Concept Reduction Methods

Miklós F. Hatwagner

Abstract <To be prepared>

1 The Motivating Problem

The title of Adrienn Buruzs's PhD thesis [1] is "Evaluation of Sustainable Regional Waste Management Systems with Fuzzy Cognitive Map". As the title suggests, she analyzed the internal driving forces, dynamic behavior and sustainability of Integrated Waste Management Systems (IWMSs), which are very complex systems including many aspects (environmental, economic, social, institutional, legal and technical) and stakeholders. Even at an early stage of her investigations became apparent that Fuzzy Cognitive Map (FCM) is an appropriate tool to describe the large number of interacting and coupled entities and it copes with the inherent uncertainties of the system.

At first, a new FCM model [3] was created, which contained six concepts. These concepts were identified on the basis of the literature. The strength of relationships among concepts were defined by the results of a survey filled out by 75 stakeholders. The simulation results provided by FCM were validated later in [5]. Time series data were collected based on the relevant literature and it served as the input of a Bacterial Evolutionary Algorithm to learn the connection weights among the already specified concepts and parameter λ of the threshold function. The goal of optimization was to find an FCM that generates as similar time series as possible. Unfortunately, a strong contradiction was explored between the models created by experts and machine learning.

In order to resolve the experienced problem the concepts of the original model were decomposed to further 4-7 sub-concepts according to the System-of-Systems approach, which led to a very detailed, completely new model of IWMS [4]. A

Miklós F. Hatwagner
Széchenyi István University, Győr, Hungary e-mail: miklos.hatwagner@sze.hu

workshop was organized with the help of 12 stakeholders who decided the sub-concepts and their interconnections. The result of their work is a FCM containing 33 concepts in total (Fig. 1). Unfortunately such an extremely complex model is often confusing for the experts (Fig. 2), and to work with them may be very laborious. Note that the number of connections is a quadratic function of the number of concepts.

That is why, in general, the following approach is suggested to follow in practice: start with an obviously oversized model. Experts are often uncertain about the importance of system components thus it worth include all of them in the preliminary model. Then start reducing the model automatically, in an algorithmic way until the balance of model size and required accuracy is found. In the following sections several possible ways of model reduction is presented.
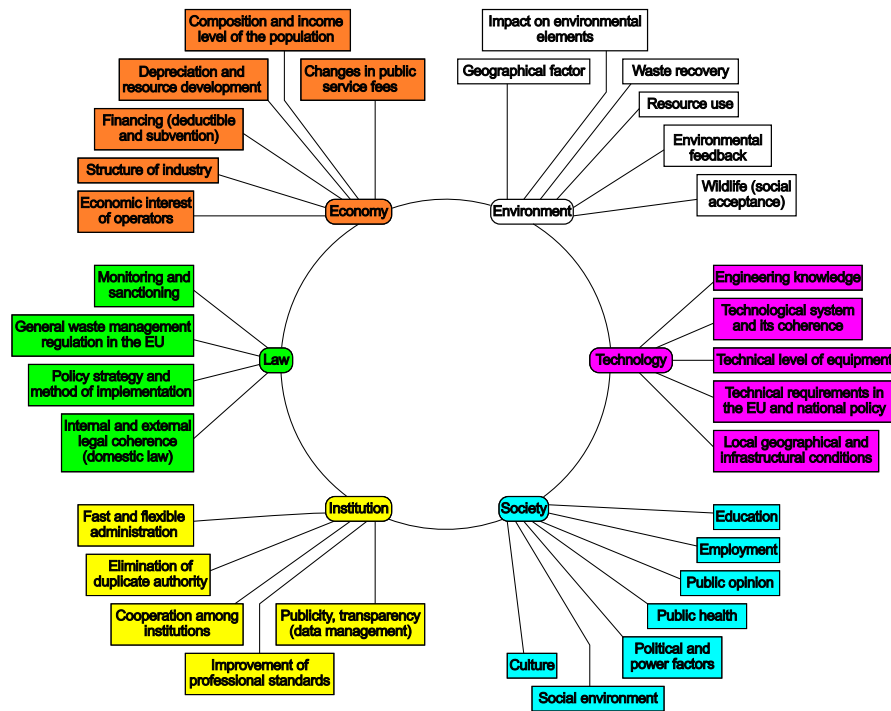

Fig. 1 The main concepts and their sub-concepts of regional IWMS.
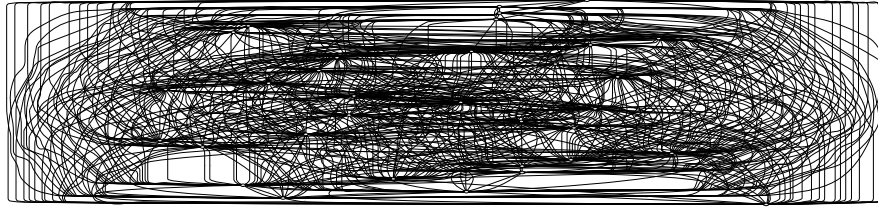
Fig. 2 The 33 concept model of regional IWMS and the relationships among its concepts. It is hard to illustrate complex FCM models appropriately and graphical model visualizations often confuse experts.

2 Early model reduction methods

Several methods had been suggested to solve the problem of oversized models before the complex model of IWMS saw the light of the day. These approaches are based on different perspectives.

In [2] an FCM is learned using historical data and its concepts are grouped into clusters in a unique way. The clustering is based on the DEMATEL [6] method. The concepts are arranged on a two dimensional plot. The vertical axis classifies concepts to cause and effect groups, the position of concepts along the horizontal axis expresses the importance of them. Based on this rearrangement of concepts two clustering methods are suggested. The first one uses K-Means clustering to create clusters according to the cause-effect behavior of concepts. The cluster centers replace the original concepts in the reduced model. The second method contains two consecutive steps, and takes also the importance of concepts into consideration. Regardless of the methods applied, experts have to define the number of clusters and they must be disjoint.

FCM was used to predict and discover knowledge about the HIV-1 drug resistance in [8]. The protease protein was modeled by FCM and causalities among sequence positions were estimated by Particle Swarm Optimization. Furthermore, Ant Colony Optimization was also applied in order to find the strongest sequence positions related to the resistance target. After that, some concepts and their connections were removed to decrease the complexity of the model until the quality of inference remained acceptable.

Another reduction method was introduced in [7]. Weak concepts and their connections were removed, then the inference capabilities of the simplified model were tested with time series data collected from real-world applications. The reduced model replaced the original if its estimation errors were acceptable.

References

1. Buruzs Adrienn. *Fenntartható regionális hulladékgazdálkodási rendszerek értékelése fuzzy kognitív térképpel*. PhD thesis, Doctoral School of Multidisciplinary Engineering Sciences (MMTDI), Széchenyi István University, Győr, Hungary, 2015.
2. Somayeh Alizadeh, Mehdi Ghazanfari, and Mohammad Fathian. Using data mining for learning and clustering fcm. *International journal of computational intelligence*, 4(2):118–125, 2008.
3. A Buruzs, RC Pozna, and LT Kóczy. Developing fuzzy cognitive maps for modeling regional waste management systems. In *3rd International Conference on Soft Computing Technology in Civil, Structural and Environmental Engineering, CSC 2013*. Civil-Comp Press, 2013.
4. Adrienn Buruzs, Miklós F Hatwágner, László T Kóczy, et al. Modeling integrated sustainable waste management systems by fuzzy cognitive maps and the system of systems concept. *Czasopismo Techniczne*, 2013(Automatyka Zeszyt 3-AC (11) 2013):93–110, 2013.
5. Adrienn Buruzs, Miklós F Hatwágner, RC Pozna, and László T Kóczy. Advanced learning of fuzzy cognitive maps of waste management by bacterial algorithm. In *2013 Joint IFSA World Congress and NAFIPS Annual Meeting (IFSA/NAFIPS)*, pages 890–895. IEEE, 2013.
6. A. Gabus and E. Fontela. Perceptions of the world problematique: Communication procedure, communicating with those bearing collective responsibility (dematel report no.1). Technical report, Battelle Geneva Research Centre, Geneva, Switzerland, 1973.
7. W. Homenda, A. Jastrzebska, and W. Pedrycz. *Computer Information Systems and Industrial Management: 13th IFIP TC8 International Conference, CISIM 2014, Ho Chi Minh City, Vietnam, November 5-7, 2014. Proceedings*, chapter Time Series Modeling with Fuzzy Cognitive Maps: Simplification Strategies, pages 409–420. Springer Berlin Heidelberg, Berlin, Heidelberg, 2014.
8. Gonzalo Nápoles, Isel Grau, Rafael Bello, and Ricardo Grau. Two-steps learning of fuzzy cognitive maps for prediction and knowledge discovery on the hiv-1 drug resistance. *Expert Systems with Applications*, 41(3):821 – 830, 2014. Methods and Applications of Artificial and Computational Intelligence.