beyond modalities," *IEEE Robotics and Automation Letters*, Vol. 5, No. 4, pp. 5229-5236, 2020.

[15] K. Ok, W. N. Greene, and N. Roy, "Simultaneous tracking and rendering: Real-time monocular localization for mavs," in *Proc. International Conf. of Robotics and Automation*, 2016, pp. 4522-4529.

[16] K. Qiu, T. Liu, and S. Shen, "Model-based global localization for aerial robots using edge alignment," *IEEE Robotics and Automation Letters*, Vol. 2, No. 3, pp. 1256-1263, 2017.

[17] R. W. Wolcott and R. M. Eustice, "Visual localization within lidar maps for automated urban driving," in *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Sep. 2014, pp. 176-183.

[18] A. Melbourne, G. Ridgway, and D. J. Hawkes, "Image similarity metrics in image registration," in *Medical Imaging 2010: Image Processing*, Vol. 7623. SPIE, 2010, pp. 962 -971.

[19] P. Viola and W. M. Wells, "Alignment by maximization of mutual information," in *Proc. International Conf. on Computer Vision*, 1995, pp. 16-23.

[20] A. L. Majdik, C. Till, and D. Scaramuzza, "The zurich urban micro aerial vehicle dataset," *The International Journal of Robotics Research*, Vol. 36, No. 3, pp. 269-273, 2017.

[21] M. Ramezani, Y. Wang, M. Camurri, D. Wisth, M. Mattamala, and M. Fallon, "The newer college dataset: Handheld lidar, inertial and vision with ground truth," 2020.

[22] M. Quigley, K. Conley, B. P. Gerkey, J. Faust, T. Foote, J. Leibs, R. Wheeler, and A. Y. Ng, "Ros: an open-source robot operating system," in *ICRA WS on Open Source Software*, 2009.

[23] A. Hornung, K. M. Wurm, M. Bennewitz, C. Stachniss, and W. Burgard, "OctoMap: An efficient probabilistic 3D mapping framework based on octrees," *Autonomous Robots*, 2013.

Optimizing train rescheduling with reinforcement learning

Lehel Csató[1], **Sándor Kolumbán**[2], **Anikó Kopacz**[3] and **Ágnes Mester**[4]

[1]Babeş-Bolyai University, Cluj-Napoca, Romania,
[234]Babeş-Bolyai University, Cluj-Napoca, Romania and Széchenyi István University, Győr, Hungary,

[1]E-mail: lehel.csato@ubbcluj.ro
[2]E-mail: sandor.kolumban@ubbcluj.ro
[3]E-mail: aniko.kopacz@ubbcluj.ro
[4]E-mail: agnes.mester@ubbcluj.ro

1 Executive Summary

Train scheduling and traffic flow optimization are major challenges in the transportation world. As railway networks are facing a need to increase their transportation capacity and train rescheduling is still mostly done by human experts, the research in automated traffic optimization is more prominent than ever.

In order to efficiently test different approaches for automated train scheduling, the Swiss Federal Railways has developed a railway traffic simulator, called the Flatland environment. Then, a series of competitions under the name Flatland Challenge has been organized in order to study various traffic optimization approaches, promoting especially the use of reinforcement learning methods. As the industrial partner, Accenture, provides consulting services for many railway companies, it initiated the investigation of current reinforcement learning techniques for approaching the vehicle rescheduling problem.

In this paper, we present preliminary results in the context of the Flatland Challenge. We propose a new observation method by characterizing local conflict situations, then we use a reinforcement learning based neural network for rescheduling. We test our approach on a series of simulations, in accordance with the Flatland competition. We stipulate that the proposed technique holds promising ideas, worth developing further, in the context of train traffic management.

2 Challenge overview

Automated traffic management and train rescheduling are complex and challenging optimization problems, which are becoming increasingly pressing issues nowadays. These are real-world problems faced by many transportation and logistics companies around the world such as the Swiss Federal Railways (SBB) or the Deutsche Bahn.

Currently, train rescheduling is done by human dispatchers, which is of course suboptimal and unstable in terms of scalability and individual dependencies. As the problem is NP-complete, the proposed approaches to its solution are diverse and many of them rely on heuristic techniques, see e.g. [3, 4, 10, 11] and references therein.

Recently, there has been an increasing interest in reinforcement learning (RL) and, in particular, its application to traffic management and train rescheduling, see e.g. [1, 5, 12] and [7].

In 2020, NeurIPS – in collaboration with SBB and Deutsche Bahn – proposed a train routing competition called the Flatland Challenge [2], with the specific objective of using reinforcement learning methods to tackle train traffic management on complex railway networks. The goal is to create an adaptive controller which plans the trains' routes and actions, and has the ability to adapt to possible delays and malfunctions.

In the light of the above challenges, we investigate a potential approach to the train traffic management problem in the context of the Flatland Challenge, by proposing a reinforcement learning method featuring Q-learning.

3 Implementation of the project

The industrial partner is Accenture, a software development partner to many railway institutions. As the company has a strong inclination for incorporating artificial intelligence into its main resources, Accenture proposed the investigation of the vehicle rescheduling problem formulated by the Flatland Challenge, in order to leverage state of the art machine learning techniques.

The academic partners developed a potential approach to the optimization problem, by using a combination of analytical methods and reinforcement learning techniques. The algorithm was implemented in the framework provided by the Flatland Challenge, then tested and evaluated in accordance with the competition's evaluation metrics.

4 The research problem, methods, and results

The vehicle rescheduling problem (VRSP) formalized by the Flatland Challenge is a highly complex optimization problem. On the one hand, the large number of trains, tracks and switches yields an exponential-growth of rerouting options, which makes the global optimization of the full problem infeasible. On the other hand, the stochasticity introduced by the disruption incidents requires constant observation and efficient dynamic rescheduling.

Our approach to the VRSP relies on the identification and characterization of local conflicts, which is then used as data for a model-free RL algorithm featuring a Q-network in order to handle global scheduling.

4.1 The research problem

The objective of the Flatland Challenge [2] is to address the problem of efficient train traffic management featuring a complex railway infrastructure and dense traffic. The simulation framework is provided by the Flatland library (see [6] and [8]), which is an open-source toolkit that implements a multi-agent grid world environment (see Figure 1), designed specifically to facilitate the development of multi-agent reinforcement learning algorithms in Python. In this section we outline the basic elements of the Flatland framework, for more details see [6], [8] and [9].

Fig. 1. Visualisation of a simple Flatland environment

Flatland is a discrete time simulator featuring a two-dimensional grid world environment, which consists of different cells and supports multiple trains and train stations. The trains represent the agents, each of them having an associated speed profile, which is a fraction in the $(0, 1]$ interval indicating the cell percentage a given train travels under one environmental time step.

A Flatland test environment consists of a randomized railway network and a set of trains, which are assigned a starting position, a predetermined moving direction and a target station (see Figure 1). Given this configuration, the goal is to *minimize the time* it takes for all agents to arrive to their respective destination. In order to achieve this objective, Flatland offers a public benchmark for experimenting with multi-agent RL algorithms, where the agents move by executing different actions, request environmental information under the form of observations, and get feedback by receiving rewards.

Generally, an agent can perform a limited set of actions: move forward, left, or right, stop moving or do nothing (i.e. continue the previous action). However, in most cases only a few actions are valid, as the trains' feasible movements are determined by the given railway system. This highly constrained action space implies a vast imbalance between the impact factor of different actions, as a single wrong decision can have critical effects.

Similarly, each agent can request information about the current state of the environment. The Flatland library provides three basic observation models, which are as follows: local grid, global grid and tree observation. However, a main objective of the challenge is to *develop suitable observations* containing the most relevant information in order to implement an efficient reinforcement learning approach. Therefore, Flatland promotes the possibility of designing custom observations.

During a simulation episode, at every time step t, each agent i receives a local reward $r_l^i(t) = -1$ if it is moving or stopped along the railway, and $r_l^i(t) = 0$ if it has reached its target station. Furthermore, the global reward signal is usually $r_g^i(t) = 0$ for each agent, unless every train has reached its respective destination, then it becomes $r_g^i(t) = 1$ and the episode terminates; an episode also ends if the maximum number of time steps is reached. The primary objective of the Flatland Challenge is to maximize the cumulative reward signal

$$r = \sum_{i=1}^{N} \sum_{t=1}^{T} r_l^i(t) + r_g^i(t), \tag{1}$$

for every episode over an open-ended sequence of test environments, where T and N are the total number of time steps and the number of agents of a given episode.

A significant aspect of the challenge is the effect of stochastic events manifested in the form of train malfunctions, which are simulated using a Poisson process. During a disruption period the malfunctioning train is unable to move, blocking other agents'

trajectories and causing delays. The stochastic events generate the necessity to observe the environment at all times and, in case, call for the rescheduling of trains in order to avoid prolonged delays or deadlocks.

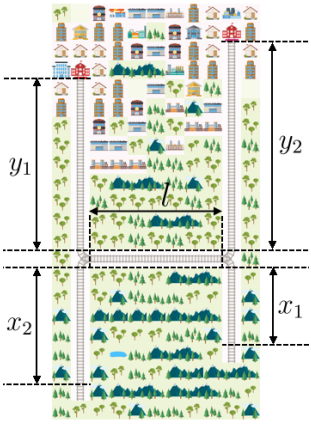
4.2 The applied methods

Our proposed approach relies on identifying and extracting a family of local conflicts which arise between train pairs, referred to in the following as *H-shaped problems*. Using these conflicts, we create a covering of the simulation environment, then we solve these problems individually, using a straightforward analytical method. The solutions of the aforementioned problems then generate feature vectors, which become input of the neural network incorporated by the reinforcement learning model.

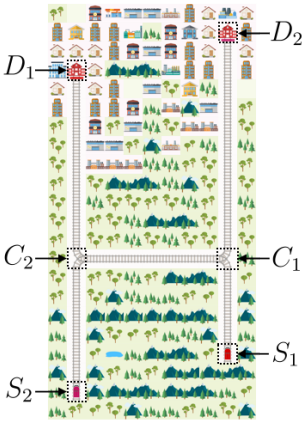
4.2.1 The H-shaped Flatland model

Let us consider the railway infrastructure given by Figure 2, in the sequel referred to as the *H-shaped Flatland*.

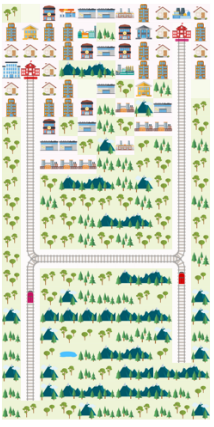
This railway system is uniquely determined by the parameters $x_1, x_2, y_1, y_2, l \in [0, \infty)$, which are arbitrary natural numbers indicating the number of cells between the sections illustrated in Figure 2a. In this environment we study two trains denoted by T_1 and T_2, each having a given speed v_1 and $v_2 \in (0, 1]$. Now consider the scenario illustrated in Figure 2b, where agent T_i is allocated the starting cell S_i and the target station D_i, $i = 1, 2$. This setup is prone to conflict, as both trains need to use the single-track section in the middle, coming from opposite directions.



(a) The infrastructure of the H-shaped Flatland



(b) Trains T_1 and T_2 starting from cells S_1 and S_2



(c) Agent T_1 is required to make a critical decision, entering cell C_1 in the next time step

Fig. 2. The H-shaped Flatland model

Notice that in this H-shaped infrastructure, the route of the trains is uniquely determined, thus the set of decisions of the agents is reduced to the actions **Move**

forward along the given railway, or **Stop**. Moreover, the default action of an agent can be considered **Move**, unless the agent is facing a switch. Therefore, the number of situations when decision making is required drastically reduces to the cases when T_1 or T_2 encounters its first switch on its route (see Figure 2c). Furthermore, these states can be fully characterized by the velocities and positions of the trains.

More precisely, let d_i denote the distance between train T_i and its destination D_i at a given time step t, i.e. $d_i \in [0, x_i + l + y_i + 3]$, $i = 1, 2$. First, let us disregard the possibility of disruptions. Then, given the pair (d_1, d_2), one can determine the optimal policy explicitly, which maximizes the cumulative reward signal (1) of trains T_1 and T_2. Namely, the optimal action of the trains is to move forward, unless T_1 (or T_2, respectively) arrives to the cell preceding the switch C_1 (or C_2). In the case of T_1, this is equivalent with the condition $d_1 = y_1 + l + 3$. Now, if $d_2 \in [0, y_2 + v_2]$, then the optimal action of T_1 is $a_1 =$ **Move**, whereas if $d_2 \in (y_2 + v_2, y_2 + l + 3 - v_2]$, then $a_1 =$ **Stop**. Finally, if $d_2 > y_2 + l + 3 - v_2$, the optimal action a_1 is given by $\operatorname{argmin}(t_1, t_2)$, where

$$t_1 = \left\lceil \frac{d_1 - y_1 - v_1}{v_1} \right\rceil + \max\left(\left\lceil \frac{d_1 - y_1 - v_1}{v_1} \right\rceil, \left\lceil \frac{d_2 - y_2 - l - 3}{v_2} \right\rceil \right) + \left\lceil \frac{l + 3 - v_2}{v_2} \right\rceil$$

and

$$t_2 = 2 \left\lceil \frac{d_2 - y_2 - v_2}{v_2} \right\rceil + \left\lceil \frac{d_1 - y_1 - v_1}{v_1} \right\rceil,$$

$\lceil \cdot \rceil$ being the ceiling function. Namely, if $t_1 \leq t_2$, then $a_1 =$ **Move**, otherwise $a_1 =$ **Stop**. The optimal action a_2 in the case of T_2 can be determined in a similar manner.

It turns out that the majority of conflicts between two trains in the Flatland environment can be modelled and solved optimally using the aforementioned H-shaped policy, when we set aside the possibility of stochastic events.

4.2.2 The observation space

We propose to construct the agents' observations by applying the optimal policy of the H-shaped model. Namely, for every agent T_i we construct a 6-dimensional vector $v^i \in \mathbb{R}^6$ encoding numerical information about the desirability of its 5 actions (see Section 4.1), and we add a 6^{th} entry indicating the case of an inevitable deadlock. These features are determined by the H-shaped algorithm in the following way.

Let us consider a given episode with the set of agents $\mathcal{T} = \{T_i\}_{i=\overline{1,N}}$ present in the environment. For every train T_i one can construct the ordered set of paths $\mathcal{P}(T_i) = \{P_k^i\}_{k=\overline{1,3}}$, which gives the 3 shortest paths between the agent's current location and its target station. These paths are determined by applying the A^* algorithm, which is provided in the Flatland library. We adopt the common policy that the default action of T_i can be considered **Move** forward along the shortest path P_1^i, unless the agent's next cell is a switch. This approach significantly reduces the decision making process and enables us to isolate and focus on key actions having a high impact factor.

In particular, when an agent T_i is facing a switch, we construct a so called decision vector $v^i = (v_r^i), r = \overline{1,6}$, by pairing the 3 shortest paths of T_i with the 3 shortest paths of each train $T_j \in \mathcal{T}$, $j = \overline{1,N}, j \neq i$. Then, for every pair $(P_k^i, P_l^j), k, l = \overline{1,3}$,

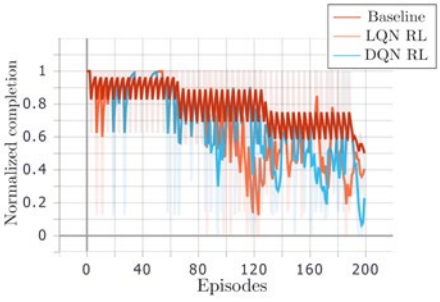
we determine whether the two paths are in conflict, and if so, we fit the H-shaped model to the given setting, as described in section 4.2.1. Finally, based on the optimal action of T_i determined by the H-shaped algorithm, we increment the corresponding entry of the decision vector v^i by 1. In the end, we normalize the vector by the number of votes, so that for each coordinate we have $v^i_r \in [0, 1]$.

4.2.3 The neural network

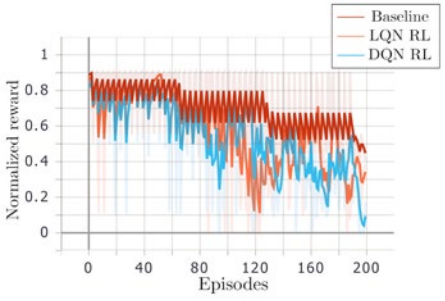
We adopt an off-policy reinforcement learning technique featuring a neural network, which provides the centralized controller for all the agents. We test our method by applying a linear Q-network (LQN) and a deep Q-network (DQN), as well. The input of these neural networks is given by the agents' observation vectors described in Section 4.2.2. These feature vectors then determine the optimal actions of the agents which maximize the reward signals given by equation (1).

4.3 The achieved research results

We compare our method with the greedy baseline policy (which follows every agent's shortest path determined by the A^* algorithm, disregarding other trains) based on the accumulated reward signals obtained during the training process, the length of the generated episodes, the number of agents and the percentage of agents that successfully arrived to their target station. Figure 3 summarizes the performance of the three algorithms over a series of 28×16 test environments, where the number of agents varies between 5 and 11, the number of cities is 4 or 5, and train malfunctions are not permitted. The described model yields promising results for small-scale problems, identifying and solving the majority of critical situations successfully. However, it seems to fail at adapting efficiently to environments with high agent density.



(a) The completion rate of the episodes, given by the number of trains which reached their destination divided by the total number of agents of a given episode



(b) The normalized accumulated reward values of the episodes, given by $\overline{r} = 1 + \frac{r}{T \cdot N}$, where r is the reward signal from (1), while T and N are the total number of time steps and the number of agents of a given episode

Fig. 3. Evaluation of training results for the proposed method applying a linear Q-network and a deep Q-network, respectively, in comparison with the baseline algorithm

5 Solution of the industrial problem and its benefits

The simulations indicate that our method approaches the baseline algorithm both in terms of episode completion and reward values. The proposed model identifies key decisions effectively and shows intelligent behavior when avoiding deadlocks. Therefore, we believe that this paradigm can be utilized with promising results in the train rescheduling problem after further refinement and development.

Additionally, the adaptation of RL techniques has several advantages over the classical methods. First, it generalizes well for new, previously unseen problems and it is computationally more efficient. Secondly, it scales well to more complex problems by the modification of the observation space and in case, the reward function.

6 Conclusions

We proposed a train rescheduling algorithm based on a reinforcement learning method in the context of the Flatland Challenge. We implemented an analytical algorithm for identifying and solving local conflicts arising between two trains, then we designed a corresponding observation space featuring the most relevant information considering these conflicts. This data then translated to actions in the context of the reinforcement learning framework, maximizing the reward function given as the primary evaluation metric of the Flatland Challenge.

Regarding future work, the potential directions of improvement include:

- designing a more complex observation model which features more sophisticated information such as traffic flow, train speed, distance from the conflicting location and probability of malfunctioning;
- introducing a stochastic model in order to adapt to agent malfunctions;
- experimenting with higher-performance reinforcement learning techniques.

The problem of automated train traffic management is a novel challenge having major practical benefits, therefore remaining an important issue to be explored in future research.

References

[1] B. Abdulhai, R. Pringle, and G. J. Karakoulas, Reinforcement learning for true adaptive traffic signal control. *Journal of Transportation Engineering*, 129(3):278-285, 2003.

[2] AIcrowd, Flatland AMLD 2021 Competition. `https://www.aicrowd.com/challenges/neurips-2020-flatland-challenge`, 2020.

[3] A. D'Ariano, M. Pranzo, and I. A. Hansen, Conflict resolution and train speed coordination for solving real-time timetable perturbations. *IEEE Transactions on Intelligent Transportation Systems*, 8(2):208-222, 2007.

[4] S. Dündar and I. Sahin, Train re-scheduling with genetic algorithms and artificial neural networks for single-track railways, *Transportation Research Part C: Emerging Technologies*, 27:1-15, 2013.

[5] X. Liang, X. Du, G. Wang, and Z. Han, A deep reinforcement learning network for traffic light cycle control, *IEEE Transactions on Vehicular Technology*, 68(2):1243-1253, 2019.

[6] S. Mohanty, E. Nygren, F. Laurent, M. Schneider, C. Scheller, N. Bhattacharya, J. Watson, A. Egli, C. Eichenberger, C. Baumberger, G. Vienken, I. Sturm, G. Sartoretti, and G. Spigler, Flatland-RL: Multi-agent reinforcement learning on trains. arXiv: 2012.05893, 2020.

[7] M. Obara, T. Kashiyama, and Y. Sekimoto, Deep reinforcement learning approach for train rescheduling utilizing graph theory, *2018 IEEE International Conference on Big Data (Big Data)*, pages 4525-4533, 2018.

[8] The Flatland Community, Flatland documentation, `https://flatlandrl-docs.aicrowd.com/index.html`, 2019.

[9] The Flatland Community, Flatland Environment, `https://flatland.aicrowd.com/getting-started/env.html`, 2020.

[10] J. Törnquist, Design of an effective algorithm for fast response to the rescheduling of railway traffic during disturbances, *Transportation Research Part C: Emerging Technologies*, 20:62-78, 2012.

[11] J. Törnquist and J. Persson, Train traffic deviation handling using tabu search and simulated annealing, *Proceedings of the 38th Hawaii International Conference on System Sciences*, pages 1-10, 2005.

[12] D. Šemrov, R. Marsetič, M. Zura, L. Todorovski, and A. Srdic, Reinforcement learning approach for train rescheduling on a single-track railway. *Transportation Research Part B: Methodological*, 86:250-267, 2016.

Artifact reduction in CT images

Tamás Dózsa[1], **Sándor Fridli**[2] and **Péter Kovács**[3]

[1] Eötvös Loránd University, Budapest, dotuaai@inf.elte.hu
[2] Eötvös Loránd University, Budapest, fridli@inf.elte.hu
[3] Eötvös Loránd University, Budapest, kovika@inf.elte.hu

1 Executive Summary

Computer tomography provides medical experts with invaluable information on the patient's condition by accurately describing tissue attenuation properties in and around the investigated organs. Because of imperfect hardware and non-ideal working conditions, many different types of errors can appear on reconstructed CT images. These errors are often referred to as artifacts and can be divided into several subcategories [1, 5]. Artifact reduction is still an active area of research [1, 2, 4]. The aim of this project was to provide the industrial partner with an efficient algorithm to mitigate the effects of beam hardening, which causes so-called streak and cupping artifacts [5, 3] on reconstructed images.
The two major challenges of the project were efficient modelling and optimization of the model parameters. Our optimization method is more simple and faster than the state-of-the-art methods [1], with same or better precision of the approximation. The algorithms developed can directly be translated to an industrial edge, the proposed method supports medical experts to obtain relevant information more efficiently.

2 Challenge overview

Cupping artifacts are caused by the hardening of X-ray beams passing through the object. Lower energy (high wavelength) components of the beams are more easily absorbed than high energy (shorter wavelength) components, thus as the beam traverses away from the emitter, its intensity profile changes. For homogeneous materials, this results in artifacts resembling Gauss-curves. There are a number of methods for reducing cupping artifacts including dual-energy methods [12, 13], linearization based approaches [2] and iterative methods [1]. The aim of this project was to develop an efficient iterative algorithm for the correction of cupping artifacts on CT images of homogeneous materials. This method can directly be used during the spectral calibration of the CT machine with artificial phantoms. We note that the proper generalization the proposed methods will be suitable for arbitrary CT images.

3 Implementation of the project

The separation of tasks among the academic and industrial partners was as follows. The academic partner was responsible for elaborating the existing methods, devising suitable models, developing corresponding methods to correct cupping artifacts,

and testing the proposed algorithms. The industrial partner specified the objective of the research, provided measurements, data for testing purposes, and experts for consultation and evaluation.

4 The research problem, methods, and results

Different materials tend to attenuate X-ray beams to different extents, therefore one can learn about the internal structure of an object by measuring attenuation along the beam paths. The linear relationship between the measured beam intensities and the attenuation is given by Beer's law [5]

$$I = I_0 e^{-\int_L \mu},$$

where I stands for the measured intensity, I_0 for the emitted intensity and $\mu : \mathbb{R}^2 \to \mathbb{R}$ is the attenuation function describing the object along path L of a single X-ray beam. The combination of X-ray measurements taken along every angle is called a sinogram [5, 3]. The sinogram is a function of two variables, whose value at a single point is the Radon–transform of the attenuation function:

$$p(r, \varphi) = -\ln \frac{I}{I_0} = \int_{L(r,\varphi)} \mu(t)dt, \tag{1}$$

where r is the distance of the line from the origin and φ is for the direction of its normal vector. Cupping artifacts appear on reconstructed CT images as a result of the physical phenomenon of beam hardening. In Beer's law the X-ray beams are assumed to be monochromatic, which is not the case in real life. Let $\rho(E)$ denote the probability that a photon is on energy level E. Then, the sinogram value for an angle φ and detector r can be expressed as

$$p(r, \varphi) = -\ln \frac{I}{I_0} = -\ln \left(\int_0^{E_{max}} \rho(E) e^{-\int_{L(r,\varphi)} \mu(t,E)dt} dE \right). \tag{2}$$

Equation (2) shows that in the presence of beam hardening, the relationship between the beam intensities and the attenuation function is no longer linear.

4.1 The research problem

The reduction of cupping artifacts is a well researched problem with several possible solutions. These include hardware-based methods [5, 3], dual-energy approaches [12, 13] and linearization methods [2]. Each of these approaches have their own advantages and drawbacks. In this project our aim was to directly estimate the attenuation and related probability parameters of a discretized version of equation (2) through some optimization methods. Such approaches are usually referred to in the literature as iterative methods [1]. They have the advantages of requiring only few assumptions on the object and providing efficient cupping artifacts reduction. Their most cited drawback is their slowness. This is the area in which our results show significant improvement compared to the state-of-the-art.

4.2 The applied methods

First we considered the special case, when the object contains a single material. This is an important sub-problem, since single material phantoms are often used for the spectral calibration of CT machines [5, 8]. Then we continued with the multiple material case.

4.2.1 Objects containing a single material

For objects containing a single material we assumed that
a) the scan region is the unit circle of $\mathbb{R}^2$,
b) $\mu(\cdot, E)$ is constant for a given energy level E,
c) E_{max} and the number of discretization points N in $[0, E_{max}]$ are known.
Then the equation (2) can be simplified as follows

$$p(r,\varphi) = -\ln\frac{I}{I_0} = -\ln\left(\sum_{k=1}^{N}\rho_k \cdot e^{-\mu_k\cdot 2\cdot\sqrt{1-r^2}}\right), \quad r\in[-1;1],\ \varphi\in[0,\pi]. \tag{3}$$

The task is then to estimate the so-called relative bin intensities ρ_k and the associated attenuation coefficients μ_k. Once these parameters have been estimated, artifact correction can be carried out by for example taking their median [1]. By (3), each point on the sinogram depends only on the detector id r, therefore it suffices to estimate the parameters for a single angle φ.

Let φ be fixed, M denote the number of emitter-detector pairs, and $\boldsymbol{p_\varphi} := [p(r_1,\varphi),\dots,p(r_M,\varphi)]^T \in \mathbb{R}^M$. Estimation of the parameters ρ_k and μ_k is to minimize

$$u(\boldsymbol{\rho},\boldsymbol{\mu}) = \|\boldsymbol{p_\varphi} - \boldsymbol{Q}(\boldsymbol{\rho},\boldsymbol{\mu})\|_2^2, \tag{4}$$

where $\boldsymbol{Q}(\boldsymbol{\rho},\boldsymbol{\mu})_j := -\ln\left(\sum_{k=1}^{N}\rho_k\cdot e^{-\mu_k\cdot 2\cdot\sqrt{1-r_j^2}}\right)$ $(\boldsymbol{\rho},\boldsymbol{\mu}\in\mathbb{R}^N,\ j=1,\dots,M)$. This is a nonlinear optimization task involving $2\cdot N$ unknown parameters. In order to simplify the optimization task, we can reformulate (4) as a separable nonlinear least squares (SNLLS) problem [6]. This can be done by introducing $\boldsymbol{y} := \left[1/e^{p(r_1,\varphi)},\dots,1/e^{p(r_M,\varphi)}\right]^T \in \mathbb{R}^M$. Then the minimum of $v(\boldsymbol{\mu}) = \|\boldsymbol{r}_2(\boldsymbol{\mu})\|_2^2$, where

$$\boldsymbol{r}_2(\boldsymbol{\mu}) := \boldsymbol{y} - \boldsymbol{\Phi}(\boldsymbol{\mu})\boldsymbol{\rho} = (I - \boldsymbol{\Phi}(\boldsymbol{\mu})\boldsymbol{\Phi}^+(\boldsymbol{\mu}))\boldsymbol{y} \tag{5}$$

matches the minimum of $u(\boldsymbol{\rho},\boldsymbol{\mu})$. In (5), $\boldsymbol{r}_2$ is called the variable projection operator and $\boldsymbol{\Phi}^+$ stands for the Moore-Penrose inverse of the matrix $\boldsymbol{\Phi}\in\mathbb{R}^{M\times N}$, where $\boldsymbol{\Phi}(\boldsymbol{\mu})_{jk} = e^{-\mu_k\cdot\sqrt{1-r_j^2}}$, $(k=1,\dots,N\ j=1,\dots,M)$.

The gradient of v can be derived from the partial derivatives of $\boldsymbol{\Phi}(\boldsymbol{\mu})$ [6], therefore the above expression successfully halved the number of parameters to be estimated by gradient method. This reformulation is usually referred to as variable projection [6] and our approach is based on it. We implemented three different variations of the optimization algorithm for the single material case: variable projection using the gradient method, Tikhonov-regularized variable projection method (based on [7]) and a variable projection algorithm using nonlinear constraints to disallow negative components in $\boldsymbol{\rho}$.

4.2.2 Objects containing N materials

The so called IGR method, introduced in [1], minimizes the function

$$u(\boldsymbol{\mu}, \boldsymbol{\rho}, \boldsymbol{d}, \boldsymbol{s}) = \frac{1}{D}\sum_{i=1}^{D}\left(\ln\left(\frac{I_i}{I_{0,i}}\right) - \ln\left(\sum_{m=1}^{Q}\rho_m \cdot e^{-\sum_{n=1}^{N}\mu_{n,m}\cdot\sum_{j=1}^{J} l_{i,j}d_j s_{n,j}}\right)\right)^2, \tag{6}$$

where D stands for the total number of X-ray beams used for the scan, N is the (estimated) number of materials present in the object, J is the number of pixels on the reconstructed image, Q is the estimated number of available energy levels and l_{ij} is the length of the intersection between the i-th ray and the j-th pixel. The IGR method tries to find the optimal attenuation coefficients $\boldsymbol{\mu}$, relative bin intensities $\boldsymbol{\rho}$, relative densities $\boldsymbol{d}$ and segmentation parameters $\boldsymbol{s}$. The relative densities account for small differences in density between the pixels, and $s_{n,j} = 1$ if the j-th pixel contains material n. The IGR method optimizes the segmentation, the relative density and the relative bin intensity-attenuation parameters in separate steps. The original algorithm prescribes the use of a gradient method to find the optimal relative bin intensities and the associated attenuation coefficients. One of the main contributions of this project is the simplification of this step. For a fixed φ angle of the detectors, the sinogram in the general case can be written as

$$p(\varphi, r_k) = p(r_k) := -\ln\left(\sum_{i=1}^{Q}\rho_i \cdot e^{-\sum_{j=1}^{N}\mu_{ij}L_j^{\varphi}(r_k)}\right),\ (k = 1, \ldots, M), \tag{7}$$

where M stands for the number of detectors (rays) for a single angle, and $L_j^{\varphi}(r_k)$ denotes the distance the k-th beam traversed through material j. Given a segmentation of the image and a known distance between the detector emitter pairs, this distance can be considered to be known. Similarly to the single material case, one can reformulate the problem of finding the optimal $\boldsymbol{\mu}$ and $\boldsymbol{\rho}$ coefficients as an SNLLS problem. Let $\boldsymbol{y} := [1/e^{p(\varphi,r_1)}, \ldots, 1/e^{p(\varphi,r_M)}]^T \in \mathbb{R}^M$, then for a single angle and fixed $\boldsymbol{d}$ and $\boldsymbol{s}$ parameters one can reformulate (6) as $v(\boldsymbol{\mu}) := \|\boldsymbol{r}_2(\boldsymbol{\mu})\|_2^2$, where $\boldsymbol{r}_2(\boldsymbol{\mu}) := \boldsymbol{y} - \boldsymbol{\Phi}(\boldsymbol{\mu})\boldsymbol{\rho} = (I - \boldsymbol{\Phi}(\boldsymbol{\mu})\boldsymbol{\Phi}^{+}(\boldsymbol{\mu}))\boldsymbol{y}$ and $\boldsymbol{\Phi}(\boldsymbol{\mu})_{k,j} = e^{-\sum_{i=1}^{N}\mu_{i,j}\cdot L_i^{\varphi}(r_k)}$. In this case the matrix $\boldsymbol{\Phi}(\boldsymbol{\mu})$ changes with the angle φ. This, along with the fact that the cost function is non convex mean that the optimal $\boldsymbol{\mu}$ and $\boldsymbol{\rho}$ parameters for each angle are different. In order to overcome this issue, let $U \in \mathbb{N}^+$ be such that $0 = \varphi_1 < \varphi_2 < \ldots \varphi_U = \pi$ holds. Furthermore let $\boldsymbol{\Phi}_m(\boldsymbol{\mu})_{k,j} = e^{-\sum_{i=1}^{N}\mu_{i,j}\cdot L_i^{\varphi_m}(r_k)}$, $(m = 1, \ldots, U)$. Then one can apply variable projection to solve the system of equations

$$\boldsymbol{\Phi}(\boldsymbol{\mu})\cdot\boldsymbol{\rho} := \begin{bmatrix}\boldsymbol{\Phi}_1(\boldsymbol{\mu})\\ \vdots \\ \boldsymbol{\Phi}_U(\boldsymbol{\mu})\end{bmatrix}\cdot\begin{bmatrix}\rho_1\\ \vdots\\ \rho_Q\end{bmatrix} = \begin{bmatrix}\boldsymbol{y}^{\varphi_1}\\ \vdots\\ \boldsymbol{y}^{\varphi_U}\end{bmatrix}, \tag{8}$$

which gives a global estimate for the optimal $\boldsymbol{\mu}$ and $\boldsymbol{\rho}$ parameters. Finally, in order to exploit the physical characteristics of the parameters to be estimated, let us consider

the following. The vector of relative bin intensity parameters $\boldsymbol{\rho}$ constitutes a compound vector in the sense that $\rho_k \geq 0$ and $\sum_{k=1}^{Q} \rho_k = 1$ hold. Notice that the variable projection operator $\boldsymbol{r}_2$ yields the difference between $\boldsymbol{y}$ and its orthogonal projection $P_{\boldsymbol{\Phi}} y := \boldsymbol{\Phi}(\boldsymbol{\mu})\boldsymbol{\Phi}(\boldsymbol{\mu})^{+}\boldsymbol{y}$ onto the column space of $\boldsymbol{\Phi}(\boldsymbol{\mu})$. Replacing this projection with the projection

$$P^{C}_{\boldsymbol{\Phi}(\boldsymbol{\mu})}\boldsymbol{y} := \sum_{k=1}^{Q} \rho_k \cdot \boldsymbol{\Phi}(\boldsymbol{\mu})_k, \ \forall \rho_k \geq 0, \ \sum_{k=1}^{Q} \rho_k = 1, \tag{9}$$

produces a solution which satisfies the above mentioned physical requirements. In our implementation we used linear programming methods from [9, 10] to acquire the relative bin intensities. This solution guarantees certain properties of the relative bin intensities, however it slows down the optimization process greatly and does not guarantee a better approximations of the sinogram.

4.3 The achieved research results

The proposed algorithms were tested on several simulated phantoms. In the single material case we tested three different versions of the variable projection based optimization. The table below shows the true attenuation coefficients for 5 energy levels of a simulated homogeneous phantom, and the estimates produced by the different variations of the proposed method.

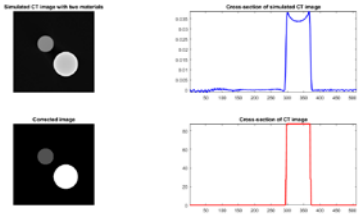
$\boldsymbol{\mu}_k$	VarPro	VarPro - nonlinear conditions	VarPro - Tikhonov regularized
0.999	0.9543	**0.994**	0.965
0.595	0.713	**0.602**	0.550
0.416	0.513	**0.498**	0.534
0.265	**0.268**	0.284	0.250
0.208	0.233	**0.187**	0.100

All three variations provide reasonable estimates, however, in this case disallowing negative relative bin intensities showed the best performance.

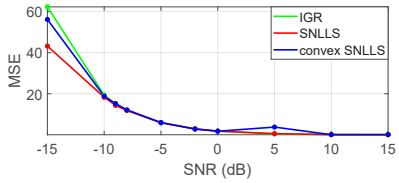
Several different simulations were used to test the modified IGR algorithm. In the simplest case, the scanned object was composed of two distinct materials with different attenuation characteristics. We compared the errors of the approximation, the errors of the estimated parameters and the running times of the original IGR, and the reformulated SNLLS (separable nonlinear least squares) IGR variations. The results are displayed in the table below. The column SNLLS contains results acquired using the variable projection method, and convex SNLLS column refers to the convex projection variant. The notations $\|\boldsymbol{\mu}^{(i)} - \tilde{\boldsymbol{\mu}}^{(i)}\|_2 \ \ (i = 1, 2)$ denote the error of the estimated attenuation coefficients for the first and second material respectively, while $\|\boldsymbol{\rho} - \tilde{\boldsymbol{\rho}}\|_2$ refers to the error of the estimated relative bin intensities. The tests were run by using 3 IGR iterations and a noise-free phantom.

	SNLLS	convex SNLLS	IGR
MSE	$\mathbf{9.0 \cdot 10^{(-5)}}$	$7.4 \cdot 10^{(-4)}$	$1.4 \cdot 10^{(-4)}$
$\|\boldsymbol{\mu}^{(1)} - \tilde{\boldsymbol{\mu}}^{(1)}\|_2$	71.2	28.7	**27.9**
$\|\boldsymbol{\mu}^{(2)} - \tilde{\boldsymbol{\mu}}^{(2)}\|_2$	2314	47.2	**32.6**
$\|\boldsymbol{\rho} - \tilde{\boldsymbol{\rho}}\|_2$	1.53	**0.81**	0.89
Execution time (sec)	**44.18**	91.40	64.54

The results show that all three algorithms result in good approximations, however neither of them can be used to accurately estimate the true attenuation coefficients and bin intensities in the multi material case. This issue was discussed in [1] and can be attributed to the fact, that the cost function (6) is non convex. Therefore, the algorithms can find only local minimums. Nevertheless, a good approximation can be applied for the correction of the cupping artifacts as presented on figure 1a. The proposed SNLLS method in this case not only produced the best approximation of the measured sinogram, but also significantly improved the running time of the algorithm. The same, unfortunately, can not be said about the convex projection variant. Namely, it produces good approximation and a slight improvement on the estimation of relative bin intensities, but it is slower than IGR.



(a) Removed cupping



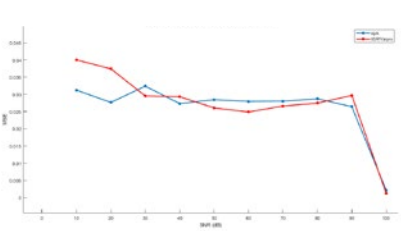
(b) SNR curve of each method

Figure 1: Artifact correction and SNR curves

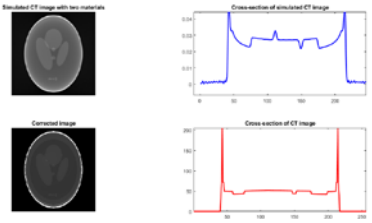
Tests were also run in noisy environments to simulate real life conditions. Figure 1b shows the errors of the approximations for the different methods, when the SNR (signal to noise ratio) is increased gradually. Random noise with normal distribution was used to generate the images. The SNR curves show that the proposed variants of the IGR algorithm match the state-of-the-art's performance in terms of approximation accuracy even in noisy environments. Finally, we simulated the beam hardening effect on the modified Shepp-Logan [11] head phantom. The phantom contains several different overlapping materials, thus presenting a difficult challenge for each algorithm. For a noiseless measurement, the approximation error and running times of each algorithm variant are given in the below table.

	SNLLS	convex SNLLS	IGR
MSE	**0.0012**	0.0293	0.0020
Execution time (sec)	**34.43**	165.16	45.47

Again, the proposed SNLLS variant of the IGR algorithm provides the best approximation of the sinogram, with the shortest running time. Figure 2 shows the results of the noisy tests with an increasing SNR. The last SNR value represents the noiseless phantom. The convex variant was omitted from the noisy tests, as it showed significantly worse results both in terms of MSE and running time. The SNR curve shows that the proposed method matches the state-of-the-art in terms of approximation accuracy even when the CT scan is noisy. Finally, Figure 2b shows how the approximation of the measured sinogram can be used to remove cupping artifacts.



(a) IGR and SNLLS SNR curves



(b) Cupping artifact removed from Shepp-Logan phantom

Figure 2: Artifact correction and SNR curves

5 Solution of the industrial problem and its benefits

Efficiently addressing the effects of artifacts appearing during CT scans can give an edge to the company on the competitive market. The usual method of mitigating the effects of beam hardening is by fine tuning the machine in the setup phase known as spectral calibration. Technicians take scans of simple homogeneous objects with well-known attenuation properties, such as a phantom filled with water to adjust the system and the correction methods used. The proposed algorithms can support such approaches by the single material correction methods. The general case correction algorithms described in section 4.2.2 go further than this. Namely, they can be used to correct artifact-ridden scans while the system is in day-to-day use.

6 Conclusions

In this project we proposed several novel algorithms to address the unwanted effects of cupping artifacts on CT images. Based on the physical relationship between the attenuation properties of the material and the beam intensities, we developed an efficient method to remove cupping artifacts from CT images showing a homogeneous material. Then we extended a state-of-the-art iterative method capable of removing cupping artifacts from images containing arbitrary number of materials. The introduced algorithms depended on the mathematical concept of variable projection, through which we managed to simplify and speed up the complicated nonlinear optimization needed for the artifact reduction task. We tested the developed algorithms using various simulations. Test results match the state-of-the-art in terms of approximation accuracy and show a significant improvement in terms of running speed. The final steps of the project include the handover of the prototype codes to the industrial partner and the implementation of interfaces to enable the use in real circumstances.

References

[1] Van Gompel, G., Van Slambrouck, K., Defrise, M., Batenburg, K. J., de Mey, J., Sijbers, J., Nuyts, J., *Iterative correction of beam hardening artifacts in CT.*

Medical Physics, 38, S36-S49, 2011.

[2] Levi J., Eck BL., Fahmi R. and et al., *Calibration-free beam hardening correction for myocardial perfusion imaging using CT*. Medical Physics, 46(4):1648-1662, 2019.

[3] Gabor T. Herman, *Fundamentals of Computerized Tomography*. Springer-Verlag London, 2009.

[4] G. Bognár, *A No-Reference Image Quality Metric with Application in Low-Dose Human Lung CT Image Processing* Int. J. Adv. Telecom. Elect. Sign. Syst., 5.1, 2016.

[5] Jiang Hsieh, *Computed Tomography Principles, Design, Artifacts, and Recent Advances*, SPIE press, 2003.

[6] G.H. Golub, V, *Pereyra The differentiation of pseudo-inverses and nonlinear least squares problems whose variables separate*, SIAM J. Numer. Anal., 10, 413-432, 1973.

[7] G. Chen, M. Gan, C. L. P. Chen, H. Li, *A Regularized Variable Projection Algorithm for Separable Nonlinear Least-Squares Problems*, IEEE Transactions on Automatic Control, 64, 526-537, 2019.

[8] Kachelriess M, Sourbelle K, Kalender WA, *Empirical cupping correction: a first-order raw data precorrection for cone-beam computed tomography.*, Medical Physics, 33(5):1269-74, 2006.

[9] Michael Grant, Stephen Boyd, *CVX: Matlab software for disciplined convex programming, version 2.0 beta.*, http://cvxr.com/cvx, 2013.

[10] Michael Grant, Stephen Boyd, *Graph implementations for nonsmooth convex programs*, Lecture Notes in Control and Information Sciences, Springer, pages 95-110, 2008.

[11] Shepp, Larry A., Logan, Benjamin F., *The Fourier Reconstruction of a Head Section*, IEEE Transactions on Nuclear Science, NS-21 (3): 21-43, 1974.

[12] L. Yu, S. Leng, C.H. Mccollough, *Dual-energy CT-based monochromatic imaging*, AJR Am. J. Roentgenol. 199 (5), 2012.

[13] Y.B. Chang, D. Xu, A.A. Zamyatin, *Metal artifact reduction algorithm for single energy and dual energy CT scans*, Nuclear Science Symposium and Medical Imaging Conference, IEEE, pp. 3426-3429, 2012.

Epileptic seizure detection via Hawkes processes

László Gerencsér[1], **György Perczel**[2] and **Zsuzsanna Vágó**[3]

[1]Institute for Computer Science and Control, (ELKH SZTAKI), Budapest, and Széchenyi István University, Győr,
[23]Faculty of Information Technology and Bionics, Pázmány Péter Catholic University, (PPKE ITK), Budapest and Széchenyi István University, Győr,

[1]E-mail: gerencser.laszlo@sztaki.hu
[2]E-mail: perczel.gyorgy.miklos@itk.ppke.hu
[3]E-mail: vago@itk.ppke.hu

1 Executive Summary

The industrial problem is the exploration of the possibility of improvements in an emerging technology for the treatment of epileptic patients resistant to medication or surgery, comprising 30% of all. It can be rephrased as the problem of short-term prediction of seizures for the use in *neuro-stimulation,* an emerging moderately invasive therapy in epileptology.

The targeted research carried out in this mini-project lead to *internationally published results*, see [5], providing mathematically rigorous and computationally tested components with a potential application in the development of responsive neurostimulators (RNS). In addition to initiatives in the US, this emerging technology for the therapy of epilepsy has attracted significant attention in the National Institute of Clinical Neurosciences (OKITI). We think that *front-line mathematical technologies*, partially surpassing those of the leaders of the field, see [9] and [6], and the services made accessible to local industry may considerably increase their competitive edge.

2 Challenge overview

With a prevalence of 0,5-1 % epilepsy is one of the most common neurological disorders. Numerous anti-epileptic drugs are available and in some cases even surgical options exist. However, for approximately 30 % of the patients sufficient seizure control cannot be achieved. An alternative emerging technology is delivering electrical stimulation to the area thought to be responsible for the seizure initiation, [1]. In controlling these responsive neurostimulators (RNS) a key element is the reliable forecasting of epileptic seizures based on information extracted from EEG recording.

For the modeling of firing patterns we propose to use a mathematical model reflecting the assumed dynamics of neuronal firing that has been developed in the literature under the name of self-exiting point processes or Hawkes processes, and has been extensively used for earthquake prediction. The main challenges of the project are the articulations of the interaction of neuro-sciences and mathematical modeling and

statistical analysis, in particular advanced techniques of system identification, and the development of an environment for testing our methods for a large variety of simulated and real data.

3 Implementation of the project

Participating Institutions:

- Research group: HU-MATHS-IN SZTAKI, formed within a major Hungarian research institute, the Institute for Computer Science and Control (SZTAKI), Eötvös Loánd Research Network (ELKH).
- Company: National Institute of Clinical Neurosciences (OKITI). The third largest neurosurgical centre of Europe and regional epilepsy center.
- Pázmány Péter Catholic University, Faculty of Information Technology and Bionics (PPKE ITK), accommodating our junior participant in its PhD program within the Roska Tamás Doctoral School of Sciences and Technology.

The principal investigator for this project was *László Gerencsér*, D.Sc. in mathematics, a scientific advisor of SZTAKI, the PhD co-supervisor of our junior participant, with a record of more than three decades of research in stochastic systems, in particular system identification and control. His main role was to explore the relevant literature on point processes, to capitalize on his professional contacts internationally, and to invest serious efforts in looking for possibilities for mathematical innovations.

The junior researcher was *György Perczel*, MD, having been enrolled as a PhD student at PPKE ITK since 2015. His main task was to develop and use a platform for large scale testing of statistical methods in modeling firing patterns via self-exiting point processes both on simulated and real (animal and human) data, and to explore the interaction of the relevant literature in neurosciences and mathematical modelling.

Finally, *Zsuzsannak Vágó*, Ph.D. in mathematics, an associate professor at PPKE ITK was cooperating with the principal investigator in exploring the potentially relevant mathematical methodologies, and served as an internal consultant of our PhD student, assigned by the university, contributing to his add-on mathematical training.

The project has been carried out with the cooperation of National Institute of Clinical Neurosciences (OKITI), in particular with *Loránd Erőss*, Director, Head of the Department of Functional Neurosurgery, PhD co-supervisor of Görgy Perczel, and *Dániel Fabó*, Principle investigator of the Juhász Pál Epilepsy Center.

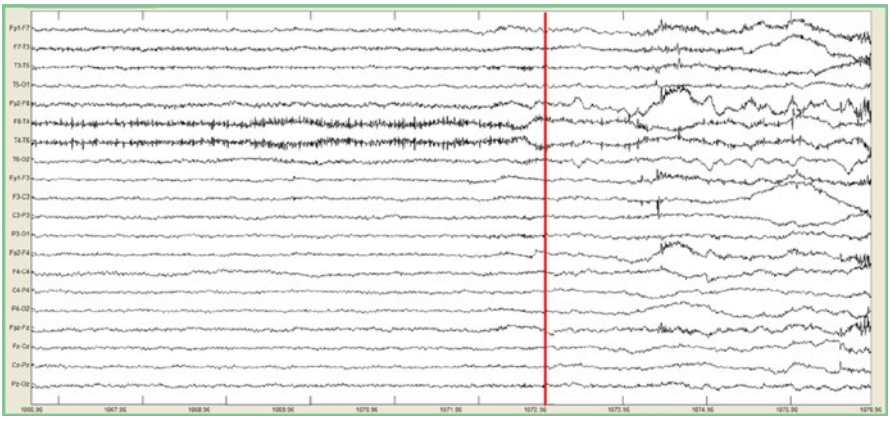
4 The research problem, methods, and results

A brief overview of the *challenges and goals* is as follows:

- To model the brain's electrical activity
- To adapt methods from seismology
- To simulate and identify Hawkes processes
- To provide empirical and theoretical justification

4.1 The research problem

In spelling out the above program we first note that in modeling the brain's electrical activity the traditional set of data we can have is provided by a scalp EEG with as many as 20 channels, reflecting the aggregated electrical activity or firing pattern of specific regions, see below:



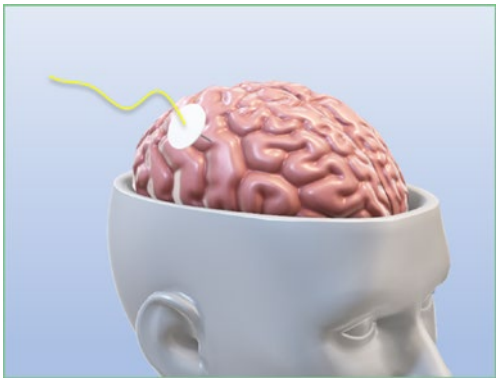
The vertical red line in the EEG signal clearly indicates the onset of a seizure.

Alternatively, we may have an implanted thumbtack electrode collecting data from smaller regions, but providing much higher resolution, see below:

In both cases the challenge is to *predict the onset of seizure* in order to provide a mathematical tool for the new therapy of responsive neuro-stimulation.

Following a widely used analogy between firing patterns around the onset of a seizure and shock waves of an earth quake we adapt methods from seismology. Borrowing a standard mathematical model of the latter, see [2], we are faced with the problem of *simulating and estimating or identifying* so-called Hawkes processes.

Finally, we had a definite program to provide *empirical and theoretical justification* for the maximum likelihood (ML) method that we were going to implement, systematically test and analyze.



4.2 The applied methods

Let us briefly give a few highlights of the brain's electrical activity. The electrical activity of the brain can be described by the firing patterns of more than 10^{10} neurons, arranged in networks. Between neurons, there are interactions, which may be either excitatory or inhibitory. An epileptic seizure is then characterized by a sudden hyper-synchronous neuronal activity. A key observation is that this synchronization can be prevented by electric stimulation of the brain, inspiring efforts to develop alternative therapies in epileptology based on neuro-stimulation, [1].

Now, let us see what can we learn from seismology. Standard tools in modelling shocks in seismology are the so-called self-exiting point processes or Hawkes-processes. A point process is a random sequence of time moments signalling the occurance of specific events, such as level crossings in the seismograph.

Let A be a time-interval and $N(A)$ be the number of events in A and let us define the intensity of the point process at some time t as the average density of occurances around this point:

$$\lambda(t) = \lim_{\Delta \downarrow 0} \frac{1}{\Delta} \mathbf{E}\,[\, N\,[t, t+\Delta) \,|\, \mathcal{F}_t^N\,]. \tag{1}$$

Alternatively, we say that $\lambda(t)$ is the intensity of the point process $N(.)$ is for any $a < b$ we have

$$\mathbf{E}\,[\, N((a,b]) \,|\, \mathcal{F}_a] = \mathbf{E}\,[\int_a^b \lambda(t)\ \mathrm{d}t \,|\, \mathcal{F}_a] \qquad \text{a.s.} \tag{2}$$

Now, a linear self-exiting or Hawkes process is a point process the intensity of which is influenced by past occurances via a linear feedback, more specifically, which satisfies

$$\lambda(t) = \nu + \int_{-\infty}^{t-} h(t-u)\ \mathrm{d}N_u \tag{3}$$

with some baseline intensity $\nu \geq 0$ and impulse response function $h(.) \geq 0$. The above integral should be read as a Riemann-Stieltjes integral. A necessary and sufficient for the existence of the above Hawkes process is the following stability condition:

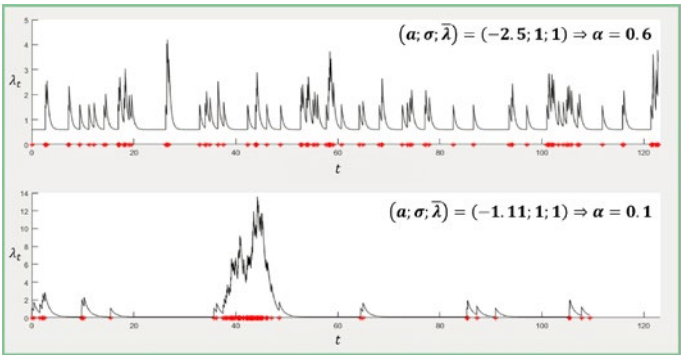
$$\alpha = 1 - \int_0^\infty h(u)\ \mathrm{d}u > 0. \tag{4}$$

Here α is called the stability margin. A Hawkes process can be used to model the activity of a single neuron interacting its neighborhood. A key technical problem is to estimate ν and the impulse response $h(.)$.

4.3 The achieved research results

The *mathematical and computational methods* that have been used can be summarized as follows. *First*, our mathematical models are based on the theory of point processes, in particular Hawkes-processes, which seem to be appropriate to capture the main characteristics of neuronal firing patterns, see [9], [6]. We have developed an elaborate set of tools for simulating and estimating Hawkes-processes.

In the figure below to we present apparently familiar intensity functions of two simulated Hawkes processes with exponential response function. What is surprising though, is that they have the same mean intensities $\bar{\lambda}$, the dramatic difference in their morphology is due to the change in the stability margin from 0.6 down to 0.1. The latter figures have been obtained from modeling real human data by Hawkes processes with exponential response function: *Second*, the estimation or identification



problem is handled via the maximum likelihood (ML) estimation method that have been developed, and analyzed, using partially heuristic arguments, in [7] and [8].

To give a glimpse of the maximum likelihood estimation method for a Hawkes-process let us assume that $h(.)$ belongs to a parametric family, such as the family of exponential impulse responses: $h(u,\theta) = \sigma e^{au}$, $a < 0$, where θ stands for the pair (a,σ). For any feasible, tentative $(\theta,\nu) =: \eta$ define an assumed intensity

$$\lambda(t,\eta) = \nu + \int_0^{t-} h(t-u,\theta)\ \mathrm{d}N_u. \tag{5}$$

Then the negative log-likelihood function, assuming empty history for $t \le 0$ is

$$L(T,\,\eta) = \int_0^{T-} \lambda(t,\,\eta)\ \mathrm{d}t - \int_0^{T-} \log\lambda(t,\,\eta)\ \mathrm{d}N_t. \tag{6}$$

This complicated looking expression is in fact nothing but a formal extension of the log-likelihood ratio of two sets of independent Poisson distributions into continuum,

which has a rigorous foundation. The global minimum of $L(T, \eta)$, within the feasible set, is achieved at $\hat{\eta}_T$ while the true parameter is denoted by η^*.

A serious challenge in applying the ML method is the issue of computability. A key observation presented back in [8] and [7] is that on the case of exponential impulse response function $L(T, \eta)$ is computable. It has been noted in [8] that the sum of exponential functions can be treated similarly. A rigorous elaboration of this remark has been presented in [5] relying on the state-space theory of linear systems.

The theoretical justification and an advanced analysis of the maximum likelihood method has been outlined in [5] by establishing a link between the estimation of Hawkes processes and system identification, in particular the identification of linear stochastic systems. A key technical tool in this is the verification of the remarkable fact that the process $N[t, t+\Delta)$ is L-mixing, as stated in Theorem 6 [5]:

Theorem 1 *Assume that the impulse response function $h(.)$ decays exponentially, $0 \le h(t) \le ce^{at}$ with some $a < 0$. Then for any $\Delta > 0$ the process $N((t-\Delta, t])$ is L-mixing w.r.t. a pair of σ-algebras $(\mathcal{G}_t^-, \mathcal{G}_t^+)$ in the sense of [3].*

Here $\mathcal{G}_t^-$ is monotone non-decreasing, $\mathcal{G}_t^+$ is monotone non-increasing, and ($\mathcal{G}_t^-$ and $\mathcal{G}_t^+$) are independent for all $t \ge 0$. In fact, it can be shown that for any integer $q \ge 1$ and any $a' > a$ we have for all $\tau \ge 0$, with some constant $C = C(a')$,

$$\sup_t \mathbf{E}^{1/q} |N((t-\Delta, t]) - \mathbf{E}[N((t-\Delta, t]) \,|\, \mathcal{G}_{t-\tau}^+]|^q \le Ce^{a'\tau/q}. \tag{7}$$

The condition $h(t) \le ce^{at}$ is not very restrictive and it is certainly satisfied if the support of h is bounded, which is realistic for neuronal networks, see [6].

The above result ensures that the enormous arsenal of system identification is applicable, leading to results such as the strong approximation theorem for the estimation error $\hat{\eta}_T - \eta^*$ stated below, implying standard limit theorems of probability theory.

Assume that the linear dynamics of a Hawkes process is defined via an ARMA-system, with numerators and denominators having the same degree, and leading coefficents 1, parametrized by the remaining coefficients of the numerator and the denominator. Then methodology of [4], developed for ARMA processes, yields the following result:

Theorem 2 *Under the setup given above we have, with error terms $O_M(T^{-1/2})$ such that their L_p-moments are decreasing like $T^{-1/2}$ for any $p \ge 1$,*

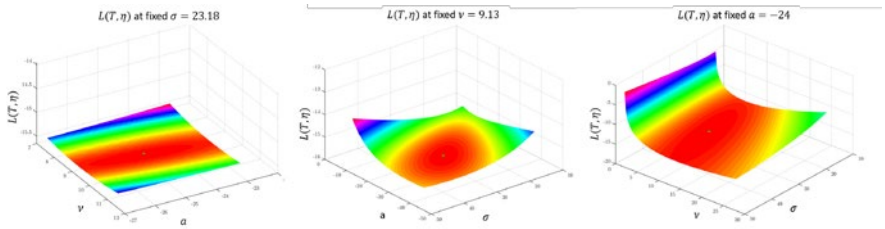
$$\hat{\eta}_T - \eta^* = -(R^*)^{-1} \int_0^T \frac{\partial}{\partial \eta} \lambda(\eta^*, t) \left(\mathrm{d}t - \frac{\mathrm{d}N(t)}{\lambda(\eta^*, t)} \right) \mathrm{d}t + O_M(T^{-1/2}).$$

Thirdly, a definite conibution of the present research is the extensive and systematic empirical study of the ML method with simulated data in a number of scenarios exceeding 1000, followed and paralleled by modeling animal and human data. The sampling rates and the sample periods were: 20kHz and 100 seconds, resp.

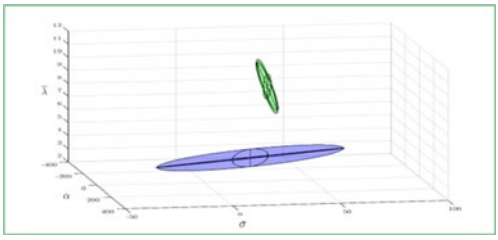
To give an idea of the associated models for a single patient, the estimated parameters prior to the onset a seizure, for the so-called preictal versus the ictal periods, were

found to be $\eta^0 = (-27; 15; 8.89)$ and $\eta^1 = (-470; 430; 8.51)$, respectively. Let us bring your attention to the drastic change in the impulse response, and more importantly, in the stability margins moving down from 0.6 to 0.1, approaching instability!

There is empirical evidence that the cost function associated with the ML method has a unique local minimum, resonating on a similar theoretical result on the identification of ARMA processes. Visual inspection even suggests that the cost function $L(T, \eta)$ is convex. A typical cost function, with one parameter fixed, and the two others varying, is depicted in the figure below, exhibiting shallow valleys, and indicating the existence of a unique stationary point:



Another remarkable fact is that the associated cost function has typically shallow valley, which may result in numerical instability. On the other hand, the same phenomenon causes the associated 95% confidence ellipsoids to have a definite, and highly variable, orientation in $\mathbf{R}^3$, which is in fact favorable for detecting changes in the dynamics, see the figure below:



5 Solution of the industrial problem and its benefits

The *targeted research* carried out in this mini-project lead to published results providing mathematically rigorous and computationally tested components with a potential application in the development of reliable responsive neuro-stimulators. In addition to initiatives in the US, this emerging technology has attracted significant attention in OKITI, and may be favorable also for some local industry.

6 Conclusions

In the currrent project we have extended our previous works for the development of a mathematically sound method for predicting epileptic seizures, relying on a method-

ology borrowed from seismology. In particular, we considerred the problem of fitting self-exiting point processes, or Hawkes processes to point proccesses, representing neuronal firing patterns, derived from EEG data.

The main mathematical advance we have achieved is the elaboration of a key technical tool by showing that the process $N[t, t+\Delta)$ is L-mixing. This observation has lead to a fruitful interaction with the mathematical technology of system identification, leading to among others a strong approximation theorems for $\hat{\eta}_T - \eta^*$.

Our approach provides a solid foundation for the statistical analysis of Hawkes processes the impulse response functions of which are defined via a finite dimensional linear system mapping from $\mathrm{d}N_u$ to λ_t, leading to computable procedures. Based on the current work we have recently developed, tested, and partially analyzed a real-time or on-line version of the ML method for exponential impulse response functions.

The results presented in this summary and their relevance in epileptology have been partially published in the invited paper [5], dedicated to the founding father of non-linear filtering, T.E. Duncan. We think that the front-line mathematical technologies made accessible to OKITI and local industry may increase their competitive edge.

References

[1] Antal Berényi, Mariano Belluscio, Dun Mao, and György Buzsáki. Closed-loop control of epilepsy by transcranial electrical stimulation. *Science*, 337(6095):735-737, 2012.

[2] Daryl J Daley and David Vere-Jones. *An introduction to the theory of point processes: volume II: General theory and structure*. Probability and Its Applications. Springer, New York, 2007.

[3] László Gerencsér. On a class of mixing processes. *Stochastics*, 26(3):165-191, 1989.

[4] László Gerencsér. On the martingale approximation of the estimation error of ARMA parameters. *Systems & Control Letters*, 15(5):417-423, 1990.

[5] László Gerencsér, György Perczel, Loránd Erőss, Dániel Fabó, and Zsuzsanna Vágó. Hawkes processes: some key ideas, links to neuroscience and system identification. *Special Issue in Communications in Information and Systems, dedicated to Tyrone Duncan's 80th birthday*, 2021.

[6] Regis C Lambert, Christine Tuleau-Malot, Thomas Bessaih, Vincent Rivoirard, Yann Bouret, Nathalie Leresche, and Patricia Reynaud-Bouret. Reconstructing the functional connectivity of multiple spike trains using Hawkes models. *Journal of Neuroscience Methods*, 297:9-21, 2018.

[7] Yosihiko Ogata. The asymptotic behaviour of maximum likelihood estimators for stationary point processes. *Annals of the Institute of Statistical Mathematics*, 30(1):243-261, 1978.

[8] Tohru Ozaki. Maximum likelihood estimation of Hawkes' self-exciting point processes. *Annals of the Institute of Statistical Mathematics*, 31(1):145-155, 1979.

[9] Wilson Truccolo. From point process observations to collective neural dynamics: Nonlinear Hawkes process GLMs, low-dimensional dynamics and coarse graining. *Journal of Physiology-Paris*, 110(4):336-347, nov 2016.

One class classifier development for PAP smear image screening

Balazs Harangi[1], **Agnes Baran**[2], **David Kupas**[3] and **Andras Hajdu**[4]

[1] University of Debrecen, Department of Computer Graphics and Image Processing,
[2] University of Debrecen, Department of Applied Mathematics and Probability Theory,
[34] University of Debrecen, Department of Computer Graphics and Image Processing,

[1]E-mail: harangi.balazs@inf.unideb.hu
[2]E-mail: baran.agnes@inf.unideb.hu
[3]E-mail: kupas.david@inf.unideb.hu
[4]E-mail: hajdu.andras@inf.unideb.hu

1 Executive Summary

The main focus of our project was a problem of the cytological industry. During a widely-known cytological examination process, many working hours are invested in the task to determine whether a patient is diseased, or not. Our main target is to assist the industrial partner and build a solution that can optimize the working hours needed for these kinds of examinations. The major challenge regarding this topic is the so-called class imbalance problem, which is a concept from the field of machine learning. To overcome this difficulty, a method had to be researched, that can provide competitive performance while solving the class imbalance problem. Our results suggest, that we have successfully solved this issue. This is a major step for reaching our vision of an automated ranking system, that would help in optimizing the examination process and save many working hours in the field of cytology.

2 Challenge overview

The Papanicolaou smear test (Pap test) [1] is a cervical screening method, where cells are sampled from the outer opening of the cervix. Using special scanners (3D Histech Pannoramic 1000 in our case), digital smears are collected using this process. By examining the cells on these smears, cervical cancer can be detected even in earlier stages.

The cell examination is an extremely time-consuming process, in which cytological experts manually assess each cell on the smear, in order to pinpoint the diseased ones. As such, this is an expensive and relatively tedious task. Our main goal is to develop a screening system driven by machine learning algorithms, that can automatically extract the cells on these smears and classify them into healthy and diseased cases, respectively. The main intent of the planned system is to sort the digital smears based on their severity. In this way, the system would be able to select those cases, where urgent intervention and probable second grading is needed. During the solution of

this challenging task, we are faced with many different obstacles, that we overcome by the application of state-of-the-art machine learning algorithms.

3 Implementation of the project

In this project, the role of the academic team was to find and develop novel solutions relevant to the challenges faced by the industrial partner. Work between the academic partners was split into many different sub-tasks to optimize the workflow. Results that were achieved are a conjunction of the work of each team member. There was an active cooperation with the industrial partner regarding every decision. Furthermore, data was provided completely by the industrial partner as well as some of the annotation steps were also carried out by them.

4 The research problem, methods, and results

During our research, we mainly focused on solving a major problem regarding the industrial case. All the applied methods and proposed solutions will be presented in the next subsections.

4.1 The research problem

The digital smears can contain over 10,000 cells for a single patient. The main issue is that a small number of these cells are genuinely diseased in the case of a diseased patient. This means, that if our goal is to classify the extracted cells, the natural dataset that we can apply for training would be highly imbalanced.

In the field of machine learning, the class imbalance problem is a known and difficult challenge to overcome. In general, an even distribution between the classes is preferred when working with binary classification. Many methods were proposed in order to solve this problem, such as undersampling, oversampling, and cost-sensitive learning. [2].

4.2 The applied methods

The input data that we work with is the digital smear that has been collected with the help of special microscopes. An example of this can be seen in Figure 1. In one of our previous works [3], we solved the problems regarding the segmentation of cells on these smears. Our solution for the task consists of the following: first, we needed to make smaller, 2000x2000 sized images, then we built an ensemble system, applying different segmentation methods based on either fully convolutional networks [4] or super-pixel based segmentation [5]. In this way we have successfully developed an algorithm that has the potential to accurately segment different cell regions from this input image. Using the output of this ensemble system, we can move on to the classification of each cell.

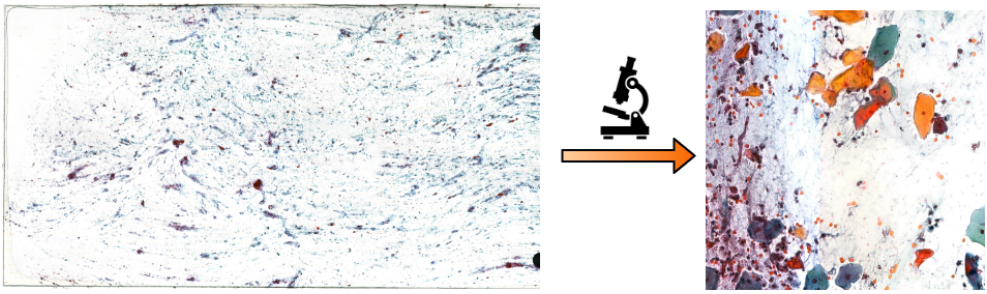


Figure 1: Collection of smaller digital images from Pap smear

In order to build the necessary dataset, after the histological examination of the cells, an annotation was done by experts. They labeled the individual cells into two classes depending on the fact whether they were healthy or diseased cells. In this way, 2527 cells were annotated in total, 2164 of which were healthy, while 363 were diseased. Before training the models, we have split these images into training and test sets. The training set consists of 1727 normal and 294 diseased cells, while the test set is made of 437 normal and 69 diseased cells (20% of the whole dataset). The training set was used to train the models, while the test set was used to measure the model performance. Some examples of healthy and diseased cells can be seen in Figure 2.

As a result of experimenting with many convolutional neural networks (CNN) based solutions, we have selected the appropriate model for the classification of the cells. With additional modifications to the base network, we have managed to classify the cells collected from the smears into classes of healthy and diseased.

A method that we applied to solve the class imbalance problem is the modification of the loss function. The cross-entropy loss can be customized by adding a weight mask. We set it so that the network is penalized more when it misclassifies diseased cells. For this we have used the Real-World-Weight Cross-Entropy (RWWCE) [6], which can be described with the following formula:

$$J = -\frac{1}{M}\sum_{k=1}^{K}\sum_{m=1}^{M}\left[w_{fn}^{k}\cdot y_{m}^{k}\cdot\log\left(h_{\theta}\left(x_{m},k\right)\right)+\sum_{\substack{k'=1\\k'\neq k}}^{K}w_{fp}^{k,k'}\cdot y_{m}^{k}\cdot\log\left(1-h_{\theta}\left(x_{m},k'\right)\right)\right] \tag{1}$$

where M denotes the number of training samples. K is the number of classes, while y_m^k is the target label for the m-th training sample regarding class k. h_θ denotes the model with neural network weights of θ. x_m stands for the m-th input sample. w_{fn}^k indicates the cost of a false negative over a true positive, and $w_{fp}^{k,k'}$ is the cost of a false positive of class k' over a true negative when the true positive is k.

With the goal to further increase the classification performance, we propose to balance out the dataset (detailed in the ongoing paper in the 43rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society) in order to make an equal distribution between the two classes. To achieve this we have used a method based on generative networks. After successfully training the network with the original dataset, we applied it to generate synthetic images. In this way, we generated 1433 diseased images which were added to the diseased cells in the training set, thus obtaining an expanded dataset, that has in total 3454 images, with equally distributed classes. An example of generated synthetic images can be seen in Figure 3. We have experimented with different setups regarding this method, and results are included in the measurements (see later) as "Generative method".

4.3 The achieved research results

Each evaluation mentioned in this section was measured using the same test set, introduced in the previous section. The results of the overall evaluation can be seen in Table 1, where the values regarding the different measurements mean the average and standard deviation derived from 5-fold runs. For the discussion of the results, two standard metrics were selected: recall and precision. In order to focus on the performance regarding the diseased cells, we have also used the balanced accuracy [7], to show a complex view about the performance and to avoid the misleading simple accuracy in case of an imbalanced test set. The balanced accuracy is defined as the mean of recalls calculated on the different classes. It can be formulated as follows:

$$\frac{1}{C} \times \sum_{i=1}^{C} \frac{p_i}{n_i} \tag{2}$$

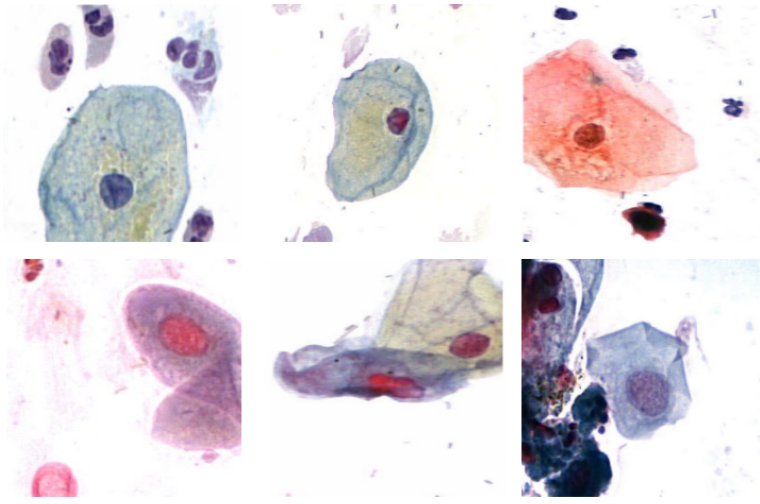


Figure 2: Samples of healthy cells can be seen in the first row, diseased in the second

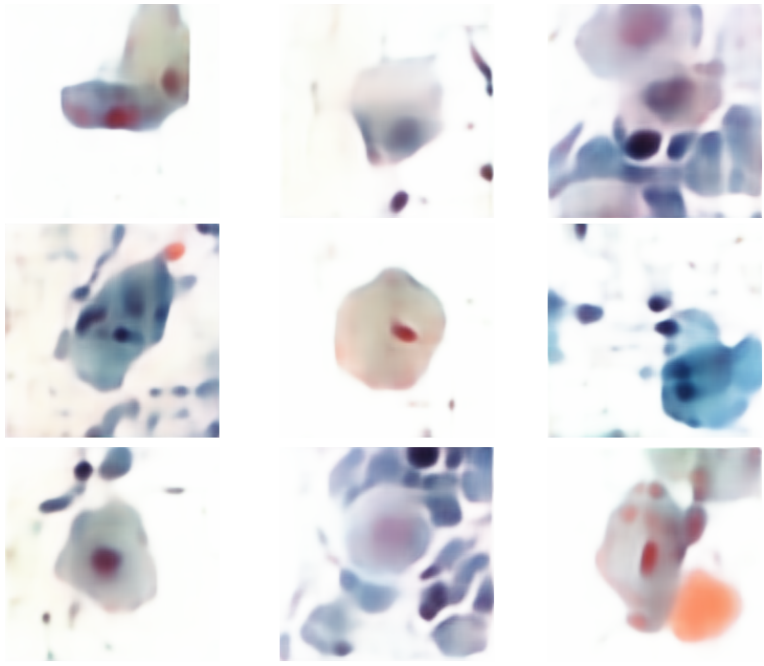


Figure 3: Sample of synthetic images produced with generative networks

where C denotes the number of classes, p_i indicates the number of correct prediction of class i, and n_i is the number of data items in the dataset for class i.

As it can be seen in Table 1, we experienced an increase in the value of the balanced accuracy, when we had used our proposed method. In order to give a thorough overlook of the performances of the different solutions, we also present the corresponding confusion matrices which can be seen in Figure 4. Based on these results, it can be noted that our method correctly classified more diseased cells than the original solution. Since the main goal of these kinds of examinations is to find the diseased cells, we can consider this as a successful development.

	Healthy		**Diseased**		**Total**
	Precision	***Recall***	***Precision***	***Recall***	***Balanced Accuracy***
Benchmark	0.924 ± 0.005480	0.970 ± 0.007070	0.726 ± 0.041590	0.484 ± 0.053670	0.7150 ± 0.021210
Modified loss	0.934 ± 0.008940	0.924 ± 0.020740	0.546 ± 0.047750	0.596 ± 0.054130	0.7580 ± 0.022800
Generative method	**0.944 ± 0.005477**	**0.924 ± 0.026077**	**0.588 ± 0.069785**	**0.650 ± 0.044159**	**0.7880 ± 0.028284**

Table 1: Results of the final evaluation

5 Solution of the industrial problem and its benefits

With our method of synthetic data generation, we have successfully overcome the main challenge of the project. Using the expanded dataset we have managed to get competent classification performance, especially regarding the diseased cells. This is highly relevant since, in the industry, the examination of Pap smears is done manually which is a time-consuming task.

6 Conclusions

To conclude our research project, we need to mention, that the solution for the class imbalance problem is a big step towards our final goal of developing an automatic screening system, which is able to rank the smears. The successful development of such a system could minimize the number of working hours needed to evaluate the digital smears, and provide a way to focus on the most critical cases.

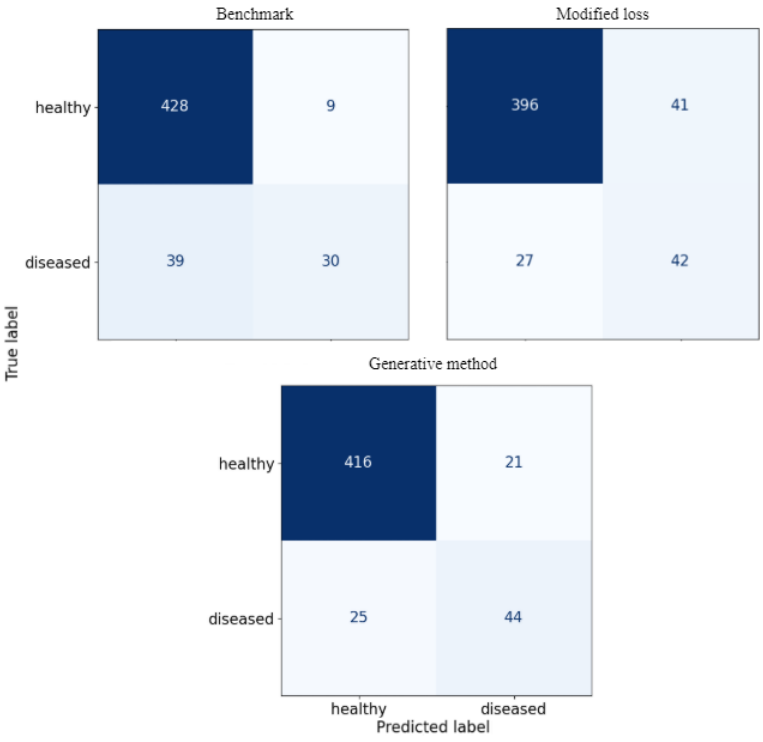


Figure 4: Confusion matrices of the final evaluation

References

[1] G. N. Papanicolaou, H. F. Traut, *The diagnostic value of vaginal smears in carcinoma of the uterus.***this study has been aided by the commonwealth fund. Presented before the new york obstetrical society, march 11, 1941." American Journal of Obstetrics and Gynecology, vol. 42, no. 2, pp. 193–206, March 1941.

[2] S. Wang, W. Liu, J. Wu, L. Cao, Q. Meng, P. J. Kennedy, *Training deep neural networks on imbalanced data sets.* 2016 International Joint Conference on Neural Networks (IJCNN), Vancouver, BC, Canada, 2016, pp. 4368–4374.

[3] B. Harangi, J. Toth, G. Bogacsovics, D. Kupas, L. Kovacs, A. Hajdu, *Cell detection on digitized Pap smear images using ensemble of conventional image processing and deep learning techniques.* 2019 11th International Symposium on Image and Signal Processing and Analysis (ISPA), Dubrovnik, Croatia, 2019, pp. 38–42, doi: 10.1109/ISPA.2019.8868683.

[4] J. Long, E. Shelhamer, T. Darrell, *Fully convolutional networks for semantic segmentation.* The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2015.

[5] Z. Lu, G. Carneiro, A. P. Bradley, *Automated nucleus and cytoplasm segmentation of overlapping cervical cells.* International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2013, pp. 452–460.

[6] Y. Ho, S. Wookey, *The real-world-weight cross-entropy loss function: Modeling the costs of mislabeling.* IEEE Access, pp. 4806–4813, Dec 2019.

[7] A. Gupta, N. Tatbul, R. Marcus, S. Zhou, I. Lee, J. Gottschlich, *Class-Weighted Evaluation Metrics for Imbalanced Data Classification.* unpublished.

Markov chain methods in healthcare II

Ayman Hijazy[1] and **Balázs Dobi**[2]

[1]Eötvös Loránd University, Department of Probability Theory and Statistics, Budapest, Hungary.
[2]University of Debrecen, Faculty of Informatics, Debrecen, Hungary.

[1]E-mail: aymanhijazy@caesar.elte.hu
[2]E-mail: dobibalazs@inf.elte.hu

1 Executive Summary

The first problem we tackled is dealing with the identifiability in chronic disease progression models, these progression is governed by two main parameters: the preclinical intensity and sojourn time. The estimation of these parameters helps in optimizing screening programs (with an additional parameter: sensitivity of the screens), and examine their effect in improving survival. Classic disease progression models are based on strong underlying assumptions. Our main aim is to investigate the effect of these assumptions. We first examine the performance of the model under different parameterizations and investigate the effects of different models for the sensitivity, the inter-screening intervals and misspecification of the used parametric distributions using simulations. Our results indicate that the underlying assumptions have a strong effect on the overall performance of the model. The next step was to investigate the identifiability of a convolution of two independent, non-identically distributed and individually not observed random variables from a theoretical point of view. Our results shed light on the complexity of this problem as well as provide some methods to check the identifiability of the parameters.

The second aim was to update the mathematical background (and the related **R** package) of the Markov chain-based control chart framework. The new theoretical development aimed to solve problems emerging from the modelling of continuous treatment costs due to chronic illnesses. The new methods were successfully incorporated into the framework and implemented in the `Markovchart` package.

2 Challenge overview

The industrial problem can be divided into two parts, the first deals with chronic disease progression processes where screening programs are organised. The aim is the estimation of the embedded parameters within such a framework, this allows dealing with the left censored survival data one usually faces in such a setup. However, the parameters governing the process are not always identifiable.

The second challenge is to create an **R** package that implements Duncan's cycle model for multiple chart types, carries out optimisation with respect to the expected cost per hour and provides visualisation. The aim is to allow sample size, sampling interval and the critical value to be free parameters in the package.

3 Implementation of the project

We worked in close cooperation with our industrial partner the Healthware Consulting Ltd. They provided data, but even more importantly they scrutinized our research and gave many important advice about the practical aspects of the problem. Their contribution was especially important in the analysis of the diabetes treatments.

The young researchers shared the subjects between themselves: A. Hijazy worked on the disease progression topic, while B. Dobi was responsible for the control chart approach and the creation of the **R** package. We consulted regularly about the progress and exchanged ideas, e.g. about the implementation issues.

4 The research problem, methods, and results

4.1 Part I disease progression setup

4.1.1 The research problem

Statistical modeling of natural disease progression aids in understanding its dynamics and forecast its incidence rates. This allows better prevention and treatment plans which improves survival. However, in many cases, some data is not observable, as some diseases have an asymptomatic phase in which the patient does not know he has the sickness yet.

Classic chronic diseases progression models [1, 2] are built by assuming that the times spent in the disease free (S_f) and the asymptomatic state (S_p) are random variables following specified distributions. However, there is a some discrepancy in the results when using the classical approach. We aim to show that these models are sensitive to the chosen distributions and the underlying assumptions.

The parameters of interest in such a process are the preclinical intensity which is the probability of moving from S_f to S_p in $(t, t + dt)$, the sojourn time (amount of time spent in the asymptomatic state S_p) and screening sensitivity. The estimation of these parameters is essential to optimize screening intervals and to correct lead time bias, that is defined as the apparent increase in survival due to early detection by means of screening. However, the underlying assumption are quite strong, and the parameters are not always identifiable. More generally, if we consider two random independent variables X and Y and suppose one can only observe $Z = X + Y$. The main question is if one is able to get unique estimates of the parameters of these random variables when solely observing Z.

4.1.2 The applied methods

Consider the disease natural progression process [1], where individuals propagate between three states: the disease free state S_f, the asymptomatic state S_p and the symptomatic state S_c. Let X and Y be the times spent in state S_f and S_p respectively, which are assumed to be independent. In other words, X is the distribution of the disease onset and Y is the distribution of the sojourn time (amount of time needed for the disease to show symptoms). Furthermore, suppose that X and Y have their

distribution in the families $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$ and $\mathcal{Q} = \{Q_\omega : \omega \in \Omega\}$ respectively. The parameters to be estimated are $\eta = \{\theta, \omega\}$.

Suppose that X and Y have densities f_X and f_y respectively with respect to the Lebesgue measure. Denote by Q_Y the survivor function of Y. Furthermore, suppose that screens occur at time (age) $t_i = t_0 + (i-1)\Delta$ $i = 1, \ldots, K$, where t_0 is the age at the beginning of followup, Δ is the inter-screening time and K is the total number of screens. Assume everyone is disease free at t_0 and denote by $\xi(t_i)$ the sensitivity at a screening exam at age t_i.

Denote the probability of being detected on screen i by D_i and the probability of moving to S_c (showing symptoms) between t_i and t_{i+1} by I_i, these probabilities can be computed as follows [2]:

$$D_k = \sum_{i=1}^{k-1} \xi(t_k)\Big[(1-\xi(t_i))\cdots(1-\xi(t_{k-1}))\Big] \cdot \int_{t_{i-1}}^{t_i} f_X(x)Q_Y(t_k - x)dx + \xi(t_k)\int_{t_{k-1}}^{t_k} f_X(x)Q_Y(t_k - x)dx,$$

and:

$$I_k = \left[\sum_{i=0}^{k-2}(1-\xi(t_i))\cdots(1-\xi(t_{k-1}))\int_{t_{i-1}}^{t_i} w(x)[Q_Y(t_{k-1}-x) - Q_Y(t_k - x)]dx\right] + \int_{t_{k-1}}^{t_k} w(x)[1 - Q_Y(t_k - x)]dx.$$

Using these probabilities, we get a likelihood function similar to the multimomial distribution, therefore:

$$L \propto \prod_{k=1}^{K} D_k^{s_k} I_k^{r_k}(1 - D_k - I_k)^{n_k - s_k - r_k},$$

Our important suggestion is that we propose to incorporate the exact dates of diagnosis of symptomatic patients (z_i) in the likelihood function if they are available, as they carry important information. Then the likelihood is of the form:

$$L_2 = \prod_{t_1=t_{min}}^{t_{max}} \prod_{k=1}^{K} \left[D_k^{s_k}(1 - D_k - I_{k,})^{n_k - s_k - r_k} \prod_{i=1}^{r_k} f_{Z_i}^k(z_i)\right].$$

where $f_{Z_{t_1}}^k(z) = \sum_{i=1}^{k} f_{Z_{t_1}}^{(i,k)}(z)$ and

$$f_{Z_{t_1}}^{(i,k)}(z) = \begin{cases} \int_{t_{k-1}}^{z} f_X(x)f_Y(z-x)dx \text{ if } k = i. \\ \int_{t_{i-1}}^{t_i} f_X(x)f_Y(z-x)dx \prod_{j=i}^{k-1}(1-\xi(t_j)) \text{ if } k > i. \end{cases}$$

Based on these two likelihoods, we ran simulations based on known distributions and studied the performance of the model and examined the effect of the assumptions. A serious concern that was raised was the identifiability of the parameters.

The next step was to give theoretical insight into the identifiability of convolutions of random variables. Namely, let $Z = X+Y$ and suppose that one can only observe Z, where X and Y are independent, individually not observed but usually not identically distributed. Furthermore, let Θ and Ω be subsets of Euclidean r-space and s-space respectively, let $\mathcal{F} = \{F_{\theta,\omega} = P_\theta * Q_\omega\}$, where P_θ and Q_ω are parametric families of distributions, with $\theta \in \Theta$ and $\omega \in \Omega$. Let $Z = (Z_1, \cdots, Z_N)$ be N independent identically distributed random variables with their common distribution belonging to the class $\mathcal{F}$. Then the identifiability under convolution can be defined as:

Definition 1. The family $\mathcal{F} = \{F_{\theta,\omega} : F_{\theta,\omega} = P_\theta * Q_\omega\}$ is said to be identifiable under convolution if $F_{\theta,\omega} = F_{\theta',\omega'}$ implies $\theta = \theta'$and $\omega = \omega'$.

Within this setup, we were able to reach some notable theoretical results, these can be used to check and prove the identifiability in practice. Namely:

Theorem 1. Suppose that the real line is partitioned into $r + s + 1$ intervals A_i such that $P(Z \in A_i) > 0$, then the convolution is identifiable if the transformation from $\Theta \times \Omega$ onto the set:

$$S = \{(h_1, \ldots, h_{r+s}) : (\theta, \omega) \in (\Theta, \Omega),\ h_i(\theta, \omega) = E_{\theta,\omega}(\mathcal{I}(A_i))\}$$

is one-to-one.

Corollary 1.1. Suppose X and Y are arbitrary random variables distributed over known sets $x_1, \ldots, x_r$ and $y_1, \ldots, y_s$ respectively. If $\nexists x_i, y_j, x_k, y_l$, $i \neq k$, $j \neq l$ such that $x_i + y_j = x_k + y_l$ then the convolution Z is identifiable.

Theorem 2. Suppose that X is an arbitrary distribution over a known subset $x_1, \ldots, x_r$ where $x_1 < x_2 < \cdots < x_r$ with positive probabilities p_i for all $i = 1, \ldots, r$ and Y is a random variable with a known exact lower bound $M = \inf\{x : F_Y(x) > 0\}$. We may suppose that $M = 0$ (it can be achieved by translating the distribution of Y and the points x_i in a way that Z will not be changed). Let us consider the support of $X + Y$ and partition it into $A_i = [x_i, x_{i+1})$ $(i = 1, \ldots, r$; let $x_{r+1} = \infty)$. Suppose that the interval $[0, x_2 - x_1)$ can be divided into $s + 1$ intervals denoted by $A_{1,j}$ with $P(A_{1,j}) > 0$ for all $j = 1, \cdots, s + 1$ in a way that the distribution of Y can be determined by these probabilities, i.e. the transformation from $(0, 1) \times \Omega$ onto the set:

$$S = \{(h_1, \ldots, h_{s+1}) : (\theta, \omega) \in ((0, 1) \times \Omega),\ h_i(\theta, \omega) = E_{p_1,\omega}(\mathcal{I}(A_i))\}$$

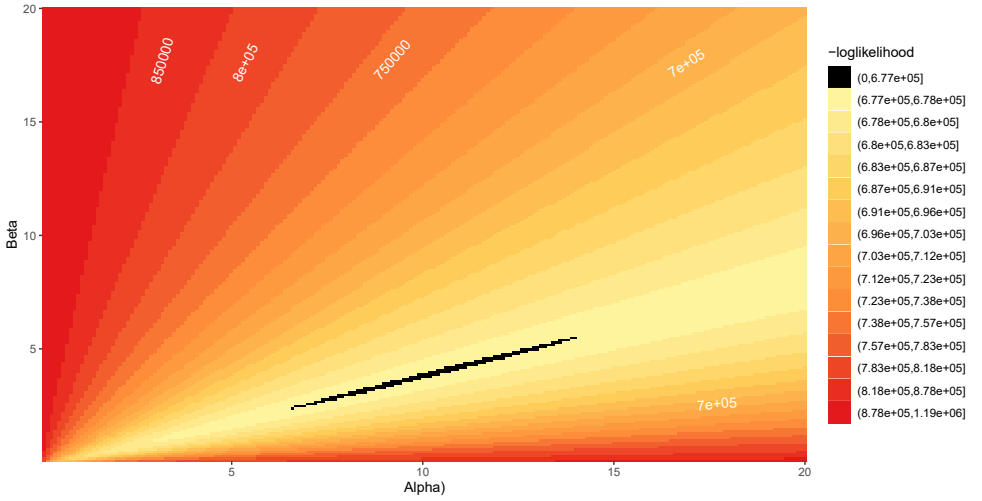
is one-to-one, then the convolution is identifiable.

4.1.3 The achieved research results

The theorems provide a new way to prove the identifiability and to estimate the parameters of convolutions of random variables. It was possible to apply this result to random variables, which are bounded from below (and by symmetry the same holds for distributions bounded from above), where one of the terms is discrete and is distributed over a known subset.

On the other hand, the simulations showed that the classical models are very sensitive to the underlying assumptions. One should take great care when using such an approach and multiple trials with different models are needed before in order to get reliable results. One way to solve problem might be to include more information in the model to stabilize the results such as tumor growth shape and tumor size [5, 6].

Figure 1: Contour plot of the loglikelihood using a lognormal onset and a gamma sojourn time with parameters α and β. The black region represents the likelihood based confidence region.



We have written two papers about our results, the first [8] includes our simulation results and has already been accepted and will be published soon. The second paper [9] deals with the identifiability of convolutions and has been submitted.

4.2 Part II: Extending the `Markovchart` R package

The description of the Markov chain-based framework and the research problem below is just a brief introduction and summary necessary for understanding the mechanics of the Markov chain-based control chart method. For further reading and more detailed descriptions see Zempléni et al. (2004) [10]; Dobi and Zempléni (2019a,b) [11, 12].

In the above-cited papers we have introduced random shift sizes, random repairs and random sampling times to be used in single-sample, one-sided, cost-optimal Shewhart X-charts. We investigated the effect of different parameters with a focus on different shift size distributions. The motivation behind these generalisations and investigations was the application of control charts in patient health monitoring. The methods were implemented in the `Markovchart` **R** package.[13] We have - together with our industrial partner, the Healthware Consulting Ltd. - successfully applied the method on the data of diabetic patients from Hungary.

4.2.1 The research problem

At least part of the repair (i.e. treatment) cost may also occur during out-of-control (OOC) operation (i.e. when the health of the patient is deteriorated but it is not detected), and not just during targeted interventions. These costs are often overlooked, and originally were not taken into account in our theoretical results. However, this is a necessary part of the model since in many medical applications treatment is continually present (e.g., medicine has to be taken daily).

4.2.2 The applied methods

The default function for the repair cost in the random shift size case is a linear function:

$$c_r(v) = c_{rb} + c_{rs}v,$$

where c_{rb} is the base TA (true alarm/repair) cost and c_{rs} is the distance scaling TA cost. v is the distance from the target value, i.e. the desired health level of the patient by some characteristic (e.g. blood sugar level).

Thus the total cost per unit time associated with an alarm state v distance from the target value is

$$c_{v,TA}(v) = \frac{c_s}{h}T_h(v) + c_o(v) + \frac{c_r(v)}{h},$$

and with an OOC state is

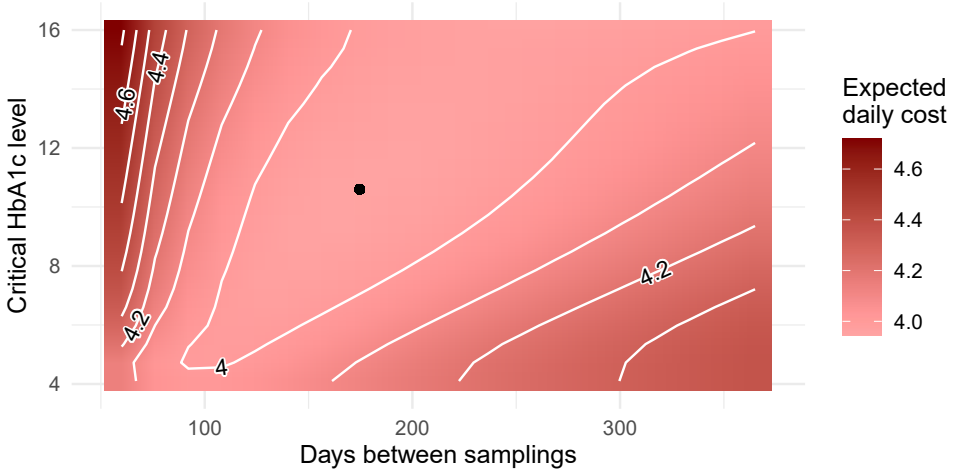
$$c_{v,OOC}(v) = \frac{c_s}{h}T_h(v) + c_o(v) + \frac{p_r c_r(v)}{h}. \tag{1}$$

p_r is the amount (proportion) of repair cost occurring during OOC operation, c_s and c_o are the sampling and OOC costs respectively. As we mentioned before, this is necessary because in many medical applications treatment is continually present. In classical, industrial settings p_r would be 0 in most applications. The sampling and repair costs need to be divided by the length of the sampling interval and the OOC cost should be given by default as per unit time. Whether or not the repair cost is given per unit time or as the total cost of treatment is optional and is parametrisable in the updated `Markovchart` package. $c_o(v)$ and $c_r(v)$ need to be distinguishable in Equation 1 because they may have completely different connections with the process (i.e., distances from the target value) which needs to be reflected in the calculations. The cost of sampling needs to be weighted by the probability of the sampling ($T_h(v)$). This probability is incorporated into the stationary distribution of the Markov chain that describes the transitions between health-states - this is why it is not needed for the repair cost in the equation above.

4.2.3 The achieved research results

The above cost function was incorporated into our framework and its **R** implementation enabled us to model health deterioration-treatment cycles more accurately. Consequently, we updated our diabetes analysis accordingly. A figure summarising the results can be seen below.

Figure 2: Contour plot of expected costs (EUR) related to insulin analogue therapy with $p_r = 0.67$ parameter; HbA1c: glycated haemoglobin (blood sugar) level.



The most important result is the fact that there is a clear optimum. This is not a trivial thing in health care processes, as often trivial solution would be to monitor and intervene as soon as possible to minimise costs. However, it makes no sense in such cases as the blood sugar level, since patients take medications often daily. This problem is solved by including part of the emergency intervention cost (here the 2/3) into the "regular" daily costs.

5 Solution of the industrial problem and its benefits

In the disease progression setup, the simulations shed a light on the seemingly discrepant results obtained by different authors using the same data sets but different assumptions. The theoretical results provide new methods to examine identifiability in practice. Moreover, the updated `Markovchart` package is useful for refining cost-effectiveness models for our industrial partner, as such problems often arise at the company, be it diabetes or any chronic illness-related projects.

6 Conclusions

The new Markov chain-based control chart methods for continuous treatment cost modelling were successfully incorporated into the framework and implemented in the `Markovchart` package.

In the disease progression methods, we now have more accurate methods to estimate the parameters governing the process and correct the lead time bias. We also give a theoretical insight into the identifiability of the parameters.

References

[1] Zelen, M. and Feinleib, M. (1969), On the theory of screening for chronic diseases. *Biometrika*, **56(3)**, 601-614.

[2] Wu, D., Rosner, G. and Broemeling, L. (2005), MLE and Bayesian inference of age-dependent sensitivity and transition probability in periodic screening. *Biometrics*, **61(4)**: 1056-1063.

[3] Abdel-Hameed, M. (1975), A gamma wear process. *IEEE Transactions on Reliability*, **24** (2): 152-153.

[4] Burbea, J. and Rao, C.(1982), On the convexity of some divergence measures based on entropy functions. *IEEE Transactions on Information Theory*, **28**(3): 489-495.

[5] Hijazy, A., Zempléni, A. (2021), Gamma Process-Based Models for Disease Progression. *Methodol Comput Appl Probab*, **23**: 241-255. https://doi.org/10.1007/s11009-020-09771-4

[6] Hijazy, A, Zempléni, A. (2020), Optimal inspection for randomly triggered hidden deterioration processes. Qual Reliab Engng Int. ; 36: 1-16. https://doi.org/10.1002/qre.2707

[7] Sclove, S., Van Ryzin, J. (1969), Estimating the Parameters of a Convolution. *Journal of the Royal Statistical Society.* Series B (Methodological), **31**(1), 181-191.

[8] Hijazy, A., Zempléni, A. (2021), How well can screening sensitivity and sojourn time be estimated (accepted manuscript)

[9] Hijazy, A., Zempléni, (2021), Identifiability of convolutions (submitted manuscript)

[10] Zempléni, A., Véber, M., Duarte, B., Saraiva, P. (2004), Control charts: a cost-optimization approach for processes with random shifts. Appl Stoch Model Bus Ind. **20**(3): 185-200. https://doi.org/10.1002/asmb.521

[11] Dobi, B. and Zempléni, A. (2019a), Markov Chain-Based Cost-Optimal Control Charts for Health Care Data. Quality and Reliability Engineering International, **35**(5): 1379-1395. doi: 10.1002/qre.2518.

[12] Dobi, B. and Zempléni, A. (2019b), Markov Chain-Based Cost-Optimal Control Charts with Different Shift Size Distributions. Annales Universitatis Scientiarum Budapestinensis de Rolando Eötvös Nominatae, Sectio Computatorica, **49**: 129-146.

[13] Dobi, B., Zempléni, A. (2020), Markovchart: Markov Chain-Based Cost-Optimal Control Charts, version 1.1.1, The Comprehensive R Archive Network

Copula-based anomaly scoring and localization for large-scale, high-dimensional continuous data

Gábor Horváth[1], **Edith Kovács**[1], **Roland Molontay**[1], **Szabolcs Nováczki**[2]

[1] Budapest University of Technology and Economics
[2] Nokia, Bell Labs

1 Executive Summary

An anomaly detection method was developed with a special feature: it does not only indicate whether an observation is anomalous or not but also tells what exactly makes an anomalous observation unusual.Hence, it provides support to localize the reason of the anomaly. We demonstrated the usability of the method by analyzing a large data set consisting of the performance counters of a real mobile telecommunication network. Since such networks are complex systems, the signs of sub-optimal operation can remain hidden fora potentially long time. With the proposed procedure, many such hidden issues can be isolated and indicated to the network operator. The anomalies reported by our method were reconfirmed by real network operators.

Our method was awarded with the prestigious Innovation Award of the Pro Progressio Foundation of the Budapest University of Technology. The results have been published in ACM TIST [1], this report is heavily based on that paper.

2 Challenge overview

Mobile networks are complex systems with many parameters, where discovering that the system performance is sub-optimal is difficult for a human operator.

For decades, telecommunication operators have relied on network domain experts to report problems that affect the network performance and customer experience, with the aim of performance monitoring tools. The traditional tools are mainly based on human knowledge, and that the complex and hidden rules in the telecommunication network are not easy to fully detect, capture and utilize.

Thus there is a strong need to revolutionize network management with AI/machine learning technology. This is essential as today's communication networks are extremely complex systems consisting of hundreds of thousands network elements organized in cooperating, coexisting and overlapping technology layers. The network elements generate huge amounts of versatile data for performance monitoring, optimization and troubleshooting purposes. It is already quite cumbersome to tackle these tasks with traditional approaches and the effects will be even more emphasized in case of 5G. Automated solutions are needed that are capable to analyze the raw data and draw conclusions, generate actionable insights using AI technology. One important area in this field is the predictive detection of anomaly patterns that appear in the data.

3 Implementation of the project

The academic and industrial partners have worked tightly on the project with regular project meetings. The academic partner was responsible for the algorithm design and mathematical considerations, while the industrial partner was responsible for providing the data, evaluating and interpreting the results and supervising the project from a business point of view.

4 The research problem, methods, and results

We proposed a novel anomaly detection method based on probabilistic modeling. The main idea is to determine the joint distribution of the random variables corresponding to the features, and to assign anomaly scores to the observations based on the density of the observations. The subspaces where an observation deviates from the majority of the samples indicate which variables are involved in the anomaly. However, the realization of this high-level description of the algorithm faces some challenges that the proposed approach aims to overcome:

- Obtaining the joint distribution of a high-dimensional dataset is difficult. Simple solutions for this problem aim to fit the dataset by a multivariate Gauss distribution or by a mixture of multivariate Gaussians. However, Gaussian density functions have a fast decay, they fail to capture the heavy-tailed behavior where most anomalies take place, making them inappropriate to use for anomaly detection.
- Even if the joint density function of the dataset is available, translating the density (a number between zero and infinity) to an anomaly score (between zero and one) having a physical interpretation, is not straightforward.
- Providing aid to localize the reason of the anomaly in a high-dimensional space, such that a human can interpret it and take the necessary actions, is also challenging.

4.1 The research problem: high-dimensional anomaly detection

The research goal is to find anomalies in a dataset originating from the live commercial LTE network of a Western European mobile operator. Mobile networks are complex systems with many parameters, where discovering that the system performance is sub-optimal is difficult for a human operator.comprises a large number of performance counters collected at thousands of base stations, recorded in every hour. The original set of counters has been filtered based on data quality requirements, which resulted in 54counters and 1,500,000observations. These counters represent a wide variety of measurements: related to the PDCP (Packet Data Convergence Protocol), RLC(Radio Link Control), MAC (Media Access Control) and PHY (Physical) layers, to the handovers, the activity of the schedulers, the UL (upload) and DL

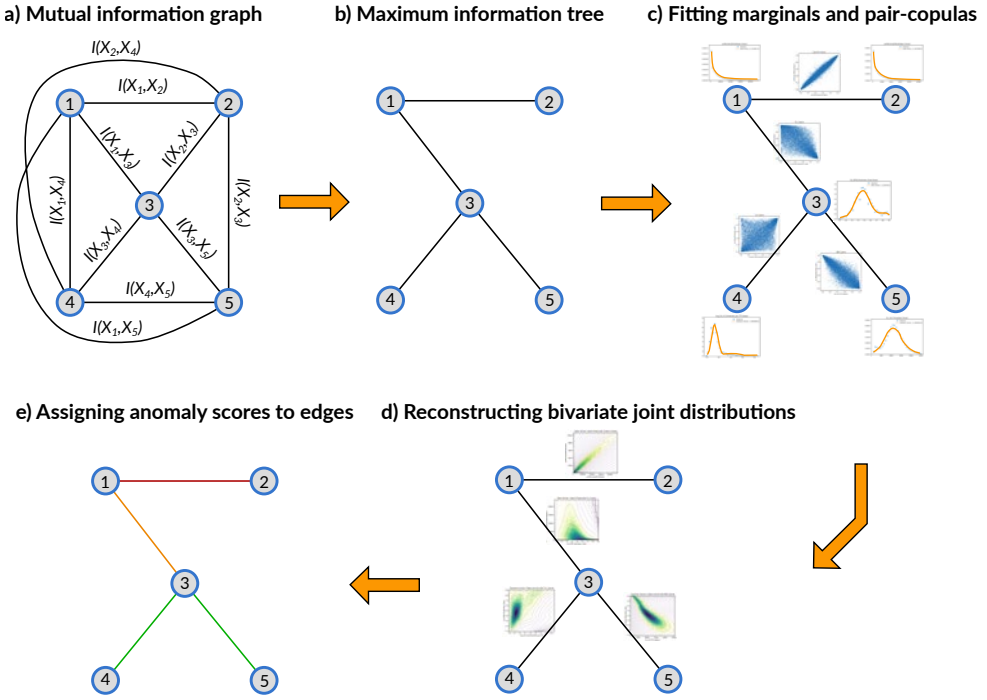


Figure 1: Concept of the presented anomaly scoring method

(download) throughput , delay, UL interference etc. Anomaly detection for such high dimensional data is a challenging problem.

4.2 The applied methods

In our approach (see Figure 1), the high-dimensional problem is decomposed into several two-dimensional ones. As the first step, we identify the most relevant *pairs of variables* whose joint behavior retains as much information from the high-dimensional joint behavior as possible. This method is inspired by the work of Chow and Liu [2], where a similar problem has been solved for discrete joint probability distributions based on the maximum information tree.

In the next step, the joint distributions of the selected pairs of variables obtained by the maximum information tree are characterized by a copula-based method. Copula-based methods are widely used in modelling joint probability distributions (mainly in the field of financial mathematics) due to their benefits in representing heavy tails [3]. Since anomalies can be considered as events falling in the tail of the probability distribution, copulas are very useful in modeling anomalies.

According to Sklar's theorem, the joint distributions are uniquely decomposed to the marginal distributions and a so called copula function, that characterizes only the dependency structure between the variables, and is independent from the marginal

distributions. Consequently, copulas enable the separate modeling of the dependence structure and the marginals; we are going to rely heavily on this feature in the presented method.

Finally, the anomaly scores are calculated from the joint densities. For each observation, our procedure is able to report both an overall anomaly score, and individual anomaly scores for the selected variable pairs. Knowing the variable pairs affected helps to localize the reason of the anomaly.

4.3 The achieved research results

Our novel copula-based anomaly detection technique for high-dimensional data that does not only assign an anomaly score to the observations but it also localizes the reason of the anomaly. The proposed approach relies on the modeling of the multivariate probability distribution associated with the instances. Our procedure is able to handle large-scale, high-dimensional data since the intractable high-dimensional problem is broken into smaller tractable ones by using two dimensional projections (bivariate marginal distributions) of the joint probability distribution in such a way that it retains maximum information and reduces redundancy. Since rare events occur in the tails of the probability distribution, copulas were used to model the bivariate marginals. Another advantage of the copula approach that the univariate marginal probability distributions and the bivariate copulas can be fitted separately.

Some joint distributions are depicted in Figure 2. The color of the points reflect the anomaly score, light shades correspond to more rare (hence, more anomalous), dark shades to more typical observations. Observe that the joint distributions in these examples are very far from a bivariate Gaussian distribution, and are in fact difficult to fit with a mixture of bivariate Gaussians, too. However, based on the figures it is clear that the the copula-based modeling of the relevant variable pairs can be applied successfully.

Besides an overall anomaly score, our approach also reports individual anomaly scores for the selected variable pairs that can be illustrated on an anomaly tree that enables the users to get an overview on the current state of the system, and to observe it evolving in time.

5 Solution of the industrial problem and its benefits

Figure 3 illustrates an anomaly tree, where the color indicates the severity of the anomaly. For example, the red color of a link means that the anomaly score is above 0.996. Orange shades mean lower anomaly score, and blue shades are assigned to non-anomalous relations.

In Figure 3 we can see that the number of PRBs (Physical Resource Blocks) is extremely low, compared to the number of times a user equipment is selected for transmission (affecting `pmSchedActivityCellDl - Num_PRBs_used_by_PDSCH` relations). Given the transport block size (`pmRadioTbsPwrUnrestricted`), the high usage of the 16-qam modulation scheme (`pmMacHarq- UlSucc16qam`) and the large number of negative acknowledgements in case of the 64-qam modulation scheme

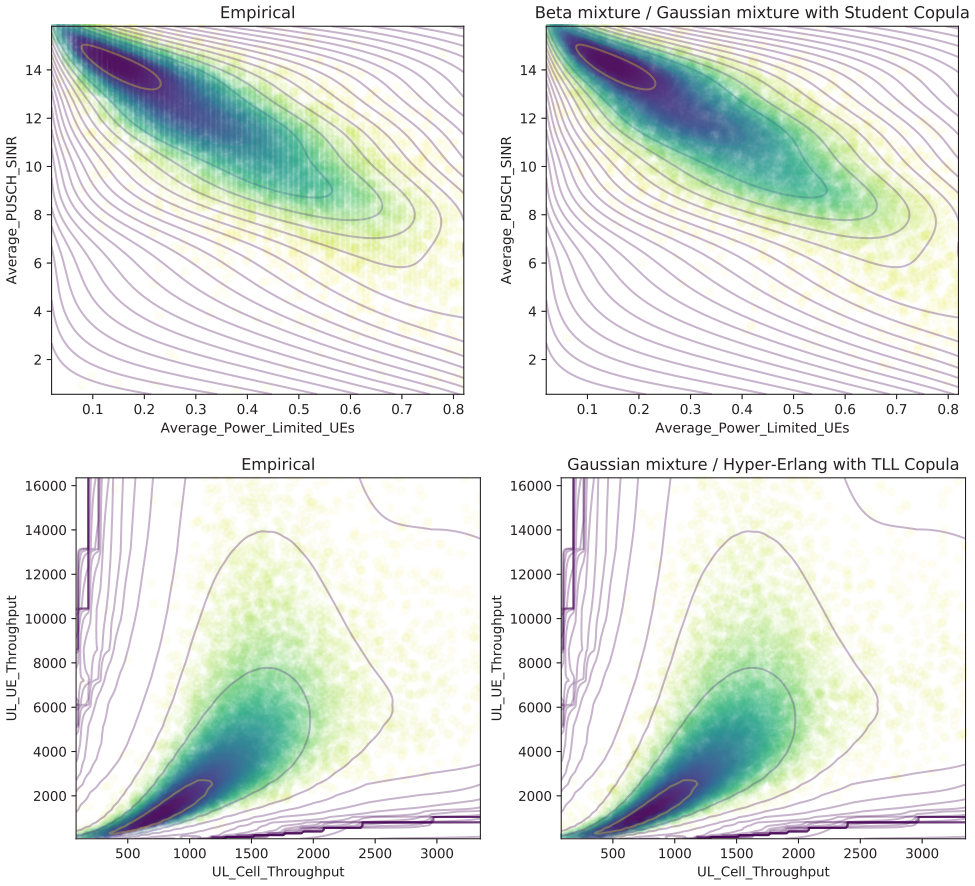


Figure 2: Reconstructing the joint distributions. Left: original data points. Right: random points simulated from the model.

(`pmMacHarqDlNack64qam`) implies that there could be something wrong with the radio channel, possibly with the antenna.

The manufacturer of the base station hardware has released guidelines for performance management and optimization, and also a troubleshooting guide, that the operators can use in order to resolve the problems detected by the presented anomaly detection algorithm.

6 Conclusions

The anomalies reported by our method were reconfirmed by real network operators. Moreover, we compared the anomaly score of the proposed approach with other standard anomaly detection methods. We can conclude that the overall anomaly score returned by our approach is similar to the scores returned by other methods. On

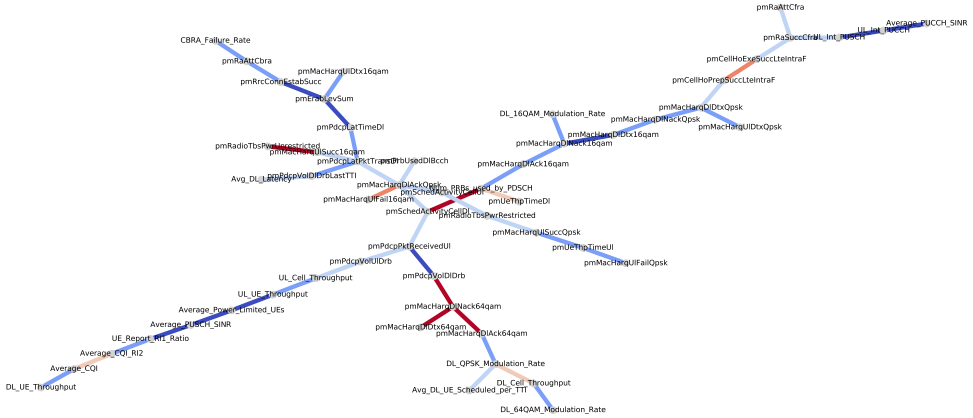


Figure 3: An anomaly tree

the other hand, our method also provides support to localize the anomaly, can be parallelized, can cope with a large number of high-dimensional observations and is also capable of handling missing data.

In the future, we aim to develop an improved method that that can also overcome the limitations of the proposed method. The most important limitation of the proposed approach is that it only works well with large scale continuous data. Categorical or discrete variables must be omitted since our method is not able to handle them.

References

[1] G. Horváth, E. Kovács, R. Molontay, S. Nováczki, *Copula-based anomaly scoring and localization for large-scale, high-dimensional continuous data*. ACM Transactions on Intelligent Systems and Technology, 11(3), 1–26, 2020.

[2] C. Chow, C. Liu, *Approximating discrete probability distributions with dependence trees*. IEEE Transactions on Information Theory, 14(3), 462–467, 1968.

[3] U. Cherubini, E. Luciano, W. Vecchiato, *Copula methods in finance*. John Wiley & Sons, 2004.

Highly accurate numerical solution of the Navier-Stokes equations in pipes

Zoltán Horváth[1], **Mátyás Constans**[1], and **Ádám Bácsi**[1]

[1] Széchenyi István University, Department of Mathematics and Computational Sciences

[1]E-mail: {horvathz, matyas.constans, adam.bacsi}@math.sze.hu

1 Executive Summary

Accurate and fast simulation tool has been developed for the computation of the exhaust flow sound pressure level (SPL), validated to measurements and compared with industrial state-of-the-art methodology. For the exhaust flow simulation, first a commercial code was used with Large Eddy Simulation (LES) method, which is an accurate method to reproduce the industrial state-of-the-art methodology and get benchmark values for accuracy and running time. Then a fully inhouse solver was developed, the Fluid-Solver by creating and applying numerical methods for the Navier-Stokes equations for unsteady, compressible gases in 3 spatial dimensions. The resulting code is parallel in the sense that it runs on multicore CPU and NVIDIA GPU computing platforms. The algorithm behind Fluid-Solver is of second-order accuracy in space and uses the finite volume method in which a cell-based upwind scheme with linear reconstruction was applied for the convection part and vertex based central scheme was applied to the viscous part. The computational mesh was an unstructured tetrahedral mesh with 70 million cells, for a direct numerical simulation (DNS). The SPL was then computed from the pressure time series in a standard way for both solvers.

Codes were tested under several exhaust geometrical configurations and inlet conditions, started with a simple straight pipe and suitably large acoustics domain. The two solvers used different computational meshes: the commercial software used its own mesh generator and Fluid-Solver was preprocessed by ANSA. Meshes were refined several times to get a stable and reasonably accurate result.

The overall comparision says that for roughly the same accuracy, the running time of the Fluid-Solver is within one day on an NVIDIA A100 GPU, while the commercial software needed two weeks on 30 processor cores with 2.27 GHz.

2 Challenge overview

Prediction of exhaust flow sound pressure level (SPL) by computational simulation is an important step of the development of vehicles (and other machines like expansion devices), which involves many challenges. The core of the challenges relies in that the air flow has to be simulated very accurately since SPL is a measure of the pressure fluctuation, which can be considered as an unsteady small perturbation of the mean pressure field. Thus a very accurate simulation of the air flow in a large domain is requested, in particular under industrially relevant conditions, i.e., short running time on available hardware.

3 Implementation of the project

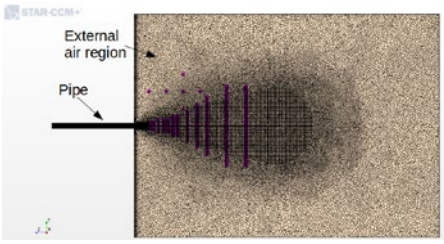
The industrial partner of the project, Audi Hungaria formulated the specification of different geometrical models of the exhaust system (straight pipe, bended pipe, simplified muffler on straight and bended pipes) and the operational conditions (inlet characteristics: mass-flow rate, temperature, pressure) for the simulation use cases. The research partner, the Mathematical Simulation and Optimization Research Group of the Széchenyi István University (SZE) developed a solver from the scratch, the Fluid-Solver on GPU for air flow simulation and performed two series of simulations for a selected set of the industrial use cases: one with the industrial state-of-the-art solver and one with developed solver. Also, from the SZE FIEK project, the industrial partner sponsored the costs of the measurement rigs and contributed in conducting reference measurements for a representative set of the use cases to provide data for validation of the simulations. All of these descriptions were compiled into the research plan with milestones and other project details.
Results were evaluated by the industrial team at meetings with the research team from time to time.
The performance of the code has been increased by the research team within the HiDALGO-project, supported by the H2020 programme.

4 The research problem, methods, and results

The main research problem is to construct a fast and accurate numerical algorithm for the exhaust air flow, implement it efficiently, and test the resulting code on prescribed very large scale computational problem. Here "fast" and "accurate" refer to the industrially requested running time, which is 1-2 days maximum, and be able to resolve pressure oscillations stemming from small-scale turbulent eddies.
First we shall introduce the industrial benchmark solution, which uses the Large Eddy Simulation with a commercial code, then own method with implementation details, and finally the computational results for both ways.

4.1 Large Eddy Simulation with commercial CFD software

The problem of turbulent flow simulation in the exhaust pipe system has been modelled with the commercial software STAR-CCM+ (version: 13.06) using LES. The geometry of the simulated region consists of the exhaust pipe and an external air region. The exhaust pipe is a straight cylindrical pipe with the length of 1 m and radius of 57 mm. The external air region has also the shape of a cylinder with the radius of 1.15 m and length of 3.1 m.
For the simulation, polyhedral mesh has been used with prism layers at the pipe wall. The mesh density is larger in the turbulent region of the flow and the cell size ranges from 2 mm to 30 mm. At the wall of the pipe, prism layers are applied. The resulting mesh consists of 6.5 million cells and is shown in Fig. 1. The cell sizes have been estimated based on that the waves up to the frequency of 2 kHz are resolved with at least 10 cells inside the jet. Outside the jet, the cell size has been chosen to result in a managable total number of cells.

a) 

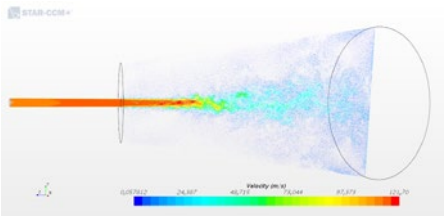
b) 

Figure 1: a) Vertical cross-section of the simulated area. Purple dots indicate measurement points. b) Snapshot from velocity field.

The simulation has been initialized with steady Reynolds averaged Navier-Stokes equations (RANS) and its results have been used as initial condition for the LES. The boundary condition at the entrance of the pipe is a constant mass-flow of 839 kg/h in accordance with the measurement and at the edge of the external air region, pressure outlet was used.
The LES was performed with the time step of $4 \cdot 10^{-5}$ s which has been estimated by taking into account a CFL number of 1 on the underlying mesh and the ability to resolve the time evolution of the largest frequency of 2 kHz. To achieve proper convergence, 12 internal iterations have been used for each time step. The total, simulated physical time was 1 s in order to simulate at least 10 periods of the smallest frequency (20 Hz) oscillations after transients.
The total simulation time took approximately two weeks on the available architecture (30 processor cores with 2.27 GHz). Due to the long running time, the sensitivity on mesh density and temporal resolution has not been studied completely and should be performed in future research.
During the simulation, the three components of the velocity vector and the pressure has been monitored in 786 points which are denoted by purple dots in Fig. 1. Based on the pressure values, the noise levels are obtained. At a specific point, the discrete pressure values p_j have been measured in each time step. The Fourier transform is defined by

$$\tilde{p}_k = \frac{1}{N} \sum_{j=1}^{N} e^{i f_k t_j/(2\pi)} p_j \tag{1}$$

where N is the number of timesteps and $f_k = k/(1\,\mathrm{s})$ are the relevant frequencies with integer values of k. The frequency dependent sound pressure level (SPL) is defined as

$$\mathrm{SPL}(f_k) = 10 \log_{10} \left(\frac{|\tilde{p}_k|^2}{p_0^2} \right) \tag{2}$$

where $p_0 = 20\mu$Pa is the threshold of human hearing which is commonly used in acoustics [1].
The overall SPL (OASPL) describes the total noise level at a specific position. The OASPL is defined by

$$\mathrm{OASPL} = 10 \log_{10} \left(\frac{\sum_k |\tilde{p}_k|^2}{p_0^2} \right) . \tag{3}$$

The noise levels are summarized in Table 1 and Fig. 2. The OASPLs are higher in the simulation than in the measurement which may be explained by reflecting sound waves from the boundary of the external air region.

Point ID	Angle and distance to oriface		Measurement [dB]	Simulation [dB]
1	90°	0.354 m	80.1	81.3
2	63°	0.396 m	81.0	86.7
3	45°	0.5 m	83.0	90.7
4	34°	0.637 m	86.8	93.0
5	56°	0.637 m	78.5	85.7
6	27°	0.396 m	99.7	106.2

Table 1: Comparison of the OASPL of the measurement and the simulation performed with STAR-CCM+.

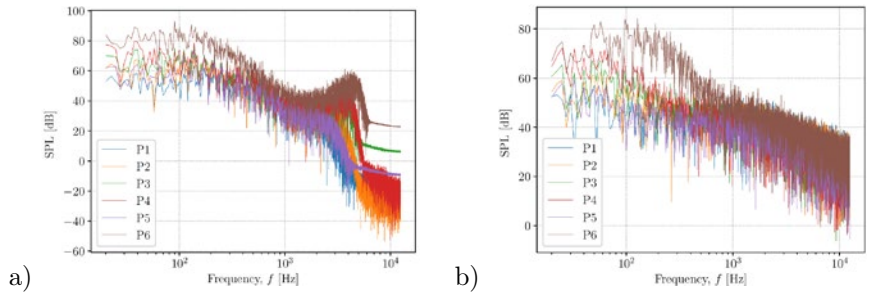


Figure 2: Frequency dependent SPL calculated from a) LES simulation results and b) measurement results.

The frequency dependent SPL exhibit an extra contribution at around 5 kHz in the simulation which is most probably an effect of finite cell size and has no physical meaning. The trend of the frequency dependence show remarkable similarities with the measurement results.
From the velocity values, time-averaged values $\bar{v} = \sum_j v_j/N$ and the root means squares $\sigma_v = \sqrt{\sum_j (v_j - \bar{v})^2/N}$ are derived. In the formulas, v_j is the velocity component perpendicular to the x axis. This component has been measured with hot wire probe in the measurements as well.
The velocity results of the simulation are evaluated along the same vertical sections as in the measurments, see Fig. 1. For the section located at the distance of 150 mm

from the oriface, the average velocity and its root mean square are plotted in Fig. 3. In the figure, the measurement results are also plotted.

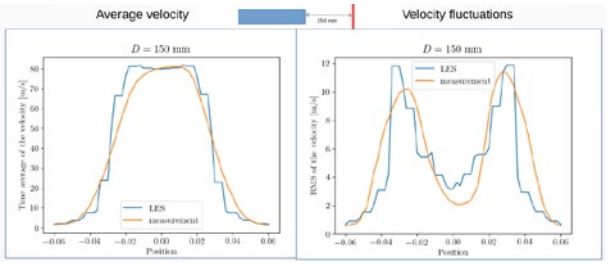


Figure 3: Average velocity and root mean square in LES simulation and measurement. The velocity values are plotted along the section at the distance of 150 mm from the oriface.

The figure shows that the basic features of the measurements are captured by the LES simulation.

4.2 Direct Numerical Simulation with the Fluid-Solver in-house software

4.2.1 The numerical method

To address the challenge, the unsteady, compressible Navier-Stokes equations on 3D domain were solved numerically. Turbulence was not modelled, instead the mesh was chosen very fine, in the magnitude of the Kolmogorov length scale near the outflow (within 1 millimeter) and some other more rapidly varying parts. Purely tetrahedral meshes were used.
The finite volume method was applied that was once already successfully applied to industrial problems (see [4]). This means that for the inviscid part of the Navier-Stokes equations the Vijayasundaram flux vector splitting upwind scheme was applied while the viscous part was treated in a central way. Due to the requirement of high accuracy, second order schemes were applied, for which vertex values were computed by avereging corresponding cell values of the conservative variables and linear reconstruction functions were applied. More details of the methods can be found in [5] and in references.

4.2.2 Implementation of the algorithm

The Fluid-Solver is a finite-volume, cell-based tetrahedral solver, algorithms can mostly be broken down into vertex-iterative, face-iterative, and cell-iterative algorithms. Since we are using the Explicit-Euler method for time-stepping, the current method for our solver consists of computing flux values by iterating through all vertices, faces, and cells respectively. (Iterating through vertices as well is necessary since a 2nd order space discretization is applied). Our flux variable has 5 components: a density (1 component), velocity (3 components), and energy (1 component).

Fluid-Solver also has a CUDA version, that can run on GPU-s via nvidia's CUDA technology. Since GPU-s are heavily memory limited compared to CPU-s (an nvidia V100 GPU for instance, only has 32GB of VRAM), one of the main challenges with the GPU version was to make everything as memory efficient as possible; furthermore, since the AUDI mesh we simulated on was well over 70 million cells, it became clear memory layout and memory usage optimization was our primary objective. Choosing between SOA (structure of arrays), and AOS (array of structure) memory representations is also critical, since L1 cache misses have a huge impact on performance since almost all of our algorithms are based on iterating over a set of vertices/faces/cells, according to the second order cell based finite volume code. We ended up storing flux states in a linear, pre-allocated array, stored as an AOS. The reasoning behind this is that whenever we need to access one flux component (ex. velocity), we almost always need to either read/write to the other values (density, momentum), since these values all heavily depend on each other.

4.2.3 Results

Computed SPL values were compared to measurements in the same way as in the previous section. The results of simulations and measurements can be seen on Figure 3 below.

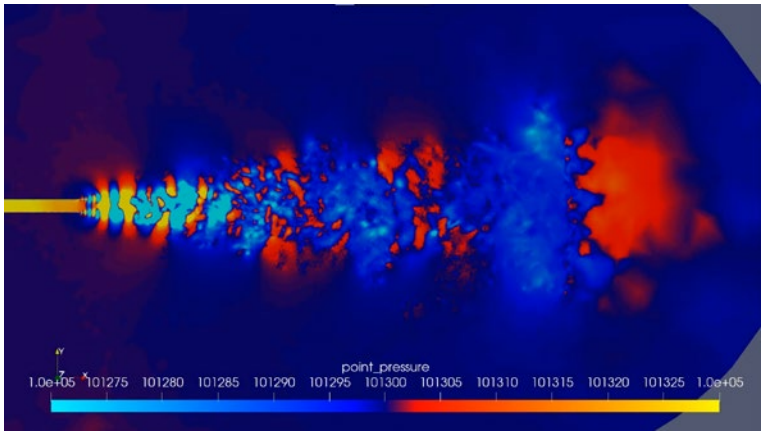


Figure 4: Cross section of the pressure field from the simulation.

It is observable that SPL measurements and simulations are close to each other with the range of frequency domain $[10^2, 10^4]$.

5 Solution of the industrial problem and its benefits

The research results presented in the previous sectoin concluded with propositions on the sufficient meshing and method configuration parameters for the air flow and SPL simulation of exhaust systems. These all constitute two methodologies for the industrial problem, i.e. for the computational prediction of the SPL of exhaust systems.

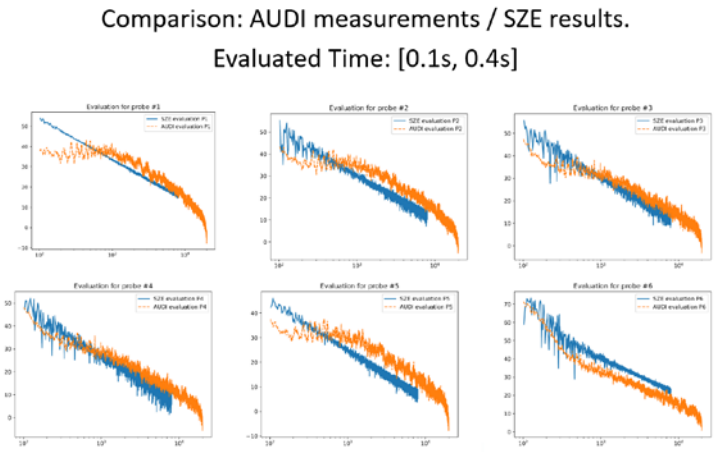


Figure 5: Comparison between the measured SPL values for the straight-pipe provided by the industrial partner (see [3]), and the simulated SPL values by SZE.

These two solutions, the LES-STAR CCM+ and the Fluid-Solver-DNS provide the following benefits to the industrial partner.

- The computational time for the pipe acoustics is 5 day on a cluster with the LES-simulation and 1 day on an NVIDIA A100 GPU with Fluid-Solver.
- Both methods provide suitable accuracy for the pipe SPL values. For the exhaust configurations with muffler, only the Fluid-Solver was applied, due to the long expected running time of the commercial software. With Fluid-Solver

The Fluid-Solver saves several working days for each pipe scenario at the same accuracy as the present industrial state-of-the-art.

6 Conclusions

The research team has developed an industrial benchmark and found sufficient parameters for the prediction of the SPL of pipes, though the running time is on the edge of acceptability.
Own developed GPU-code, based on mathematical algorithms seems very promising to introduce into operations since at good validation properties it runs for 1 day (pipe) and 3 days (muffler), which is far within the initial expectation range.
We remark that during the research different model order reduction methodologies were investigated aiming at development of fast codes. According to our numerical results, the classical POD-DEIM method is not stable for the compressible Navier-Stokes equations when Fluid-Solver is applied. However, a new variant of the POD-DEIM model order reduction method has been created, which is implemented and runs for medium scale external flow problems very well. We shall publish the results elsewhere with use cases less demanding than the SPL computation.

References

[1] Ross Roeser, Michael Valente, *Audiology: Diagnosis*, Thieme Medical Publishers, Inc., 2007

[2] C. Freitag, D. Feszty, Á. Kovács, B. Kuti, Z. Horváth, Á. Bácsi: *Method development for exhaust pipe noise with fast simulation of turbulent flows with LES-DNS.* Research plan, 2018.

[3] B. Vehovszky, K. Horváth, B. Kuti, D. Feszty: *Measurement protocol and report for flow induced noise.* Technical Report, SZE FIEK-project, 2021.

[4] A. Horváth, Z. Horváth: *Application of CFD numerical simulation for intake port shape design of a diesel engine.* Journal of Computational and Applied Mechanics 4 (2), 129-146

[5] S. Chaturantabut, DC. Sorensen: *A state space error estimate for POD-DEIM nonlinear model reduction.* SIAM Journal on numerical analysis 50 (1), 46-63

Securing Cloud Authentication

Andrea Huszti[1] and **Norbert Oláh**[2]

[12]University of Debrecen, Facultics of Informatics,

[1]E-mail: huszti.andrea@inf.unideb.hu
[2]E-mail: olah.norbert@inf.unideb.hu

1 Executive Summary

Security issues today are critical for organizations and governmental institutions where authentication is one of the critical aspect. Our goal with the CCLab Company was to design and develop an effective solution that increases the security level of systems and is provably secure. We rejected the traditional centralized authentication methods and designed an authenticated key exchange (AKC) protocol which takes advantage of distributed systems. Our protocol would fit into these systems and takes advantage of the capabilities of these systems like robustness, scalability and greater availability. If one or more servers break down or become corrupt, the service provider will still be able to serve and authenticate the users securely. We focused heavily on making the protocol effective and we achieved promising efficiency results, which can be led back to the facts that the session key is generated by ECDH key exchange, moreover MAC, xor operations and symmetric encryption are applied. We prove that the proposed protocol is a secure AKC protocol in the random oracle model, assuming the ECCDH assumption holds in the elliptic curve group if MAC is universally unforgeable under an adaptive chosen-message attack and the symmetric encryption scheme is indistinguishable under chosen plaintext attack.

2 Challenge overview

Nowadays usually healthcare data is stored and processed in clouds. In practice, OpenStack ([26]) is one of the most popular cloud computing softwares. OpenStack Identity service supports multiple methods of authentication, including user name and password, Lightweight Directory Access Protocol (LDAP), and external authentication methods (i.e. Kerberos). There are concerns about these systems due to their centralized structure and the most frequently used authentication methods are based on short secrets. Services that store password information for a large number of enterprises in a central database are primary targets for hackers (e.g. Golden Ticket Attack [20, 24], OneLogin attack ([25])). The recent hack of OneLogin, an IAM (Identity and Access Management) provider proves that the credentials may not be as safe as we are led to believe. The weakest point of Kerberos authentication model is the key distribution center KDC. The entire authentication system depends on the trustability of the KDC, so anyone who can compromise system security on a KDC system can theoretically compromise the authentication of all users of systems depending on the KDC. On the other hand, Kerberos requires the continuous availability of the KDC server, if the server is not available, no one can log in. If a server

is compromised, the secrets are vulnerable to theft which reveals the weaknesses of this solution. If this happens, data leakage and data misuse are considered emerging security threats to organizations, especially when carried out by insiders.

3 Implementation of the project

The CCLab Company would like to apply a new provably secure auhentication system which can be used in several environments (healthcare system, organizations, etc.) and they required certain properties and features for the scheme. These properties include scalability, robustness, effectiveness. They prepared the specification of this scheme and monitored the developing of the protocol. They would like to apply the proposed scheme in industrial applications. Andrea Huszti and Norbert Oláh designed a new protocol which provides secure authentication between the participants. They proved that the proposd solution is provably secure. They created a Java implemention and tested the recommendation on Microsoft Azure environment. The results are promising for the company.

4 The research problem, methods, and results

In the following sections we describe the research problem, the applied methods, and summarize achieved research results. We demonstrate the research problem, the scientific solutions and the disadvantages of these solutions. After that we introduce our proposed protocol which takes advantage of distributed systems and applies multiple-server authentication. Our solution meets the following requirements: mutual authentication, key secrecy, robustness, scalabilty and greater availability. We prove the security of the protocol and show that our protocol is provably secure. We could minimize computational costs by using fast cryptographic operations which results in an effective authentication process.

4.1 The research problem

Our research problem was to provide secure mutual entity authentication applying a distributed system which consists of multiple, autonomous computers that communicate through a network even during completing their task. The goal of a distributed system is to solve a single problem by breaking it down into several tasks where each task is computed by a computer of the system. Distributed systems can run in a cloud infrastructure as well. Secure user authentication is an important issue of cloud services. If it is breached, confidentiality and integrity of the data or services may be compromised. In the case of Software as a Service model, the cloud service provider takes responsibility for securing all the data from unauthorized access.

In scientific literatures, usually centralized, one-factor [13, 19] or two-factor identity verification protocols [5, 6] are proposed. Xavier Boyen presented a Hidden Credential Retrieval protocol [3], where the protocol applies one-server solution and a blind signature technique with a user password. However, the concept of distributing authentication to multiple servers enhance the security level. The advantage of a distributed

system is that external attackers have to attack multiple servers simultaneously, which increases the attack cost.

Several protocols are proposed for a multi-server environment. Sood, Sarje and Singh [28] and Brainard et.al. [4] suggested authentication protocols in a two-server environment, where two servers together decided on the correctness of the password submitted for authentication. Katz et al. [17] demonstrated the first provably-secure two-server protocol for the important password-only setting (in which the user needs to remember only a password, and not the servers' public keys), and was the first two-server protocol (in any setting) with a proof of security in the standard model. Acar et al. constructed a solution [1] in which the corresponding secret key is securely store and blinded by some function of user password at storage provider(s) different from the login server. In this proposition [9] a multiple-server authentication protocol is designed, where one-time passwords are shared among the cloud servers. A Merkle tree or a hash tree [22] is applied for verifying the correctness of the one-time password. In most of the cases an encrypted communication channel is established after successful authentication. Our goal is also to establish a symmetric encryption key.

Passwords are often used as authentication information in key exchange protocols. In case of a password-authenticated key exchange (PAKE) after user registration a user and a server establish a session key for a secure channel [21, 18]. In a multiple-server environment usually threshold password-authenticated key exchange protocols are designed. Devriş Işler and Alptekin Küpçü introduced a scheme [15] where the protocol ensures that multiple storage providers can be employed, and the adversary must corrupt the login server and threshold-many storage providers to be able to mount an offline dictionary attack.

4.2 The applied methods

We propose a multi-server password-based authenticated key exchange scheme. It is similar to the (k, n) threshold PAKE systems, however we don't apply secret-sharing algorithms. In contrast to other threshold password-based protocols applying secret-sharing algorithms [27, 14, 16, 21, 18], we share the password information among the servers, but it is not reconstructed from the shares to verify it. Ford et al. [7] proposed a protocol that securely generates a strong secret from a weak secret (password), based on communications exchanges with two or more independent servers. We demonstrate a new way of generating a strong secret (e.g. long-lived key) from a password, which is also suitable for scalability. Like other schemes we also consider the scalability property that is one of the main requirements for clouds. Proving the security requirements of cryptographic protocols is a challenging issue. Several protocols had been thought to be secure, until a simple attack was found. Provable security aims to show mathematically that a specific property of a scheme cannot be broken by a class of attackers. Provable security has several approaches. Reductionist proof introduced by Goldwasser and Micali [8] proceeds by presenting an explicit reduction showing how to convert any efficient adversary that succeeds in breaking the construction with non-negligible probability into an efficient algorithm that succeeds in solving the problem was assumed to be hard. Unlike [7] we also provide a detailed security analysis based on the Bellare and Rogaway model.

4.3 The achieved research results

We assume that there are thousands of servers in a cloud system therefore we reject the single-server authentication (e.g. Kerberos) and instead we propose the multi-server authentication. It is important to note that a single point of failure occurs typically in single-server solutions. If the server is unavailable, the provider usually needs to ensure replication to tackle the failure of their servers. Our scheme consists of n servers and the user randomly selects $k \leq n$ ones for each authentication. Besides the randomly chosen k servers, a server called authentication server is also chosen randomly in our solution. In the authentication phase, the k servers use their long-lived keys and the user's authentication is verified by the chosen server via the correct MAC values. Even if one or more servers fail out of the n servers, the client can still choose k servers randomly. So if one or more servers break down or become corrupt, the service provider will still be able to serve and authenticate the users securely. During authentication, instead of securely constructing the secret password from its shares, a random challenge generated by the client is constructed and verified by the authentication server. The randomly chosen participating servers are able to compute their challenge shares with the help of the password-based long-lived key set during registration and send them to the authentication server. In this way, the confidentiality of the password is assured. We focused heavily on making the protocol effective and achieved promising efficiency results. In several cases the related protocols in the literature employ asymmetric cryptographic primitives, which are considered slow compared to symmetric solutions and hash functions. The results of our protocol can be led back to the facts that the session key is generated by ECDH key exchange, moreover MAC, xor operations and symmetric encryption are applied. We also prove the security of the protocol and demonstrate that the protocol is provably secure. For our protocol, we extended the Bellare and Rogaway security model in [2] to prove that our multi-device scheme is secure and we introduced the threshold hybrid corruption model.

Definition 1 *We call a model* threshold hybrid corruption model *if we assume that the client is uncorrupted and there are at least* $n - k + 1$ *uncorrupted servers out of the* n *servers, if* k *servers are chosen randomly for AKC. Moreover, the server chosen to communicate with the client is*

1. *uncorrupted, or*
2. *corrupted weakly and among the remaining servers there is at least one uncorrupted.*

We prove that the proposed protocol is a secure AKC protocol in the random oracle model and we determine the definition of *secure AKC protocol*:

Definition 2 *A protocol is a* secure AKC protocol *if*

1. *In the presence of the benign adversary the client and the server oracle communicating with the client always accept holding the same session key* ssk*, and this key is distributed uniformly at random on* $\{0,1\}^{\kappa}$.

and if for every adversary $\mathcal{A}$

2. *If in a threshold hybrid corruption model there is a server oracle* $\prod_{I,J}^{i}$ *having matching conversations with a client oracle and if* $\prod_{I,J}^{i}$ *is weakly corrupted,* $\prod_{I,J}^{i}$ *has matching conversation with an uncorrupted server oracle, then the client oracle and oracle* $\prod_{I,J}^{i}$ *both accept and hold the same session key ssk.*
3. *The probability of* $\texttt{No-Matching}^{\mathcal{A}}(\kappa)$ *is negligible.*
4. *If the tested oracle is fresh, then* $Adv^{\mathcal{A}}(\kappa)$ *is negligible.*

Finally we formalize the theorem which states that our protocol is a secure AKC protocol and we can lead this back to the hard problems:

Theorem 1 *The proposed protocol is a secure AKC protocol in the random oracle model, assuming MAC is* universally unforgeable under an adaptive chosen-message attack *and symmetric encryption scheme is* indistinguishable under chosen plaintext attack*, moreover* ECCDH assumption holds *in the elliptic curve group.*

You can find our research results in Presentations and publications within the section References ([10, 11, 12]).

5 Solution of the industrial problem and its benefits

In addition to the theoretical results, we also managed to achieve practical results. The company has a prototype of multi-server mutual entity authentication system for clouds which is a new mutual entity authentication protocol with key agreement. Identity verification is executed by multiple servers on the provider's side and the protocol is provably secure in the threshold hybrid corruption model. Java simulation was tested with six servers, one of which received the client authentication request and the remaining five servers were required for authentication. Time means the time interval from the first step of the protocol to the successful connection. The average connection time was then calculated based on the received time results. For the cloud environment, we tested the standard package and the p3v2 package. The Standard Package contains 100 Azure compute units [23] and the p3v2 package includes 840 Azure compute units. The concept of the Azure Compute Unit (ACU) provides a way of comparing compute (CPU) performance across Azure Stock Keeping Units (SKUs). This will help you to identify easily the SKU which is the most likely to satisfy your performance needs. ACU is currently standardized on a Small (Standard_A1) VM being 100 and all other SKUs then represent approximately how much faster that SKU can run a standard benchmark.

The average connection time is 0,1092 second per connection when we use the p3v2 package. In the standard package the result of average connection time is 0,1537 second per connection. This time contains the mutual authentication including the

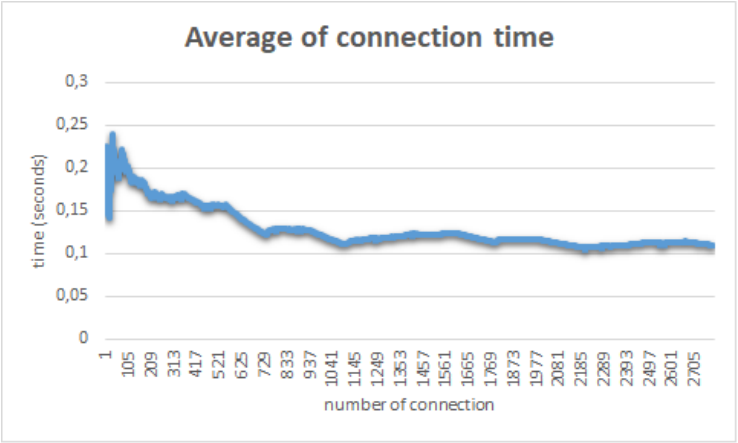


Figure 1: Average connection times of successful authentications between the client and the servers.

authentication of selected servers and generating a session key. It is an important issue in our protocol that the user gives his/her password and after that the long-lived keys are available. The static password is comfortable for the user and the long-lived keys provide the appropriate security level. Since in each authentication the values are random and fresh, the key freshness holds and the protocol execution cannot be successfully finished with old, already used values and keys.

6 Conclusions

Instead of centralized authentication we have designed an Authenticated Key Agreement with Key Confirmation protocol for a distributed environment. Our system is robust against server breakdown and applies multiple-server identity verification. We introduce the treshold hibrid corruption model and we model when the different servers become corrupt on the protocol. We give a detailed security analysis, and we prove that our proposed protocol is a secure AKC protocol in the random oracle model, assuming MAC is universally unforgeable under an adaptive chosen-message attack and symmetric encryption scheme is indistinguishable under chosen plaintext attack, moreover ECCDH assumption holds in the elliptic curve group. It is also important to note that during the authentication phase fast ECDH key exchange, MAC, xor operations and symmetric encryption are used, hence we achieve promising results in computational time. We successfully developed a professional relationship with the CCLab company which can be beneficial in the realization of our future plans.

References

[1] Acar, T., Belenkiy, M., & Küpçü, A, *Single password authentication* Computer Networks, 57(13), 2597-2614, 2013

[2] S. Blake-Wilson, D. Johnson and A. Menezes, *Key agreement protocols and their security analysis* Proceedings of the sixth IMA International Conference on Cryptography and Coding, LNCS 1355, pp 30-45, 1997.

[3] X. Boyen, *Hidden credential retrieval from a reusable password.* In Proceedings of the 4th International Symposium on Information, pp 228–238. ACM, 2009.

[4] J. Brainard, Ari Juels, Burt Kaliski, and Michael Szydlo, *A New Two-Server Approach for Authentication with Short Secrets*, Proceeding SSYM'03, Proceedings of the 12th conference on USENIX Security Symposium - Volume 12, pp 1-14, 2003.

[5] N. Chen, R. Jiang, *Security Analysis and Improvement of User Authentication Framework for Cloud Computing*, Journal of Networks, 9(1), pp. 198-203, 2014.

[6] A. J. Choudhury, P. Kumar, M. Sain, *A Strong User Authentication Framework for Cloud Computing*, Proceedings of IEEE Asia -Pacific Services Computing Conference, pp. 110-115, 2011.

[7] W. Ford and B. S. Kaliski, *Server-assisted generation of a strong secret from a password.* In Enabling Technologies: Infrastructure for Collaborative Enterprises,(WET ICE 2000). IEEE, 2000.

[8] Shafi Goldwasser and Silvio Micali, *Probabilistic encryption*, Journal of Computer and System Sciences, 28(2):270–299, 1984.

[9] A. Huszti, N. Oláh, *A simple authentication scheme for clouds*, Proceedings of IEEE Conference on Communications and Network Security (CNS), (2016), Pages: 565 - 569.

[10] Andrea Huszti, Norbert Oláh, *Identity-Based Cloud Authentication Protocol*, Proceedings of The 11th Conference of PhD Students in Computer Science, Szeged, (2018), 33-36.

[11] Andrea Huszti, Norbert Oláh, *Provably Secure Authenticated Key Agreement with Key Confirmation for Distributed Systems*, Proceedings of the 14th International Conference for Internet Technology and Secured Transactions, (ICITST - 2019), London, UK, 2019.

[12] Andrea Huszti, Norbert Oláh, *Provably Secure Scalable Distributed Authentication for Clouds*, In International Conference on Cryptology and Network Security, Springer, Cham. (2020, December), pp. 188-210.

[13] M. S. Hwang, L. H. Li, *A new remote user authentication scheme using smart cards*, IEEE Transactions on Consumer Electronics, 46(1), pp. 28-30, 2000.

[14] Jarecki, S., Kiayias, A., & Krawczyk, H, *Round-optimal password-protected secret sharing and T-PAKE in the password-only model*, In International Conference on the Theory and Application of Cryptology and Information Security (pp. 233-253). Springer, Berlin, Heidelberg. 2014.

[15] Işler, D., & Küpçü, A, *Threshold single password authentication.* ESORICS Data Privacy Management, Cryptocurrencies and Blockchain Technology, pp. 143-162, 2017

[16] Işler, D., & Küpçü, A, *Distributed Single Password Protocol Framework.* IACR Cryptol. ePrint Arch., 2018, 976.

[17] Katz, J., MacKenzie, P., Taban, G., & Gligor, V, *Two-server password-only authenticated key exchange*, In International Conference on Applied Cryptography and Network Security (pp. 1-16). Springer, Berlin, Heidelberg. (2005)

[18] J. Katz, R. Ostrovsky, and M. Yung, *Efficient password-authenticated key exchange using human-memorable passwords.* In EUROCRYPT 2001. Springer, 2001.

[19] W. C. Ku, S. M. Chen, Weaknesses and improvements of an efficient password based remote user authentication scheme using smart cards, IEEE Transactions on Consumer Electronics, 50(1), pp. 204-207, 2004.

[20] George Kurtz, Dmitri Alperovitch, Elia Zaitsev, *Hacking exposed: Beyond the Malware*, RSA 2015 (slide deck), https://www.rsaconference.com/writable/presentations/file_upload/expt10_hackingexposedbeyondthemalware.pdf, (2015).

[21] P. MacKenzie, T. Shrimpton, and M. Jakobsson, *Threshold password-authenticated key exchange.* In CRYPTO 2002. Springer, 2002.

[22] Ralph C. Merkle, *A Digital Signature Based on a Conventional Encryption Function*, Advances in Cryptology - CRYPTO '87, Lecture Notes in Computer Science, **293**, (1987), pp. 369-378.

[23] https://docs.microsoft.com/en-us/azure/virtual-machines/acu

[24] Miguel Soria-Machado, Didzis Abolins, Ciprian Boldea, Krzysztof Socha, *Kerberos Golden Ticket Protection, Mitigating Pass-the-Ticket on Active Directory*, CERT-EU Security Whitepaper 2014-007, 2016.

[25] https://www.entrepreneur.com/article/295831.

[26] https://www.openstack.org/.

[27] Di Raimondo, Mario, and Rosario Gennaro, *Provably secure threshold password-authenticated key exchange*, International Conference on the Theory and Applications of Cryptographic Techniques. Springer, Berlin, Heidelberg, 2003.

[28] Sood, Sandeep K., Anil K. Sarje, and Kuldip Singh, *A secure dynamic identity based authentication protocol for multi-server architecture* Journal of Network and Computer Applications 34.2: 609-618, 2011.

Improving inventory management and transport efficiency using innovative market demand prediction

István Kolumbán[1] and **Sándor Kolumbán**[2]

[1]Explorer Consulting Ltd.,
[2]Babeş-Bolyai University, Cluj-Napoca, Romania and Széchenyi István University, Győr, Hungary,

[1]E-mail: info@explorer-consulting.ro
[2]E-mail: sandor.kolumban@ubbcluj.ro

1 Executive Summary

For industrial companies, such as installation or automotive companies, it is important to properly forecast product demand. Correct forecasting helps managing stocks and deliveries. In the course of our work, we had access to data provided by multiple industry partners, such as user order data and billing data. We used these data to create aggregated time series for each product and predicted demand for products that had adequate statistical information. One of the main expectations was that a product had to have at least three years of sales history. We were also looking for connections between products. We evaluate the data of *IMPAR Ltd.*, an automotive company and *Melinda Instal Ltd.*, an installation company. The mentioned companies' own forecasts and our results were compared. We proposed to embed our own algorithms in the forecasting process of the companies. The study of inter-product relationships yielded less promising results.

2 Challenge overview

The major difficulty is the lack of enough historical sales data for the products. We could only forecast demand for products where at least three years worth of data was available. The situation is exacerbated by the quickly changing life cycle of products, being replaced by other products. Another challenge is that future sales of products are not necessarily related to the number of past sales. Rather it can be influenced by the price of the product, the current trend, new other products or other economic conditions. We identified two major product groups. The first group included products that often had a lot of sales, the prices of which were generally low. The second group includes products that are rarely sold, these products often cost a lot. The products of the second group were more challenging to forecast and the material loss calculated from the forecasts was higher than in the case of the first group. A big step forward could be a better forecast of the products in the second group.

3 Implementation of the project

The industrial partners provided the data and the domain knowledge. Academic partners performed statistical analysis on the data using state-of-the-art algorithms, libraries and technologies. These algorithms were evaluated and suggestions were made.

The completed project works as a pipeline that implements the following steps:

- Creates aggregated time series from data
- Calls Python script to forecast
- Evaluates predicted result
- Examines auto and cross correlations

Most of the project is written in .NET, C#. The classes responsible for the logical part of the project are adequately covered by both unit and integration tests. To improve verification, we also used a code coverage measurement application. Scripts that are used for prediction are written in Python.

4 The research problem, methods, and results

In this section, we present the pipeline of data processing, forecasting, examining relationships between products, the results and evaluations.

4.1 The research problem

The first step is processing the data provided by the partners. The data included a list of sales. Every sale among others contains the product id, the sell date, the quantity and the price. The first step of our pipeline was to create aggregated time series one for each product. These time series aggregated sales quantities over months, and average prices were also calculated. In the following, we examined only those products for which statistics on the number of sales were available for at least three years. Predictions for other products proved unreliable.

4.2 The applied methods

Several of the forecasting methods were discussed, PyTorch[1], TensorFlow[2] and Prophet[3], among others. In the end we chose the Prophet[3] library. Prophet is a procedure for forecasting time series data based on an additive model where non-linear trends are fit with yearly, weekly, and daily seasonality, plus holiday effects. It works best with time series that have strong seasonal effects and several seasons of historical data. In order to check the accuracy of the forecasts, we compared the results with the number of products sold.

Suppose that array p contains all predicted values of a given product from month to month, and array x contains the number of products currently sold from month to

month. Let the length of both arrays be m, which is the number of months predicted. The following formulas were used for comparison.

$$RoS = \frac{\sum_{i=1}^{m} p_i}{\sum_{i=1}^{m} x_i} \tag{1}$$

$$Norm = \sum_{i=1}^{m} \frac{|p_i - x_i|}{\sum_{i=1}^{m} |x_i|} \tag{2}$$

$$NormS = \sum_{i=1}^{m} \frac{(p_i - x_i)^2}{\sum_{i=1}^{m} x_i^2} \tag{3}$$

$$SoL = \sum_{i=1}^{m} |p_i - x_i| \cdot price_i \tag{4}$$

The $price_i$ in the 4 formula is the average price for the i-th month. The results of the comparisons are plotted on the figure 1. Mind that both scales are logarithmic. The values on the SoL axes are removed due to confidentiality reasons.

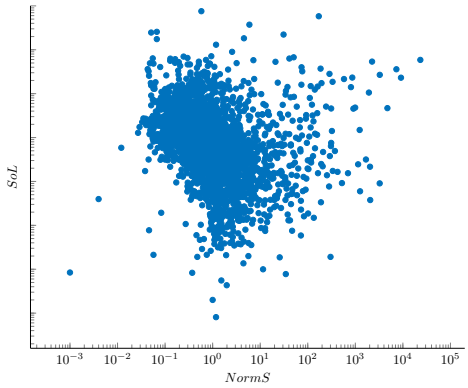


Figure 1: Comparison result

To uncover relations between products we calculated linear correlation was calculated. We examined how well sales of a given product correlates with that of another, shifted by a number of months.

4.3 The achieved research results

We had the opportunity to compare the predicted results with the companies' own forecasts. For greater efficiency, companies update their forecasts until the last minute, so they represent a strong baseline for comparison. The forecasts we produced showed approximately 18% weaker results than the forecasts manually adjusted by the companies.

We found many high correlation relationships between products that were subsequently manually verified by company experts. We concluded that there are many

product pairs where the correlation is very high, yet there is no real relationship between sales. This was due to the strong seasonality of company sales and all products being in sync with it. Due to the high number of products, just by chance we found many strongly correlated ones.

5 Solution of the industrial problem and its benefits

We created a data processing pipeline, based on the Prophet library[4]. Our results increased firms 'insights into the importance and effectiveness of forecasts, and highlighted the effectiveness of research on relationships between products.

We identified a set of products where our prediction pipeline produced significantly better results than the current practice at the companies. The prediction process for these products will include our methods.

Finding true inter-product relations in large set of products with highly correlated sales data requires more sophisticated techniques.

6 Conclusions

The presented results form the initial stage of a longer research collaboration. We will focus on integration of external explanatory variables (other than sales) into the prediction process. As mentioned in section 2, product sales cannot be clearly predicted from past sales. There are internal and external influencing factors that greatly influence the number of sales. Among the internal factors are, for example, promotions and product price changes. External factors may include public data issued by the government, such as building permits, zoning plans. By processing this extra information, we could make a stronger forecast.

References

[1] A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, A. Lerer, *Automatic differentiation in PyTorch.* 2017.

[2] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, X.Zheng, *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems.* 2015.

[3] Facebook's Core Data Science team, *Prophet: Automatic Forecasting Procedure.* 2017.

[4] Michael Su, *Prophet Documentation.* Release 0.1.0, 2018.

A patient registry database and application of mathematical models to analyze psychiatric disorders

András London[1], **Balázs Bánhelyi**[1], **Andor Kanka**[2] and **János Kálmán**[2]

[1]University of Szeged, Institute of Informatics,
[2]University of Szeged, Department of Psychiatry

[1]E-mail: london@inf.u-szeged.hu

1 Executive Summary

Integrated patient data constitute a computable collection of fine-grained longitudinal phenotypic profiles. Methods from data mining, machine learning and text mining to investigate, for example, disease comorbidities, patient stratification, drug interactions and clinical outcome, can be utilized. The proposed program focused on the mathematical analysis of a wide-range good quality medical data of psychiatric patients collected in a patient registry being continuously developed by our team. Towards descriptive analytic, the literature of examining medical data considering its temporal (or longitudinal) nature has still been rare and not much effective results has been found in the past years. Our main goal was to investigate the applicability of several mathematical methods and their modified versions on temporal medical data, summarize and review what the best practice solutions for analyzing certain types of data are. More concretely, highlight the main differences between the methods when applying them, regarding advantages and disadvantages. The main benefit is that researchers, R&D experts and decision makers might get a clearer picture about the applicability, advantages and limitations of currently available advanced mathematical solutions applied on psychiatric medical data.

2 Challenge overview

Over the past few years, 'big data' has become a frequently used catchall phrase for research approaches involving the use of complex, large-scale datasets. There are many types of data that may fit this description, but within the sphere of clinically oriented research this term is often considered synonymous to electronic health record (EHR) data, or electronic medical record data. From a research perspective, integrated patient data constitute a computable collection of fine-grained longitudinal phenotypic profiles, facilitating cohort-wide investigations and knowledge discovery on an unprecedented scale. Processing and utilizing this data methods from data mining, machine learning and text mining will become applicable to investigate, for example, disease comorbidities, patient stratification, drug interactions and clinical outcome.

Mental disorders are commonly occurring and often seriously impairing in many countries throughout the world. Many mental disorders begin in childhood-adolescence and have significant adverse effects on the subsequent life of the patients

[1]. As a group, mental disorders have a high prevalence, in general, compared with many other health conditions. Even schizophrenia and bipolar disorder, with relatively lower prevalence among the mental disorders have a higher prevalence than many other diseases and health conditions [2]. A significant part of mental illnesses (e.g. schizophrenia) are life-long chronic diseases. In addition, information on the history of the disease, the efficacy and side effects of previous treatments are important for actual treatment decisions. Therefore, in case of mental disorders it is especially important that the patient data should be structured, easily manageable and accessible to the clinicians.

In case of chronic diseases, it would be particularly important to monitor and evaluate the effect of drugs on the long-term course of the disease, on the possible occurrence of late side effects, including effects from possible interactions, etc. In other words, the treatment and research of mental diseases is an area that can gain the most from the use of integrated health records and patient registries. Therefore earlier we have started to develop a patient registry with the goal of efficiently collecting, storing and analyzing high-quality data of psychiatric patients (for an illustration, see Fig. 2). Our first experiments in a pilot project concerning symptoms (positive and negative) of schizophrenic patients and their relation with medication were promising. As we have utilized the novel patient registry to collect data that is ready for analysis, and based on previous experiments, our main goal and challenge was to investigate the applicability of several mathematical methods and their modified versions on medical data, summarize and review what the best practice solutions for analyzing certain types of data are. More concretely, highlight the differences between the methods when applying them on psychiatric data, regarding applicability, advantages and disadvantages.

3 Implementation of the project

The project team was led by the Institute of Informatics of University of Szeged (András London, Balázs Bánhelyi). The main collaborator in this project was the Department of Psychiatry at the University of Szeged (Andor Kanka, János Kálman). The partner provided its long-term experience in diagnostic and treatment procedures, in clinical aspects of medical research and medical documentation systems, in long-term genetic and epidemiological research, quality supervision of clinical trials and in management of regional health care development projects. The consulting industrial partner was Takeda pharmaceutical company. Their vision is patient-centricity through understanding the current patient journey and experience, the need for developing better scales and to better understand psychiatric disorders by using deep phenotyping provides a platform to collaborate with academic institutions towards fulfilling these goals. Takeda's vision for patient-centricity is to incorporate direct patient input in R&D in order to develop their medicines with patients rather than the traditional approach of developing medicines for patients.

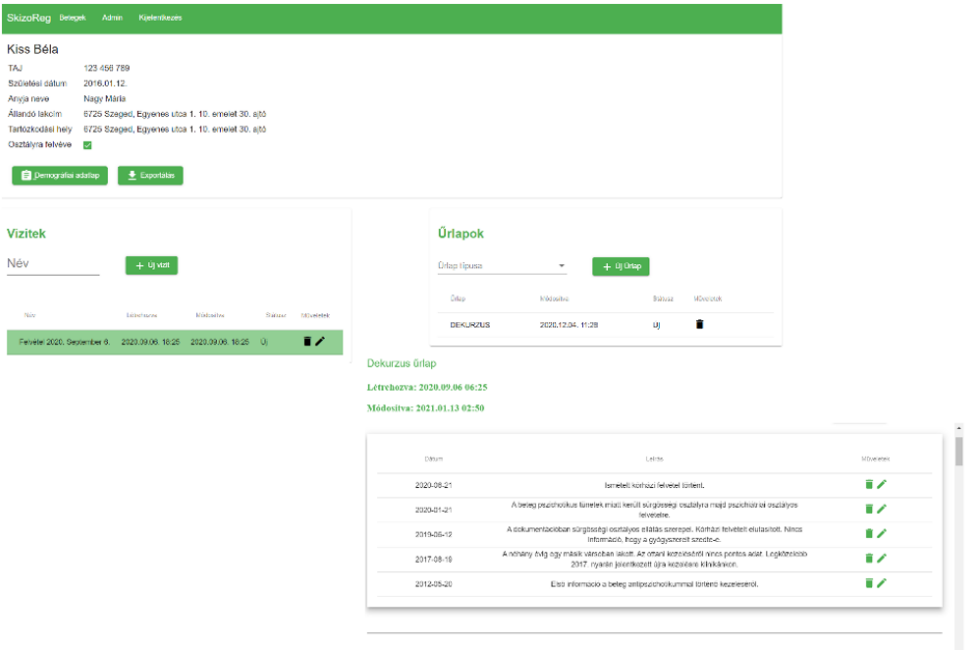


Figure 1: Psychiatric patient registry database developed by project team

4 The research problem, methods, and results

The proposed program focused on the deep mathematical analysis of a wide-range good quality medical data of psychiatric patients. The discipline of statistical machine learning (and in general data mining and analysis) offers a framework which allows researchers to decipher what the computed characteristics reveal, and how these characteristics could be used to offer a decision support tool. A further aim was to investigate whether additional characteristics, which may have been previously ignored, could or should be taken into account [3]. The main goals were to investigate, among others, disease comorbidities, patient stratification, drug interactions and clinical outcome of psychiatric patients. Performing that, apply methods from data mining, machine learning (ML) and text mining on high quality patient data collected in a patient registry database. We should note, that literature of examining medical data considering its temporal (or longitudinal) nature has still been rare and not much effective results has been found in the past years. On the other hand, in the literature many methods and theoretical results are available in the topic such as Bayesian learning methods, trajectory mining using non-Markovian or semi-Markovian time series analytic, functional data analysis, deep learning methods, clustering in different levels of data, just to mention a few approach, see e.g. [4-6].

4.1 The research problem

Our main goal is to investigate the applicability of the above mentioned methods on medical data in detail, summarize and review what the best practice solutions are for analyzing certain types of data, especially focusing on temporal medical data of psychiatric patients. As a benefit, researchers, R&D experts and decision makers would get a clearer picture about the applicability, advantages and limitations of currently available mathematical models applied on medical data.

4.2 The applied methods and some results

For psychiatric patients involved in this study the data collected included PANSS (medical scale used to measure symptom severity of schizophrenic patients) scores, CGI (measures illness severity (CGIS), global improvement or change (CGIC) and therapeutic response) scores, medication data (previous drug history and medication during hospitalization), detailed demographic data and historical disease diagnosis data for each patient. The methods we applied used the features, target variables (for supervised learning) and feature similarities (for unsupervised learning, e.g. clustering) extracted from the collected data. Supervised methods consisted of multivariate logistic regression, decision tree, random forest algorithms and time series prediction algorithms, while unsupervised methods consisted of linear regression, K-means clustering and similarity graph-based clustering procedures.

The following table shows some of the applied methods and their results.

Target variable (to predict)	Method	Finding
CGI scores	Linear regression / Moving average	modest predictability both short (1-2 days) and long term (>1 week)
Diagnosis (Schizophrenia / depression / other)	Multivariable Logistic regression	65% prediction accuracy
Schizophrenic status (avg)	K-means clustering	well separated clusters
Schizophrenic status (avg)	Graph-based clustering	high correlation btw clusters and PANSS means
CGI status change	ARIMA time series model	modest predictability both short (1-2 days) and long term (>1 week)
CGI status change	Markov-chain based time series model	better predictability than ARIMA, interesting patterns

Figure 2: Summary of findings. Experiments were performed in a cohort of 41 patients hospitalized in the Psychiatric Clinic, University of Szeged, in 2019.

Briefly summarizing, overall we could see that some advanced supervised machine learning methods worked well in case of certain target variables (like diagnosis) and some unsupervised models (like clustering) provided some interesting insights and patterns that worth to further investigate by experts. Graph-based methods (such as clustering nodes of a patient similarity graph), due to their strength in visualization, provided some interesting information for clinicians. As a consequence it is worth to

continue the investigations in that direction on a much larger cohort.

We further performed detailed statistical analysis on accurate drug utilization data. We found that changes in negative symptoms of schizophrenia over time correlated with drug medication. We observed that PANSS positive symptoms mean is significantly smaller than negative mean, and for schizophrenic patient clusters (based on CGI and demographic data) it highly correlates with drug medication protocols.

5 Main benefits

For R&D and industry EHR integrated patient register could provide more reliable and analyzable data. One of the main benefits to the pharmaceutical industry is a collection of accurate drug utilization data. Just to mention other potential benefits, significant cost and time reduction in clinical trials (more accurate feasibility assessment, shorter recruitment, no or less source data verification, reduced need for on site monitoring, easier safety reporting) is expected. Using such patient registries post-marketing clinical follow up studies can also be much easily performed.

For clinicians such system (and performing the analytic in the background) provides easily available instant information about the patients that support clinical decisions. It helps to apply clinical guidelines and planned care and facilitates the introductions of quality outcomes and quality improvement. We believe that the performed study and computed characteristics could be used to offer some interesting insights of psychiatric data and may provide a decision support tool, as well.

From research and mathematical modeling perspective access to disease specific patient data and more specific, detailed and customizable data then typical "big data" (e.g. insurance claims, medical records) becomes available.

6 Conclusions

In order to be able to use the huge amount of data generated during the medical care in research and in the routine care of the patients we need structured data that can be statistically analyzed. Systems that meet the needs both of research and patient care must be designed to record data in an appropriate form and should not increase the administrative workload of healthcare personnel. On the other hand, the systems that will have been developed in this way may underpin the almost unimaginable development of health care and medical research. Utilizing such a register system the main goal and challenge of the project was to investigate the applicability of several mathematical methods and their modified versions on medical data. Our main conclusions are summarized as follows:

1. For R&D and industry EHR integrated patient register could provide more reliable and analyzable data, and usability of such systems are not restricted to psychiatric R&D.

2. Machine learning and network approaches on the analysis of medical data are promising, however hard to collect or mine enough reliable data. Withut being exhaustive, a few of our findings:
   (a) Accurate drug utilization data and its detailed statistical analysis.
   (b) PANSS positive symptoms mean is significantly smaller than negative mean.
   (c) Changes in negative symptoms over time correlated with drug medication.
   (d) Schizophrenic patient clusters (based on CGI and demographic data) correlates with drug medication protocols.
3. Investigating major depression using the listed machine learning and graph-based tools may be a better target (than schizophrenia) due to the availability a larger cohort and more reliable data.

References

[1] Kessler, R. C., et al. (2011), *World Mental Health Surveys*, 18(1), 23-33.

[2] Eaton, W. W., et al. (2008), The burden of mental disorders. *Epidemiologic Reviews*, 30(1), 1-14.

[3] Tsanas, A., et al. (2013), A methodology for the analysis of medical data. In: *Handbook of Systems and Complexity in Health*. Springer, New York, NY, 2013. p. 113-125.

[4] Sutton, A. J., et al. (2000), *Methods for meta-analysis in medical research* (Vol. 348). Chichester: Wiley.

[5] Jensen, A. B., et al. (2014), Temporal disease trajectories condensed from population-wide registry data covering 6.2 million patients. *Nature Communications*, 5(1), 1-10.

[6] Ullah, S., & Finch, C. F. (2013), Applications of functional data analysis: A systematic review. *BMC Medical Research Methodology*, 13(1), 1-12.

Thermal Modeling of Electric Machines

Abdelhakim Lotfi[1], **Dániel Marcsa**[2] and **Zoltán Horváth**[3]

[123]Széchenyi István University, Department of Mathematics and Computational Science
[1]E-mail: lotfi@sze.hu
[2]E-mail: marcsa@maxwell.sze.hu
[3]E-mail: horvathz@sze.hu

1 Executive Summary

The development of electric vehicles and hybrid electric vehicles for urban transportation in automotive industry contributes to reduce air pollution and greenhouse gas emissions from metropolitan areas. In electrical machines, both permanent magnets and insulation materials are sensitive to temperature variations. The high temperature effects can lead to performance degradation of electric engine and reduction of its life. In order to prevent degradation of the sensitive parts, the prediction of the temperature distribution inside an electric motor is required at machine design stage in order to control the temperature rise and avoid overheating.

The main contribution of this work is to develop numerical techniques to analyse the thermal effects in electric machines, implement the developed algorithm in multiprocessor environments and apply them to real-world cases. In this study, we take into account the coupling among thermal and electromagnetic problems. Both thermal and electromagnetic processes are considered transient and time-discretization is realized using time operator splitting.

The thermal model has been developed and validated with experimental results. The studied examples have shown that the developed algorithm can be used to better understand the thermal behaviour of the electrical motors and would allow the development of new and more efficient motors.

2 Challenge overview

This work develops a framework for coupling transient thermal and electromagnetic fields to predict the evolution of temperature distribution and magnetic fields inside electric motors. The approach employed is based on finite elements simulations, considering weak coupling between magnetic and thermal models.

The accuracy of the thermal model depends on the material properties and the knowledge of losses in electrical machine. In order to obtain reasonable computation times, the model has been simplified by assuming the following assumptions:

1. To avoid the large computational effort associated with solving a full 3-D electromagnetic problem, the EM part of the model was carried out with the aid of 2D finite element simulations on the cross-section of the PM motor with temperature-dependent material properties.

2. In order to reduce the number of elements used in the 3D thermal model, the windings and the stator were treated as homogeneous medium with equivalent thermal parameters and the effective properties to characterize the thermal behaviour are calculated based on the volume-weighted average over all constituents.

3. For heat transfer through the external surface of the machine, natural convection was considered. On the other hand the internal air gap wos defined as solid and the effective conductivity to characterize the thermal behaviour of the air gap is calculated from empirical correlations.

3 Implementation of the project

The project is the result of cooperation between two departments at the Széchenyi István University, Győr:

1. The academic group: Department of Mathematics and Computational Sciences.
2. The industrial partner: Research Center of Vehicle Industry.

The academic group is composed of three researchers: Abdelhakim Lotfi, Dániel Marcsa and Zoltán Horvath. Their role was to develop the mathematical models, numerical techniques and computer program for coupled magneto-thermal analysis and also to take care of the implementation and validation of the developed algorithm. The thermal model was developed in multiprocessor or multi-core environments using the open source Feel++ software [8] and the electromagnetic losses have been calculated in Maxwell, ANSYS software. The last step was to perform a sensitivity analysis the proposed coupling scheme with respect to time increment and also investigate the mesh size effect in numerical simulation. For validation purposes, the industrial partner provided us with the electric and thermal measurements data made on a 18-slots/16-poles outer-rotor permanent magnet synchronous machine (BMW C1 11 kW) .

4 The research problem, methods, and results

4.1 The research problem: Electromagnetic-thermal coupling model for permanent magnet synchronous machines

This work intends to develop modelling techniques and simulation algorithms to support thermal modeling of electric machines using mathematical modelling and computer simulation. The approach used in this work considers the entire system and takes into account the coupling among thermal and electromagnetic problems.

In order to predict the temperatures of the critical parts in permanent magnet electric machines, a coupled mathematical models of magneto-thermal problem is considered. The mathematical model involves the solution of two submodels and the coupling is established by the temperature dependent materials and the electromagnetic losses.

The process of EM-thermal coupled analysis starts from the (EM) simulation that determines the losses for a given operating point. The losses are then fed into the thermal simulation to estimate temperature under such loss conditions. Since the

material properties are temperature-dependent, they are updated according to the estimated temperature at the end of thermal simulation, and fed back to the (EM) simulation for another iteration. This iterative process stops when the temperatures converge.

4.2 The applied methods

Let $I = [0, t_f]$ be the time interval of interest. The mathematical model involves the solution of two submodels and the coupling is established by the temperature dependent materials and the electromagnetic losses. The two sub-models are presented as follows, [1, 2, 3, 6]:

(HT) 3D Thermal equation	(EM) 2D Magnetic equation
$\rho c_p \partial_t T + \nabla \cdot \left(- \lambda \nabla T \right) = Q,$	$\sigma \partial_t \overrightarrow{A} + \nabla \times \left(\nu \nabla \times \overrightarrow{A} \right) = \overrightarrow{J}_S + \nabla \times \overrightarrow{H}_c,$

(1)

where T is the temperature, c_p is the specific heat capacity, ρ represents the density, $[\lambda]$ is the thermal conductivity matrix, $\overrightarrow{A}$ is the magnetic vector potential, $Q(\overrightarrow{A})$ is the heat source, $\omega \in \mathbf{R}^2$ is the cross-section of the electric motor, $\sigma(T)$ is the electrical conductivity which is dependent on the temperature, the magnetic reluctivity $\nu = \frac{1}{\mu(H)}$ is the inverse of the magnetic permeability which is dependent on the magnetic field intensity, $\overrightarrow{J_S}$ is the external current source and $\overrightarrow{H}_c$ is called the coercive field strength of the permanent magnet. The following boundary and initial conditions can be considered:

$\overrightarrow{n} . \left([-\lambda] \nabla T \right) = h(T - T_0), \quad \text{on} \quad \delta\Omega \times I,$	$\overrightarrow{A} \times \overrightarrow{n} = \overrightarrow{0}, \quad \text{on} \quad \delta\Omega \times I,$
$T(., 0) = T_0, \quad \text{on} \quad \Omega$	$\overrightarrow{A}(., 0) = \overrightarrow{A}_0, \quad \text{on} \quad \Omega$

(2)

where $\overrightarrow{n}$ is normal vector to the boundary, h is heat transfer coefficient and T_0, A_0 are the initial temperature distribution and the vector field at $t = 0$. The thermal and electromagnetic problems are interconnected by the following relations [4, 5]:

$\sigma(T) = \sigma_0 \frac{1}{1+\alpha(T-T_0)},$	$Q(\overrightarrow{A}) = K_h f(B_m)^2 + K_c (f B_m)^2 + K_c (f B_m)^{1.5},$

(3)

where σ_0 is the electrical conductivity at reference temperature T_0, α is temperature coefficient, K_h, K_c, K_e are the static hysteresis loss, eddy current loss and excess loss coefficients and B_m is the peak flux density where $\overrightarrow{B} = rot(\overrightarrow{A})$.

5 Solution Techniques for the coupled problem

The electromagnetic problem (EM) can be analysed as (2D) problem if we assume that the physical behaviour of the machine is the same in axial direction. In this case, the currents circulate according to the longitudinal direction (Oz) and the magnetic vector potential have only z component: $\overrightarrow{A} = (0, 0, A_z)$. Standard continuous finite elements

are used for the discretization of the magnetic potential A_z and the temperature T. In three-dimensional space (3D) we use tetrahedral nodal element (P1) and in two-dimensional space (2D) we use triangular elements (P1). The FEM discretization results in two systems of differential-algebraic equations:

$$[M_{HT}]\partial_t[T] + [K_{HT}][T] = [F_{loss}(A_z)] + [F_{ext}], \tag{4}$$

$$[M_{EM}(T)]\partial_t[A_z] + [K_{EM}(\nu)][A_z] = [F_{J-H}]. \tag{5}$$

where $[M_{HT}]$, $[M_{EM}]$, $[K_{HT}]$ and $[K_{EM}]$ are mass and stiffness matrices. $[F_{loss}]$ and $[F_{ext}]$ are the right hand side load vectors.

The equations (4) and (5) are coupled because the temperature resulting from the thermal computation is used for evaluation of the material characteristics in the magnetic equations and the power losses resulting from the magnetic computation are used as heat sources in the thermal model. The proposed numerical method to solve the coupled problem is based on independently and successively solving each sub-problem and then transferring the results between the two sets of the equations. The coupling is realized weakly due to the different time scales for (HT) and Maxwell equations. Since the the thermal phenomena is slowly varying in comparison with electromagnetic field, the coupled problem is discretized in time with a time step reflecting the thermal scale to avoid unreasonable computational times.

Let us consider that the simulation time period is split into time intervals: $t_{initial} = t_0 < t_1 < ... < t_n < ... < t_N = t_f$, t_n is the time where the data is exchanged between the two sub-problems. The size of each interval is determined based upon the variation of material properties in the electromagnetic problem. In this approach we assume that the temperature variations are not significant, meaning that the electromagnetic properties remain approximately the same in each sub-interval. The main interest of such a decomposition is that the (HT) problem can be solved several times over the sub-interval $[t_n, t_{n+1}]$. The algorithm for the time period between t_n and t_{n+1}, is composed of 2 steps [1, 2, 3]:

1. The iteration process start from an (EM) simulation and the recalculation of the electromagnetic losses. The average of all temperatures calculated over the previous time period is used to determine the electrical conductivity $\sigma(T)$ in (EM) problem.

2. The losses are then fed into the thermal simulation and the (HT) problem continues to be solved without correcting the heat sources during certain times until the variation of the electrical conductivity not exceed certain limit. As soon as the electrical conductivity variation becomes significant, the calculated temperatures are fed back to the (EM) simulation for another time period.

5.1 The achieved research results

The developed numerical technique is used to simulate distributions of electromagnetic and thermal fields in Toyota Prius 2004 electric motor, as shown in the left figure in Fig. (1). The 3D mesh used for the for the thermal analysis has more than 7 500 000 cells and it is shown in Fig. (1). The electromagnetic analysis is performed on a periodic section of the PM considering a soft ferromagnetic material with the B(H) curve shown in Fig. (1). The thermal simulation was executed on 60 cores using the FEEL++ library, [8]. The electromagnetic simulation was performed with the help

of the finite element package ANSYS Maxwell. The figures in Fig. (2) represent the

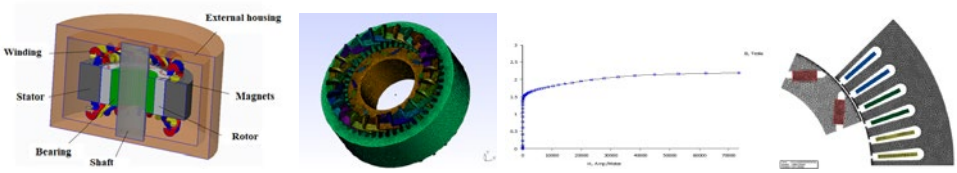


Figure 1: Toyota electric motor; 3D mesh for (HT) analysis; B(H) curve; 2D mesh

temperature distribution, magnetic flux density distribution with equipotential lines and Ohmic loss in magnets with magnetic flux density vectors. These results are obtained at maximum current of 100A and angular velocity 3000 rpm. The first two

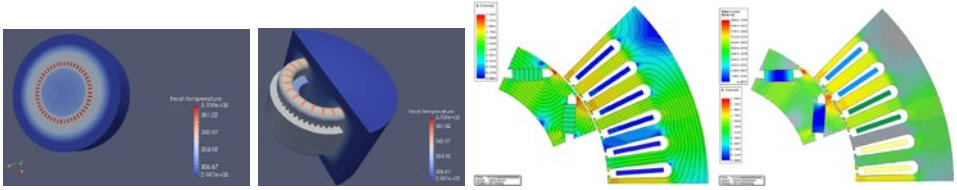


Figure 2: Contours of temperature, magnetic flux density and Ohmic loss.

figures in Fig. (3) show the calculated time evolution of the temperature and losses in different parts of the PM motor and the last two represent the convergence history of the relative errors.

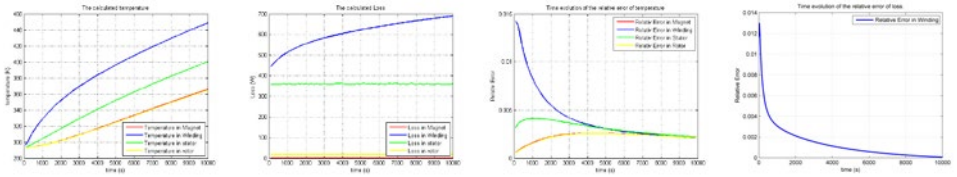


Figure 3: Time evolution of the temperatures, losses and relative errors

In order to study the effect of the coupling period on the simulation, the developed coupled algorithm has been applied to the previous example using 5 different values for fixed time period (from 1 sec to 30 sec). The temperature and losses at selected points in the critical parts of the electric machine are listed in the following tables at time $t = 4000s$:

$\Delta\tau$ (s)	Magnet temp	Winding temp	Stator temp	Rotor temp
$\Delta\tau = 1$	343.0430	377.3619	360.9835	342.3611
$\Delta\tau = 5$	343.0035	377.3375	360.9363	342.3219
$\Delta\tau = 10$	342.9131	377.1765	360.8151	342.2326
$\Delta\tau = 20$	342.7220	376.8425	360.5590	342.0442
$\Delta\tau = 30$	342.4257	376.3866	360.1886	341.7508

$\Delta\tau$ (s)	Magnet loss	Winding loss	Stator loss	Rotor loss
$\Delta\tau = 1$	1.4011	579.9021	354.6399	18.5773
$\Delta\tau = 5$	1.4031	580.9592	354.5072	18.5649
$\Delta\tau = 10$	1.4027	580.5475	354.5347	18.5645
$\Delta\tau = 20$	1.4030	580.9519	354.5126	18.5628
$\Delta\tau = 30$	1.4026	580.5139	354.5379	18.5664

From the results presented in previous tables, we notice that the relative errors are less than 0.1% in all parts of the PM motor. The results show that a large coupling period can be used in coupled algorithm and the coupled algorithm can predict accurate temperature fields and losses at a reasonable computational cost.

To test mesh size sensitivity for the developed coupled algorithm, a series of models with the same geometry but different element sizes were used. The thermal problem was resolved using these different meshes. On the other hand the electromagnetic losses have been calculated using the very fine mesh because the sensitivity analysis of the mesh applied to the electromagnetic problem showed that the electromagnetic losses increase as the grid size decreases in all parts of the PM motor while the losses remain constant in the winding, [1]. The meshes were obtained by refining only the critical parts of the electric machine. The number of elements increases from the coarsest mesh of 2063369 cells to the finest mesh of 11333661 cells. The temperature at selected points in the critical parts of the electric machine are listed in the following tables at time $t = 8000s$:

Nb of elements	Winding temp	Magnet temp	Stator temp	Rotor temp
2063369	422.4599	372.6781	409.2707	374.5502
3796499	423.3494	372.7312	409.3348	374.5674
6796499	424.0866	372.9473	409.4043	374.6465
11333661	424.8905	372.9612	409.4050	374.5605

6 Solution of the industrial problem and its benefits

To validate the effectiveness of the coupled thermo-magnetic model and code, the second example examines the transient thermal analysis of a 18-slots/16-poles PMSM with exterior rotor (BMW C1 11 kW) manufactured by the Vehicle Developing Centre of the Széchenyi István University, [7].

The measurements were performed on the BMW C1 11 kW motor presented in Figure (4). The temperature was measured using a thermocouple type PT100 in

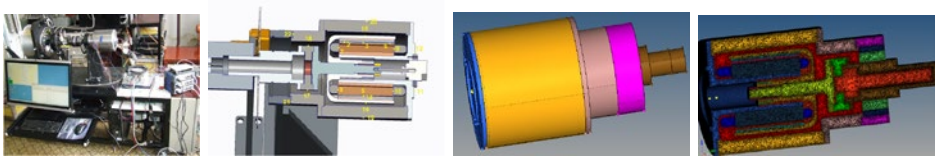


Figure 4: Test set up for the prototype machine, sensors position, geometry, 3D mesh.

contact with the surface of the tested motor. As shown in Figure (4), the electric motor model has a very complicated geometry. The stator coils are supplied with a maximum current of 100A for duration of 1800 seconds and then the current is turned off; the rotor is locked in a stand still position. During the temperature test, the most significant heat source in the tested motor is copper losses. The left figure in (5) shows the instantaneous Joule losses calculated from experimental data, used in thermal analysis. The computational grids used for this simulation, shown in Fig (4), has approximately 4 000 000 tetrahedral elements. According to the model mentioned

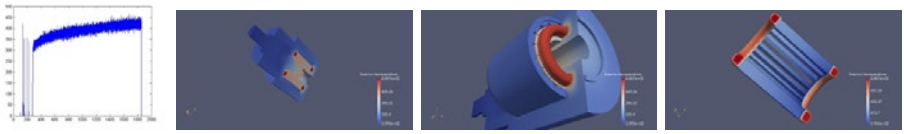


Figure 5: Joule losses from experimental data, contours of temperature.

above, a transient 3D FE thermal was carried out using the FEEL++ library [8] and it was executed on 30 processors. Figures in (5) show the temperature distribution of the motor at time t = 1600 s. Fig. (6) shows a comparison of the simulated and

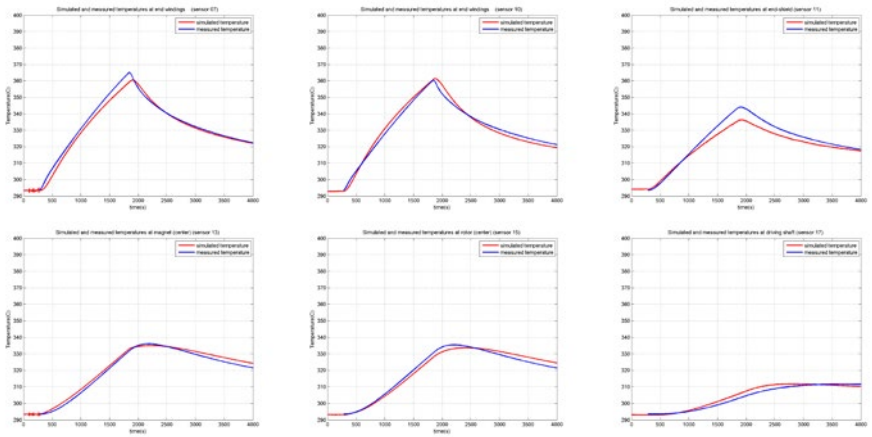


Figure 6: Simulated and measured temperature variation at different position.

measured temperature variation obtained by a few thermocouples [7]. The simulated temperatures were very accurate for locations in the motor obtained by thermocouples 1, 2, 3, 4, 5, 6, 9, 10, 11, 12. The temperature results showed a small difference but they had the same trends with the measured temperatures obtained at the remaining thermocouples. The difference between transient temperature simulation and the results of the experiments can be attributed to the following: the glue layers between the thermocouples and the prototype motor can decrease heat transfer, the material properties were assumed to be independent of temperature.

7 Conclusions

In this work, a detailed 3D thermal model of (PM) motors has been presented and demonstrated the effectiveness of this approach by comparisons of the FE-based simulation and experimental results.

The developed method gives a fast and accurate way to evaluate magnetic field intensity, the core losses in the material and the temperature distribution of PM motor for different currents and for different geometrical parameters. The results obtained can contribute to a better understanding of thermal behaviour in electric motors and can help the designer to develop a new and more efficient motors.

The developed model enables to predict temperature distribution with good accuracy of the critical parts of the electric machine such the winding, the rotor and the magnets without using the time-consuming CFD and 3D electromagnetic simulations. Convection heat-transfer problems are treated with dimensionless numbers and empirical correlations are used to determine heat-transfer coefficient.

Acknowledgements

This work was supported by the (EFOP-3.6.2-16-2017-00015) program (Center for cooperation between higher education and the industries at the Széchenyi István University).

References

[1] A. Lotfi, D. Marcsa, Z. Horvath, C. Prudhomme, and V. Chabannes 'Coupled physics solvers for analysing the magneto-thermal behaviour of a permanent magnet synchronous machines', Lecture Notes in Computational Science and Engineering: Numerical Mathematics and Advanced Applications ENUMATH 2019, Volume 139 Chapter 68 (2021)

[2] A. Lotfi, D. Marcsa, Z. Horvath, 'Modeling the Thermal Behavior of Permanent Magnet Synchronous Motors' I.J.E.M.S. 4 : 1 pp. 466-477, 12 p. (2019)

[3] A. Lotfi 'Numerical solver for analyzing the thermal behaviour of permanent magnet synchronous machines' AIP Conf Proc 2116 Paper: DOI: 10.1063/1.5114557, 6 p. (2019)

[4] J. Gieras, R. Wang and M. Kamper 'Axial Flux Permanent Magnet Brushless Machines', Springer (2008).

[5] P. H. Mellor, D. Roberts and D. R. Turner 'Lumped parameter thermal model for electrical machines of TEFC design' IEE Proceedings-B, 138 (1991) 5, p. 205-218.

[6] A. Arkkio, T. Jokinen, and E. Lantto, 'Induction and permanent-magnet synchronous machines for high-speed applications,' in Proc. 8th ICEMS, Sep. 27-29, 2005, vol. 2, pp. 871-876.

[7] Kuslits M. 'BMW C1 motor rovidzárási hőmérsékletmérése: Mérési jegyzokonyv', 2014. (in Hungarian language)

[8] The Feel++ Book: https://book.feelpp.org/

Traffic modelling and simulation by sparse data recovery and refinement

Attila Lovas[1] and **Balázs Gerencsér**[2]

[1] Alfréd Rényi Institute of Mathematics and Budapest University of Technology and Economics, email: lovas.attila@renyi.hu
[2] Alfréd Rényi Institute of Mathematics and and Eötvös Loránd University, email: gerencser.balazs@renyi.hu

1 Executive Summary

Getting an insight on emission by traffic with high precision is a key element of future smart cities, and its precision goes hand in hand with the precision of the knowledge on traffic. The project is on improving the latter, taking a step further the challenge of giving a high resolution, realistic dynamics that can recover changes and fluctuations on the streets despite having only sparse and aggregated measurements.

We propose a mathematical model for the time evolution of the traffic, encompassing ideas from literature, physical intuition and novel additional destination parametrization. Change in time is formulated for road segments and junctions separately, as is needed for their different structure, this also reveals the potential for parallelization in execution.

The model is also coded and proven to work on moderate examples. Simple data fusion is provided in code and demonstrated with the supplied data, with the proposal of certain refined data assimilation techniques in theory.

The outcome of the project allows a reasonable emission estimation once used as an input source. Moreover, there is even more potential when further data is added, or when data assimilation is upgraded. The code is implemented in a modular form waiting for this future.

2 Challenge overview

The insiration for the project is to support emission prediction and simulation for urban areas which could serve as a tool for related decision makers. An intermediate, but well defined target is to estimate and simulate fluctuations of traffic and its characteristics, consistent with available measurement data, and separate emission calculation as a future project, where traffic can be taken as input to work with.

Given as initial data the map of the city is available, and aggregated measurements from loop-detectors, that is, the number of cars passing certain lanes in total during each hour. A big challenge comes from dimension mismatch: the number of instantaneous parameters of speed, density on every road segment is incomparably larger to the dimension of the above, time-aggregated data. Therefore a model has to be composed that is both realistic in terms of traffic dynamics, and also stable for fitting to the sparse data available.

Depending on the future utilization of the project - say, for analysis or for on-field feedback control - a necessity of performance might be added to the priority list.

3 Implementation of the project

The role of Magyar Közút Nonprofit Zrt. was on supporting the project with the measurement data. The testing area was for the city of Győr, a major but controllable size city of 100k+ inhabitants. This was a good scale to have not only smaller partial tests but increase the challenge to somewhat more complexity, without getting into the metropolitan regime of 1M+ with immense complications out of range currently.

Attila Lovas and Balázs Gerencsér were representing the academic area. The research carried out shows both the complexity and the potiential of the question being investigated, a model has been built, prototyped and tested for small parts of Győr. Open questions have been posed for potential follow-up.

4 The research problem, methods, and results

The research aiming at the challenge splits into two directions. First, a modeling, simulation framework, that can provide an analytical - and then numerical - evolution of the state of the traffic as time passes. Second, fitting this dynamical system to the data points that appear once in a while in space-time in the best possible way.

For the first, we construct a system of partial differential equations, using knowledge from literature and physical representation. This is extended by having different "types" of vehicles depending on their destination. For the second, a simple adaptive resetting approach already numerically demonstrates realistic simulations. This is hot-swappable with elaborate data assimilation techniques, but the increased computational power was not available for testing and tuning during the project.

We have our code written in a structured, modular way so that it can be easily upgraded, even partially. This is a realistic case we should be prepared for in the mid-term future if additional sensors are deployed and more, different type of measurements are added to the input data stream, when traffic light signals are handled or when the data filtering algorithm is replaced.

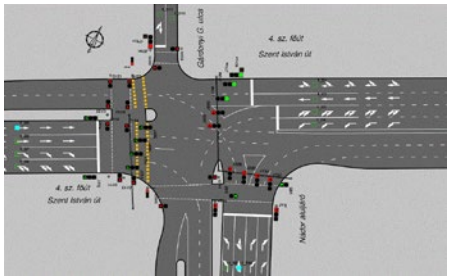
4.1 The research problem

In the first part, the goal is to formalize a dynamical sytem consistent with the map as a constraint space for capturing the traffic with its proper state space. It has to have a well defined time-evolution that is numerically tractable and reasonably stable. We should seek the good compromise for conceptual resolution: if it is too coarse, we might miss key phenomena, if too detailed, dimension might inflate and make the next part much harder.

The second part is about data fitting. The goal is to have the dynamical sytem developed emulate city traffic through the data points acquired. Critical traffic is expected at major roads, thus sensors are located accordingly as it can be seen in Figure 1a. Zooming in we see an example image at Figure 1b, showing exact arrangement

of lanes, traffic lights. Note however, the cyan markings for the loop-detectors show that not all lanes are measured at every intersection. This is an additional noise to be aware of. The disparity of the dimension of the dynamical system and the data makes this heavily non-trivial.

(a) in the city



(b) at intersections

Figure 1: Sensor locations in Győr (Hungary)

4.2 The applied methods

We use a macroscopic description of traffic in view of dimension reduction, leading to differential equation descriptions, both along edges and at nodes. Highlighting a novelty, we label vehicles (densities, in the end) based on their destination. This gives room for a more refined description, allowing to better capture e.g., morning rush to the city center.

4.2.1 System dynamics

For data fitting, we have built our dynamical system with easy interfacing to be able to test multiple variants. Until now testing the simplest merging methods, to testing recent Kalman-style algorithms, even more is possible in the future.

Transport network modeling

The directed graph $G = (V, E)$ describes the spatial network permitting vehicular movement. Besides the topological structure, for simulations, we need further information about the vehicular transport on road segments (edges) and at intersections (nodes). We used the Python package called OSMnx to access and manipulate freely

Edges	**Nodes**
Number of lanes	Priority rules
Length	Allowed transitions
Speed limit	Traffic light states

Table 1: Edge and node fields required for traffic simulation.

available data from OpenStreetMap [2]. In Table 1, we collected edge and node data used as input by our mathematical model. Most of these are available directly through the OSMnx API but for example, traffic light states are provided by our industrial partner. Furthermore, precise information about priority rules and allowed transitions at intersections are not available directly and thus we need to infer these information from OpenStreetMap data. We applied semi-heuristic considerations to retrieve priority rules.

We applied a clustering by the mean shift algorithm to annotate suburbs, giving K clusters (Figure 2), where cluster centers serve as destinations of vehicles. In the sequel, $c(v)$ denotes index of the cluster to which $v \in V$ belongs. Actually, apart

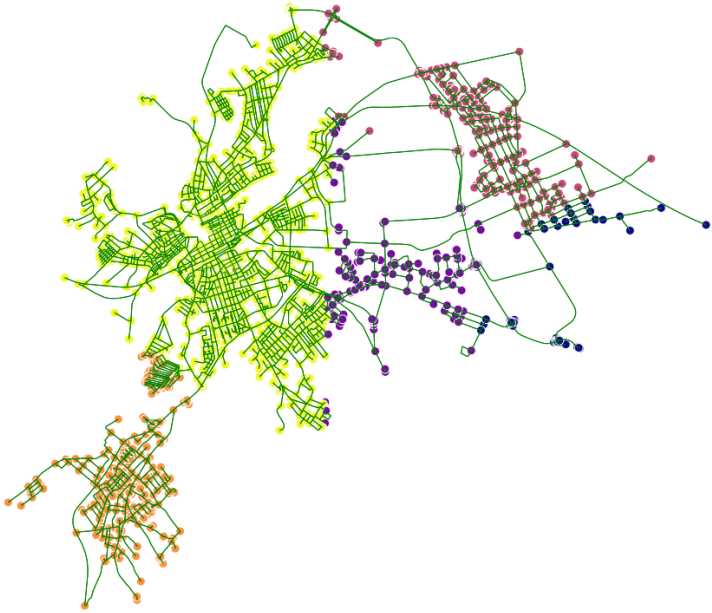


Figure 2: Clustering of Győr (Hungary)

from clustering, we can use other more simple methods to define the destinations of vehicles. For example, we can select either randomly or deterministically boundary points of the graph and so-called POIs (Points Of Interests) with their location also available through OSMnx.

We take on a macroscopic approach, analyzing densities of vehicles instead of particles, but refine this by dissecting it for destinations. We discretize edges in a way that we fix the desired cell size Δ and for $e \in E$, we cut the edge e into N_e equal cells such that the length of each cell Δ_e approximately equals Δ.

As a state descriptor, we introduce $(\rho^e, v^e)_{e\in E}$, where $\rho_i^e \in [0,\infty)^K$ is the vector valued density and $v_i^e \in [0,\infty)$ stands for the traffic speed in the i-th cell, $i = 1,\dots,N_e$. In this formalism, for $i \in \{1,\dots,N_e\}$ and $j \in \{1,\dots,K\}$, $\rho_{i,j}^e$ refers to the density of those vehicles in the i-th cell which have their destination in the j-th cluster. To make the notation compact, we denote the total vehicular density in the

i-th cell by $\rho^e_{i,\cdot}$ that is $\rho^e_{i,\cdot} = \sum_{j=1}^{K} \rho^e_{i,j}$. Furthermore, $\rho^e_{-1,\cdot}$, $\rho^e_{-1,j}$, $j \in \{1,\dots,K\}$, and respectively v^e_{-1} refers to the traffic density and speed in the last cell of the edge $e \in E$.

Evolution along edges

Let us now present the dynamical system that evolves the traffic along the edges. Let us consider an edge $e \in E$. The vehicular flow on e is described by a system of ordinary differential equations which we obtained through the semi-discretization of the Aw-Rascle-Zhang partial differential equation [1, 6, 3]. The only difference is that in our model, the traffic density is vector valued and thus the evolution equations for all $j = 1,\dots,K$ take the form

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} \rho^e_{i,j} \\ y^e_i \end{bmatrix} &= -\frac{1}{\Delta_e}\left(\begin{bmatrix} v^e_i \rho^e_{i,j} \\ v^e_i y^e_i \end{bmatrix} - \begin{bmatrix} v^e_{i-1}\rho^e_{i-1,j} \\ v^e_{i-1} y^e_{i-1} \end{bmatrix}\right) - \frac{1}{\tau}\begin{bmatrix} 0 \\ y^e_i \end{bmatrix}, \ i \geq 2 \\ \frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} \rho^e_{1,j} \\ y^e_1 \end{bmatrix} &= -\frac{1}{\tau}\begin{bmatrix} 0 \\ y^e_1 \end{bmatrix} \end{aligned}, \tag{1}$$

where $\tau > 0$ is a model parameter describing the relaxation to the equilibrium speed, which is given by

$$V^e_{eq}(\rho) = v^e_{\max}\left[1 - \left(\frac{\rho}{\rho^e_{\max}}\right)^a\right]^b, \tag{2}$$

and $y^e_i = \rho^e_{i,\cdot}\left(v^e_i - V^e_{eq}(\rho^e_{i,\cdot})\right)$ is the relative flow, i.e. the difference between the actual flow and the equilibrium flow. In the definition of $V^e_{eq}(\rho)$, model parameters $a, b > 0$ are positive dimension free constants, $v^e_{\max}$ is the speed limit on $e \in E$, and the maximum traffic density is defined by

$$\rho^e_{\max} = \frac{\text{Number of lanes}}{\text{Safety distance}}.$$

In simulations, we set $\Delta = 10\,\mathrm{m}$, $a = 2.0$, $b = 6.0$, $\tau = 10\,\mathrm{s}^{-1}$, and for the safety distance, we use $7.5\,\mathrm{m}$.

Evolution at nodes

We turn to address the kinetic model we refined to describe the vehicular transport at nodes. In order to present this model, we need some further notations. For the in- and out-neighbors of a node $v \in V$, we introduce $\mathcal{N}_{in}(v)$ and $\mathcal{N}_{out}(v)$. Furthermore, we define the so-called "next node matrix" $P(v) \in \mathbb{R}^{\mathcal{N}_{out}(v)\times K}$ which can be considered as a transition matrix, where $P(v)_{i,j}$ gives the ratio of vehicles moving towards the j-th cluster and choosing the i out-edge compared to the total amount of vehicles with destination in the j-th cluster. Roughly speaking, $P(v)_{i,j}$ is the probability of choosing $i \in \mathcal{N}_{out}(v)$ under the condition that the destination is in the j-th cluster.

In our model, priority rules are represented by a partial order on the set $\mathcal{N}_{in}(v) \times \mathcal{N}_{out}(v)$. For, $(i_1, o_1), (i_2, o_2) \in \mathcal{N}_{in}(v) \times \mathcal{N}_{out}(v)$, we say (i_1, o_1) is inferior to (i_2, o_2), $(i_1, o_1) \prec (i_2, o_2)$, if the transition $i_2 \to o_2$ has higher priority than $i_1 \to o_1$. The traffic light matrix $L(v) \in \mathbb{R}^{\mathcal{N}_{in}(v)\times\mathcal{N}_{out}(v)}$ is a 0-1 matrix, where $L(v)_{i,o} = 1$ if the

transition $i \to o$ is allowed and $L(v)_{i,o} = 0$ otherwise. By a slight abuse of notation, we use the same symbol for neighboring nodes and edges connecting v to its neighbors.

The following system of differential equations describes the dynamics governing the vehicular flow at the intersection $v \in V$:

$$\left.\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\rho^i_{-1,j} &= -C_{i,j}[AB]_{i,j} - \alpha\rho^i_{-1,j}\mathbf{1}_{c(v)=j},\ i \in \mathcal{N}_{in}(v) \\ \frac{\mathrm{d}}{\mathrm{d}t}\rho^o_{1,j} &= B_{o,j}[A^\top C]_{o,j},\ o \in \mathcal{N}_{out}(v) \end{aligned}\right\},\ j = 1, \dots, K. \tag{3}$$

In this model, matrices $A \in \mathbb{R}^{\mathcal{N}_{in}(v)\times\mathcal{N}_{out}(v)}$, $B \in \mathbb{R}^{\mathcal{N}_{out}(v)\times K}$, and $C \in \mathbb{R}^{\mathcal{N}_{in}(v)\times K}$ are given by

$$\begin{aligned} A_{i,o} &= L(v)_{i,o} \times \exp\left(-\sum_{l\in\mathcal{I}^v(i,o)} \frac{q^l_{-1}}{q^l_{\max}}\right) \\ B_{o,j} &= \left(1 - \frac{\rho^o_{1,\cdot}}{\rho^o_{\max}}\right) P(v)_{o,j} \\ C_{i,j} &= \frac{\max(v^i_{-1}, v_{\text{start}})}{\operatorname{diam}(v)}\rho^i_{-1,j}, \end{aligned}$$

where $\alpha > 0$ is a kinetic constant describing vehicles leaving the traffic flow, $c(v)$ provides the cluster id of node v. $\mathcal{I}^v(i,o) \subseteq \mathcal{N}_{in}(v)$ denotes the subset of in-edges superior to (i,o), that is $\mathcal{I}^v(i,o) = \{v' \in \mathcal{N}_{in}(v) \mid \exists\, o' \in \mathcal{N}_{out}(v)\ (i,o) \prec (i',o')\}$. Furthermore, $q^i_{-1} = \rho^i_{-1,\cdot}v^i_{-1}$ stands for the actual, and $q^i_{\max}$ for the maximal possible in-flux from $i \in \mathcal{N}_{in}(v)$. In the definition of $C_{i,j}$, the kinetic constant is defined by the ratio of $\max(v^i_{-1}, v_{\text{start}})$ and $\operatorname{diam}(v)$, where this latter is the characteristic diameter of the section, and v_{start} denotes the minimum speed of vehicles passing through the section. Note that, $\frac{\mathrm{d}}{\mathrm{d}t}\left(\sum_{i\in\mathcal{N}_{in}(v)}\rho^i_{-1,j} + \sum_{o\in\mathcal{N}_{out}(v)}\rho^o_{1,j}\right) = -\alpha\mathbf{1}_{c(v)=j}\sum_{i\in\mathcal{N}_{in}(v)}\rho^i_{-1,j}$ and thus the dynamics satisfy Kirchoff's first law with sinks at nodes at destination. In simulations, we set $\operatorname{diam}(v) = 20\,\mathrm{m}$, $v_{\text{start}} = 2.0\,\mathrm{ms}^{-1}$, and $\alpha = 10\,\mathrm{s}^{-1}$.

4.2.2 Data assimilation – state estimation

Data Assimilation (DA) is a standard technique in meteorology, oceanography and in geology for combining partial or even sparse observations of a very high $(10^4 - 10^6)$ dimensional system, with estimates of that system from a dynamical model, typically arising as the semi-discretization of certain partial differential equations, to obtain a new and more accurate description of the system [5].

In recent years, the new trend in traffic estimation is to develop data driven approaches. Especially, researchers turned their attention to DA modeling methodologies. For example, a particle filter based approach with Kriging interpolation was proposed in [4] to give more accurate prediction of traffic from sparse sensor data.

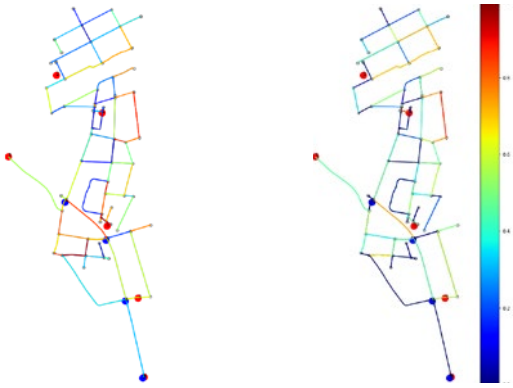
Due to the absence of pointers and other serious limitations of the Python language, we decided to implement a primitive DA algorithm employing a simple Bayesian estimation. At each time step between the observations, the state vector is propagated using the dynamical model presented in the previous sections. At observation time,

the traffic density on measured edges are updated using the sensor data to a more accurate estimate of the traffic.

4.3 The achieved research results

The main innovations of our modeling work without the sake of detailed analysis is a vector-valued vehicular transport model describing the traffic on the whole road network taking into account the effect of traffic lights and priority rules. The operator splitting approach we applied alleviated the problem of coupling between equations allowing to use parallel solvers and thus taking the advantage of multi-core architectures.

We proposed a simple Bayesian data assimilation algorithm extensively using the "next node matrix" structure of the graph for fitting the state vector to data. On edges



(a) initial traffic (b) after 3 minutes

Figure 3: Simulation of congested traffic in Győr (Hungary). Cluster centers are marked by red dots and blue dots were used for loop detector positions. Edges are colored according to their relative saturation i.e. $\rho^e/\rho^e_{\max} \in [0,1]$.

equipped with traffic sensors, predictions fits well to measurement data, moreover simulation results show realistic behavior on non-measured edges as well (See Figure 3). If traffic camera data is available, by using our Bayesian update model, we can gain more accurate estimation about the destination of each vehicles.

5 Solution of the industrial problem and its benefits

For the long-term goal of getting an insight on emission the current project serves as a step forward to provide an input source for the next phase, when the question is modeling emission, given high resolution traffic estimate.

6 Conclusions

During the research involved, an algorithm has been developed with multiple innovative ingredients. Although the theoretical foundation is solid, for the numerical testing we have reached our limitations. For instance, data was not conveniently preprocessed, importing became lagging. Also, we have reached the plateau with Python in terms of performance, which was a perfect choice for prototyping, but for application we will need a parallelism-oriented HPC aware re-implementation to scale it up.

References

[1] A. Aw and M. Rascle, *Resurrection of" second order" models of traffic flow*, SIAM journal on applied mathematics, 60 (2000), pp. 916-938.

[2] G. Boeing, *Osmnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks*, Computers, Environment and Urban Systems, 65 (2017), pp. 126-139.

[3] S. Fan and B. Seibold, *Data-fitted first-order traffic models and their second-order generalizations: Comparison by trajectory and sensor data*, Transportation research record, 2391 (2013), pp. 32-43.

[4] K. J. Offor, L. Vaci, and L. S. Mihaylova, *Traffic estimation for large urban road network with high missing data ratio*, Sensors, 19 (2019), p. 2813.

[5] S. Vetra-Carvalho, P. J. Van Leeuwen, L. Nerger, A. Barth, M. U. Altaf, P. Brasseur, P. Kirchgessner, and J.-M. Beckers, *State-of-the-art stochastic data assimilation methods for high-dimensional non-Gaussian problems*, Tellus A: Dynamic Meteorology and Oceanography, 70 (2018), pp. 1-43.

[6] H. M. Zhang, *A non-equilibrium traffic model devoid of gas-like behavior*, Transportation Research Part B: Methodological, 36 (2002), pp. 275-290.

Modeling General and Geographical Health Risk Factors for Health Insurance Premium Calculation

László Márkus[1], **Miklós Arató**[2], **Ashish Kumar**[3]

[12]Department of Probability Theory and Statistics, Loránd Eötvös University, and Department of Mathematics and Computational Sciences, Széchenyi István University,
[3]CITIBANK Hungary and Department of Mathematics and Computational Sciences, Széchenyi István University,

[1]E-mail: laszlo.markus@ttk.elte.hu
[2]E-mail: miklos.arato@ttk.elte.hu
[3]E-mail: ashish.kumar@citi.com

1 Executive Summary

Segmented disease incidence data is not available for insurance companies in sufficient amounts and detail because of legal reasons and the lack of resources. The research objective is to make a segmented risk analysis using publicly available general population data and locality attributes for the 3153 localities in Hungary. The company initiating the research is particularly interested in whether gender and age are risk factors and, if so, how their effect can be quantified. Founding and quantifying other risk factors are also of interest. The answers are given using a Poisson generalized linear model with additive and multiplicative explanatory variables and their parameter estimation by conditional likelihood and iteration. The research concluded that gender is not a significant risk factor when controlling for age and quantified the risk represented by age cohorts. These findings are useful for the company in reserves calculations and facilitate correct premium calculations for health insurance products while providing for premium advantage. Budapest stands out as an outlier, which necessitates extended research with a proper database as a natural continuation of the present research.

2 Challenge overview

The gross premium income of all commercial health insurances in Hungary totaled 12.4 billion Hungarian Forints in 2016. Although a significant amount, it still hardly exceeds 1% of the overall premiums in Hungarian insurances, while the same proportion in the EU context exceeds 10%. Based on these data, a significant development in commercial health insurance can be anticipated in Hungary, but it has never materialized as yet. The fundamental reasons lay in legal and tax regulations, but in part, the lack of analyzed data may also contribute to Hungary's lagging behind the rest of the EU.

The insurance companies' current assortment of health insurance packages consists only of fixed amount insurances, meaning the insurer pays an advance-set-compensation for hospitalization days, surgeries, tumors, etc. This way, the insurers

groundlessly limit their liabilities and risks in these products, because in fact, the payout depends only on claim numbers and does not depend on the severity of the cases. The newly built private hospitals, however, advertise, e.g., insurances fully covering surgery treatment costs. Here the severity of the cases plays an important role, especially when recovery and recreational treatments are also covered. Therefore, the insured party's overall health status and its exposure to health risk factors heavily influence claim occurrence and amount and hence is of utmost importance in premium calculation.

So far, commercial insurance companies could collect only very limited amount of Hungarian health insurance claim history and data. Therefore, they have to rely on the publicly available Hungarian data and their experience gained in foreign countries. These, however, are not sufficiently segmented to distinguish possible risks. Insurers are not allowed to differentiate in premium calculations according to all possible risk factors, in particular, according to the gender of the insured. Nevertheless, exact knowledge about the risk associated with a given contract is essential in business analysis and reserves calculations.

3 Implementation of the project

The industrial partner provided the problem description consulted on the possible data to be used and provided consultation with medical experts. The academic party created the model, implemented it and analysed the reliability of the obtained results. LM and MA has built the model for risk factors, LM computed the conditional likelihood estimator, MA collected, suitably modified and implemented the methods for risk factor selection. AK implemented the estimators analyzed their spatial dependence and created the probability maps.

4 The research problem, methods, and results

In the context of what has been said before the research problem is

1) to identify the nature and quantify the extent of influence of the risk factors of diseases and mortalities based on
    i) spatially (by location) referenced disease data (number of prescriptions and patients associated with predefined disease groups)
    ii) town and village attributes (number of inhabitants, incomes, unemployment, transportation availability/time, drink water quality, spatial alignment, etc.),
2) to explore the difference between the expected and observed number of incidences of deaths or diseases after filtering out the effects of explanatory variables.

Based on the company's request particular attention was payed to gender and age as risk factors.

4.1 The research problem

The objective of the research is to determine the effect of gender, age, income, number of divorces (indicating stress), and the presence of the nitrification cycle traces (ammonia, nitrite, and nitrate concentration) in drinking water as potential risk factors for cardiovascular, gastrointestinal, and musculoskeletal diseases. As no segmented disease incidence data is available according to these factors, we can only use the number of overall incidences in a locality together with locality attributes such as prescription numbers, number, gender, age and income of the inhabitants, and water quality data in the locality. These data are publicly available for scientific use for all the 3153 settlements (from the capital Budapest to the smallest hamlet) in Hungary in the National Area Development and Management Information System "TeIR" database.

The disease types were selected because these belonged to the interest of the industrial partner and consultations with medical experts revealed that the data at our disposal suited best for these. Age cohorts were also selected by company interest and medical expert suggestions, no mathematical tool has been utilized in that.

4.2 The applied methods

The number of disease incidences is an integer in every locality. Therefore a simple linear model, i.e., a classical regression with risk factors as explanatory variables and the incidence numbers as response variable observed in every locality, is not applicable. Taking into account the integer nature of the response variable, we build a transformed quasi-Poisson generalized linear model on disease incidence intensities. The quasi-Poisson model's choice is justified by the eventual difference in the mean and the standard deviation of the response variable; this over-dispersion is disease-specific. In the case of cardiovascular, gastrointestinal, and musculoskeletal diseases, considered in detail in the present study, the choice of the quasi-Poisson or the Poisson distribution having no overdispersion makes no significant difference in the results; therefore, to consider the latter is satisfactory. On the other hand, in the case of tumors, overdispersion probably needs to be taken into account. However, we do not consider the case of tumors because only aggregated incidence data has been available, and the various tumor types depend on very much different risk factors. As a result, throughout the current study, the number of incidences of the considered disease in locality i is Poisson distributed with intensity λ_i.

It is meaningful to suppose that a population-related risk factor – like age – influences disease incidences linearly: e.g., doubling the number of elderly doubles their effect on incidence numbers. However, to suppose the same on the effect of, say, a substance concentration in the drinking water does not make sense. Therefore, in our model, we choose incidence intensity to be additive on population- related and multiplicative on supplementary risk factor data.

The intensity λ_i of the Poisson distribution depends on the number of inhabitants $P_{i,k}$ belonging to the k-th of the C age cohort (or, with $C = 2$, the two genders) in the population of the i-th locality and the values $X_{i,j}$ of the R risk factors in locality

i, as

$$\lambda_i = \left(\sum_{k=1}^{C} \alpha_k \cdot P_{i,k}\right) \cdot \exp\left(\sum_{j=1}^{R} \beta_j X_{i,j}\right)$$

where the constants $\alpha_k,\ k = 1, \ldots, C,\ \beta_j,\ j = 1, \ldots, R$ are to be estimated.

First we address the estimation of the first type term. Consider only two risk classes, say, two age cohorts, the young (younger than 60) and the elderly (60 or older) ones. Let N_i denote the Poisson distributed disease incidence number, $P_{i,y}$, and $P_{i,o}$ the number of young and old inhabitants in the i-th locality. Then $N_i = P_{i,y} + P_{i,o}$, and $P(N_i = k) = \frac{\lambda_i^k}{k!} \cdot e^{-\lambda_i}$ where λ_i is the intensity. Refraining from considering any other risk factors but age we have

$$\lambda_i = \alpha_y P_{i,y} + \alpha_o P_{i,o} \tag{1}$$

where α_y es α_o are unknown constants – weights – to be estimated. From here, the log-likelihood function has the form

$$l(.) = \sum_{i=1}^{n} (N_i \cdot \log \lambda_i - \lambda_i) + const, \tag{2}$$

with n denoting the number of localities, i.e., $n = 3153$. This is to be maximized in α_y, α_o. Taking the corresponding derivatives of the log-likelihood, substituting here λ_i from (1) and introducing the notations $\gamma = \frac{\alpha_o}{\alpha_y}$, $\pi_i = \frac{P_{i,o}}{P_{i,y}}$ the likelihood equations are

$$\frac{\partial l}{\partial \alpha_y} = 0: \qquad \frac{1}{\alpha_y} \cdot \sum_{i=1}^{n} N_i \frac{P_{i,y}}{P_{i,y} + \gamma P_{i,o}} = \sum_{i=1}^{n} P_{i,y} \tag{3}$$

$$\frac{\partial l}{\partial \alpha_o} = 0: \qquad \frac{1}{\alpha_y} \cdot \sum_{i=1}^{n} N_i \frac{P_{i,o}}{P_{i,y} + \gamma P_{i,o}} = \sum_{i=1}^{n} P_{i,o} \tag{4}$$

Taking the fraction of the two after some algebra we arrive at the equation

$$\sum_{i=1}^{n} N_i \cdot \frac{1 - \pi_i Y}{1 + \pi_i \gamma} = 0, \tag{5}$$

to be solved for γ, where Y is the proportion of the young and elderly in the pooled population

$$Y = \frac{\sum_{i=1}^{n} P_{i,y}}{\sum_{i=1}^{n} P_{i,o}}.$$

This equation is analytically intractable, but an R solver provides an approximate numeric solution to it, and importantly, it can be shown to be non-negative. When three or more risk classes are considered, negative solutions can be produced, and those are meaningless for our application. Therefore, conditioning on the previous risk classes, one class can be split into two with two relative risk weights to be estimated, and the conditional likelihood maximized the way it was done for two classes. This

results in two estimated conditional risk weights, and the unconditional weights can be straightforwardly computed from them. Since the conditional weights are always positive, we obtain positive unconditional weights as a result. However, the joint optimality of the weights cannot be guaranteed by this procedure.

Multiplying in the equation (3) by α_y and dividing by $\sum_i P_{i,y}$ we get the risk weights as

$$\alpha_y = \frac{\sum_{i=1}^{n} N_i \frac{P_{i,yf}}{P_{i,y}+\gamma P_{i,o}}}{\sum_{i=1}^{n} P_{i,y}}, \quad \text{and} \quad \alpha_o = \gamma \cdot \alpha_y.$$

The case of genders as risk factors can be analyzed in a completely identical way.

Turning now to the multiplicative term in risk factors, and their weight estimation, for the time being, we suppose the lack of additive risk factors except for the population size P_i of the i-th locality. However, this risk factor cannot have a meaningful weight, it acts as a constant multiplier, i.e., as the *offset* in the model. This way (1) reduces to

$$\lambda_i = P_i \cdot exp(\sum_{j=0}^{R} \beta_j X_{i,j}). \tag{6}$$

For estimating the unknown β parameters we again start form the log-likelihood (2). This is to be maximized in the $\beta_j, j = 0, \ldots, R$ parameters. The derivatives by the parameters are

$$\frac{\partial l}{\partial \beta_j} = \sum_{i=1}^{n} \frac{\partial l}{\partial \theta_i} \frac{\partial \theta_i}{\partial \beta_j} = \sum_{i=1}^{n} (N_i - \lambda_i) \frac{\partial \log(\lambda_i)}{\partial \beta_j}.$$

Since $\log(\lambda_i) = \sum_{j=0}^{R} \beta_j X_{i,j}$, we have

$$\frac{\partial l}{\partial \beta_j} = \sum_{i=1}^{n} (N_i - \lambda_i) X_{i,j}$$

This function is the so called score function and the maximum can be found by solving it to zero. It is again intractable analytically , therefore, the Newton-Raphson algorithm or the Fisher scoring method can be used for obtaining the risk weights.

When considering only multiplicative risk factors the offset in (6) was the population size. After obtaining the weights for age cohorts without multiplicative risk factors in the model, we can introduce a new offset by weighting the population according to age as

$$\lambda_i = (\alpha_y \cdot P_{i,y} + \alpha_o \cdot P_{i,o}) \cdot exp \sum_{j=0}^{r} \beta_j X_{i,j}.$$

Keeping α-s fixed, with the new offset we can estimate the multiplicative risk factors just like above. Doing this way, an independence of the additive and multiplicative risk factors is supposed implicitly, that does not correspond to reality. Therefore, we go further by re-estimating the additive factor weights in the presence of the

multiplicative factors. Using the notation $w_i = exp \sum_{j=0}^{R} \beta_j X_{i,j}$ and regarding this estimated quantity as known for every locality we can introduce the new intensities as

$$\lambda_i' = (\alpha_y' P_{i,y} + \alpha_o' P_{i,o}) \cdot w_i$$

with the new weights α_y', α_o'. The log-likelihood can be computed similarly to (2), and substituting $\lambda_i' = (\alpha_y' P_{i,y} + \alpha_o' P_{i,o}) \cdot w_i$, and introducing $\gamma' = \frac{\alpha_o'}{\alpha_y'}$ we obtain the likelihood equations in a slightly changed form

$$\begin{aligned} \frac{\partial l}{\partial \alpha_y'} &= 0: \qquad \frac{1}{\alpha_y'} \cdot \sum_{i=1}^{n} N_i \frac{P_{i,y}}{P_{i,y} + \gamma' P_{i,o}} = \sum_{i=1}^{n} w_i \cdot P_{i,y} \\ \frac{\partial l}{\partial \alpha_o'} &= 0: \qquad \frac{1}{\alpha_y'} \cdot \sum_{i=1}^{n} N_i \frac{P_{i,o}}{P_{i,y} + \gamma' P_{i,o}} = \sum_{i=1}^{n} w_i \cdot P_{i,o}. \end{aligned}$$

From here some algebra leads to the same equation as (5) with $\gamma = \gamma'$ and $Y = Y' = \frac{\sum_{i=1}^{n} w_i \cdot P_{i,y}}{\sum_{i=1}^{n} w_i \cdot P_{i,o}}$. and it can be solved by the same R solver as (5).

The procedure can be iterated further improving the result. As of our experience after 3-4 iterations the results do not change significantly.

4.3 The achieved research results

The methodology described in the previous subsection has been applied to data on cardiovascular, gastrointestinal and musculoskeletal diseases. For example, in the case

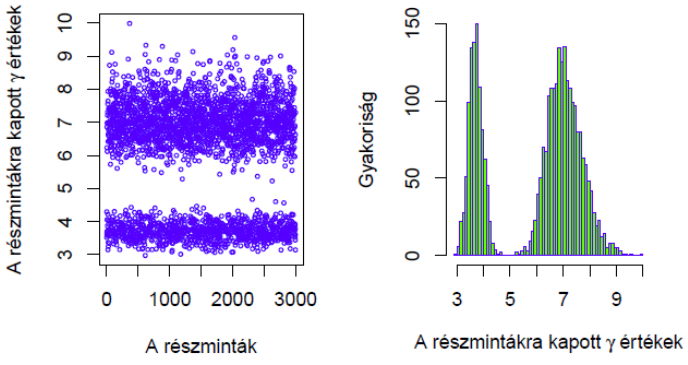


Figure 1: Gamma values and their histogram from 3000 subsamples of 1000 localities

of cardiovascular diseases and two age cohorts: 0-59, 60- we obtain $\gamma = 5.096$, meaning that the risk for the elder group is roughly 5 times higher than the younger. The stability of this value, obtained by analyzing the data of all the 3153 localities, raises a question. To address it, we randomly choose 1000 out of the available 3153 localities and compute the γ value for this subsample. We repeat this random subsampling 3000

times, collect the computed γ values and display them together with their histograms in the Figure 1. It turns out that Budapest causes the strong bimodal feature of the distribution as a single outlier. It points to the great difference between the capital and the rest of the country in terms of disease risk. It also highlights the unsatisfactory nature of the TeIR database that the capital, Budapest, is one single entity in it. As a result, we were forced to drop the data for Budapest from the analysis.

Significant effect of gender as a risk factor cannot be established based on the given data for any of the considered disease types. This may be explained by the controlling for age and the fact that the average life expectancy of women is higher.

Age is significant for all three disease groups and particularly significant for cardiovascular diseases. Table 1 shows the estimated proportion of the age cohorts affected by the disease.

Age / Disease type	0-49 y.o	50-59 y.o.,	60+ y.o.
cardiovascular	9.5%	47.9%	84.6%
gastrointestinal	12.3%	31.2%	30.8%
musculoskeletal	11.1%	39.1%	44.6%

Table 1: estimated proportion of the age cohorts affected by the disease

Income, divorce and ammonium contamination in drink-water are other significant factors for cardiovascular diseases. Their effect is probably indirect, divorce e.g. may serve as indicators of stressed life, while ammonium may signal satellite contaminants.

A peculiar outcome is that nitrate contamination of drinkwater is significant in *decreasing* the cardiovascular risk. An explanation may be that nitrate is known to ease certain cardiovascular disorders and it is even used in medicines. After correcting for the mentioned risk factors a further analysis reveals spatial risk components. Its areal distribution for cardiovascular diseases is displayed in Figure 2.

5 Solution of the industrial problem and its benefits

The industrial partner followed the project closely, providing consultation opportunities and access to medical expertise when necessary. The primary goal set forth has been achieved, and the industrial partner has acknowledged that. The research topic fits the company's broader research objectives, and in their opinion, the current results can serve as a starting point for more detailed analysis. We have serious indications for a continuing collaboration in the field. We managed to address the most important questions raised by the company:

- Gender is not a risk factor when controlling for age, and hence it is not necessary to consider it when calculating reserves.
- We could quantify the increased disease risk with age that can be taken into account in both premium calculations giving a premium advantage to the company

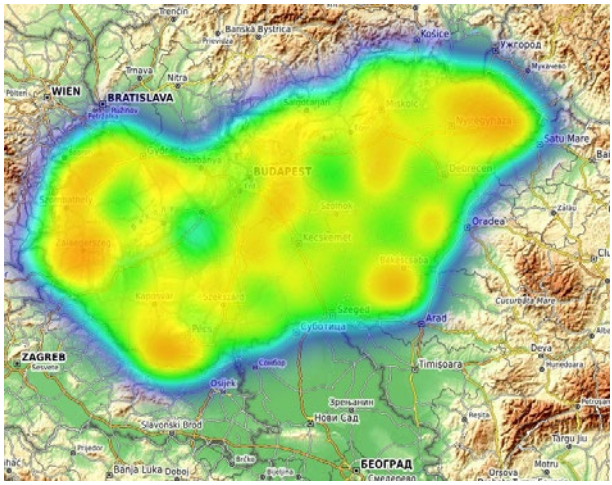


Figure 2: Gamma values and their histogram from 3000 subsamples of 1000 localities

and portfolio cleaning.

- We pointed to the extraordinary position of the capital that needs further study and, most importantly, a more detailed database.
- We have shown the territorial dependence of the risk, which later may serve for differences in premiums.

Note here that reserves calculation is a highly sensitive confidential field in the insurance businesses, so further details on the project's benefits cannot be given.

6 Conclusions

Using non-segmented disease incidence data for the 3153 localities in Hungary completed with general population data and locality attributes, the research results made it possible to make segmented risk analysis according to gender, age, and income. Additional risk factors such as drink water quality could also be incorporated into the analysis. The research came to the conclusion that gender is not a significant risk factor when controlling for age. This has serious consequences for the company, initiating the research in reserves calculations. Another result of the research quantifies risk represented by age cohorts, determined by medical expertise. As a benefit for the company, it facilitates correct premium calculations for health insurance products and provides a premium advantage in this product segment's competition. The capital, Budapest stands out as an outlier; its populations' risk characteristics differ from the rest of the country. Therefore, it had to be excluded from the research because of the lack of sufficiently detailed data. Given that roughly one-fifth of the country's population lives in the capital, extended research with a proper database seems to be a natural continuation of the research.

References

[1] R.A. Davis, W.T.M. Dunsmuir, S.B. Streett, *Maximum Likelihood Estimation for an Observation Driven Model for Poisson Counts*, Methodology and Computing in Applied Probability, Vol. 7, pp 149-159, (2005)

[2] Prokaj Vilmos, Általánosított lineáris modell (GLM) https://prokajvilmos.web.elte.hu/16-17ii/biztmat/GLM.pdf

Developing risk predicting methods and corresponding depersonalized databases

István Miklós[1]

[1]Rényi Institute, ELKH,

[1]E-mail: miklos.istvan@renyi.hu

1 Executive Summary

Risk prediction methods aim to substitute time and money consuming clinical inferring methods. Nevertheless, some of these methods are invasive, cause pain to the patients and/or might have their own side effects. In a health care program, people are involved when they are still healthy and with suggestions on changing life style, can be prevented to have a severe disease. Obviously, risk prediction methods are useful only if they have high accuracy. Our aim in this project is to develop more accurate risk prediction methods predicting cardiovascular and tumour diseases.

The standard model for risk prediction has been the multivariate logistic regression model, where the input are casual and risk factor parameters and the correlated features are predicted from this input data. However, there might be non-linear correlations in the system that cannot be modelled with linear models. In that case, more advanced methods have to be used. These methods include Support Vector Machines, prediction tree methods, and artificial neural network methods. In case of genetic data is available, standard sequence analysis methods might be also applicable including (Bayesian) alignment methods, SNP analysis, etc.

2 Challenge overview

In risk prediction methods, one of the central question is which variables contribute to the risk and how. There are several potential models: linear regression, non-linear regresion, machine learning methods, methods based on Bayesian statistics. Different models need different fileds in mathematics. While linear and non-linear regression needs theoretical results from function analysis, linear algebra and geometry, machine learning and Bayesian methods very frequently work with discrete mathematical objects: for example with trees or more complex networks. In many cases, the background bioinformatics models also have a natural discrete mathematical description. Biological macromolecules, like DNA, RNA or protein sequences can be described as abstract mathematical sequences. Evolutionary relationships can be described with rooted binary trees. Sometimes evolution happens at micro-level with an extremely high speed. We can see that in case of virus evolution, when different strains of viruses can evolve in the body of a single persion in time frames measurable in weeks or months. Cancer cells might also escape immune surveliance, during that process their evolution is also measurable.

It is a natural rule of thumb that more data provides more reliable prediction (though that is also a natural phenomena in biological sciences that any theory has a counterexample). Combining different methods needs a team of mathematicians working on different fields. Their collaboration must also be carefully managed, and the balance between basic and applied research is extremely important.

Another industrial challange is to collect data to analyise. Data are collected either in a large scale study that needs a huge investment many of the start-up companies cannot afford or collected from previous studies performed by others. In the latter case, the data is expected to be incomplete. Even when data is collected systematically like the one in the Hungarian Electronic Healthcare Service (Elektronikus Egészségügyi Szolgáltatási Tér), some of the data migh be missing. For example, some of the blood tests failed or had not been measured. Risk prediction from such data frequently needs data purification, imputation on missing data, etc.

The research focused on the following topics

1. How to build a searchable database from everyday health data?
2. What kind of risk prediction methods could be obtained from such a database?
3. How to find unexpected signals in databases?
4. Specific risk prediction methods. Here we considered two main questions: how to predict risk of diabetes from blood samples? How to predict the risk of cancer from human genomic data?

3 Implementation of the project

Together with our industrial partner, the Hospitaly IT Develper Ltd., we set up a database collecting data from the Elektronikus Egészségügyi Szolgáltatási Tér, Győr region. An unforseenable challenge was that during the implementation of the database, the General Data Protection Regulation (GDPR) came out, therefore, our data collection had to be fulfill the regulatory rules. That needed a huge effort to depersonalize the data wile still providing a searchable database. We also collected antigen and immunogenetic data for predicting risk of cancer.

Above the specific industrial questions, some of our research focused on background mathematical problems that could find their application in applied mathematics. We focused on the following topics: *a)* random walks, periodic walks. Random walks are essential in Monte Carlo methods, that are widely used in Bayesian statistics, machine learning and data imputation. Periodic walks are important to predict (quasi)periodical behaviour of certain data. Just to mention a recent example: weekly periodicity in reporting Covid cases and related deaths can be found worldwide both in national and international databases. *b)* network problems. We focused on patterns in networks, as this is expected to be searched in real life data. In pure mathematical terms, these are orderings, colorings, subgraphs, etc.

4 The research problem, methods, and results

Here we focus on two specific problems where we had publised outcomes: building the depersonalized, searchable database of the Elektronikus Egészségügyi Szolgáltatási Tér, Győr region and cancer risk prediction.

4.1 The research problem I. Depresonalized searchable database

Our aim was to build a depersonalized, searchable database from which the healthcare lifeway of a person or a cohort of persons can be retrieved, yet without personal information like TAJ number (social security number).

4.2 The research problem II. Cancer risk prediction methods

A relevant biological question is to infer the relationship between the risk of getting cancer and inherited immune system. We developed a risk prediction method that considers the immune responds to 48, mainly melanoma specific, tumor-associated antigens. The antigens have been selected based on literature data. The teaching set of our statistical learning algorithm was built from the immunonetic data (HLA data, Human Leucocyte Antigens) of non-American subjects. We inferred what are those HLA alleles which significantly help subjects carrying them to develop immune response against these 48 tumor-associated antigens than those subject not possessing these HLAs. We give a log-p value to these antigens. The test population consisted of American melanoma and non-melanoma subjects.

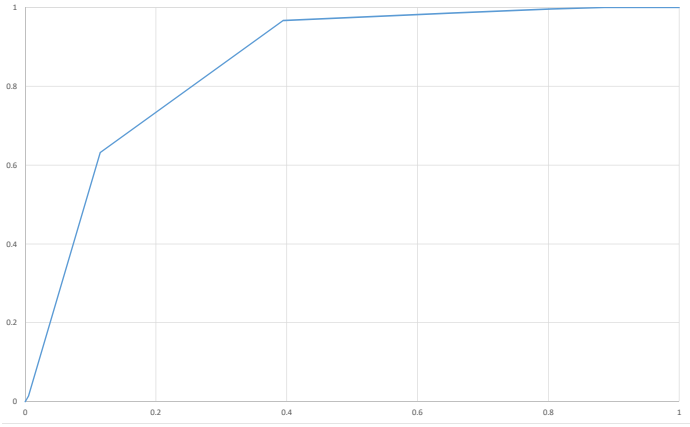
4.3 The applied methods

Depersonalization was achieved by one-way functions well-known in cryptography. Using such functions data coming from the same person can be retrieved without getting personal data. The vision was to build a database from which the healt care data from the same persion in different time points can be retrieved. In this way, it is possible to collect data from a person agiven time period before the symptoms were detected. For example, a person was detected having diabetes, then his/her blood samples a given time before the diagnosis can be colleted. This provides a nice, structured dataset on which risk prediction methods can be taught.

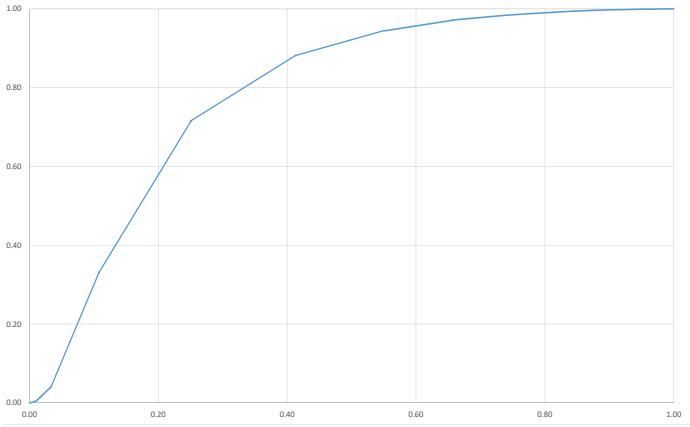
4.4 The achieved research results

Here we report preliminary results on predicting risk of diabetes, developed on the searchable database of the Elektronikus Egészségügyi Szolgáltatási Tér, Győr region. The method was our proof-of-concept that the developed databse can be used for risk prediction method. We present the ROC curves of two predictors, one is based on blood sugar level, the other is based on HbA1c, that is, glycolated hemoglobin (see Figure 1) [1].

Some of the large scale analysis showed unexpected correlation between accidents in nuclear power plants and cases of thyroid cancer, see Figure 2.



b)



a)

Figure 1: The ROC curves for predicting risk of diabetes **a)** based on blood sugar level, **b)** based on HbA1c. See the text for details.

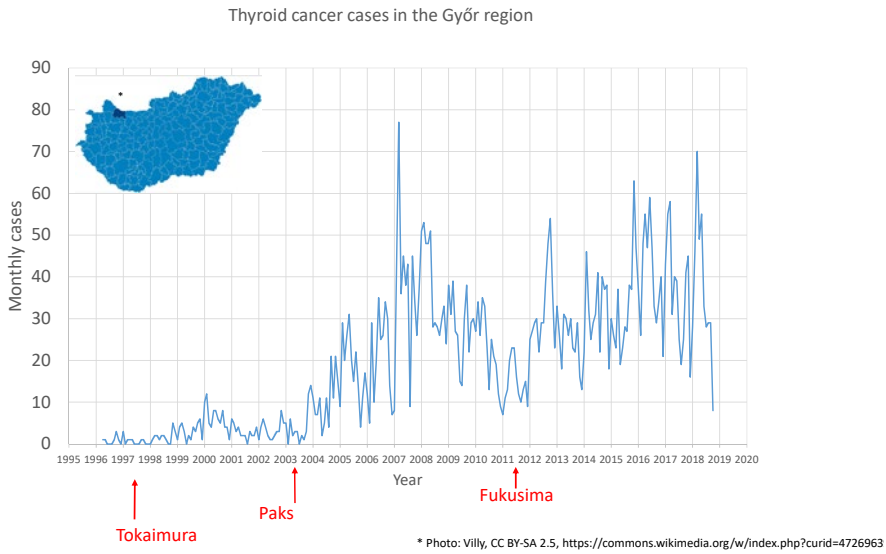


Figure 2: Temporal correlation between accidents in nuclear power plants and thyroid cancer cases in the Győr region. See text for details.

Regarding cancer risk prediction methods, we found that the log-p value of the HLA alleles is a good predictor for predicting the risk of melanoma. The ROC-AUC value was 0.68. This value is statistically highly significant, considering the size of the test population. Although the AUC value is not extremely large, we have to mention that our predictor considered only the HLA data of subjects, and did not considered environment factors like Sun exposure or other genetic/morphological factors that are known to be associated to melanoma like freckles and/or red hair, etc.

We also showed that the the average value of population computed from the log-p values of HLA allels correlates to the prevalence of melanoma in different populations. We presented our findings at the two world-leading oncology conferences, ASCO 2019 and ESMO2019, and has been publised in leading oncology journals [2,3].

5 Solution of the industrial problem and its benefits

The online, depersionalized database is on and working. With the depersonalization method, the life pathways of individual patients is still searchable. This allows that once any disease is developed and registered in the database, the antecedent data can be searched. This definitely help researchers to develop better risk models.

6 Conclusions

The XXI. century should be the century of health-care, when the emphases are on preventing the diseases and not to cure them. For that, data should be collected to reveal the relationship between environment, life-style, clinical measurement and the developing diseases. We developed a depersonalized database based on the Elektronikus Egészségügyi Szolgáltatási Tér. Biologists and clinical doctors started using this database and exploring the possibilities. We hope that many of the findings will be soon published in medical journals.

We also developed a risk model to predict risk of cancer based on immunonetic data (HLA data, Human Leucocyte Antigens). The achieved precisity is promising, and in the ongoing work we are looking for developing a combined method that predicts the risk of cancer based on genetic as well as environmental data.

References

[1] Király, Gy., Skaliczky, Z., Páli, T. (2020) *Kutatószoba projekt, azaz mesterséges intelligencia az egészségügyben*, Infokommunikáció, 19(2):39–42.

[2] Miklós, I. et al. (2019) *Determination of the immunogenetic risk of developing cancer.* Journal of Clinical Oncology, 37:15 suppl, e13132-e13132

[3] Miklós, I. et al. (2019) *Inferring the correlation between incidence rates of melanoma and the average tumor-specific epitope binding ability of HLA class I molecules.* Annals of Oncology (2019) 30 (suppl 5): v574-v584.

Numerical Investigation of Dynamic Contact Problems using Finite Element Method and Model Reduction Techniques

Balázs Pere[1] and **Dániel Serfőző**[2]

[12]Széchenyi István University, Department of Applied Mechanics,

[1]E-mail: pere.balazs@ga.sze.hu
[2]E-mail: serfozo.daniel@ga.sze.hu

1 Executive Summary

Engineering structures often contain parts which are in contact. Even today, solving contact problems is one of the most challenging tasks in numerical analysis of structures. On the one hand, the difficulty of such problems results from its non-linear nature. On the other hand, inaccuracy raising from the discretization of a continuum as a set of distinct points can also cause problems. The problem discussed here is that when bodies are getting into touch the solution begins strongly to oscillate. This oscillation has no physical background, but it comes from the disctretization itself. One possible way to treat this problem is the usage of penalty method for solving the contact problem and applying a properly chosen dashpots connected in parallel with contact stiffnesses. Another approach is the creation of a special damping matrix which suppresses undesirable oscillations. Finally, the model reduction technique will be presented as a possible method to eliminate the modal frequencies of unwanted oscillations.

2 Challenge overview

Engineering structures often contain parts which are in contact. Even today, solving contact problems is one of the most challenging tasks in numerical analysis of structures. On the one hand, the difficulty of such problems results from its non-linear nature. On the other hand, inaccuracy raising from the discretization of a continuum as a set of distinct points can also cause problems. The first contact problem was solved by Hertz in 1881 [1]. He determined the contact area and the distribution of the contact pressure of two elastic spheres analytically. There are very few problems involving contact which can be solved in an analytical way. With the growing power of modern computers, the Finite Element Method (FEM) evolved and has become capable to give approximate solutions of problems arising from elasticity. FEM was firstly used for solving contact problems in the mid 70's (see [2]). The application of finite element (FE) analysis for contact problems has a vast literature [3]. The simplest, fastest, but less accurate method for solving mechanical contact problems with FE codes is the *penalty method*, where a relatively big penalty parameter ensures that penetration between the contacting bodies remains small compared to their sizes [3]. The *Lagrange multiplier method* [4] uses the contact pressure as an additional

unknown field. It accurately satisfies the contact boundary conditions, i.e., there is no penetration at the contact interface, but the computation requires more resources, even in the static case. Combination of the aforementioned methods results in the *augmented Lagrangian method* [5], which is a good compromise between the penalty and the Lagrange multiplier methods. The main drawback of the augmented Lagrangian method is the applied double loop iteration in the algorithm. All of the aforementioned methods can be utilized for dynamic simulations of structures.

Solution of a dynamic problem usually means time integration of the semi-discretized equation of motion. This semi-discretized equation is discretized in space, but continuous in time. Discretisation and integration in time of such an equation can be performed with either explicit or implicit methods [6]. In implicit methods, such as the Newmark method, the time step can be much larger than the period of the highest eigenfrequency, while the solution remains numerically stable. This method is not suitable to follow sudden changes (like collision of bodies) and to solve vibrations whose frequency is close to the highest eigenfrequency [7]. In contrast, explicit methods, such as central difference method, can solve contact-impact problems within reasonable time. Here time step size has to be as small as or smaller than the period of the highest eigenfrequency, but unlike the implicit method, computation of one time step could be much faster. Various time-stepping methods keep being developed nowadays. Numerous publications provide techniques to increase the accuracy of solutions for contact-impact problems [8, 9]. The main problem is that in case of collision of bodies the solution strongly oscillates. This oscillation has no physical background, but it comes from the discretisation itself.

Another way of improving the solution process of the dynamic contact problem might be the application of model reduction techniques [10]. Those frequencies of vibration the amplitude of which is much smaller than other amplitudes can be neglected form the computation of the motion. In this way, the size of the model is reducible and ignoring frequencies which are not of interest can speed up the calculation.

3 Implementation of the project

The aim of the research is to develop a more accurate and faster algorithm for solving dynamic contact problems. As a first step, a simple one dimensional problem will be analyzed, a rod with constant velocity collides to a rigid wall. This problem has an analytical solution, which can be compared to various existing and self-developed numerical methods. In order to reduce oscillations in the results, two approaches will be used. In the first approach, we attempt to create a special damping matrix which suppresses undesirable oscillations (methods to create an arbitrary damping matrix can be found in [11].). Secondly, by applying model reduction technique an attempt will be made to eliminate the frequencies of undesirable oscillations. After drawing the consequences, the results will be applied to solve simple two dimensional problems.

4 The research problem, methods, and results

4.1 The research problem: elastic collision in one dimension

Consider a one dimensional problem where an elastic body (rod) collides with a rigid obstacle. This is a very simple problem but the phenomena observed here are also present in higher dimension cases. It is assumed that the displacements and strains are small and the rods exhibit linear elastic behavior. Let the centroidal axis of the rod coincides the x axis of the coordinate system (see Fig. 1). The strain-displacement

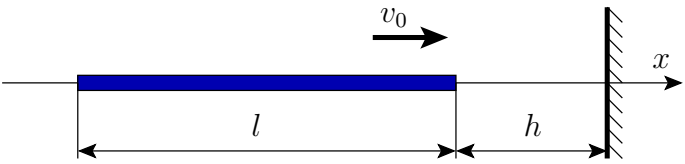


Figure 1: The rod moving towards the obstacle with constant v_0 velocity.

relation, equation of motion and the constitutive equation can be written as follows

$$\varepsilon_x(x) = \frac{\partial u(x,t)}{\partial x} \tag{1}$$

$$\frac{\partial N(x,t)}{\partial x} + f_x(x,t) = \rho A \frac{\partial^2 u(x,t)}{\partial t^2} \tag{2}$$

$$N(x,t) = AE\varepsilon_x(x,t) \tag{3}$$

where ε_x is the normal strain in the rod, u is the displacement in the direction of x axis, N is the normal force in the rod, f_x is the distributed force in the direction of x, ρ is the mass density and A is the area of the cross section. Let the rod move with constant v_0 velocity towards the obstacle. The rod is assumed to be unstretched at the beginning. The left end of the rod is unloaded while the right end cannot penetrate the obstacle and its distance from the obstacle at $t = 0\,\mathrm{s}$ is h. The movement of the rod can be divided into three stages. In the first stage the rod is moving towards the obstacle but does not touch it. It means that

$$h - u(l,t) > 0 \tag{4}$$

$$N(0,t) = 0 \tag{5}$$

$$N(l,t) = 0 \tag{6}$$

$$u(x,0) = 0 \tag{7}$$

$$\left.\frac{\partial u(x,t)}{\partial t}\right|_{t=0} = v_0 \tag{8}$$

where $0 \leq t < t_1$ and $0 \leq x \leq l$. Equation (5)-(6) are the boundary conditions (BC) and (7)-(8) are the initial conditions (IC) for the system of partial differential

equations (PDE) (1)-(3). In the second stage the rod touches the obstacle (at time t_1), deforming and separates from the obstacle (at time t_2). In this stage the BCs and ICs can be written as follows

$$h - u(l,t) = 0 \tag{9}$$

$$N(0,t) = 0 \tag{10}$$

$$N(l,t) \leq 0 \tag{11}$$

$$u(x,t_1) = h \tag{12}$$

$$\left.\frac{\partial u(x,t)}{\partial t}\right|_{t=t_1} = \begin{cases} v_0 & 0 \leq x < l \\ 0 & x = l \end{cases} \tag{13}$$

where (9)-(10) are the BCs and (12)-(13) are the ICs. The third stage is almost the same as the first one, except for the ICs where initial displacement inherits its value from the previous stage and the velocity which has the opposite direction. Equations (4) and (6) in pairs with (9) and (11) form the contact conditions (CC). It can be observed that their product is always zero.

$$(h - u(l,t))\, N(l,t) = 0 \tag{14}$$

The above given problem has basically two ways of solution. The first is the analytical solution of system of PDEs (1)-(3) with the appropriate BCs and ICs respect to the actual stage. The second way is the numerical solution with FEM.

4.2 The applied methods

After semi-discretisation of system of PDEs (1)-(3) with its BCs, ICs and CCs the equation of motion can be written in the form of

$$\mathbf{M}\ddot{\mathbf{u}} + \mathbf{K}\mathbf{u} + \mathbf{G}^T p = \mathbf{f} \tag{15}$$

$$h - \mathbf{G}\mathbf{u} \geq 0 \tag{16}$$

where $\mathbf{M}$ is the mass matrix, $\mathbf{K}$ is the stiffness matrix, $\mathbf{G}$ is the surface contact displacement constraint matrix, $\mathbf{f}$ is the load vector, $\mathbf{u}$ is the nodal displacement vector, h is the initial gap between the rod and obstacle and p is the contact pressure. The latter one can be solved from the CC (11) by multiplying it with minus one. The contact pressure is never less than zero.

$$p = -N(l,t) \geq 0 \tag{17}$$

Equation (16) is the discretized form of CC (4) and (9).

For solving the contact problem the Lagrange multiplier technique was applied. The time integration of equation (15) and inequality (16) was performed using various methods such as central difference method, implicit Euler method, Newmark method, HHT-α method.

4.3 The achieved research results

A numerical computation was performed using the previously mentiond methods. Some of the methods exhibit spurious oscillations after sudden changes (jumps) in the exact velocity and contact pressure (see Fig. 2). Figure 3 shows the surroundings

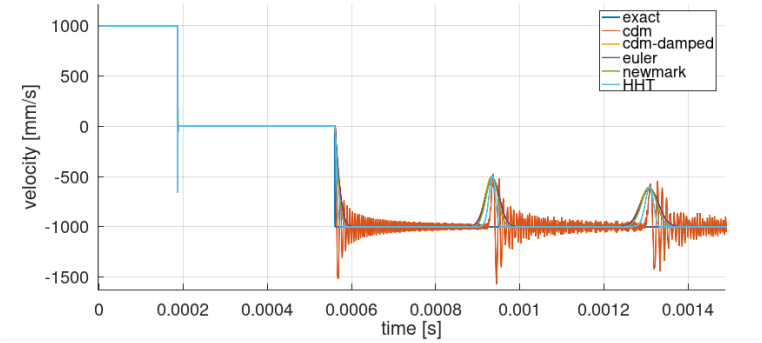


Figure 2: Exact and approximate velocities of the right end point of the rod. Applied numerical methods: central difference method, central difference method with viscous damping, implicit Euler method, Newmark method, HHT-α method.

of the jump. The implicit Euler, Newmark and HHT-α methods contain numerical

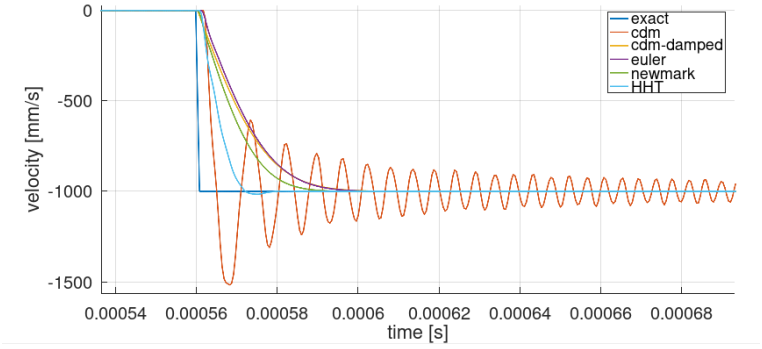


Figure 3: Velocities at the moment of separation.

damping, hence the oscillation there is much less than in case of central difference method. Despite the implicit methods are numerically stable with larger time step sizes, large step size causes inaccuracy in detecting the contact and generates extra dissipation. The stability of explicit methods requires smaller time step size, but needs less calculation during a time step, hence it is more advantageous to use it. In order to eliminate the oscillations a viscous damping matrix was constructed. The aim was to use critical damping, i.e., the system is neither overdamped nor underdamped. Figure 4 shows the contact pressure against time and figure 5 highlights the magnified region

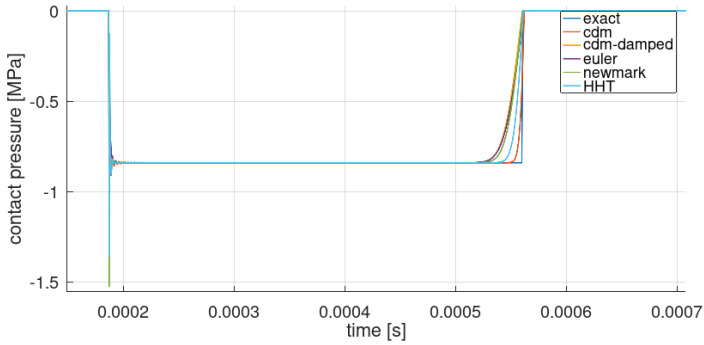


Figure 4: Contact pressure versus time. Applied numerical methods: central difference method, central difference method with viscous damping, implicit Euler method, Newmark method, HHT-α method.

around the time when the rod is getting into contact to demonstrate the efficiency of critical damping. Here the proposed method has as sharp jump as the exact solution

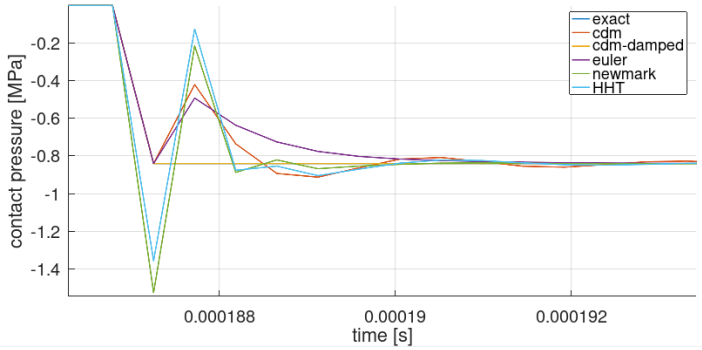


Figure 5: Contact pressure at the impact.

but without oscillations. In figure 6 the decay of total energy can be seen. The more energy is dissipated, the less oscillation can be observed in the displacement, velocity and contact pressure diagrams. The proposed method can be easily implemented with commercial FE softwares, like e.g. Abaqus. According to the knowledge of authors the procedure applied here was not yet publish in the literature. An article about the assemblage of such a damping matrix is under preparation.

5 Solution of the industrial problem and its benefits

At this level, the method presented has no direct applicability in the industry but further development is in progress to use it in two-dimensional problems. Modern

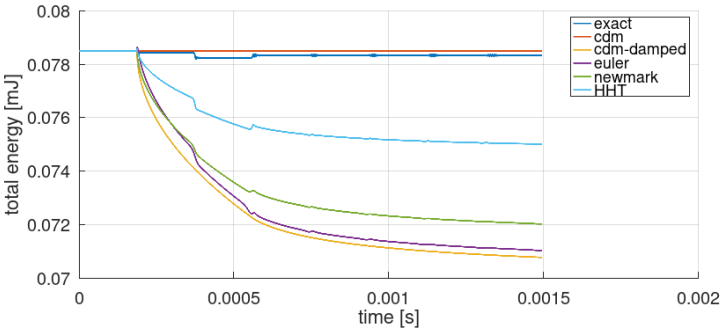


Figure 6: Total energy versus time. Applied numerical methods: central difference method, central difference method with viscous damping, implicit Euler method, Newmark method, HHT-α method.

FE softwares are capable to import externally assembled damping matrices. In the knowledge of a given problem the damping matrix can be solved with self developed codes, e.g. in Matlab.

6 Conclusions

This paper has presented a method which used special damping matrix for eliminating spurious oscillations in case of dynamic contact problems. The efficiency of the method was demonstrated in a one dimensional problem. Although the results are very encouraging, new questions may arise on the application of the method. Can it be usable for non-equidistant meshes? Can this method applicable for problems where model reduction technique is also employed? What is the optimum between the accuracy and elimination of oscillations? The authors try to answer these questions during further research work.

References

[1] H. Hertz, *Über die Berührung fester elastischer Körper*, Journal für die reine und angewandte Mathematik, **92**, 156–171, 1881

[2] A. Francavilla and O. C. Zienkiewicz, *A note on numerical computation of elastic contact problems*, International Journal for Numerical Methods in Engineering, **9** (4), 913–924, 1975

[3] P. Wriggers, *Computational Contact Mechanics*, John Wiley & Sons Ltd, 2002

[4] P. Papadopoulos and J. M. Solberg, *A Lagrange multiplier method for the finite element solution of frictionless contact problems*, Mathematical and computer modelling, **28** (4-8), 373–384, 1998

[5] J. C. Simo, and T. A. Laursen, *An augmented Lagrangian treatment of contact problems involving friction*, Computers & Structures, **42** (1), 97–116, 1992

[6] T. Belytschko and T. J. R. Hughes, *Computational methods for transient analysis*, Amsterdam North-Holland (Computational Methods in Mechanics), **1**, 1983

[7] N. J. Carpenter and R. L. Taylor and M. G. Katona, *Lagrange constraints for transient finite element surface contact*, International journal for numerical methods in engineering, **32** (1), 103–128, 1991

[8] W. Kim, *A new family of two-stage explicit time integration methods with dissipation control capability for structural dynamics*, Engineering Structures, **195**, 358–372, 2019

[9] W. Kim, *An accurate two-stage explicit time integration scheme for structural dynamics and various dynamic problems*, International Journal for Numerical Methods in Engineering, **120** (1), 1–28, 2019

[10] B. Starc and G. Čepon and M. Boltežar, *A mixed-contact formulation for a dynamics simulation of flexible systems: an integration with model-reduction techniques*, Journal of Sound and Vibration, **393**, 145–156, 2017

[11] S. Adhikari, *Damping modelling using generalized proportional damping*, Journal of Sound and Vibration, **293**, 156–170, 2006

Developing a sensor fusion based robust and reliable position estimation method in non-stationary indoor environment

Szilveszter Pletl[1], **Zoltan Kincses**[2], **Laszlo Schaffer**[3] and **Adam Kalo**[4]

[1234]University of Szeged, Department of Technical Informatics,

[1]E-mail: pletl@inf.u-szeged.hu
[2]E-mail: kincses@inf.u-szeged.hu
[3]E-mail: schaffer@inf.u-szeged.hu

1 Executive Summary

The localization of facilities in internal conditions is a very current area in the fields of industry, culture, sports and trade. The completed development offers a competitive solution in the field of localization in these areas. The developed algorithm, in relation to the already existing market solutions, provides additional information regarding the expected accuracy of localization. And for this reason, it has a certain potential in the market. The developed system is open to such an extent that its interactions into a larger whole are possible. The localization application in the museum premises is offered directly. There has been very positive feedback from economies regarding the potential use of the proposed solution. The main challenge of the research is to find an adequate solution for localization in indoor conditions. The research is of an applied nature, where the developed algorithms are directly applied in practice.

2 Challenge overview

Visually impaired people lack the information of their indoor position, which makes their indoor movement very hard. Using a position estimation system, people with vision problems can get reliable information about their position. There are some solutions for the problem, e.g. WiFi signal based positioning using already installed routers, but the solution are not robust, because in a well populated area with routers in every room, the signals interfere with each other.

It is well known that micro-electromechanical system (MEMS) inertial sensor (3D accelerometers and 3D gyroscopes) measurements are obtained at high sampling rates and can be integrated to obtain position and orientation information. These estimates are accurate on a short time scale, but suffer from integration drift over longer time scales. To overcome this issue, inertial sensors are typically combined with additional sensors and models. There are several signal processing aspects of position and orientation estimation using inertial sensors with different modeling choices and a selected number of important algorithms. These algorithms include optimization-based smoothing and filtering as well as computationally cheaper extended Kalman filter and complementary filter implementations. Integration of the gyroscope measurements provides information about the orientation of the sensor. After subtraction

of the earth's gravity, double integration of the accelerometer measurements provides information about the sensor's position. To be able to subtract the earth's gravity, the orientation of the sensor needs to be known. Hence, estimation of the sensor's position and orientation are inherently linked when it comes to inertial sensors.

One of the open problems is the integration drift phenomena, which ideally consists of measuring a quantity that is constant and equal to zero. The integrated and double integrated signals are therefore also equal to zero. If we measure this quantity using a non-perfect sensor the measurements are corrupted by a constant bias, integration of these measurements will lead to a signal which grows linearly with time. Double integration leads to a signal that instead grows quadratically with time. If the sensor instead measures a zero-mean white noise signal, the expected value of the integrated measurements would be zero, but the variance would grow with time. Hence, integration drift is both due to integration of a constant bias and due to integration of noise. Inertial sensors provide pose estimates at high sampling rates which are accurate on a short time scale but drift over longer time scales. They are therefore very suitable for being combined with sensors with a lower sampling rate but with information that does not drift over time (e.g. UWB).

Given an indoor environment with some rooms, our task is to determine the position of the people wearing the localization unit using well-placed stationary anchors (transmitters with known position) in different rooms with time-of-flight (TOF) measurements fused with microelectromechanical systems (MEMS) based sensor data (accelerometer, gyroscope, magnetometer). Sensor fusion uses the reliability value (often a priori knowledge) of the different sensors combined with the mathematical model of the system and the measurements. The mathematical model and the placing of the anchors have to be well designed to estimate position reliably.

3 Implementation of the project

In this project 3 academic researchers and one MSc student were involved from University of Szeged Institute of Informatics and two electrical engineers from the TEConncept Hungary Ltd. The academic researchers developed an UWB based position estimation system both the hardware and the software parts including the sensor fusion algorithm. The industrial partner tested and validated the proposed system.

4 The research problem, methods, and results

Indoor position estimation is an important part of any indoor application which contains object tracking or environment mapping. Many indoor localization techniques (Angle of Arrival - AoA, Time of Flight - ToF, Return Time of Flight - RToF, Received Signal Strength Indicator - RSSI) and technologies (WiFi, Ultra Wideband - UWB, Bluetooth, Radio Frequency Identification Device - RFID) exist which can be applied to the indoor localization problem.Based on the measured distances (with a chosen technique), the position of the object can be estimated using several mathematical methods. The precision of the estimated position crucially depends on the placement of the anchors, which makes the position estimate less reliable. In this paper a

simulation framework is presented, which uses genetic algorithm and the multilateral method to determine an optimal anchor placement for a given pathway in an indoor environment. In order to make the simulation more realistic, the error characteristics of the DWM1001 UWB ranging module were measured and implemented in the simulation framework. Using the proposed framework, various measurements with an optimal and with a reference anchor placement were carried out. The results show that using an optimal anchor placement, a higher position estimation accuracy can be achieved.

4.1 The research problem

The location of a device or user can be effectively obtained outdoor using the Global Positioning System (GPS), but it could be challenging in an indoor environment. During the last decade, indoor localization has been investigated mainly for wireless sensor networks and robotics. However, nowadays, the wide-scale usage of mobile phones and wearable devices has enabled localization in a wide range of applications like health-care, industry, surveillance and home management.

In the literature many localization technologies and techniques are available [**Loc**]. A Received Signal Strength Indicator (RSSI), which is the strength of the signal received usually measured in decibel-milliwatts (dBm), and a wireless Ethernet based localization approach is used in [**RSS**]. Using a path-loss model and the RSS, the distance between the sender and receiver can be estimated. In [**AOA**] an Angle of Arrival (AoA) and Wireless LAN (Wifi) based method is applied using an antennae array for the estimation of the angle by computing the difference between the arrival times at the individual elements of the array. A Time of Flight (ToF) and 2.4 GHz radio based approach is presented in [**ToF**],using signal propagation time to compute the distance between the transmitter and the receiver. A similar technique, the Return Time of Flight (RToF) is used in conjunction with RSSI in a Wifi-based method in [**RToF**].

RToF is a two-way ranging method where the transmitter sends a ranging message to the receiver at t_1 time. The receiver sends it back with a delay of t_{proc} time and it arrives to the original transmitter at t_2 time. The time of flight is $t_2 - t_1 - t_{proc}$, and the distance can be calculated with the speed of the signal, depending on the technology. The accuracy of the measurement highly depends on t_{proc}.

The UWB is a recently researched communication technology providing more accurate ToF and RToF estimations. It uses ultra-short pulses with a time period less than a nanosecond, resulting in a low duty cycle which leads to lower power consumption. Its frequency range is from 3.1 to 10.6 GHz with a bandwidth of 500 MHz. Since the UWB usually operates at a low energy level, typically between -40 and -70 dB, most of the other technologies detect it as background noise. This makes it practically immune to interference with other systems since it has a radically different signal type and radio spectrum. Moreover, the signal (especially in its lower frequencies) can penetrate through walls because signal pulses are very short. Utilizing this attribute, it is easier to differentiate the main path from the multi-paths, providing more accurate estimations [**UWB_base**].

Once the point-to-point distances between the objects are measured, the unknown position of the object can be estimated. There are various algebraic methods to estimate the position from the point-to-point distances like triangulation or multilateration. Most of them require a few devices with known fixed positions (anchor nodes) to calculate the actual position of the moving device (mobile node). In case of error-free distance measurements, these methods theoretically give an exact position. But real distance measurements contain errors which depend on the relative position of the devices, the orientation of the antennas of the devices, and also the technique and technology used. These result in a varying reliability of the position estimation.

In this work, a genetic-algorithm-based simulation framework is presented, which takes the specific error characteristics of the chosen localization system (DWM1001 a commercially available UWB localization system using RToF) into consideration and uses a 2-dimensional (2D) version of the multilateral method to determine an optimal anchor placement for a given pathway in a given environment. This framework can also be extended to 3-dimensional (3D) space in the future.

4.2 The applied methods

During the research, several methods were used together. To calculate the position of the object from the measured point-to-point distances, a 2D version of the original 3D multilateral algorithm [**multi**] was used. The evaluation of the calculated positions was based on the Root-Mean-Squared Error (RMSE) of distances between the original and the calculated positions.

To minimize this error, genetic algorithm with a special fitness function was used [**ga**].

The environment of our experiment is a corridor in the University of Szeged. During the simulation and also in the validation a belt-type topography is used as can be seen in Fig. 1. The boundary conditions are derived from the physical capabilities of the corridor. For details of the anchors and mobile nodes used.

Two sets of measurements were carried out. The first set consisted of two measurements with 4 anchors using equidistant and optimised placement. The second set of measurement differed only in the number of anchors, since 8 anchors were used.

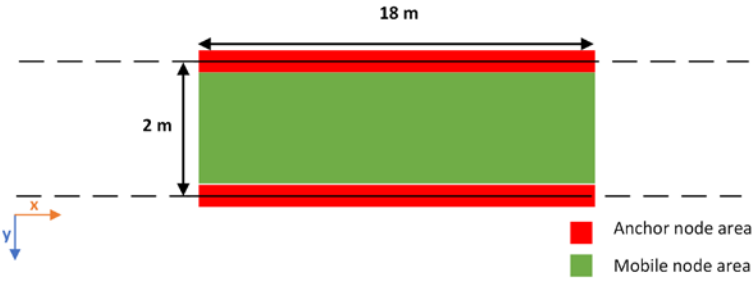


Figure 1: Measurement setup

During the measurements, the exact position of the discrete grid points (mobile

node position) is determined using a professional laser rangefinder. In each grid point, the mobile node collected the distances from the anchors many times (at least 100 samples per point) and the aggregated data was sent to the data collector.

The UWB based distance measurements were carried out by a commercially available localization system framework (MDEK1001), which consists of DWM1001 modules. The anchors were configured via the official mobile application (DecaWave DRTLS manager). The 8 anchors were organized in 2 networks, each network consisted of 4 anchors. On the measurements with 4 anchors, only one network was used.

The DWM1001 was connected to an STM32F746ZGT6 nucleo-144 development board (STM) via UART. The original program of the DWM1001 was modified to gather the 4 distances in both networks, and send the collected data towards the STM. The STM sent the data towards a mobile computer via UART trough an USB cable. On the mobile computer, the data was captured and saved via PuTTY, and were later evaluated with MATLAB.

The simulation framework is based on the GA implementation of the optimtool toolbox of MATLAB. The implementation was slightly modified to run the multi-population version of GA. The genetic representation of the system (the phenotype) was the x coordinates of the anchors. The y coordinates were along the walls of the simulated area, and the distribution of the anchors between the two walls was equal, or its difference was 1. The phenotype had a lower limit of 0, and an upper limit of 18, and they were free to move within this interval. The initial population range was 100 candidates. For each generation, the 20 highest-ranked candidates of the previous generation survived and further 60 were generated with cross-breeding and 20 new were randomly generated. The maximum number of generations, and the maximum number of stall-generations were both 100. The evaluation of the candidates was done by the special fitness function. During the optimization process the number of anchors was 4 and 8.

The aim of the fitness function in a GA is to order the candidates of the population based on their phenotype. The input of the function is the x coordinates of the candidate anchors, and the output is a corresponding RMSE value. After that the exact positions of the mobile node have been generated along a given path in the simulated area, the point-to-point distances from these points and the anchor nodes can be calculated. For each distance value, a unique error is added. Using the erroneous distance values and the anchor positions the simulated mobile node positions can be determined. Since the physical hardware device distance horizon is limited to 10 metres, any distance value over this range is discarded. The RMSE value can be calculated from the exact and simulated positions, resulting a fitness value of the given candidate.

In order to implement a more realistic simulation, the UWB distance sensor calibration was performed using a more accurate ($\pm 1mm$) class of laser rangefinder. Based on the measurements, the 2D error characteristics of the DWM1001 module with the built-in on-board antenna was determined. The measurement consisted of placing two modules in a known distance from each other, and performing a measurement with the UWB technology using RToF technique. The measured distance ($\hat{d}$) is the sum of the exact distance measured with a laser rangefinder (d) and the RToF

measurement additive error (Δd) as follows:

$$\hat{d} = d + \Delta d. \tag{1}$$

The Δd was calculated in discrete points from 0.3 to 10 metres in steps of 0.3 meters. At each discrete point at least 100 measurements were conducted with varying antenna orientation (0°, 90°, 180°) and the mean value was calculated.

After the measurements, at each measured distance, an error value was interpolated for every degree between $0° - 180°$. Since the error characteristic is symmetrical in this plane [**dwm**], the corresponding error values for degrees between 180° and 360° can be used from the interpolated ones. Each point of the 2D error characteristic is loaded into a Look-Up-Table (LUT), since in the optimization process the genetic algorithm reads the corresponding error value addressing the LUT by the distance and orientation parameters.

4.3 The achieved research results

The main results of the proposed method are creating a realistic (hardware- and environment-specific) 2D error characteristic of DWM1001, determining an optimal placement of anchor nodes and experimentally validating the results of the simulation.

The visualization of the realistic 2D error characteristic can be seen in Fig. 2, where one slice of the surface represents the distance error of one degree of rotation of the DWM1001 module. In the data sheet of DWM1001 [**dwm**] it is claimed that the ranging accuracy of the module is within $10cm$. But the results in Fig. 2 showed that the accuracy is a nonlinear function of distance and orientation. Furthermore, between 0-1m and 9-10m, the error is significantly higher than $10cm$ ($17cm \pm 2cm$). The highest accuracy can be reached around 2 meters ($7cm \pm 2cm$).

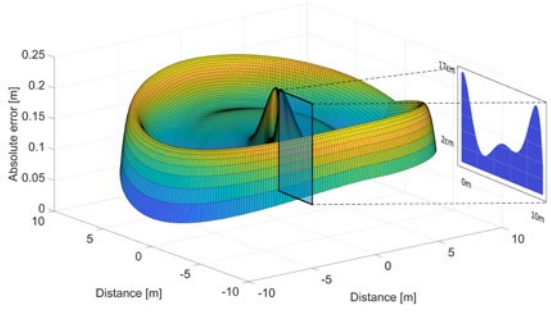


Figure 2: Realistic 2D error characteristic of DWM1001

The goal of the optimization process was to find an optimal anchor placement in case of 4 and 8 anchors. During this process, linear and non-linear paths of the mobile node was used. The results (see in [**ICAI**]) showed that the optimization significantly increases the accuracy of the position estimation in case of 4 anchors, but in case of 8

anchors, it has a lower impact. However, using only 4 anchors, the robustness of the system is lower.

In Fig. 3 the placement of the anchor nodes can be seen with and without optimization. The optimization process provides significantly different anchor placement in case of 4 anchors but just a slightly different in the 8-anchor case compared to the equidistant placing. Analyzing the results of the optimization, it can be stated that the algorithm places the anchors considering two main conditions. It tries to cover the whole area to have at least 3 anchor nodes in the range of the mobile node and places as many anchors as possible in the border of the mobile node area since the accuracy of the DWM1001 is lower nearby its horizon.

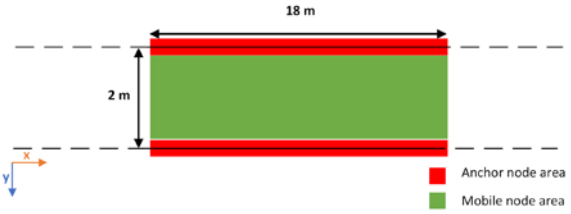


Figure 3: Optimized and reference anchor placements for 4 anchors (A) and 8 anchors (B)

The purpose of the validation is to verify the localization accuracy by real measurements using the proposed anchor placement in case of 4 and 8 anchors. The results show that using an optimized anchor placement (see in **[ICAI]**), the localization accuracy can be increased in case of 4 anchors but there is no significant improvement using 8 anchors compared to the equidistant case. Using an optimal anchor placement in case of limited number of anchors the space of the mobile nodes can be effectively covered. Furthermore, the validation shows that using the simulation framework, the same positioning results can be achieved as with the real measurement.

5 Solution of the industrial problem and its benefits

We developed a hardware and software that is capable of correcting the measurement based on the short-range radio distancing and calculate the exact indoor position using sensor fusion. This system can be used by the company in their future projects like a museum guide. The result was summarized in Annales Mathematicae et Informaticae, May 19, 2020.

6 Conclusions

In this work, a genetic-algorithm-based simulation framework is presented to determine an optimal anchor placement in an indoor environment. To implement a realistic and precise simulation environment, the 2D error characteristics of the DWM1001 module was measured and implemented in this work. Using the proposed framework,

various measurements with an optimal and with a reference anchor placement were carried out. The results show that the optimal anchor placement is crucial when the number of anchors is limited. It can also be concluded that if the number of anchors are increasing, their placement becomes less relevant. Furthermore, the validation shows that there is no significant difference between the simulation and the real experiments.

Also a sensor-fusion based algorithm was implemented on the hardware, which is capable of estimating the position of the user in real-time using other sensors, such as gyroscope and accelerometer.

References

[ICAI] Kalo Adam et al. "Indoor localization simulation framework for optimized sensor placement to increase the position estimation accuracy". In: *Annales Mathemat-icae et Informaticae* 51 (2020), pp. 29-39. Doi: 10.33039/ami.2020.07.002.

[dwm] DecaWave. *DWM1001 Datasheet, Version 1.10*. 2017.

[ga] John H. Holland. "Genetic Algorithms". In: *Scientific American* 267.1 (1992), pp. 66-73. Doi: 10.1038/scientificamerican0792-66.

[RSS] A.M. Ladd et al. "Robotics-Based Location Sensing Using Wireless Ethernet". In: *Wireless Networks* 11 (2005), pp. 189-204. Doi: https://doi.org/10.1007/s11276-004-4755-8.

[ToF] S. Lanzisera, D. T. Lin, and K. S. J. Pister. "RF Time of Flight Ranging for Wireless Sensor Network Localization". In: *International Workshop on Intelligent Solutions in Embedded Systems*. 2006, pp. 1-12. Doi: 10 . 1109/WISES . 2006.329127.

[multi] Abdelmoumen Norrdine. "An algebraic solution to the multilateration problem". In: *International Conference on Indoor Positioning and Indoor Navigation*. 2012, pp. 1-4. Doi: 10.13140/RG.2.1.1681.3602.

[AOA] Jie Xiong and Kyle Jamieson. "ArrayTrack: A Fine-Grained Indoor Location System". In: *Symposium on Networked Systems Design and Implementation*. 2013, pp. 71-84.

[UWB_base] Keqen Yu and Ian Oppermann. In: *UWB Theory and Applications*. John Wiley and Sons, Ltd, 2005, pp. 175-196. Doi: 10.1002/0470869194.ch8.

[Loc] F. Zafari, A. Gkelias, and K. K. Leung. "A Survey of Indoor Localization Systems and Technologies". In: *IEEE Communications Surveys Tutorials* 21.3 (2019), pp. 2568-2599. Doi: 10.1109/COMST.2019.2911558.

[RToF] Zheng Sun et al. "Cortina: Collaborative context-aware indoor positioning employing RSS and RToF techniques". In: *2011 IEEE International Conference on Pervasive Computing and Communications Workshops* (*PERCOM Workshops*). 2011, pp. 340-343. Doi: 10.1109/PERCOMW.2011.5766901.

Optimization methods for the implementation of clinical decision support systems

György Terdik[1], **András Hajdu**[2], **Henrietta Tomán**[2], **János Tóth**[2], **Attila Tiba**[2] and **László Kovács**[2]

[1] University of Debrecen, Faculty of Informatics, Department of Information Technology, email: terdik.gyorgy@inf.unideb.hu
[2] University of Debrecen, Faculty of Informatics, Department of Computer Graphics and Image Processing, email: {hajdu.andras, toman.henrietta, toth.janos, tiba.attila, kovacs.laszlo}@inf.unideb.hu

1 Executive Summary

Sightspot Network Ltd. (Hungary) is interested in methods for optimizing medical decision support systems that rely on ensembles of image processing algorithms. The aim of the project was to develop methods to find the optimal components and parameter settings of ensemble-based systems using stochastic methods.

A sampling-based stochastic search method was developed for accelerating parameter optimization with simulated annealing over large image datasets. The proposed method significantly reduces the time requirement of the optimization of image processing ensembles while preserves convergence in probability to the global optimum.

Furthermore, we investigated the cost functions appearing in particular optimization tasks and developed a stochastic method that takes into consideration the restricted resources during the optimization. This problem is formulated as a knapsack one, where the cost is the ensemble accuracy formed by some aggregation rules.

2 Challenge overview

Ensemble-based systems are rather popular in several application fields and are employed to increase the accuracy of individual approaches. We encounter such methodologies for pattern recognition purposes [3], using models based on e.g., neural networks [4, 5], decision trees or other principles [6, 7, 8]. In the most recent results, we can recognize this attitude in the design of state-of-the-art convolutional neural networks or in the direct combination of them [9]. Since contemporary techniques typically require a lot of time and capacity, optimization methods to facilitate the more efficient use of resources is required.

The industrial problem was to attenuate the increasing demand on resources when using ensembles for pattern classification and object detection tasks in medical decision support systems.

The two main optimization challenges were the following:

- Using the individually optimal parameter settings of the members not necessarily maximizes the accuracy of an ensemble. Therefore, ensemble-level parameter optimization is required, which is a resource-demanding problem, especially considering large datasets.

- In the practice, the available resources strictly limit the number of ensemble members. Therefore, optimal selection of the ensemble members under an additional cost constraint (time, power consumption, etc.) is required.

3 Implementation of the project

During the project two stochastic methods were developed. Simulated Annealing with Sampling-based Evaluation (*SA-SBE*) [1], which is a search method for the parameter optimization of ensembles over very large datasets, and Stochastic searcH for EnsembLe Creation (*SHErLoCk*) [2], which is a search method for the optimal selection of ensemble members under time constraints.

György Terdik was the principal investigator of the project. He supervised the project and took part in developing the method SHErLoCk. András Hajdu participated as a senior researcher. He contributed to the development of both methods with his expertise in discrete, continuous, and stochastic optimization and medical image processing. Henrietta Tomán also contributed to the development of both ensemble optimization methods with her expertise in discrete optimization. János Tóth worked on the development of SA-SBE using his expertise in optimization and medical image processing. Attila Tiba contributed to the development of SHErLoCk with his expertise in stochastic methods. Finally, László Kovács dealt with some of the implementation issues using his expertise in high performance computing.

4 The research problem, methods, and results

Ensemble-based approaches are routinely considered to aggregate the outputs of medical image processing algorithms, usually by some variant of the majority voting rule.

The industrial problem was to attenuate the increasing demand on resources when using ensembles for pattern classification and object detection tasks in medical decision support systems while maintaining the required accuracy.

We developed novel and efficient optimization techniques for selecting the algorithms of decision support systems and optimizing their parameters. In the selection of the right algorithms, their efficiency is a key issue, while in the parameter optimization the aim is to achieve the highest possible accuracy of the combined system.

4.1 The research problem

In the case of ensemble-based image processing systems, a common bottleneck of parameter optimization is the computational demand during the evaluation of the objective function, because it typically requires the application of the aggregation rule to the output of the ensemble members for a parameter setting for all images in a dataset.

To overcome this difficulty, we worked out a method that uses partial evaluation of the objective function over a certain subset of the dataset selected by random sampling.

We also worked out a method for the optimal selection of ensemble members, taking into consideration the restricted resources. This task was formulated as a knapsack problem, where the cost is the ensemble accuracy formed by some aggregation rules.

4.2 The applied methods

Considering the parameter optimization of ensembles, our main aim was to theoretically establish an accelerated, sampling-based evaluation method for Simulated Annealing (SA) that preserves the convergence properties of this stochastic search technique.

The main contribution of our approach is recognizing that sampling can be considered a specific type of noisy evaluation of the objective function. Thus, after well-designed transformations, the convergence results for noisy evaluation are suitable for sampling-based evaluation.

In the case of the optimal selection of ensemble members, the generally applied aggregation rules lead to non-separable cost functions. Therefore, the common solution tools - such as dynamic programming - was took out of action. For this reason, we proposed a novel stochastic approach that considers the cost as the joint probability function of the member's accuracy.

The knowledge originating from the joint probability function was efficiently incorporated into a stochastic search process as a stopping rule, since we have the information on the expected accuracy or, alternatively, the probability of finding ensembles that are more accurate.

4.2.1 Simulated Annealing with Sampling-based Evaluation (SA-SBE)

We proposed a method called SA-SBE [1] that is based on SA and can maintain the solution quality while reducing the runtime for objective (referred to as energy hereafter) functions that are commonly used to evaluate the average performance of object detectors and classifiers over datasets.

In particular, we proposed a sampling strategy that considers only a suitable portion of the dataset in each search step to maintain the convergence of SA, which approach can be considered a noisy evaluation of the energy function. The appropriate sample sizes required during the search process are theoretically determined by adapting the convergence results for noisy evaluation in SA.

4.2.2 Basic idea of the sampling-based evaluation in SA-SBE

For a random sample $|\Lambda_n| = n$, the noise is the difference of the energy E and its estimate $\widehat{E}_{\Lambda_n}$:

$$d_{\Lambda_n} = \widehat{E}_{\Lambda_n} - E \tag{1}$$

By considering discrete search spaces and assuming that the noise is normally distributed with mean 0 and variance $\left(\sigma_{d_n}^{(k)}\right)^2 > 0$ in the k-th ($k \in \mathbb{N}$) iteration, it can be proven [10] that SA using noisy evaluation also converges to the globally optimal solution in probability in the same manner as that when using exact energy

values if the standard deviation $\sigma_{d_n}^{(k)}$ of the noise is dominated by the temperature $T^{(k)}$ in the k-th iteration for each k, i.e., when

$$\sigma_{d_n}^{(k)} = o\left(T^{(k)}\right), \tag{2}$$

That is, we have to maximize $\sigma_{d_n}^{(k)}$ to minimize the sample size n required in each step.

A sufficiently simple general form of $\sigma_{d_n}^{(k)}$ that maximizes its value at the temperature $T^{(k)}$ can be given as follows:

$$\sigma_{d_n}^{(k)} \approx T^{(k)}(1-\epsilon)^k \text{ with } 0 < \epsilon < 1. \tag{3}$$

For an arbitrary cooling schedule, the minimum sample size $n^{(k)}$ required at the k-th iteration to maintain the convergence of the method in probability can be estimated as

$$n^{(k)} \approx \frac{N\sigma_{max}^2}{(N-1)\sigma_{d_n}^{(k)^2} + \sigma_{max}^2}, \tag{4}$$

where σ_{max} is the worst-case maximum value of the population standard deviation $\sigma_N^{\mathcal{D}(\pi)}$, and $\sigma_{d_n}^{(k)}$ can be derived using (3).

4.2.3 Stochastic searcH for EnsembLe Creation (SHErLoCk)

We proposed a stochastic technique called SHErLoCk [2] to solve Knapsack problems with complex energy functions. The model is worked out in detail for accuracy settled on the majority voting rule; however, it can also be applied to other energy functions. Our approach is based on the stochastic properties of the energy q_ℓ, providing that we have some preliminary knowledge on the distribution its parameters p_i $(i = 1, \ldots, n)$ are coming from with a special focus on *Beta* distributions that fit practical problems very well.

In other words, we estimate the distribution of q_ℓ in terms of its mean and variance. This information can be efficiently incorporated as a stopping rule in stochastic search algorithms. The main idea here is to be able to stop composing ensembles, when we can expect that better ones can be found by low probability.

4.2.4 Basic idea of ensemble member selection in SHErLoCk

The aim is to find the most efficient ensemble with an execution time limit T. The selection of the ensemble members is based on their efficiency $q(p_i, t_i)$ which is calculated as

$$q(p_i, t_i) = \sum_{k=\lfloor T/t_i \rfloor/2+1}^{\lfloor T/t_i \rfloor} \binom{\lfloor T/t_i \rfloor}{k} {p_i}^k (1-p_i)^{\lfloor T/t_i \rfloor - k} \tag{5}$$

instead of the values p_i/t_i considered in e.g., the classic greedy method, where p_i is the accuracy and t_i is the time requirement of the i-th member.

In the l-th selection step the efficiency values $q^{(l-1)}(p_i, t_i)$ of the remaining candidates are updated to the maximum time left and the i-th candidate is selected as the next member of the ensemble with the following probability:

$$(\mathrm{P}_{ens})_i^{(k)} = \frac{q^{(k-1)}(p_i, t_i)}{\sum_j q^{(k-1)}(p_j, t_j)}, \tag{6}$$

where $i, j \in \mathcal{N} \backslash \{i_1, \ldots, i_{k-1}\}$. This discrete random variable reflects that the more efficient the item is, the more probable it is selected for the ensemble in the next step.

4.3 The achieved research results

In object detection applications, it is common to optimize systems using objective functions that are calculated as an average over a dataset. Our motivation for constructing the proposed SA-SBE method was to provide a theoretically established way for reducing the time required for optimization but without any significant loss of the solution quality if the dataset considered is large. We proposed a sampling strategy to ensure that SA exhibits the same convergence in probability using sampling-based evaluation as that using complete evaluation. Our experimental results indicate that the proposed sampling-based evaluation method substantially reduced the computational time required for optimizing the parameters of the ensembles while preserving solution quality.

Ensemble-based methods are popular approaches to increase the accuracy of a decision by aggregating the opinions of individual voters. The task to find optimal ensembles under an additional constraint on the total cost of the members, can be formulated as a knapsack problem, where the energy is the ensemble accuracy formed by some aggregation rules. We introduced a novel stochastic approach (SHErLoCk) that considers the energy as the joint probability function of the member accuracies. This type of knowledge can be efficiently incorporated in a stochastic search process as a stopping rule. Experimental analyses of the created ensembles of pattern classifiers and object detectors confirm the efficiency of our approach. We compared the performance of the proposed search strategy with other stochastic, genetic, and pruning methods on binary classification problems of UCI datasets. Our stochastic search strategy was reasonably faster than the other methods and slightly dominant in accuracy, as well. Moreover, we proposed a novel stochastic search strategy that better fits the energy, compared with general approaches.

5 Solution of the industrial problem and its benefits

The industrial partner of the research (Sightspot Network Ltd., Hungary) was provided with novel methods to optimize the ensembles of the medical decision support systems they develop.

6 Conclusions

The proposed evaluation method of SA-SBE is domain independent and easy to adapt to problems where evaluation over large datasets is required. This method does not incorporate complex techniques for the determination of the required sample size (e.g., monitoring changes in energy) or sample selection (e.g., finding the critical samples in classes) to accelerate the search process.

In future research, we plan to investigate embedding sampling-based evaluation with a dynamic sample size in other stochastic search methods. In addition, we plan to generalize the proposed dataset-level sampling strategy to systematic image-level sampling, i.e., to image downsampling, where estimation of the noise originating from the sampling is required.

Natural generalizations of the stochastic approach SHErLoCk regarding multiple objectives and multiclass classification are planned. Within our framework, multiple constraints lead to a multiply-constrained knapsack problem, which can be handled e.g., with merging the constraints into a single one to make our approach directly applicable. As for multiclass classification, to realize the classic setup, we have to consider multinomial distribution instead of the binomial one to handle majority voting in the corresponding classification problem. These changes raise serious theoretical challenges over the techniques we developed during this work; they need a dedicated future research beyond the current scope.

References

[1] J. Tóth, H. Tomán, A. Hajdu, *Efficient sampling-based energy function evaluation for ensemble optimization using simulated annealing*. Pattern recognition, vol. 107, 107510, 2020.

[2] A. Hajdu, Gy. Terdik, A. Tiba, H. Tomán, *A stochastic approach to handle resource constraints as knapsack problems in ensemble pruning*. Machine Learning, under revision.

[3] L. Lam, S.Y. Suen, *Application of majority voting to pattern recognition: An analysis of its behavior and performance*. Trans. Sys. Man Cyber. Part A, vol. 27, no. 5, pp. 553-568, 1997.

[4] L.K. Hansen, P. Salamon, *Neural network ensembles*. IEEE Trans. Pattern Anal. Mach. Intell., vol. 12, pp. 993-1001, 1990.

[5] S.-B. Cho, J.H. Kim, *Combining multiple neural networks by fuzzy integral for robust classification*. IEEE Transactions on Systems, Man, and Cybernetics, vol. 25, pp. 380-384, 1995.

[6] T.K. Ho, J.J. Hull, S.N. Srihari, *Decision combination in multiple classifier systems*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 16, pp. 66-75, 1994.

[7] Y.S. Huang, C.Y. Suen, *A method of combining multiple experts for the recognition of unconstrained hand-written numerals*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 17, pp. 90-94, 1995.

[8] L. Xu, A. Krzyzak, C.Y. Suen, *Methods of combining multiple classifiers and their applications to handwriting recognition*. IEEE Transactions on Systems, Man, and Cybernetics, vol. 22, pp. 418-435, 1992.

[9] B. Harangi, Á. Baran, A. Hajdu, *Classification of skin lesions using an ensemble of deep neural networks* in 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC 2018, Honolulu, HI, USA, July 18-21, pp. 2575-2578, 2018.

[10] S.B. Gelfand, S.K. Mitter, *Simulated annealing with noisy or imprecise energy measurements*. J. Optim. Theory Appl., vol. 62, pp. 49-62, 1989.

Optimization of geometric models to support manufacturing processes

Henrietta Tomán[1], **Csaba Vincze**[2], **Zoltán Muzsnay**[2] and **János Tóth**[1]

[1]University of Debrecen, Faculty of Informatics, Department of Computer Graphics and Image Processing,
[2]University of Debrecen, Faculty of Science and Technology, Institute of Mathematics, Department of Geometry,

[1]E-mail: toman.henrietta@inf.unideb.hu
[2]E-mail: csvincze@science.unideb.hu
[3]E-mail: muzsnay@science.unideb.hu
[4]E-mail: toth.janos@inf.unideb.hu

1 Executive Summary

The company Varinex, having significant role in the development and domestic dissemination of individual and small-batch production methods based on 3D printing, was interested in designing and manufacturing of customized implants.

There are several different aspects to be considered in optimizing the designing and manufacturing technologies for bone replacement: e.g. achieving high accuracy in 3D reconstruction, finding the optimal spatial form of bone deficiency, reducing the required time for the applied techniques. For the extraction of bone structures from computer tomography (CT) images, an ensemble of segmentation algorithms with optimal parameter setting is applied, whose outputs are aggregated by some decision rule (e.g. majority voting), to outperform the individual member accuracies. The reconstruction of the defect is based on cubic Bezier curves, discretization of curvature, principal curvatures and torsion. To reduce the computational costs and time, nearest neighbor image pyramid-based noisy energy function evaluation method is proposed for the local search technique simulated annealing (SA) at different downscaling levels during the process.

2 Challenge overview

The use of Additive Manufacturing technologies, 3D printing makes one of the first steps by fabricating custom implants to enable new and essential tools for the medical industry [1]. In modern orthopedics, customized, patient-specific implants are designed and manufactured when none of the conventional implants have the particular shape and dimensions to fit the defect [2][4]. Besides the surgery time and cost, the risk of post-operative complications is also reduced by custom-made implants [3].

For the 3D reconstruction of the injured area, CT/MRI scan is performed, with the aim of observing the defect. Traditional methods of manual design of customized implants are shown to have many disadvantages (time-consuming, expensive cost,

complicated procedures). Nowadays, computer-aided design and manufacturing techniques are common tools to prefabricate patient-specific implants.

The most important aim is to optimize the designing and manufacturing process for bone replacement. In the optimization problem, among others the following objectives can be taken into consideration: to achieve the highest possible accuracy of 3D bone segmentation, to find the optimal spatial form of bone replacement, to minimize the required time/computational cost (while maintaining the accuracy) of the technologies applied in practice.

3 Implementation of the project

During the project, to increase the accuracy of the 3D reconstruction, an ensemble of segmentation algorithms with optimal parameter setting is applied. We developed a stochastic method to reduce the computational costs and time of 3D reconstruction. Moreover, to obtain a more precise joint between the implant and the region surrounding the defect, our reconstruction method is based on cubic Bezier curves, discretization of curvature, principal curvatures and torsion.

Henrietta Tomán was the principal investigator of the project. She supervised the project and took part in developing the stochastic method with her expertise in ensemble methods and discrete optimization. Csaba Vincze and Zoltán Muzsnay participated as a senior researcher. They contributed to the development of the reconstruction method based on cubic Bezier curves. Their expertise in differential and convex geometry was very beneficial in examining and approving the preciseness of the joint between the customized implant and the defect-surrounding region. János Tóth worked on the development of the stochastic method using his expertise in optimization and medical image processing.

4 The research problem, methods, and results

In medical image processing problems, ensemble-based methods are able to outperform the accuracy of each member algorithm, by applying some aggregating rule (usually a variant of majority voting) for the outputs to make the final decision. The more diverse the member algorithms are, the more beneficial the ensemble technique is. We apply an ensemble of segmentation algorithms with optimal parameter setting to achieve higher accuracy in 3D bone segmentation than the individual member algorithms.

We propose an image pyramid-based noisy energy function evaluation method for the local search technique simulated annealing. The method is primarily designed for the optimization of image segmentation algorithms, and it maintains solution quality with significantly reduced time requirement. The strategy to select the proper image pyramid levels during the search is theoretically determined via adapting results regarding evaluation in simulated annealing based on imprecise measurements. To ensure the convergence of SA, we have to control the standard deviation of the noise according to the actual temperature by appropriate downscaling level (i.e. the resolution of the images) in each iteration.

4.1 The research problem

Optimization of a system having numerous free parameters regarding a complex energy function is a very challenging task. Even if we can transform the problem to a finite discrete one via e.g. quantizing and limiting the domains of the parameters, the number of possible parameter settings grows exponentially with the number of the parameters. Accordingly, an exhaustive search for the optimal parameter setting soon becomes completely impractical. Moreover, the more efficient stochastic search approaches may still require too much time, especially if the evaluation of the energy function itself is expensive. We address this issue by suggesting a special type of noisy evaluation [6] of the energy function of simulated annealing (SA) [7], which is a popular and widely applied stochastic search strategy.

We consider an ensemble of segmentation algorithms dedicated to the extraction of bone structures from computer tomography (CT) images. The outputs of the individual segmentation algorithms are binary images containing the candidate bone regions, which are aggregated by majority voting. These algorithms have numerous adjustable parameters, whose optimal settings may differ, when a given algorithm is considered as an individual approach or a member of an ensemble. The evaluation of the ensemble performance for a given parameter setting is done by comparing the aggregated segmentation results of the ensemble for a training set with the corresponding manually annotated ground truth. The comparison is carried out using the statistic intersection over union (IoU), which is commonly considered for object detection. Thus, to compute the energy function during the optimization process, an evaluation over the whole training set should be performed at each search step. A possible speed up during the stochastic search is to apply noisy evaluation to determine the ensemble performance, which means that the energy function is just approximated instead of determined precisely. Naturally, to maintain the convergence characteristics of the search method the noise should be controlled.

4.2 The applied methods

Considering the parameter optimization of ensembles, our main aim was to theoretically establish an accelerated, sampling-based evaluation method for Simulated Annealing (SA) that preserves the convergence properties of this stochastic search technique.

In [10], we have already successfully presented how such a noisy evaluation can be provided by selecting only subsets of the training set with proper cardinalities instead of the whole training set during the search. Now, as a novel contribution, we propose an image pyramid representation of the training set, where evaluating on lower resolution levels results in noisy determination of the energy. Naturally, the lower the resolution is, the larger the noise can be, since the segmentation results using lower resolution images can be less precise. To meet the theoretical requirements, we introduce a strategy to determine the maximal allowed noise level in each iteration to control the search. We show that our method successfully reduces the time requirement of the search while preserving solution quality.

4.2.1 An Ensemble for Automatic Bone Segmentation in CT Images

Our automatic bone segmentation ensemble consists of five algorithms, each of which has many parameters. However, to gain a problem that is computationally reasonable, we selected only those parameters for the later optimization that significantly influence the ensemble output.

As the aggregation method to obtain the output of the ensemble based on the individual output of the members, we chose classic majority voting. That is, the pixel values of the ensemble output C is determined as

$$C(x,y) = \begin{cases} 1, & if \sum_{i=1}^{N} C_i(x,y) \geq \left[\frac{N}{2}\right] \\ 0, & otherwise \end{cases}, \tag{1}$$

where N is the number of member algorithms.

4.2.2 Accelerated Parameter Optimization using Image Pyramid-based Noisy Evaluation

We propose a novel image pyramid-based noisy energy function evaluation method for the parameter optimization of image segmentation ensembles that reduces the time requirement of SA while preserves its convergence properties.

Simulated annealing is a local search strategy [7] that is widely applied to address difficult combinatorial optimization problems. The main feature of SA is that it can escape from local optima by allowing moves that deteriorate the energy function value with a probability depending on a control parameter and the energy function difference of the candidate and the current solution. SA assumes that the energy of a solution can be determined exactly; however, the evaluation of a solution is often subject to noise in real-life problems.

Considering discrete search spaces and assuming that in the k-th ($k \in \mathbb{N}$) iteration the noise is normally distributed with mean 0 and variance $\left(\sigma^{(k)}\right)^2 > 0$, Gelfand and Mitter proved [6] that SA using noisy evaluation exhibits the same convergence properties as using exact energy values, if the standard deviation $\sigma^{(k)}$ of the noise is dominated by the temperature $T^{(k)}$ in the k-th iteration for each k, that is, when

$$\sigma^{(k)} = o\left(T^{(k)}\right). \tag{2}$$

We will describe how the criterion defined by (2) can be exploited to accelerate the parameter optimization of an image segmentation ensemble using image pyramids.

In the case of ensembles, the more parameters they have, the larger the training set required is to find their optimal parameter setting and to avoid overfitting. However, considering large training sets, evaluating the mean performance of the ensemble with a parameter setting can be computationally expensive even using simple cost functions. Therefore, we propose to estimate the energy function value E of SA at different downscaling levels of the original ground truth and the aggregated output of the ensemble during the SA process.

4.2.3 Nearest neighbor image pyramid

We refer to a collection of $L \in \mathbb{N}$ hierarchically downscaled versions of an image as an L-level image pyramid, in which the higher the level l ($l \in 0, 1, \dots L-1$), the smaller the image resolution is.

The most common method to construct image pyramids is the Gaussian pyramid, where levels of the pyramid are built by convolving the original image with a Gaussian-like averaging filter followed by a subsampling step [5]. However, in our case the images are binary masks, therefore to preserve sharp boundaries, we use the nearest neighbor method for simple subsampling. That is, the pixel values of a level are defined to match the original pixel whose center is the nearest to the sample position.

4.2.4 Image Pyramid-based Noisy Evaluation

Assuming that the cost of calculating E is proportional to the resolution of the input images, the calculation of the energy function estimate $\widehat{E}_l$ over the l-th level version of the input images (with an associated scaling factor γ_l) has $1/\gamma_l^2$ times lower cost than the calculation of E; however, using $\widehat{E}_l$ introduces noise in the evaluation. The noise d_l originating from using the l-th level version of the input images can be determined as

$$d_l = \widehat{E}_l - E. \tag{3}$$

This noise may cause SA to consider an inferior state as superior because of the imprecise evaluation of the energy function. That is, the stronger the noise, the more random the search. For this reason, a suitable strategy to control the noise is required.

To ensure the convergence of SA, we have to apply a strategy to select the appropriate downscaling level l in each iteration k to control the standard deviation of the noise $\sigma_d^{(k)}$ according to the temperature $T^{(k)}$, as described by (2).

First, we have to determine the maximum allowed value of $\sigma_d^{(k)}$ for a given temperature $T^{(k)}$. Using (2), we get that

$$\lim_{k\to\infty} \frac{\sigma_d^{(k)}}{T^{(k)}} = 0 \tag{4}$$

must hold.

To maintain the limit in (4), the sequence $\left\{\sigma_d^{(k)}\right\}$ has to be decreasing such that $\lim_{k\to\infty} \sigma_d^{(k)} = 0$ and $\sigma_d^{(k)} < T^{(k)}$ for each $k \in \mathbb{N}$.

Based on the above conditions, a sufficiently simple general form of $\sigma_d^{(k)}$ that maximizes its value considering a given $T^{(k)}$ can be given as:

$$\sigma_d^{(k)} = T^{(k)}(1-\epsilon)^k \text{ with } 0 < \epsilon < 1. \tag{5}$$

The next step to construct our strategy is to determine the dependence of $\sigma_d^{(k)}$ on the downscaling level l. Using a downscaled version of the images of the training set during the evaluation may result in significantly different noise for different energy functions. In some cases, it can be straightforward to determine the theoretical maximum value of $\sigma_d^{(k)}$, while for more complex energy functions it becomes a difficult

problem. However, in the case of natural images the empirical standard deviation of the noise $\sigma_{d,l}$ for a level l is expected to be much lower than the theoretical maximum. Therefore, we propose to estimate $\sigma_d^{(k)}$ for each level of the image pyramid by measuring the value of $\sigma_{d,l}$ on the ground truth used for the evaluation.

Having $\sigma_{d,l}$ measured for each level l of the image pyramid, we can determine the highest level l (i.e. the lowest resolution) where $\sigma_{d,l}$ is less than or equal to the maximum allowed $\sigma_d^{(k)}$ for each temperature level $T^{(k)}$.

We describe the methodology used to assess the performance of the proposed image pyramid-based noisy evaluation method.

4.2.5 Energy function

We have performed parameter optimization of the bone segmentation ensemble by the proposed noisy evaluation method, with the aim of efficiently finding the parameter setting that maximizes the segmentation performance of the ensemble in terms of average intersection over union $\overline{IoU}_l$, which is defined as the number of common pixels of the ensemble output $C_{m,l}$ and the corresponding ground truth $G_{m,l}$ for a given level l of the image pyramid ($l = 0, 1 \ldots, L-1, \quad m = 1, 2, \ldots, M$) over the number of pixels in either of the two:

$$\overline{IoU}_l = \sum_{i=1}^{M} \frac{|C_{m,l} \cap G_{m,l}|}{|C_{m,l} \cup G_{m,l}|}. \tag{6}$$

To obtain a minimization problem, we define the energy function as:

$$E = 1 - \frac{\overline{IoU}_0}{M}, \tag{7}$$

and the energy function estimate for the level l as:

$$\widehat{E}_l = 1 - \frac{\overline{IoU}_l}{M} \tag{8}$$

where L is the number of levels in the image pyramid and M is the number of images in the training set.

4.2.6 Realization of the noisy evaluation

For our ensemble of segmentation algorithms dedicated to the extraction of bone structures with optimal parameter setting, for each level l, we used the followings method (in more detail see [9]):

- Downscale the ground truth and ensemble output images with scaling factor γ_l to construct the corresponding level of the image pyramids.
- Upscale the images to the original size (with scaling factor $1/\gamma_l$).
- Determine d_l as defined in (3). For this, compute E and its estimate $\widehat{E}_l$ for the level l as described in 4.2.5 using the scaled versions of the m-th image (for each m) as $C_{m,l}$ and the original images as $G_{m,l}$.

- Calculate $\sigma_{d,l}$.

It can be observed that the standard deviation of the noise exhibits a similar trend in the case of both the ground truth and the corresponding ensemble output.

From the maximum value of $\sigma_d^{(k)}$, we were able to determine the required level l for each iteration (temperature level) k.

4.3 The achieved research results

To optimize the designing and manufacturing technologies for customized implants, several different aspects are considered: e.g. achieving the highest possible accuracy in 3D reconstruction, finding the optimal spatial form of bone deficiency, reducing the cost(e.g. the required time) for the applied techniques.

To increase the accuracy of the bone segmentation from CT images, an ensemble of segmentation algorithms with optimal parameter setting is applied as it outperforms the individual member accuracies. In the reconstruction of the defect, classic well-known technologies based on cubic Bezier curves, discretization of curvature, principal curvatures and torsion are applied. To reduce the computational costs and time of the methods (e.g. to find the optimal parameter setting of the ensemble members), nearest neighbor image pyramid-based noisy energy function evaluation method is proposed for the local search technique simulated annealing (SA) at different downscaling levels during the process. The strategy to select the proper image pyramid levels during the search is theoretically determined via adapting results regarding evaluation in simulated annealing based on imprecise measurements. To ensure the convergence of SA, the standard deviation of the noise is controlled according to the actual temperature by appropriate downscaling level (i.e. the resolution of the images) in each iteration. Our experimental results show that the proposed method substantially reduced the required computational time while maintaining the quality of the solution.

5 Solution of the industrial problem and its benefits

VARINEX Informatics Inc., the industrial partner of the research is provided with novel methods to optimize the designing and manufacturing process for the patient-specific implants, e.g. to reduce the costs and time of the design process. At the University of Debrecen, many interesting results have obtained in the domain of the fabrication and implantation of patient-specific implants, which is one of the leading territories of the modern orthopedic surgery. Besides a long successful cooperation with the Faculty of Medicine, Clinical Center, Varinex has the opportunity to build new collaboration with other research group of University of Debrecen on designing and manufacturing titanium-based patient-specific implants.

6 Conclusions

In this project, we examined from different aspects how the designing and manufacturing process for the patient-specific implants can be optimized. To increase the

accuracy of the 3D reconstruction, an ensemble of segmentation algorithms with optimal parameter setting is applied. We have proposed an image pyramid-based noisy energy function evaluation method for the local search technique SA. Considering an image segmentation ensemble designed to extract bone structures from CT images, we showed that using the proposed method it is possible to reach solutions with the same quality as using standard simulated annealing, but with significantly reduced time requirement. To reconstruct the defect, well-known techniques based on e.g. cubic Bezier curves are applied. In future research, we plan to adapt our proposed solutions (ensemble method, image pyramid-based noisy evaluation method with downscaling) [8] and the corresponding results in [9] in higher dimension to increase the accuracy of 3D reconstruction of the defect while reducing the cost (computational time).

References

[1] B. Berman, *3-D printing: The new industrial revolution*. Business Horizons 55(2), 155-162, 2012.

[2] V.P. Bogu, Y.R. Kumar, A.K. Khanara, *Modelling and structural analysis of skull/cranial implant: beyond mid-line deformities*. Acta of Bioengineering and Biomechanics Vol. 19, No. 1, 2017.

[3] R. Bogue, *3-D printing: The dawn of a new era in manufacturing?* Assembly Automation 33(4), 307-311, 2013.

[4] D.J. Bonda, S. Manjila, W.R. Selman, D. Dean, *Review, The recent revolution in the design and manufacture of cranial implants: modern advancements and future directions*, Neurosurgery, Vol. 77, Issue 5, 814-824, 2015.

[5] P. Burt, E. Adelson, *The Laplacian Pyramid as a Compact Image Code*. IEEE Transactions on Communications, vol. 31, 532-540, 1983.

[6] S.B. Gelfand, S.K. Mitter, *Simulated annealing with noisy or imprecise energy measurements*. J. Optim. Theory Appl., vol. 62, 49-62, 1989.

[7] S. Ledesma, G. Avia, R. Sanchez, *Practical Considerations for Simulated Annealing Implementation*. Simulated Annealing, ch. 20, 401-420, 2008.

[8] J. Toth, H. Toman, A. Hajdu, *Efficient sampling-based energy function evaluation for ensemble optimization using simulated annealing*. Pattern recognition, vol. 107, 107510, 2020.

[9] J. Toth, T.P. Kapusi, B. Harangi, H. Toman, A. Hajdu, *Accelerating the Optimization of a Segmentation Ensemble using Image Pyramids*. Proc. 11th International Symposium on Image and Signal Processing and Analysis (ISPA), 2019.

[10] J. Toth, L. Szakacs, A. Hajdu, *Finding the optimal parameter setting for an ensemble-based lesion detector*. Proc. 2014 IEEE International Conference on Image Processing (ICIP), 3532-3536, 2014.

Optimal Laser System for Motion Illumination

Tamás Vinkó[1], **Boglárka G.-Tóth**[2], **Kristóf Kovács**[3] and **Benjamin Balogh**[4]

[1234]University of Szeged, Institute of Informatics,

[1]E-mail: tvinko@inf.u-szeged.hu
[2]E-mail: boglarka@inf.u-szeged.hu

1 Executive Summary

Training in many sports includes the practice of harmonized strategic movements. The laser system aims to illuminate dynamic paths of the players, see Figure 1. Our aim was to find the optimal number and placement of the lasers, and the best possible assignment of the objects to the lasers, such that in minimal cost the illumination is feasible and satisfies the visibility constraints.

First, the optimal number and location of the lasers are determined modelling the problem as a rectangle covering problem, resulting in a Mixed Integer Linear Programming problem. In order to solve this hard problem, a new constraint generation approach is developed, that can solve any real-size problems. Next, for a given dataset describing the dynamic paths of the players the assignment of players to lasers have to be determined, such that the brightness of each point satisfy the visibility constraints. This assignment can be seen as an online problem, as each assignment and illumination have to be done in milliseconds.

A Mixed Integer Linear Programming model, including a constraint generation approach is developed to solve the real-life location problem of the lasers minimizing the cost of the system. The assignment of players to lasers to be illuminated by is

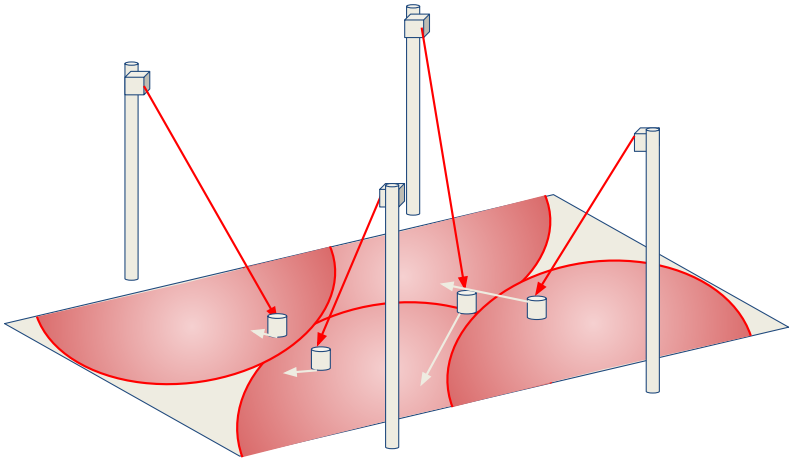


Figure 1: A possible placement of laser units with their illumination area in red and some objects with their direction and speed.

done by heuristic procedures, maximizing the minimum brightness of the illuminated points. The control of the laser beam is done by a two-mirror scanner, through a digital-analog converter, which is controlled by the developed program.

2 Challenge overview

One of the most important components in team sport nowadays is strategic thinking. Effective training of players for harmonized strategic movements is essential in world-leading coaching. Today's technology is ready for the challenge to take a big step towards this aim. The objective of this project is to investigate the possibilities and limitations of a controllable laser system, and to design a prototype for solving it. We aim to effectively show the dynamic trajectories of the players by dynamically illuminating their path to follow.

The laser system will contain some controllable laser units built from:

- a fixed laser light emitting equipment
- a two-mirror scanner for the control of the laser beam [9, 7, 11], see Figure 2.
- a digital-analog converter (DAC) so that the laser can be properly controlled by a computer

3 Implementation of the project

The project team was led by the Institute of Informatics of University of Szeged (Tamás Vinkó and Boglárka G.-Tóth). One collaborator in this project was the Department of Optics and Quantum Electronics at the University of Szeged (Béla Hopp and Tomi Smausz Kolumbán) who provided us a laser together with a galvo system and also a DAC. The partner provided some experience in laser systems, but also some data for the dynamic moves for the testing [10]. The consulting industrial partner was Bem 15 company. Their vision is tactic-oriented training through illuminating the paths for each player, which is a better way of teaching tactical movements, but can also be used for other kind of training activity.

4 The research problem, methods, and results

In this project two main problem can be defined. One is the control of the galvo system together with a software which is able to decide the assignment of the paths to illuminate to the lasers. In this problem the research question is the assigment problem, which needs to be decided in very short time.

The other research problem is the design of the laser system around the field such that the whole field can be illuminated with minimal cost. This problem is described here in more details.

4.1 The research problem

We aim to locate light sources around a rectangular field. The lasers are powerful enough to highlight multiple paths, so only their coverage of the field is important. We deal only with the location of the lights sources knowing the area that they can illuminate. The light sources can only be placed outside the field. Different types of light sources are available, characterized by their lighted areas and their costs. These lighted areas are assumed to be rectangles to simplify the model and to ignore the light areas with a too oblique angle of incidence. The whole field has to be covered by the lighted rectangles.

Mathematically the problem can be formulated as a rectangle covering problem, where the covering items are rectangles themselves and they each have a specific side that has to be on, or outside the field (the one with its light source). An example of a feasible solution with two different types of lights can be seen in Figure 3, where the light sources are represented by a numbered small circle. The blue rectangles represent the lighted areas with the darker shades being their overlapping areas.

4.2 The applied methods and some results

We have formulated a covering problem, where a rectangle has to be covered by different types of rectangles that describe the illuminated area of light sources. As

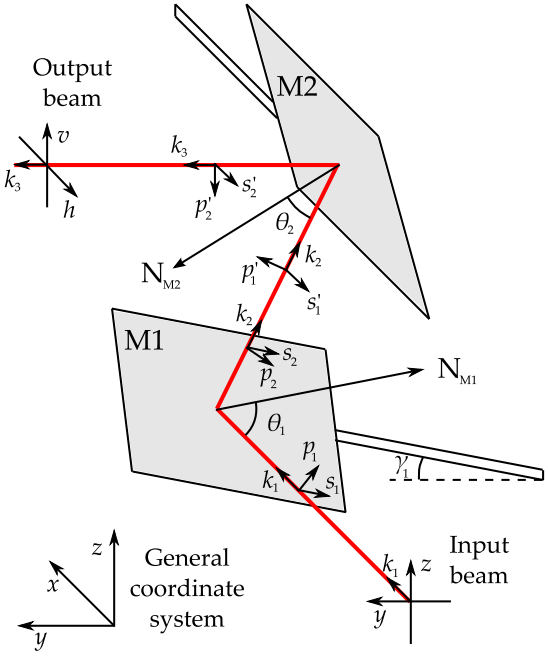


Figure 2: A galvo scanning system with the optical path of a beam indicated by the red line [1].

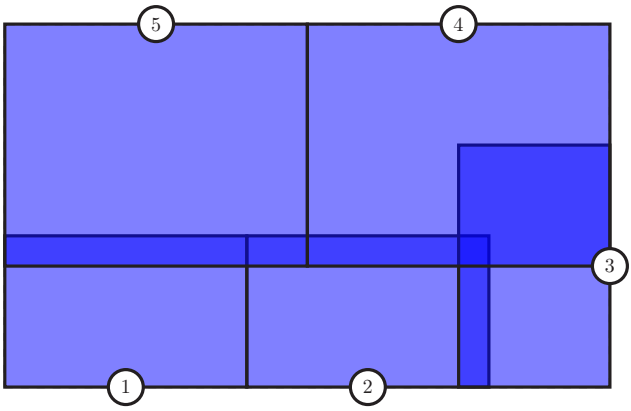


Figure 3: An example of a feasible solution to our rectangle covering problem.

the problem does not fit to standard problems, we did not find existing models or algorithms to solve this specific problem.

We propose a constraint generation approach for solving this covering problem. We formulate a MIP model to locate the light sources such that a finite number of predetermined points have to be covered. The result does not necessarily solve the original problem, i.e. it does not cover the whole field. Therefore, a constraint generation model is built to calculate a non-covered point such that the first model has to improve its previous solution to cover this new point as well. If no uncovered point is found, the result is an optimal covering, thus we stop.

We have also designed some set of additional constraints to exclude symmetrical solutions, to speed up the algorithm.

We used AMPL [4] to implement the models and solved them with CPLEX 12.8.0. The execution times (in seconds) can be seen in Table 4. The tests were run for different sets of light sources with $2, 3, 4, 5, 7$, and with $n = 10$, the maximum number of available light sources.

The problems 1-4 were generated by incrementally adding new types of light sources to the previous problem, with different attributes. Problems 5-8 were designed to test the limits of the approach.

Four configurations of the constraints were tested. The one named basic contains the basic models, OE has additionally the Order constraints as well as the Elimination of symmetric solutions. OF includes constraints on the order and on the field edge covering apart from the basic model. Lastly, OEF has every constraint we listed for the constraint generation approach.

The results make it clear that the constraint configuration using all the improvements is superior in almost all cases to the others. It can also be seen that the most important constraints apart from the order (which in itself does not lead to any improvement) are the field edge covering constraints. The elimination method improves the consistency of the algorithm, but highly increases the execution times without the field edge covering constraints.

Figure 4: Execution times (in seconds) of different sets of constraints for maximum 10 light sources.

problem	#light types	basic	OE	OF	OEF
1	2	22	209	24	38
2	3	61	346	15	20
3	5	>1800	318	19	13
4	7	>1800	738	29	27
5	2	>1800	>1800	278	187
6	4	1769	>1800	15	10
7	2	64	440	2	2
8	3	3	5	1	1
Average		>915	>708	48	38

5 Main benefits

We have formulated the problem as a covering problem by rectangles. As no general purpose solver exists to solve this problem, we have designed our solution method.

To build a useful model, a constraint generation approach is designed. The concept is to cover a set of points and check if the whole area is covered. If there is an uncovered area, we locate a new point in it and rerun the covering model requiring this new point to be covered as well. This approach is efficient for all the cases appearing in the cases generated by the known data. Still, we could improve it by introducing ordering on the rectangles, ensuring no gap between consecutive rectangles, and finally, eliminating symmetric solutions.

6 Conclusions

Apart from setting up these models to solve the given covering problem, we analyzed the model components. We checked the efficiency of the additional constraints. We have found that forcing the most structure on the rectangles, by having neighboring rectangles touch increases the model's consistency by a large margin. We also learned that eliminating symmetric solutions helps more if the model already forces a stricter structure on the rectangles.

On the implementation part, we have designed a software to demonstrate that controlling the scanner is possible, and more importantly, it is possible to illuminate more then one path at a time because all the components of the laser unit is fast enough to switch from one point to the other. The results of the implementation and assignment problem resulted in an awarded Scientific Student Thesis [2].

References

[1] G. Anzolin, A. Gardelein, M. Jofre, G. Molina-Terriza and M.W. Mitchell, Polarization change induced by a galvanometric optical scanner. *Journal of the Optical Society of America A, Optics, image science, and vision* Vol. 27, 1946-52, 2010.

[2] B. Balogh, Laser scener and assignment heuristics for the illuminated objects (in Hungarian). Scientific Student Thesis, 2018.

[3] G. Dosa, L. M. Hvattum, T. Olaj, and Z. Tuza, The board packing problem: Packing rectangles into a board to maximize profit. In *Pannonian Conference on Advances in Information Technology (PCIT 2020)*, pp. 10-16, 2020.

[4] R. Fourer, D. M. Gay, and B. W. Kernighan, *AMPL: A Modeling Language for Mathematical Programming*. Cengage Learning, 2002.

[5] D. Franzblau and D. Kleitman, An algorithm for covering polygons with rectangles. *Information and Control*, Vol. 63(3):164-189, 1984.

[6] E. Huang and R. E. Korf, Optimal rectangle packing: An absolute placement approach. *J. Artif. Int. Res.*, Vol. 46(1):47-87, 2013.

[7] S. Kaasalainen, A. Jaakkola, M. Kaasalainen, A. Krooks and A. Kukko, Analysis of incidence angle and distance effects on terrestrial laser scanner intensity: Search for correction methods. *Remote Sensing* Vol. 3(10):2207-2221, 2011.

[8] K. Kovács and B. G.-Tóth, Rectangle covering. In *AIP Conference Proceedings*, Vol. 2070, p. 020036, 2019. 10.1063/1.5090003.

[9] A. Manakov, H.P. Seidel, and I. Ihrke, A Mathematical Model and Calibration Procedure for Galvanometric Laser Scanning Systems. In *P. Eisert, J. Hornegger, and K. Polthier (Eds.), Vision, Modeling & Visualization*, pp. 207-214, 2011.

[10] S. A. Pettersen, D. Johansen, H. Johansen, V. Berg-Johansen, V. R. Gaddam, A. Mortensen, R. Langseth, C. Griwodz, H. K. Stensland, and P. Halvorsen, Soccer video and player position dataset. In *Proceedings of the 5th ACM Multimedia Systems Conference. ACM*, 2014.

[11] X. Chen, C. Li, Y. Wang and K. Yuan, *An Algorithm for Correction of Distortion of Laser Marking Systems*, 2007 IEEE International Conference on Control and Automation, Guangzhou, China, 2007, pp. 487-491, doi: 10.1109/ICCA.2007.4376404.

Markov chain methods in health care

Dr. András Zempléni[1], **Ayman Hijazy**[2] and **Balázs Dobi**[3]

Eötvös Loránd University, Department of Probability Theory and Statistics, Budapest, Hungary.
University of Debrecen, Faculty of Informatics, Debrecen, Hungary.
[1] email: andras.zempleni@ttk.elte.hu
[2] email: aymanhijazy@caesar.elte.hu
[3] email: dobibalazs@inf.elte.hu

1 Executive Summary

The first problem we tackled was to model chronic disease progression using a three state process: a disease free state S_f, a preclinical state S_p where the disease is still asymptomatic and a clinical state S_c where symptoms are finally exhibited. Estimating the parameters governing this process allows better planning of inspection (screening) programs and allows the correction of lead time bias, which is the apparent increase in survival due to the early detection of the disease. The goal was to incorporate a measure of the degradation (sickness) at diagnosis into the model. For this purpose, we assumed that the disease can be described by a gamma deterioration process, determined the distribution of the deterioration at detection and built a model accordingly. We then used the model to estimate the parameters governing the progression process in a breast cancer setup and to find the cost optimal inspection strategy.

The second challenge was to use generalized control charts so that they are suitable for health care applications. The focus was on the complete modelling of the disease progression-treatment cycle and its relationship with costs emerging during the process. To be able to successfully model this process we needed to introduce random repair and random sampling times, which are usually not present in traditional models. We also investigated the effect of different shift size distributions on the optimal time between samplings, critical value and the resulting cost expectation and standard deviation. Namely, we analysed the results of optimisations for the exponential distribution and exponential-geometric mixture distributions.

2 Challenge overview

In the first part, we have two challenges, the first is to deal with left censoring survival analysis data (correct lead time bias), which arises as a result of early detection. The second challenge is to determine the optimal inspection policy within a disease progression framework.

In the second part, we aimed to create a cost-efficient control chart framework which was specifically designed with use on health care data in mind. Specifically, we generalised control charts for the purposes of analysing and controlling a health

care characteristic of a patient over time, such as the blood glucose or lipid levels. Traditionally, minimal monitoring and process cost is achieved by finding the optimal parameters, namely the sample size, time between samplings, and critical value. Due to the monitoring of one patient at a time, we are left with two free parameters (time between samplings, and critical value).

3 Implementation of the project

We worked in close cooperation with our industrial partner (Healthware Consulting Ltd.) They provided the data, but even more importantly they scrutinized our research and gave many important advices about the practical aspects of the problem. Their contribution was especially important in the analysis of the diabetes treatments. The team has shared the workload, A. Zempléni was the PI, who proposed the methods and suggested new approaches. The young researchers shared the subjects between themselves: A. Hijazy worked on the disease progression topic, while B. Dobi was responsible for the control chart approach. We consulted regularly about the progress and exchanged ideas, e.g. about the implementation issues.

4 The research problem, methods, and results

4.1 Part I: Disease progression setup

4.1.1 The research problem

Classic chronic diseases progression models [1, 2] are built by assuming that the time spent in the disease free and the asymptomatic states are random variables following specified distributions. However, as classical approaches have shown to be sensitive to the chosen distributions and the underlying assumptions, we propose a new approach in which we model disease progression as a gamma degradation process with random starting point (onset) [5].

The model is built by assuming that the degradation is triggered at a random time and stays hidden until it is discovered either through inspection or by showing symptoms. In our model, we assume that symptoms appear after hitting a random critical threshold and that inspections may have a sensitivity less than one as well as a nonzero false positive rate.

Based on that, we derive the probabilities of cases getting detected by screens and minimize the distance between observed and calculated distributions to get estimates of the parameters of the gamma process, screening sensitivity, sojourn time. We investigate the properties of the proposed model by simulations.

After estimating the parameters, we derive the expected cost of repair and optimize the inspection rate for a cycle (which lasts from degradation-free to repaired state). This gives results for three cases: The first is for a finite observation period with no degradation recurrence, the second for infinite time horizon allowing recurrence. In the third case we derive an upper bound for the expected cost in a given constant time period. Finally, the model is applied to determine the optimal strategy for breast cancer screening with regard to the effects of different parametrizations [6].

4.1.2 The applied methods

Let us assume that a degradation (sickness) process is triggered by a random event occurring at a random time T_d. suppose that the level of degradation can be described by a non-homogeneous gamma process [3] denoted by $(X_t)_{t\geq 0}$. The considered process is assumed to have a constant scale parameter β and a time dependent shape $\eta(t)$, where $\eta(t)$ is a continuous, monotone, non-negative and increasing function for $t \geq 0$ with $\eta(0) = 0$.

Suppose that the deterioration can be detected in one of two ways: the deterioration shows symptoms and is detected when it reaches a critical threshold C or it is detected by means of preventive periodic inspections, which are arranged at times $(\tau_1, \tau_2, \dots)$ with an inter-inspection time Δ, i.e. $\tau_i = \tau_1 + (i-1)\Delta$. Suppose that the probability of detection depends on the level of degradation i.e. the sensitivity of inspection is a function of the degradation $\Lambda(x)$. Furthermore, suppose that the inspections have a false positive rate R_{FP}, which is the probability of the inspection falsely detecting a deterioration given that there is no actual deterioration.

Under these assumptions the contribution of cases moving to S_p between (τ_{j-1}, τ_j) to the probability of the detected level of degradation at inspection τ_i to be between $(x, x + dx)$ denoted by $f_{(i,j)}$ is calculated in [5].

Consequently, f_{τ_i} is given by the sum of these contributions and the probability of being detected at inspection τ_i can be derived. The probability of showing symptoms between τ_{i-1} and τ_i is determined in a similar manner, while taking the overshooting property of the gamma process into account. The parameters are then estimated using minimum divergence estimators [4], which minimize the distance between the observed and the expected distributions.

After estimating the parameters, the total expected cost for a single cycle, which is a function of the time of the first inspection τ_1 and the inter-inspection time Δ can also be determined using the probabilities of detection. Namely:

$$\mathrm{E}\left(TC^{(1)}(\tau_1, \Delta)\right) = \sum_{i=1}^{K+1} \mathrm{E}(C_E(\tau_i)) + \sum_{i=1}^{K} \Big[\mathrm{E}(C_D(\tau_i)) + \mathrm{E}(C_{FP}(\tau_i)) + \mathrm{E}(C_{ins}(\tau_i))\Big],$$

where C_E is the cost of repair of cases showing symptoms, C_D is the cost of repair of cases detected on inspections, C_{FP} is the cost of false positives and C_{ins} is the inspection cost. The total expected cost is minimized in order to find the optimal first inspection time $\hat{\tau}_1$ and the optimal inter-inspection time $\hat{\Delta}$.

4.1.3 The achieved research results

Although the proposed model is somewhat computationally expensive, it proves to be an accurate and a powerful tool to use with degradation processes triggered at a random onset. The estimators are consistent and the parameters under this setup are identifiable.

Besides, our approach gave feasible models for cost-optimised inspection programs for the described degradation phenomena. It allowed us to study the effect of each of the process governing parameters on the expected cost.

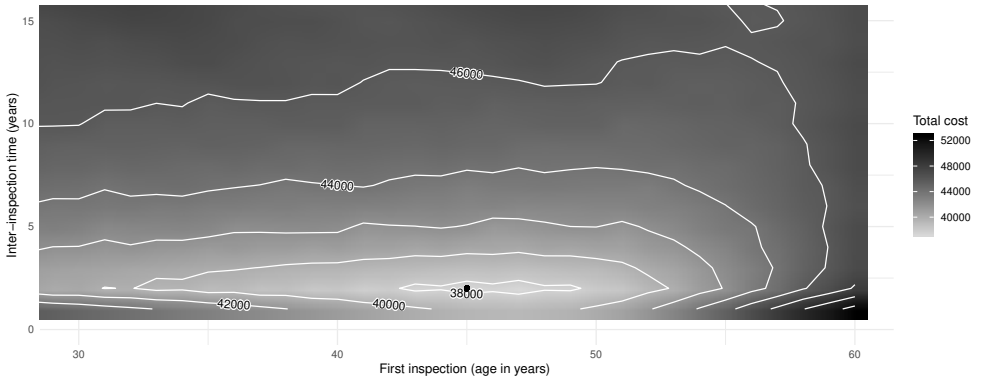


Figure 1: Contour plot of the expected total costs as a function of the first inspection time and inter-inspection time in a breast cancer scenario, the optimum is achieved at the black dot

4.2 Part II: Control charts

Control charts have traditionally been used in industrial statistics but are constantly seeing new areas of application. The research introduced a new method, which is suitable for applications in the health care sector, especially for monitoring the health characteristic of a patient. We adopted a Markov chain-based approach and developed a method in which not only the shift size (i.e. the degradation of the patient's health) can be random, but the effect of the repair (i.e. treatment) and time between samplings (i.e. visits) too. This meant that we did not use many often-present assumptions that are usually not applicable for medical treatments. The average cost of the protocol, which is determined by the time between samplings and the control limit, can be estimated using the stationary distribution of the Markov chain.

4.2.1 The research problem

Cost-efficient control charts have not been widely used in health care settings, but there are some examples which deal with cost monitoring and management. In one work, $\bar{X}$ (mean-monitoring) and R (variance-monitoring) charts were used to assess the effect of a physician educational program on the hospital resource consumption [7]. Another article is about a hospital which used control charts for monitoring day shift, night shift, and daily total staffing expense variance, among other variables [8]. Stewart and Greisler documented a case study about primary care practice performance, where control charts were used to monitor costs associated with provider productivity and practice efficiency. Further costs monitored were net patient revenue per relative value unit, and provider and nonprovider cost as a percent of net revenue [9]. Even though these studies used control charts for cost monitoring or optimisation purposes, they did not deal with the same problems as this research, as our method focuses on cost-optimisation by finding the optimal parameters of the control chart setup.

4.2.2 The applied methods

Consider a process which is monitored by a control chart. We will show methods for processes which can only shift to one (positive) direction, monitored by a simple X-chart, with sample size $n = 1$. This aims to model the monitoring of a characteristic where the shift to only one direction is possible, and the monitoring is focused at one patient at a time. Several assumptions are made for the base model. The process distribution is normal with known parameters μ_0 and σ. We will denote its cumulative distribution function by ϕ. The shift intensity $1/s$ is constant and known, and the shift size δ^* is fixed. It is also assumed that the process does not repair itself, but when repair is carried out by intervention, it is perfect. The repair is treated as an instantaneous event. All costs related to repairing should be included in the repair cost, for example, if a major repair entails higher cost, then this should also be reflected in the calculation. The time between shifts is assumed to be exponentially distributed. The above assumptions ensure that the probabilities of future transitions are only dependent on the current state. This is the so-called Markov property, and the model under consideration is called a Markov chain. The states of this Markov chain are defined at the sampling times and the type of the state depends on the measured value and the actual (unobservable) background process, namely whether there was a shift from the target value in the parameter. This way, four basic state types are defined as follows:

- No shift - no alarm: in-control (INC)
- Shift - no alarm: out-of-control (OOC)
- No shift - alarm: false alarm (FA)
- Shift - alarm: true alarm (TA)

4.2.3 The achieved research results

The previous simple framework can be used for more general designs. Zempléni et al used this method to set up economically optimal control charts where shifts are assumed to have a random size and only the distribution of the shift size is known [10]. This means that the shift size is no longer assumed to be fixed and known, which is important in modelling various processes not just in health care, but in industrial or engineering settings too. This requires expanding the two shifted states to account for different shift sizes. We expanded upon this model by implementing and comparing multiple shift size distributions, namely exponential and exponential-geometric.

The framework was further generalised by not assuming perfect repair after a true alarm signal. This meant that the treatment will not have perfect results on the health of the patient. In this case, the imperfectly repaired states act as out-of-control states. It is assumed that the repair cannot worsen the state of the process, but an imperfectly repaired process will still cost the same as an equally shifted out-of-control process, thus repaired and out-of-control states do not need distinction during the cost calculation. Different approaches were considered for modelling the repair size distribution. The one applied uses a random variable to determine the

distance from the target value after repair. A natural choice for the distribution of this random variable was the beta distribution since it has support over $[0, 1]$ - the portion of the distance compared with the current one after repair. Also, the flexible shape of its density function can model many different repair processes. Because of these considerations, we assumed that the proportion of the remaining distance from μ_0 after repair, R, which is a $Beta(\alpha, \beta)$ random variable, with known parameters.

Yet another generalisation was the random sampling time. In certain environments, the occurrence of the sampling at the prescribed time is not always guaranteed. For example, in health care, the patient or employee compliance can have significant effect on the monitoring, thus it was important to take this into account during the modelling too. Here, it was modelled in a way that the sampling is not guaranteed to take place - e.g., the patient may not show up for control visit. This means that the sampling can only occur according to the sampling intervals, for example, at every nth days, but is not a guaranteed event.

The transition probabilities can be written using the process variance distribution, the shift size distribution, the repair size distribution, and the sampling probability. This transition matrix then can be used to calculate the stationary distribution of the Markov chain, which is essentially the process distribution. By assigning costs to each probability, one can calculate the different cost moments, thus the mean and standard deviation too.

The results were published in two papers in international journals and implemented as the `Markovchart` **R** package [11, 12, 13].

5 Solution of the industrial problem and its benefits

In the disease progression framework, we were able to establish consistent estimators for the parameters governing the process using the gamma process based model. This allowed us correct lead time bias and derive cost optimal strategies. We built the necessary functions in **R** and created the simulators for chronic disease progression. The generalised control chart framework allows researchers to monitor and optimise complex health care processes and costs. The Healthware Consulting Ltd. used this model to analyse data of Hungarian diabetic patients. This analysis was incorporated into a scientific report that was sent to a client. Furthermore, the `Markovchart` package is freely downloadable in **R**, thus making the model available worldwide for academia and industry alike.

6 Conclusions

In the disease progression framework, the gamma process based model proved to be an accurate and a powerful tool to estimate the parameters governing disease progression. It can also be used in industrial setups where the degradation is triggered at a random onset. The model is very flexible, as one is free to shape of the degradation η and can include covariates. That being said, our next aim is to study the reasons behind the discrepancies in estimates when using the classical convolution based model [2], we aim to pursue that from a theoretical and a practical points of view.

Building on and expanding the work of Zempléni et al. [10] we developed three types of generalisations: the random shift size, imperfect repair, and non-compliance. The random shift size means that only the distribution of the shift size and its parameters were assumed to be known. This let us monitor processes that are potentially drifting in nature. The second generalisation - the imperfect repair - assumed that the process stays out of control even after repair, but on a level closer to the target value than before. This type of repair response is often observed in treatments in health care, but may be found in other areas too. The third generalisation was intended to help the modelling of patient or staff non-compliance. We implemented this concept in a way that allows sampling times to be skipped by a probability governed by a distribution or function with known parameters. The mathematical results were implemented as the `Markovchart` **R** package and were applied by our industrial partner on the data of diabetic patients. We are planning on substantially improving and expanding the **R** package.

References

[1] Zelen, M. and Feinleib, M. (1969), On the theory of screening for chronic diseases. *Biometrika*, **56(3)**, 601-614.

[2] Wu, D., Rosner, G. and Broemeling, L. (2005), MLE and Bayesian inference of age-dependent sensitivity and transition probability in periodic screening. *Biometrics*, **61(4)**: 1056-1063.

[3] Abdel-Hameed, M. (1975), A gamma wear process. *IEEE Transactions on Reliability*, **24** (2): 152-153.

[4] Burbea, J. and Rao, C.(1982), On the convexity of some divergence measures based on entropy functions. *IEEE Transactions on Information Theory*, **28**(3): 489-495.

[5] Hijazy, A. and Zempléni, A. (2021), Gamma Process-Based Models for Disease Progression. *Methodol Comput Appl Probab*, **23**: 241-255. https://doi.org/10.1007/s11009-020-09771-4

[6] Hijazy, A. and Zempléni, A. (2020), Optimal inspection for randomly triggered hidden deterioration processes. *Quality and Reliability Engineering International* **36**: 1-16. https://doi.org/10.1002/qre.2707

[7] Johnson, CC. and Martin, M. (1996), Effectiveness of a physician education program in reducing consumption of hospital resources in elective total hip replacement. *South Med J.*, **89**(3): 282-289. PMID: 8604457.

[8] Shaha SH. (1995), Acuity systems and control charting. *Qual Manag Health Care*, **3**(3): 22-30. PMID: 10143553.

[9] Stewart LJ. and Greisler D. (2002), Measuring primary care practice performance within an integrated delivery system: a case study. *J Health Manag.*, **47**(4): 250–261. PMID: 12221746.

[10] Zempléni, A., Véber, M., Duarte, B. and Saraiva, P. (2004), Control charts: a cost-optimization approach for processes with random shifts. *Appl Stoch Model Bus Ind.* **20**(3): 185-200. https://doi.org/10.1002/asmb.521

[11] Dobi, B. and Zempléni, A. (2019a). Markov Chain-Based Cost-Optimal Control Charts for Health Care Data. *Quality and Reliability Engineering International*, **35**(5): 1379–1395. doi: 10.1002/qre.2518.

[12] Dobi, B. and Zempléni, A. (2019b), Markov Chain-Based Cost-Optimal Control Charts with Different Shift Size Distributions. *Annales Universitatis Scientiarum Budapestinensis de Rolando Eötvös Nominatae, Sectio Computatorica*, **49**: 129-146.

[13] Dobi, B. and Zempléni, A. (2020), Markovchart: Markov Chain-Based Cost-Optimal Control Charts, version 1.1.1, *The Comprehensive R Archive Network*

Development of a comprehensive pharmacodynamics model enabling the investigation of equations suitable for practical use (nonlinear regression)

Judit Zsuga[1], **Tamas Erdei**[2] and **Rudolf Gesztelyi**[3]

[1]University of Debrecen, Department of Health Systems Management and Quality Management for Health Care,
[2]University of Debrecen, Department of Pharmacology and Pharmacotherapy,
[3]University of Debrecen, Department of Pharmacology and Pharmacotherapy

[1]E-mail: zsuga.judit@med.unideb.hu
[2]E-mail: erdei.tamas@pharm.unideb.hu
[3]E-mail: gesztelyi.rudolf@pharm.unideb.hu

1 Executive Summary

Receptor reserve is a measure of the response-inducing ability of an agonist acting on a receptor (and the downstream signaling). Thus, receptor reserve is an important tool for drug development. However, the determination of receptor reserve is challenging for agonists with short half-lives, such as adenosine. Although adenosine metabolism can be inhibited several ways (to prevent the rapid elimination of adenosine administered to construct concentration-effect (E/c) curves for the determination), the consequent accumulation of endogenous adenosine biases the results. To address this problem, we previously developed a method, by means of which this bias can be mathematically corrected. Our goals were to validate our new method to assess receptor reserve, then (starting from the first results of this study) to understand the interaction between FSCPX and NBTI (two chemicals used for our experiments). Our results provided an *in silico* validation for determining receptor reserve for adenosine, an agonist with short half-life, by means of our recently developed method. In addition, our results suggest that FSCPX and NBTI interact in a way that FSCPX blunts the action of NBTI on the interstitial endogenous adenosine (a finding that, to the best of our knowledge, has no antecedents in the literature).

2 Challenge overview

The term receptor reserve is an integrative measure of the response-inducing ability of the interaction between an agonist and a receptor system (consists of a receptor and its signaling). Although some recent models of agonist action do not use this term, the underlying phenomenon, i.e., the stimulation of a submaximal fraction of receptors apparently elicits the maximal effect (in certain cases), is well-known a generally accepted. Thus, determination of receptor reserve is based on this phenomenon, during the course of which a fraction of receptors is irreversibly inactivated in a way that the remaining fraction retains its functional integrity. The greater the receptor reserve, the greater the fraction of receptors need to be inactivated in order to achieve

detectable diminution of maximal response is. So, the measure of receptor reserve is the "resistance" (or "inertia") of the receptor system, activated with a given agonist, against an intervention that reduces the number of operable receptors, in terms of the achievable maximal response. The receptor reserve depends on three factors of the concentration-effect (E/c) relationship: agonist (exhibiting efficacy), receptor system (possessing different functional states for the receptor and a tissue-dependent signal amplification machinery), and effect (indicative of the particular signaling pathways activated).

The physiological significance of the phenomenon is that a great receptor reserve allows an endogenous agonist to produce rapid activation of its receptor, or more generally, to produce fast activation of a particular receptor system. Information about receptor reserve has high utility when predicting the behavior of an agonist in a tissue. If receptor reserve in a tissue is small, only high-efficacy agonists can evoke a significant effect (acting often as a full agonist), whereas low-efficacy agonists cannot elicit an effect (or at most they behave as a partial agonist). In turn, if receptor reserve is great, even low-efficacy agonists are able to generate a significant effect (moreover, sometimes they may act as a full agonist). Thus, application of low-efficacy agonists can ensure tissue selectivity in a sense that their effect will only be significant in tissues possessing great receptor reserve. Due to the multifactorial origin of receptor reserve, it is worthwhile assessing it for every agonist, receptor system (tissue) and effect having pathophysiological importance, especially in the point of view of drug development.

However, determining receptor reserve is challenging for rapidly metabolizing agonists, such as adenosine. Adenosine is a ubiquitous molecule of the purine metabolism, and therefore a substrate for several enzymes and carriers. All quantitative methods used to determine receptor reserve need accurate E/c data, the acquisition of which, by means of E/c curve construction, mandates the ability to reach the steady state with regard to the agonist levels at the receptors, a premise that is difficult to fulfill for an agonist with a short half-life. Moreover, reliable adenosine levels are hard to compute or measure, due to the quick adenosine turnover in most tissues including the myocardium. Although rapid elimination of adenosine, administered to construct E/c curves for the determination, can be inhibited, the consequent accumulation of endogenous adenosine biases the result. To address this problem, we have proposed a procedure, by means of which this bias can be mathematically corrected and then the receptor reserve can be qualitatively assessed [1]. This method is closely related to the receptorial responsiveness method (RRM), which has been developed to quantify agonist concentrations in the microenvironment of the receptors (under certain circumstances) [2-5].

In our own method [1], concentration-effect (E/c) curves are generated with adenosine and CPA, a stable adenosine analogue. Exogenous adenosine (administered during the E/c curve) is protected by NBTI, which blocks the adenosine transport into the cells. As NBTI also prevents the physiological elimination of endogenous adenosine, it produces an extracellular accumulation of endogenous adenosine that, being present already at the beginning of E/c curve construction, biases the E/c curves. Therefore, the biased E/c curves are corrected by means of RRM, which estimates the level of extracellularly accumulated endogenous adenosine. The corrected NBTI-

treated adenosine E/c curves (one with and another one without a pretreatment with FSCPX, an irreversible A_1 adenosine receptor antagonist) can be reliably evaluated, i.e. their maximal response to adenosine can be compared. Utilization of our new method may have practical significance [1,6].

3 Implementation of the project

The academic partner provided the scientific background including laboratory equipment, chemicals and experimental animals, furthermore skills and experiences needed to work with. The industrial partner drew attention to the impact of receptor reserve regarding drug development.

4 The research problem, methods, and results

4.1 The research problem

The determination of receptor reserve is challenging for agonists with short half-lives, such as adenosine. Although adenosine metabolism can be inhibited several ways, the consequent accumulation of endogenous adenosine makes the results unreliable. To address this problem, we previously developed a method, validation of which was our original goal. Then, a possible interaction emerged, i.e. between FSCPX (an irreversible A_1 adenosine receptor antagonist) and NBTI (a nucleoside transport inhibitor), which chemicals exhibited an unexpected (apparently paradoxical) effect pattern when added together to the *ex vivo* guinea pig heart (specifically: left atrium). In the presence of NBTI, FSCPX behaves as an A_1 adenosine receptor enhancer.

4.2 The applied methods

Two computer simulations were carried out [7,8] through *in silico* reconstruction of selected *ex vivo* E/c curves describing the direct negative inotropic response to adenosine and CPA, two A_1 adenosine receptor agonists, in isolated and paced guinea pig left atria [6,9]. As receptor function model, the operational model of agonism was chosen to link the agonist concentration to the effect. Values for concentrations and parameters of the operational model were received from biological measurements or rendered arbitrarily. To transform the *in silico* E/c curves, the known mechanisms of action of chemicals used for the original *ex vivo* experiments were translated into algebraic operations, furthermore, where appropriate, the receptorial responsiveness method (RRM) was applied. The aim of simulation was to position the *in silico* E/c curves to overlap the corresponding *ex vivo* E/c curves as perfect as possible, and, as long as possible, in line with our biological knowledge. When an algebraic operation contradicting this biological knowledge had to be introduced to reach the appropriate overlap, this was followed by making a new biological assumption that could underly the given algebraic operation.

4.3 The achieved research results

Results of our first *in silico* investigation [7] validated our method as a useful tool for the qualitative assessment of receptor reserve for adenosine, an endogenous agonist for which no reliable receptor reserve value could be determined previously. Hence, this method may contribute to the accumulation of useful data regarding receptor reserve for other endogenous agonists with short half-lives (resulting from a high exposure to the function of enzymes and transporters). The potential advantage of introducing this new method into practice comes from the significance and specific nature of receptor reserve. Efficiency of different agonists may become predictable in different tissues based on the relevant receptor reserve values, a knowledge that, if utilized for rational drug development, may help diminish untoward effects. Consistently, low-efficacy A_1 receptor agonists, proved to have a significant effect only in tissues with great A_1 receptor reserve, have been being developed in numerous indications. However, as some drugs influence the level and/or distribution of endogenous agonists, receptor reserve data related to these agonists also have therapeutical significance that indicates the raison d'etre of our method to assess receptor reserve.

In addition, findings of our first *in silico* investigation imply that FSCPX treatment weakens the inhibitory action of NBTI on the transport (and thereby elimination) of adenosine. Consequently, the adenosine E/c curve generated under NBTI and FSCPX co-treatment, after a correction with our method in the previously published manner [1,6], probably overestimates the effect of adenosine at small and medium concentrations. It is important to emphasize that this deviation affects, in a significant manner, only the starting and medium parts of this corrected E/c curve, thus it does not influence the assessment of receptor reserve for adenosine that is based on the evaluation of final parts of the corrected E/c curves. Starting from this finding, new *ex vivo* [9] and *in silico* [8] investigations were planned and carried out.

In our second *in silico* study [8], we found that E_m parameter of the operational model of agonism can influence the behavior of our E/c curves. Furthermore, we have gained *in silico* evidence for an interference between effects of FSCPX and NBTI in our *ex vivo* experimental setting used in the study yielded the E/c curves to be simulated [9]. This finding extends beyond the well-established A_1 adenosine receptor antagonist property of FSCPX, indicating an inhibitory action exerted by FSCPX on the interstitial adenosine accumulation produced by NBTI, a selective and potent blocker of the nucleoside transporter type ENT1. Regarding the mechanism of this interference, *in silico* evidence has been obtained supporting that FSCPX only inhibits the interstitial accumulation of endogenous (but not exogenous) adenosine. As an additional result, we have found that NBTI seems not to completely inhibit the inward adenosine flux in the guinea pig atrium.

5 Solution of the industrial problem and its benefits

Pharmacons acting in several tissues may possess a variety of side effects, just like adenosine receptor agonists. Moreover, numerous drugs are in use and in the pipeline that modify the level of interstitial adenosine [1,6,9]. Accordingly, receptor reserve values related to all significant adenosine-evoked effects are of interest, thus a proce-

dure, which enables acquisition of such estimates, may be a useful tool for a company interested in drug development.

6 Conclusions

Our first study [7] has provided an *in silico* validation for determining receptor reserve for adenosine, an endogenous agonist with rapid metabolism, by means of our qualitative method. Furthermore, this work pointed to the importance of *in silico* justification to disclose or exclude unforeseen interactions between treatment schedules (a problem that can occur in every study), and, in general, to shine a light on previous results. In addition, these results suggest that FSCPX and NBTI may interact when NBTI is administered after a pretreatment with FSCPX. To the best of our knowledge, this observation has no antecedents in the literature.

Our second *in silico* study dealt with the background of a paradoxical phenomenon, first described in [1], sc. the irreversible A_1 adenosine receptor antagonist FSCPX apparently increases the maximal response to adenosine in the presence of NBTI. This work used an improved approach for computer simulation by addressing the issue of E_m parameter of the operational model. Taking outcomes of this *in silico* study together with our previous *in silico* and *ex vivo* results, we have concluded that the above-mentioned paradoxical phenomenon can be ascribed to an interference between effects of two adenosine analogues, FSCPX and NBTI, in our experimental setting that was first suggested in [7] and was first evidenced (*ex vivo*) in [9]. Herein, we have provided *in silico* evidence for this interference, proposing that FSCPX, in addition to antagonizing the A_1 adenosine receptor, blunts the interstitial accumulation of endogenous (but not exogenous) adenosine produced by NBTI.

References

[1] Kiss Z, Pak K, Zsuga J, Juhasz B, Varga B, Szentmiklosi AJ, Haines DD, Tosaki A, Gesztelyi R., The guinea pig atrial A_1 adenosine receptor reserve for the direct negative inotropic effect of adenosine. Gen Physiol Biophys, 32, 325-35, 2013. (doi: 10.4149/gpb_2013041)

[2] Gesztelyi R, Zsuga J, Juhász B, Dér P, Vecsernyés M, Szentmiklósi AJ., Concentration estimation via curve fitting: quantification of negative inotropic agents by using a simple mathematical method in guinea pig atria. Bull Math Biol, 66, 1439-53, 2004. (doi: 10.1016/j.bulm.2004.03.001)

[3] Grenczer M, Zsuga J, Majoros L, Pinter A, Kemeny-Beke A, Juhasz B, Tosaki A, Gesztelyi R., Effect of asymmetry of concentration-response curves on the results obtained by the receptorial responsiveness method (RRM): an in silico study. Can J Physiol Pharmacol, 88, 1074-83, 2010. (doi: 10.1139/y10-089)

[4] Grenczer M, Pinter A, Zsuga J, Kemeny-Beke A, Juhasz B, Szodoray P, Tosaki A, Gesztelyi R., The influence of affinity, efficacy, and slope factor on the estimates obtained by the receptorial responsiveness method (RRM): a computer simulation study. Can J Physiol Pharmacol, 88, 1061-73, 2010. (doi: 10.1139/y10-078)

[5] Pák K, Kiss Z, Erdei T, Képes Z, Gesztelyi R., Új lehetőség farmakológiai agonisták receptorközeli koncentrációjának becslésére: a receptoriális válaszkészség módszer (RRM) [The receptorial responsiveness method (RRM): a new possibility to estimate the concentration of pharmacologic agonists at their receptors]. Acta Pharm Hung, 84, 38-52, 2014. (PMID: 24809165)

[6] Pak K, Papp C, Galajda Z, Szerafin T, Varga B, Juhasz B, Haines D, Szentmiklosi AJ, Tosaki A, Gesztelyi R., Approximation of A_1 adenosine receptor reserve appertaining to the direct negative inotropic effect of adenosine in hyperthyroid guinea pig left atria. Gen Physiol Biophys, 33, 177-88, 2014. (doi: 10.4149/gpb_2013079)

[7] Zsuga J, Erdei T, Szabó K, Lampe N, Papp C, Pinter A, Szentmiklosi AJ, Juhasz B, Szilvássy Z, Gesztelyi R., Methodical challenges and a possible resolution in the assessment of receptor reserve for adenosine, an agonist with short half-life. Molecules, 22, 839, 2017. (doi: 10.3390/molecules22050839)

[8] Szabo AM, Erdei T, Viczjan G, Kiss R, Zsuga J, Papp C, Pinter A, Juhasz B, Szilvassy Z, Gesztelyi R., An advanced in silico modelling of the interaction between FSCPX, an irreversible A_1 adenosine receptor antagonist, and NBTI, a nucleoside transport inhibitor, in the guinea pig atrium. Molecules, 24, 2207, 2019. (doi: 10.3390/molecules24122207)

[9] Erdei T, Szabo AM, Lampe N, Szabo K, Kiss R, Zsuga J, Papp C, Pinter A, Szentmiklosi AJ, Szilvassy Z, Juhasz B, Gesztelyi R., FSCPX, a chemical widely used as an irreversible A? adenosine receptor antagonist, modifies the effect of NBTI, a nucleoside transport inhibitor, by reducing the interstitial adenosine level in the guinea pig atrium. Molecules, 23, 2186, 2018. (doi: 10.3390/molecules23092186)

Evaluation of usefulness of the receptorial responsiveness method (RRM), a procedure based on nonlinear regression of biological data

Judit Zsuga[1], **Tamas Erdei**[2], **Adrienn Monika Szabo**[3] and **Rudolf Gesztelyi**[4]

[1] University of Debrecen, Department of Health Systems Management and Quality Management for Health Care,
[234] University of Debrecen, Department of Pharmacology and Pharmacotherapy,

[1]E-mail: zsuga.judit@med.unideb.hu
[2]E-mail: erdei.tamas@pharm.unideb.hu
[3]E-mail: szabo.adrienn23@gmail.com
[4]E-mail: gesztelyi.rudolf@pharm.unideb.hu

1 Executive Summary

The receptorial responsiveness method (RRM) is a procedure that is based on a simple nonlinear regression that uses a model with two variables (X, Y) and (upon individual curve fitting) one parameter to be determined (c_x). The model of RRM describes the co-action of two agonists that consume the same response capacity (due to the use of the same postreceptorial signaling in a biological system). While using RRM, uniquely, an acute increase in the concentration of an agonist (near the receptors) can be quantified (as c_x), *via* evaluating E/c curves that were constructed with the same or another agonist in the same system. As this measurement is sensitive to the implementation of the curve fitting, the goal of our investigations was to test RRM by combining different ways and setting options, namely: individual vs. global fitting, ordinary vs. robust fitting, and three weighting options (no weighting vs. weighting by $1/Y^2$ vs. weighting by $1/SD^2$). During the testing, RRM was used to estimate the known concentrations of stable synthetic A_1 adenosine receptor agonists in isolated, paced guinea pig left atria. The estimates were then compared to the known agonist concentrations (to assess the accuracy of RRM); furthermore, the 95% confidence limits of the best-fit values were also considered (to evaluate the precision of RRM). It was found that, although the global fitting offered the most convenient way to perform RRM, the best estimates were provided by the individual fitting without any weighting, almost irrespective of the fact whether ordinary or robust fitting was chosen.

2 Challenge overview

The receptorial responsiveness method (RRM) combines the Hill equation, the first quantitative model of receptor function [1], with a simple relationship between simultaneous effects of two concentrations of one or two agonist(s), where one concentration is known and the other one is unknown (c_x). Under certain circumstances, RRM can be used to estimate an acute increase in the level of an agonist (as c_x) in the vicinity

of its receptors in a biological system [2]. Technically, it is of great impact to possess reliable input data and ensure the most appropriate implementation of the fitting of the model of RRM.

RRM is a procedure that is based on a simple nonlinear regression while using a model with two variables (X, Y) and (upon individual curve fitting) one parameter to be determined (c_x) [2]. The fitting of this model requires two sets of concentration-effect (E/c) curves to be generated. E/c curves that are suitable for RRM are XY graphs where X is the (logarithm of the) concentration of a pharmacological agonist, while Y is a response of a biological system that is evoked by the given agonist concentration (indicated by the corresponding X value) alone (first set of curves) or together with a single extra concentration (c_x) of the same or another agonist, which was administered to the system before the generation of the E/c curve (second set of curves). Quantifying this extra agonist concentration as a c_x value is the goal of RRM [3,4].

Although an E/c curve in the second set relates the resultant effect of the two concentrations (consuming the same response capacity of a biological system) solely to the concentration that is administered for the E/c curve (that is depicted on the x-axis), the model of RRM attributes this resultant effect to two concentrations, namely to two concentrations of the agonist that was used for the E/c curve. One of these concentrations is indicated by the x-axis, while the other one, c_x, is a concentration of the agonist used for the E/c curve that is equieffective with the agonist added in a single extra dose. If the two agonists in question are the same, c_x is a real concentration, and, if not, c_x is a surrogate parameter of the single extra concentration of the other agonist [5,6]. The application of RRM might be useful when the concentration of this extra agonist is unknown and difficult to determine in any other manners [3,4]. As the single extra agonist concentration distorts (biases) the E/c curve in the second set as compared to the corresponding E/c curve in the first set (generated the same way except for the administration of the single extra agonist concentration), it will be referred to as "biasing" concentration.

3 Implementation of the project

The academic partner provided the scientific background including laboratory equipment, chemicals and experimental animals, furthermore skills and experiences needed to work with. The industrial partner drew attention to the impact of drug development in the field of cardiovascular diseases.

4 The research problem, methods, and results

4.1 The research problem

Regression analysis (especially its nonlinear form) is one of the most common ways to analyze E/c curves that might provide several useful pieces of information, e.g., about properties of receptors, receptor ligands and cell functions, which would be otherwise difficult (if possible) to gain [7-9]. The goal of regression (curve fitting) is to find

the best-fit values for the parameters of the model that is used for the regression, and thereby to create the best-fit regression function (curve), which is closest to the data points (XY data obtained usually from repeated measurements). Minimizing the sum of the squares of the vertical distances of data points (replicate Y values related to the same X value) from the curve (the Y value of the curve that relates to the corresponding X value) is the earliest and most common technique to find best-fit values. Accordingly, this procedure is called the least-squares method. The use of this method is based on two major assumptions: (i) normality, i.e., the scatter of the data points (related to the same X value) around the curve (the Y value on the curve related to the corresponding X value) follows a Gaussian distribution; and, (ii) homoscedasticity, i.e., the extent of this scatter is the same for all values of X [7-9].

Ad (i): Assuming a Gaussian distribution for the scatter of data (and performing an ordinary regression) is useful for most cases. Nevertheless, assuming a Lorentzian distribution (and carrying out robust regression) makes the curve fitting more robust to outliers, although it hinders the calculation of data characterizing the reliability of the best-fit values and the curve (e.g., standard errors, confidence intervals, confidence, and prediction bands). In some cases, it is worth considering a Poisson distribution (and performing Poisson regression), but never for normalized Y data (used in the present study as well). While the Gaussian and Lorentzian distributions (being t distributions with infinity and 1 as degree(s) of freedom, respectively) allow for the use of the least-squares method, the Poisson distribution requires an alternative way for finding best-fit values (the so-called maximum likelihood-based parameter estimation) [7-9].

Ad (ii): To counteract heteroscedasticity, the standard equation of the least-squares method can be transformed ("weighted") by a factor, a procedure that is called weighting. The most common factors are $1/Y^2$ (relative weighting) and $1/|Y|$ (Poisson weighting). The relative weighting and, to a lesser extent, the Poisson weighting reduce the influence of the higher Y values on the best-fit values and the regression curve. In addition, the Poisson weighting (performed with ordinary regression) can serve as an inferior alternative of the Poisson regression. A rarely used, although theoretically meaningful, choice is weighting by $1/SD^2$ (the inverse of the variance of Y values related to the same X value), which is expected to reduce the undue impact of Y replicates with bigger scatter [7-9].

The simplest and most common way of curve fitting is the individual regression, i.e., to find best-fit values for a single data set, e.g., a single E/c curve (or a set of E/c curves resulted from repeated measurements). An advanced way of curve fitting is when the model of regression defines a family of curves, i.e., some parameters (at least one) of the model to be fitted are (is) shared among several (at least two) data sets, called global regression. In the case of global regression, one sum of squares is computed for all of the Y replicates of all data sets [8,9].

The goal of this investigation was to explore the influence that different curve fitting ways and curve fitting settings might exert on the outcome of RRM.

4.2 The applied methods

The experiments consisted of the construction of two E/c curves. For the first E/c curve, adenosine was administered to assess the naive response of the atria to A_1 adenosine receptor stimulation. For the second E/c curve, one of three widespread, relatively stable, synthetic A_1 adenosine receptor agonists (CPA, NECA, and CHA) was used, in the absence or presence of a "biasing" concentration of the same agonist. The accuracy and precision of RRM was investigated via assessing this known "biasing" concentration in a well-established isolated and paced guinea pig left atrium model. Accordingly, c_x, estimate yielded by RRM, has been expected to directly provide the "biasing" concentration. Ways and setting options of the implementation RRM were: individual vs. global fitting, ordinary vs. robust fitting, and three weighting options (no weighting vs. weighting by $1/Y^2$ vs. weighting by $1/SD^2$).

4.3 The achieved research results

The major finding of this investigation is that the best estimates of RRM can be obtained via individual fitting without any weighting, almost irrespectively of the fact of whether ordinary (assuming Gaussian distribution) or robust (assuming Lorentzian distribution) regression is chosen. Thus, RRM is a relatively (by comparison to the possibilities it offers) easy-to-use procedure that requires neither a heavy-duty curve fitting software nor a high level of knowledge concerning regression analysis.

5 Solution of the industrial problem and its benefits

When using RRM in an appropriate way and with optimal settings, it can provide information regarding an acute increase in the concentration of an agonist in the microenvironment of its receptor. Importantly, the microenvironment of receptors is a tissue compartment that is difficult to access, especially in a moving organ [2-4]. Theoretically, RRM can be applied for each receptor; however, the A_1 adenosine receptor is uniquely suitable for this method, due to its slow and incomplete desensitization in the presence of even a full agonist [10,11].

6 Conclusions

The results of the present study show that the well-established observation, i.e., assumption of both normality and homoscedasticity is a useful (and first choice) approach when analyzing biological data, is valid for the assessment with RRM as well. Consequently, ordinary regression without any weighting is the appropriate decision, when performing RRM. Regarding only the accuracy of the fitting, robust regression can also be chosen (moreover, somewhat more accurate estimates may be obtained), but the lack of 95% confidence intervals deprives the possibility of considering the precision (and thus reliability) of the curve fitting.

However, based on the present results, the conventional individual regression is a more accurate and precise (although less comfortable) way to carry out RRM than

the global regression. This might be associated with the fact that the "intact" one of the two corresponding E/c curves has no definite best-fit value ($\log c_x$), because the expected estimate (c_x) is zero, the logarithm of which is not defined. This problem does not occur when RRM is done with individual regression, because the "intact" E/c curve is fitted to the Hill equation that can yield reliable best-fit values. Thus, the individual fitting is the appropriate choice for RRM.

References

[1] Gesztelyi R, Zsuga J, Kemeny-Beke A, Varga B, Juhasz B, Tosaki A., The Hill equation and the origin of quantitative pharmacology. Arch Hist Exact Sci, 66, 427-38, 2012. (doi: 10.1007/s00407-012-0098-5)

[2] Gesztelyi R, Zsuga J, Juhász B, Dér P, Vecsernyés M, Szentmiklósi AJ., Concentration estimation via curve fitting: quantification of negative inotropic agents by using a simple mathematical method in guinea pig atria. Bull Math Biol, 66, 1439-53, 2004. (doi: 10.1016/j.bulm.2004.03.001)

[3] Karsai D, Zsuga J, Juhász B, Dér P, Szentmiklósi AJ, Tósaki A, Gesztelyi R., Effect of nucleoside transport blockade on the interstitial adenosine level characterized by a novel method in guinea pig atria. J Cardiovasc Pharmacol, 47, 103-9, 2006. (doi:10.1097/01.fjc.0000196239.51018.a0)

[4] Karsai D, Gesztelyi R, Zsuga J, Jakab A, Szendrei L, Juhasz B, Bak I, Szabo G, Lekli I, Vecsernyes M, Varga E, Szentmiklosi AJ, Tosaki A., Influence of hyperthyroidism on the effect of adenosine transport blockade assessed by a novel method in guinea pig atria. Cell Biochem Biophys, 47, 45-52, 2007. (doi: 10.1385/CBB:47:1:45)

[5] Grenczer M, Pinter A, Zsuga J, Kemeny-Beke A, Juhasz B, Szodoray P, Tosaki A, Gesztelyi R., The influence of affinity, efficacy, and slope factor on the estimates obtained by the receptorial responsiveness method (RRM): a computer simulation study. Can J Physiol Pharmacol, 88, 1061-73, 2010. (doi: 10.1139/y10-078)

[6] Grenczer M, Zsuga J, Majoros L, Pinter A, Kemeny-Beke A, Juhasz B, Tosaki A, Gesztelyi R., Effect of asymmetry of concentration-response curves on the results obtained by the receptorial responsiveness method (RRM): an in silico study. Can J Physiol Pharmacol, 88, 1074-83, 2010. (doi: 10.1139/y10-089)

[7] McDonald JH., Handbook of Biological Statistics. Sparky House Publishing, 2009.

[8] Motulsky HJ, Christopoulos A., Fitting models to biological data using linear and nonlinear regression. A practical guide to curve fitting. GraphPad Software Inc, 2004.

[9] Prism 8 Curve Fitting Guide. GraphPad Software Inc, 2021. (https://www.graphpad.com/guides/prism/8/curve-fitting/index.htm)

[10] Fredholm BB, IJzerman AP, Jacobson KA, Klotz KN, Linden J. International Union of Pharmacology. XXV. Nomenclature and classification of adenosine receptors. Pharmacol Rev, 53, 527-52, 2001. (PMID: 11734617)

[11] Mundell S, Kelly E., Adenosine receptor desensitization and trafficking. Biochim Biophys Acta, 1808, 1319-28, 2011. (doi:10.1016/j.bbamem.2010.06.007)

HU–MATHS–IN

Hungarian Service Network of Mathematics
for Industry and Innovation

SZÉCHENYI EGYETEM
UNIVERSITY OF GYŐR

SZTE
SZEGEDI TUDOMÁNYEGYETEM
UNIVERSITY OF SZEGED

DEBRECENI EGYETEM