



ŽILINSKÁ UNIVERZITA V ŽILINE
Fakulta riadenia a informatiky



Dissertation Topics

for 2018/2019 academic year

PhD. Study

Study Programme: Applied Informatics
Branch of Study: 9.2.9 Applied Informatics
Form of Study: Full time, Part time

List of Topics

Data-centric modelling and optimisation of electric vehicles charging.....	3
Analysis of large datasets in service systems.....	5
Effective discovery of accurate and interpretable fuzzy rules in data.....	6
Selection of data from large databases for machine learning algorithms.....	7
Transformation of medical data for fuzzy data mining	8

Data-centric modelling and optimisation of electric vehicles charging

Tutor: **Assoc. Prof. Ing. Ľuboš Buzna, PhD.**

Compulsory Study Units:

Mathematical Principles of Applied Informatics
Theory and Methodology of Applied Informatics
Subject of Specialisation

Problem Area Description

If more electric vehicle drivers share the same set of public charging stations, then the need arises to coordinate them in space and time to prevent queues and delays. Another need is to match the demand of users with the time varying network capacity (e.g. due to the presence of the renewable sources of energy). In practice, the coordination may be facilitated by a booking system, social networks (social charging) or it can be self-organized. Another option is to incentivise the behaviour of users by introducing a dynamical pricing scheme that helps matching the demand and supply of electricity.

The goal of the project is to propose a method to coordinate charging of a fleet of electric vehicles in order to match available network capacity with the charging demand in time and/or space.

The proposed method will be informed by available datasets such as, dataset describing large public charging network serving users of electric vehicles.

Expected Scientific Asset:

- new methods/algorithms to support decision making in the area of charging electric vehicles.

Recommened Methods:

- formulation of mathematical models,
- design of optimisation algorithms,
- design of simulation models,
- validation of proposed models by simulation and computational experiments.

Information about Research

Type of Research:

Basic Research

Relevant to current research programs:

VEGA 1/0463/16 Economically efficient charging infrastructure deployment for electric vehicles in smart cities and communities

VEGA 1/0342/18 Optimal dimensioning of service systems

Achieved Results:

1. R. Carvalho, L. Buzna, F. B. F, M. Masera, D. K. Arrowsmith, and D. Helbing, Resilience of natural gas networks during conflicts, crises and disruptions, PLoS ONE 9, e90265 (2014)
2. R. Carvalho, L. Buzna, R. Gibbens, and F. Kelly, Critical behavior in charging electric vehicles, New J. Phys. 17, 095001 (2015)
3. M. Cebeauer, K. Rosina, Ľ. Buzna: Effects of demand estimates on the evaluation and optimality of service centre locations, International Journal of Geographical Information Science, Vol. 30, Issue 4, 2016
4. M. Cebeauer, L. Buzna A versatile adaptive aggregation framework for spatially large discrete location-allocation problems, Computers & Industrial Engineering , Vol. 111, p. 364-380, 2017

Analysis of large datasets in service systems

Tutor: **Assoc. Prof. Ing. Ľuboš Buzna, PhD., Supervisor Specialist: Ing. Michal Lekýr, PhD.**

Compulsory Study Units:

Mathematical Principles of Applied Informatics
Theory and Methodology of Applied Informatics
Subject of Specialisation

Problem Area Description

The project focuses on the application and development of methods to analyse large datasets to improve efficiency of service systems. Considered example of a service system is operation of restaurants and more broadly food industry. In this case, the dataset contains detailed records of transactions (purchases, orders, expenses and supply deliveries) for the long enough time period.

Based on the provided data, it is required to assess how the data could be used to improve the operation of a service system. The goal of the research is to develop methodology that will enable to address problems such as:

- short term estimation (e.g. one day ahead) of the number of customers, estimate of the needed personnel (What are the future trends?, What algorithms are suitable to estimate the occupancy/utilisation of the system?),
- short term estimation of the stock development (How to effectively organize the storage supply?, How to improve efficiency of the system?)

Expected Research Asset:

- new methodologies/algorithms to support decision making in the area of service systems.

Recommended methods:

- thorough analyses of the state-of-the-art,
- formulation of hypothesis and of models for data processing,
- application of computational methods to analyse large datasets,
- validation and evaluation of results.

Type of Research:

Applied Research

Relevant to current research programs:

VEGA 1/0463/16 Economically efficient charging infrastructure deployment for electric vehicles in smart cities and communities

VEGA 1/0342/18 Optimal dimensioning of service systems

Achieved Results:

1. M. Cebeauer, Ľ. Buzna: A versatile adaptive aggregation framework for spatially large discrete location-allocation problems, *Computers & Industrial Engineering*, Vol. 111, p. 364-380, 2017
2. M. Cebeauer, K. Rosina, Ľ. Buzna: Effects of demand estimates on the evaluation and optimality of service centre locations, *International Journal of Geographical Information Science*, Vol. 30, Issue 4, 2016
3. K. Rosina, P. Hurbánek, M. Cebeauer, "Using OpenStreetMap to improve population grids in Europe," *Cartography and Geographic Information Science*, s. 1-13, 2016

Effective discovery of accurate and interpretable fuzzy rules in data

Tutor: **Assoc. Prof. Ing. Ján Boháčik, PhD.**

Compulsory Study Units:

Mathematical Principles of Applied Informatics
Theory and Methodology of Applied Informatics
Subject of Specialisation

Problem Area Description:

Many organizations collect a large amount of data in their information systems nowadays. However, these cannot often be used directly for automated and intelligent decision support. Some knowledge which can be discovered through the process of knowledge discovery in data is required. A knowledge representation close to human thinking has the form of rules. The inclusion of uncertainties from the real world using the notions of fuzzy logic leads to the concept of fuzzy rules. The goal of the project is to design new or improved methods for finding a highly interpretable group of fuzzy rules in specified data and to implement a related software tool. The interpretability is judged by readability for the expert. During the development of the methods, it will be important to test the accuracy of found fuzzy rules in deployed decision support.

The proposed methods will be validated on available data such as medical data about patients with heart disease.

Expected Scientific Asset:

The project will contribute to new, more accurate, and faster algorithms for the discovery of interpretable knowledge in domain-specific data. It will also help the expert to understand the interrelationships among various input data in a way natural for humans and to support her or his decision making more precisely.

Recommended Methods:

Analysis of available literature, use of existing algorithms for the creation and deployment of fuzzy rules, development of new or improved algorithms for specific data from a defined domain, testing of achieved accuracy and interpretability.

Type of Research:

Applied Research

Relevant to current research programs:

In connection with research projects running at the Department of Informatics.

Achieved Results:

The development of a software tool with algorithms for the creation and use of fuzzy rules has been started [1], [2]. It is possible to represent knowledge in a form close to human thinking and to use it for decision support. An example of deployment is the risk estimation of malignant events for monitored patients [3], [4].

1. J. Boháčik, "Discovering fuzzy rules in databases with linguistic variable elimination," *Neural Network World*, vol. 20, no. 1, 2010, pp. 45-61
2. J. Boháčik, D. N. Davis, "Fuzzy rule-based system applied to risk estimation of cardiovascular patients," *Journal of Multiple-Valued Logic and Soft Computing*, vol. 20, no. 5-6, 2013, pp. 445-466
3. J. Boháčik, C. Kambhampati, D. N. Davis, J. G. F. Cleland, "Use of cumulative information estimations for risk assessment of heart failure patients," in *IEEE International Conference on Fuzzy Systems*, IEEE, 2014, China, pp. 1402-1407
4. J. Boháčik, M. Záborský, "Dissimilarity measure for comparison of fuzzified instances and its application in a fuzzy rule-based system for heart failure domain," in *IEEE 14th International Symposium on Applied Machine Intelligence and Informatics*, IEEE, 2016, Slovakia, pp. 339-344

Selection of data from large databases for machine learning algorithms

Tutor: **Assoc. Prof. Ing. Ján Boháčik, PhD.**

Compulsory Study Units:

Mathematical Principles of Applied Informatics
Theory and Methodology of Applied Informatics
Subject of Specialisation

Problem Area Description:

The amount of data in the databases of information systems often becomes too large for their trouble-free processing with machine learning algorithms. Because of that, various processing methods such as parallel processing with interconnected computers are being developed. Another way is to select a small subset of large data in the process of knowledge discovery in databases so that the selected data allows the discovery of knowledge corresponding with the knowledge discovered in the whole data. The consideration of representative selection leads to the concept of data selection from large databases.

The goal of the project is to design new or improved methods for a representative selection of data from chosen large databases and to implement a related software tool. During the development of the methods, it will be important to test the accuracy of the knowledge discovered in selected data.

The proposed methods will be validated on available large databases.

Expected Scientific Asset:

The project will contribute to new, more accurate, and faster algorithms and sub-algorithms for the selection of data from large domain-specific databases. It will allow the use of existing machine learning algorithms on big data of some specific domain and thus to decision-making support based on real-world data.

Recommended Methods:

Study of available literature, implementation of known algorithms for the selection of data from large databases, development of new or improved algorithms on chosen databases, testing of knowledge discovered in selected data.

Type of Research:

Applied Research

Relevant to current research programs:

In connection with research projects running at the Department of Informatics.

Achieved Results:

Research and development in the area of algorithms for the selection of representative data from large databases has been started within solved tasks [1]. There is a software tool which enables the use of selected data for knowledge discovery in databases [2], [3].

1. J. Boháčik, A. Fuchs, M. Benedikovič, "Detecting compromised accounts on the Pokec online social network," in International Conference on Information and Digital Technologies, IEEE, 2017, Slovensko, pp. 56-60
2. J. Boháčik, D. N. Davis, "Fuzzy rule-based system applied to risk estimation of cardiovascular patients," Journal of Multiple-Valued Logic and Soft Computing, vol. 20, no. 5-6, 2013, pp. 445-466
3. J. Boháčik, "Discovering fuzzy rules in databases with linguistic variable elimination," Neural Network World, vol. 20, no. 1, 2010, pp. 45-61

Transformation of medical data for fuzzy data mining

Tutor: **Assoc. Prof. Ing. Ján Boháčik, PhD.**

Compulsory Study Units:

Mathematical Principles of Applied Informatics
Theory and Methodology of Applied Informatics
Subject of Specialisation

Problem Area Description:

Uncertainties existing in almost every stage of a clinical decision making process requires a framework that can handle them by allowing variable and multiple memberships and facilitating approximate reasoning. This has made fuzzy logic a valuable tool for depicting medical concepts and preserving their imprecision without inserting distortions. The inclusion of the notions of fuzzy logic in processing data through data mining leads to the concept of fuzzy data mining.

The goal of the project is to design new or improved methods for the transformation of specified medical data into a form suitable for fuzzy data mining algorithms and to implement a related software tool. During the development of the methods, it will be important to test the accuracy of transformed data while it is used in the process of knowledge discovery in data.

The designed solutions will be validated on available medical data such as data about patients with heart disease.

Expected Scientific Asset:

The project will contribute to capturing uncertainties in clinical decision making and taking them into consideration. It will also potentially contribute to more accurate decision-making support and to a better interpretation of knowledge discovered in medical data.

Recommended Methods:

Analysis of available literature, use of known algorithms transforming medical data for fuzzy data mining methods, development of new or improved algorithms for medical data from a defined domain, testing of transformed data.

Type of Research:

Applied Research

Relevant to current research programs:

In connection with research projects running at the Department of Informatics.

Achieved Results:

The development of a software tool containing algorithms for working with medical data and their use in decision support has been elaborated [1], [2], [3], [4]. It is possible to use categorical and numerical data and their transformation for systems based on the notions of fuzzy logic.

1. J. Boháčik, D. N. Davis, "Fuzzy rule-based system applied to risk estimation of cardiovascular patients," *Journal of Multiple-Valued Logic and Soft Computing*, vol. 20, no. 5-6, 2013, pp. 445-466
2. J. Boháčik, C. Kambhampati, D. N. Davis, J. G. F. Cleland, "Use of cumulative information estimations for risk assessment of heart failure patients," in *IEEE International Conference on Fuzzy Systems*, IEEE, 2014, China, pp. 1402-1407
3. J. Boháčik, K. Matiaško, M. Benedikovič, "Linguistic variable elimination for a heart failure dataset," in *IEEE International Conference on Cybernetics*, IEEE, 2015, Poland, pp. 196-200
4. J. Boháčik, M. Záborský, "Naive Bayes for Statlog heart database with consideration of data specifics," in *IEEE 14th International Scientific Conference on Informatics*, IEEE, 2017, Slovakia, pp. 35-39