ROMAI J., v.16, no.1(2020), 1–6

CHALLENGES IN INDUSTRIAL MATHEMATICS: CASE STUDIES FROM A SWEDISH "MATHEMATICS MEETS INDUSTRY DAY"

Adrian Muntean[1], Elisabet Mellroth[1,2]
[1]*Department of Mathematics and Computer Science, Karlstad University, Sweden*
[2]*Sundsta-Älvkullegymnasiet (Karlstads kommun, Sweden)*
adrian.muntean@kau.se, elisabet.mellroth@karlstad.se

Abstract Mathematics meets industry day (MiMM® Day) offers an organized setup where companies pose challenges to applied mathematicians. The tasks are supposed to be tackled within one day. We briefly present the concept of the MiMM® day and list the three mathematical challenges that were proposed during the 4th MiMM® day organized in Karlstad, Sweden, on December 15, 2020. This note should be seen as an invitation addressed to the industrial and applied mathematics community to get actively involved in events of this type.

Keywords: industrial mathematics, mathematical modeling, computational aspects of data analysis and big data.
2010 MSC: 68T09, 97M10.

1. INTRODUCTION. CONCEPT OF THE MIMM® DAY

In recent years, one notices a steady increase in the interest of companies to use proficiently existing advanced mathematical technology to handle in a competent manner own research and development questions. This usually requires the employment of mathematically well-trained employees as well as trusted contacts with researchers active in the field of industrial and applied mathematics. Important prerequisites for this to happen include excellent communication skills as well as hands-on skills from the side of mathematicians. We refer to [5] for more on this context.

One of the main roles of a "mathematics with industry day" is to promote interactions between industry and applied mathematicians. During a MiMM® day, teams form to attack the industrial problems. Each team is composed of mathematicians with very different levels of seniority, that is: senior researchers (1-2), doctoral students (1-3), university students (2-3) and upper secondary students (2-3). To ensure success of such event, it is crucial that participants are self-motivated when joining it; this is a feature that appears to be more important than the level of mathematical expertize they bring with. The problem owners from each of the companies usually join the working teams.

2. PROPOSED PROBLEMS. APPROACH

For the 4th MiMM® day, which took place in Karlstad, Sweden, on December 15, 2020, three industrial problems have been selected and approximately 35 people from six coutries worked on them. The details of the problems are listed on the event webpage `https://www.kau.se/MIMM-day`. In this framework, we only briefly describe the posed questions, indicate a few keywords on the proposed solution strategies so that we give a flavor on the needed mathematical tools and eventual open problems. The reader should bear in mind that the industrial questions do not need *per se* to be answered fully during the day. Having intensive discussions is much more important; further follow-up joint activities are strongly encouraged. However, the ideas discussed during the day should shed light on possible answers to the posed questions.

2.1. PROPELLER IN A NOZZLE - THE KONGSBERG PROBLEM

Kongsberg Maritime AB `www.kongsberg.com/maritime/` is preoccupied with efficiency and safety throughout the whole maritime technology spectrum. The problem posed concerns the finding an explicit (analytic) mathematical formula that is able to approximate as accurate as possible the clearance between the propeller tip and the nozzle (propellers-in-nozzle is a design typical for boats that are meant to be slow).

The working assumptions are listed on `https://www.kau.se/MIMM-day`. Also, the problem owner has provided a CAD design where any discovered analytic formula can be checked for validity against realistic numerical values.

2.2. MRI RESOURCE OPTIMIZATION PROBLEM - THE REGION VÄRMLAND PROBLEM

Region Värmland `www.regionvarmland.se` is responsible for all of the county of Värmland's publicly funded regional development, health care, culture and public transport. They proposed a resource optimization problem within the field of health care. This reads as follows:

In Arvika, there is 1 MRI-machine, in Karlstad 2 such machines and in Torsby 1. The Arvika and Torsby MRI machines are recently upgraded and therefore can do some of the examinations faster than the ones in Karlstad. On the other hand, the hospital in Karlstad has more competence and therefore can do some more complicated/rare things. There is also an emergency fraction of examinations which may not be relocated and a repeated fraction that should be booked within a certain time and preferably always on the same MRI machine. All patients from each location are preferably booked in the same hospital, but they may be rearranged depending on the examination type.

The ultimate goal is to keep all machines busy, maximising the production, doing all examinations within the time designed limit (excluding the repeated and emergency fractions, the limits are roughly 1, 2, 3 and 4 weeks, depending on severity), while keeping travel distances and cancellation probabilities low. Ideally, the target solution strategy can be adapted and then applied to other similar problems, involving for example different examination types, time limits and locations.

Data containing the approximate population density for 64 places as well as the relative distance from each location to the three MRI investigation units in Arvika, Karlstad and Torsby was available.

2.3. DETECTING FORESTS WITH HIGH CONSERVATION VALUE - THE SWEDISH FOREST AGENCY PROBLEM

The Swedish Forest Agency is the government body for forests and forestry issues; see `https://www.skogsstyrelsen.se/en/`. Within this framework, their interest is in knowing where are located the forests (cca. 2/3 of the Sweden's surface is covered by forests) with high conservation value. These are forests that are important for the conservation of biodiversity. They have been able to develop naturally without the impacts of modern-day forestry such as logging and thinning, so display in general a greater variation than managed forests (for example, variation in species composition, structural diversity). Since 1993, the Forest Agency has mapped high conservation value forests though field surveys. The collected information forms a database with cca. 67000 areas, ranging from a single ancient tree to hundreds of hectares of native forest. It is however costly and time consuming to conduct field surveys, so the interest is currently in seeing to which extent can national cover datasets (such as laser scanning, satellite images and aerial photos) identify high conservation value forests.

The concrete challenge is whether the current database over forests with high conservation value could be used in machine learning, to train a model to recognise similar forests. Sub-questions include:

(i) Is this dataset appropriate to be used as training data?

(ii) If yes, how should the dataset be prepared to be optimal training data?

Further information on what means a forest with high conservation value as well as the dataset (with all sorts of information of type of forest, height of trees, type of habitat, near water *vs.* dray soils, etc.) were provided by the problem owner.

2.4. HINTS ON SOLUTION STRATEGIES

Information on the dynamics of boats can be retrieved for example in [2] and references cited therein. For the problem at hand, we notice the following basic aspects: Having a sufficient gap between the propeller tip and the nozzle is of practical importance to avoid interference. The nozzle's inner diameter has to be larger than the propeller's diameter. Two practical observations are useful to simplify the problem to a solvable situation: (i) When looking at the tip clearance, only the inner surface of nozzle and the tip section of the propeller are important. (ii) From the clearance point of view, the geometry can be simplified to the case of two concentric cylinders. Trusting (i) and (ii), one can rely on simple plane geometry arguments to derive a nonlinear algebraic equation expressing the clearance as a function of the thickness of the propeller, of the angle of tilted-ness as well as on the radii of the nozzle and propeller. The obtained formula is exact in 2D and can be still applied for a 3D case, provided the propeller is sufficiently thin. The propeller is usually titled simply to control the Coandă effect; see e.g. [2]. The validity of such formulas are checked by comparing the output of a simple Python/MATLAB code against a geometrical measurement made on a GeoGebra or CAD picture.

The MRI resource optimization problem looks, at a first sight, as a typical resource scheduling problem. However, given the number and complexity of the requirements and the fact that the applied solution strategy directly affects the quality of life of citizen, the problem at hand is a quite hard one, probably the hardest amongst the three proposed problems. If the frequency of the emergency situations is wild (i.e. disordered), then this challenge can be cast as a robust (combinatorial) optimization problem maybe using stochastic optimization techniques. In this case, it would be interesting to apply machine learning-based techniques (ML) to solve such problem fast as ML could be used to learn the structure of the problem. If the stochasticity of the emergencies is controlled, then one can simply decide not to schedule all the available slots each day. This way such free slots have a good chance to accommodate a significant part of the unplanned examinations. If the stochasticity of the situation is neglected, then a multiple objective linear programming approach (similar to [1]) is in principle applicable at least when the density of patients is not too large. The team has chosen to over-simplify the resource allocation problem to fit to a standard simplex algorithm. However, given the size of the problem such approach is computationally not feasible. However, a direct computational approach of a simplified setting is possible to be done in `Python` (via `SciPy` and `PuLP`[1].)

Considering the available very-good-quality forest data, the team believes that a proper classification of data is the right way to go. This would need to be done before any machine learning exercise. However, the number of collected features is quite large. A suitable preprocessing of the data is needed (e.g. a principal component analysis [4]) to detect at least the main trends and question the outliers. First clustering attempts have been tried (K-means and DBSCAN approaches via `pandas`; compare [3]) and results look promising. However, it is not so clear at this moment how much correlations exist in the data, and how much noise. A deep-learning approach would be recommended for instance to cluster laser data and satellite data for the high biotope areas.

3. DISCUSSION

Excepting perhaps the problem described in Section 2.1 for which a satisfactory solution has been found, the problems described in Section 2.2 and Section 2.3 are very much in search of a deeper approach. However, although not fully developed solutions were proposed, all three problem owners were given new approaches for their problems. In addition, the participants, problem owners and problem solvers, expressed gratitude for the possibility to exchange ideas which is a crucial aspects when solving complex real life problems. Potentially interested researchers may contact the authors to access more information on the problems and open data.

Acknowledgements

We express our special thanks to both the problem owners and participants in our mathematics with industry days during the years. The authors are particularly grateful to Omar Richardson (Maths & Karlstad) for his constant and very valuable input. AM thanks Erasmus+ project iTEM (grant no. 598587-EPP-1-2018-1-EL-EPPKA2-CBHE-JP) for partial support.

Notes

1. See e.g. the hands-on approach on `https://realpython.com/linear-programming-python/`.

References

[1] J. E. Annetts, E Audsley, Multiple objective linear programming for environmental farm planning, *Journal of the Operational Research Society*, **53**:9, 933–943 (2002).

[2] I. C. Clarke, *Ship Dynamics for Mariners*, The Nautical Institute, London, 2005.

[3] L. Igual, S. Segui, *Introduction to Data Science: A Python Approach to Concepts, Techniques and Applications*, Springer, 2017.

[4] W. McKinney, *Data Structures for Statistical Computing in Python*, Proceedings of the 9th Python in Science Conference (eds. S. van der Walt and J. Millman), 51–56, 2010.

[5] W. Schilders, Mathematics for industry in Europe, *EMS Newsletter*, **118**, 14–22, December (2020).