Automatic Fuzzy Cognitive Map Building Online System

Vasiliy S. Kireev, Ivan S. Smirnov, Victor S. Tyunyakov
NRNU MEPhI, Moscow, Russia
vskireev@mephi.ru, iamnotaprogrammer@yandex.ru,
tvs.private@gmail.com

Abstract
Under the present-day global crisis, global economy regionalization and technological areas redistribution there has been seen an unprecedented uncertainty growth along with various risks, which makes it harder to take managing decisions. The scenario approach appears to be one of the most popular when it comes to modelling weakly-structured subject fields and complex problems. Fuzzy cognitive maps have good prospects in the field, as they enable us to describe both the structure and the dynamic of the area under study. This paper goes along the topical automated data science and focuses on developing the OCAM (Online Cognitive Automated Mapper) system that allows to automatically-build cognitive maps without turning to experts. The cognitive map building data source of this system is website logs. The article features the main algorithms, the system architecture and some of the system work results. The current and further research is supported by the NRNU MEPhI development program.

Keywords: association rules mining, fuzzy cognitive maps, cognitive map automated building, web-mining, automated data science

1 Introduction

Modern managing processes require taking decisions within weakly-structured dynamic systems where interconnected parts ties cannot be described functionally. In order to carry out business-modelling of such systems that show qualitative nature of interconnected parts ties, analysts use special system models called cognitive maps[1]. A cognitive map is a model that shows experts' knowledge on the developing principles and features of the situation under analysis. A cognitive map can be formally represented as a directed graph G = (V, E) with aligned weights to each arc $Ei \in E$. The vertices V of a cognitive map correspond to the factors (concepts) that define the situation. The arcs E correspond to the cause-effect (casual) relations between factors. A cognitive map building process if influenced by the way an expert sets the power of cause-effect relations and values.

A fuzzy cognitive map (FCM) in general is a directed graph where the weights of oriented edges can have values of [-1; +1], or the values that characterize the influence power of a corresponding tie and belong to a linguistic scale {very weak, weak, average, strong, very strong}[5]. If the weight wij of the arc (Vi, Vj) is more than zero, then Vj increases along with the increase of Vi and Vj decreases along with the decrease of Vi. If the weight wij of the arc (Vi, Vj) is less than zero, then Vi. Fuzzy cognitive maps were first time developed by B. Kosko as the extension to symbolic cognitive maps with the use of fuzzy casual functions[4].

Latest years have seen experts and analysts researching deep into cognitive map modelling system methodology, for example[2]. The cognitive map-based automatic modelling system software is called cognitive mapper[1]. The Kanva [1] system and Mental Modeler [3] are widespread as well.

Still there are some situations when we cannot get "hand-made" fuzzy cognitive maps: there is no expert to build one; experts have diverse knowledge of the field and build different cognitive maps; the number of concepts and ties is so huge that experts can easily make mistakes. All these cases ask for algorithms that allow us to automatically single out concepts and ties, thus building cognitive maps.

2 Main approaches to the problem

The ultimate goal of the automatic algorithms that build cognitive maps is to single out concepts and to define the ties between them. Different data types (numerical, character or mixed) require different cognitive map building algorithms. If an algorithm lacks subject field knowledge, singling concepts out automatically turns into the hardest task.

Machine-learning algorithms are successfully used for the numerical data type. Character data type requires logic-based classification methods, when Bayesian networks, decision trees and fuzzy decision trees are used for fuzzy cognitive maps. In case there is numerical and character data combined, we use neuro-fuzzy systems[6], cognitive decision trees[7] and association rules[8]. The taller three are of special interest, as they are universal and enable us to get man-readable rules as a rearsult, which is vital for such fields like medicine, economy and sociology.

A fuzzy decision tree is an extended classical decision tree concept from the field of artificial intelligence. The latest researches in fuzzy decision trees development covered the ID3 algorithm modification.

Sison and Chong offered a fuzzy version of the ID3 algorithm, that automatically generated a basic set of fuzzy rules for an object. The basic fuzzy rule set was based on a set of input-output data[9]. Umano offered his ID3 version as well. The modified algorithm could generate a man-readable fuzzy decision tree based on fuzzy sets, set by the user[10]. Janikow considers a fuzzy decision tree building algorithm that used a directly entered set of data and is called FID (Fuzzy Induction on Decision Tree)[11]. This algorithm is based on different types of fuzzy rules, described here[12]. Such an approach enables us to get cause-effect ways and fuzzy rules for linguistic weights. It also helps to build a dynamic fuzzy cognitive map to support decision-making process.

Neuro-fuzzy systems combine the methods of artificial neural networks and fuzzy logic systems[13,14]. Tahmasebi looks deeper [15] into different neuro-fuzzy systems, analyzes them and other approaches to assessment prediction.

The algorithms to single out association rules are currently the most used tools of data-mining aimed at rules extraction. The field counts a lot of methods, still all of them have two common steps. The first step features the search for a frequency set of elements occurrences, which is followed by an algorithm generating association rules. The appropriate algorithm is chosen according to the structure and the size of the data under analysis, as well as to the subject field.

The fuzzy association rules-based method by F.P.Pach [16] uses a fuzzy clustering algorithm to define a trapezoidal membership function for each attribute to build a discrete value set of steady attributes. In order to generate association rules with frequency set class names we use the Apriori algorithm[20].

3 Description of the OCAM system (Online Cognitive Automated Mapper)

The chosen automatic cognitive map building algorithm is based on extracting association rules and implies six steps. The input data is a set of lines that include a CLF web-server log.

Step 1. Data pre- processing.

Input data: a CLF web-server log.

At this stage the web-server log gets parsed and the data goes to the data-base.

Output data: inserting log data to the data-base.

Step 2. Turning to fuzziness.

Input data: data from the data-base.

At this stage the obtained data is used for value clustering to move from nonfuzzy logic to fuzzy logic[27].

Output data: cluster names including membership function for each log entry.

Step 3. Searching for frequency subsets.

Input data: cluster names including membership function for each log entry, data from the data-base.

This stage features the search for frequency subsets of different size.

Output data: a list of frequency subsets.

Step 4. Generating association rules.

Input data: a list of frequency subsets.

This stage features association rules generation based on the frequency subsets.

Output data: a list of association rules with support and reliability of each rule.

Шаг 5. Cutting association rules short.

Input data: the list of association rules.

This stage features shorting the association rules list with the help of correlation rules described by Brin[28].To determine, whether the rule is r dependent, you need to check, whether strongly differ the actual number of operations of the rule r in the cell from the expected value. Using this property, you can write Chi-square statistics for the value of the deviation of the actual value from the expected:

$$\chi^2 = \sum_{r \in R} \frac{(\sup(r) - E[r])^2}{E[r]}, \tag{1}$$

where $E[i_j] = \sup(i_j)$, $E[\overline{i_j}] = n - \sup(i_j)$, $E[r] = n \times E[r_1]/n \times E[r_2]/n \times \dots E[r_k]/n$.

If the value of Chi-square is close to one, then the expected value differs from actual, then the rule can be considered correlated and discard it. Thus, the base rules remain the only independent rules that carry maximum meaning.

Fig.1 shows the high level architecture of the system. The PC Client (a web-application) sends HTTP-requests to the Web-server. Then the Web-server, depending on the HTTP-request type directs the request either to the Application server or to the Statics server. The Statics server deals with all requests linked to static content output, such as .htm files, .js files, .css styles, images etc. The Statics server takes files from the file system and sends them back to the PC Client through the Web-server. The Application server processes the requests for API data. The data often comes from a relational database. In this case the relational database is PostgreSQL. The Application server provides the data as HTTP-requests in the .json format.

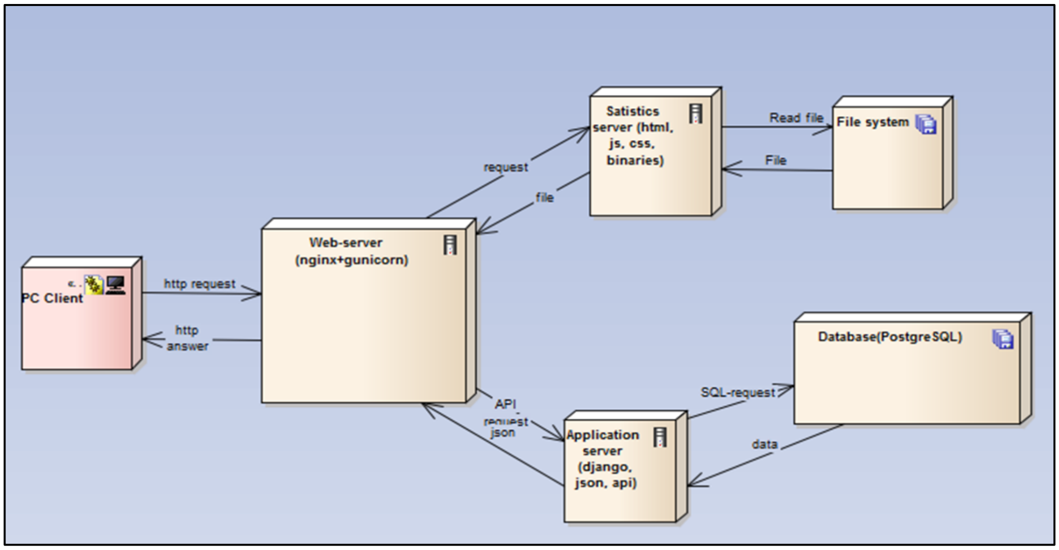

Fig. 1 High-level of OCAM architecture (UML)

Fig.2 features a cognitive map project page. To start a new project you should type in the project name, specify its type (public or private) and pick a previously uploaded web-log in case you need to automatically extraxt concepts and ties from it.

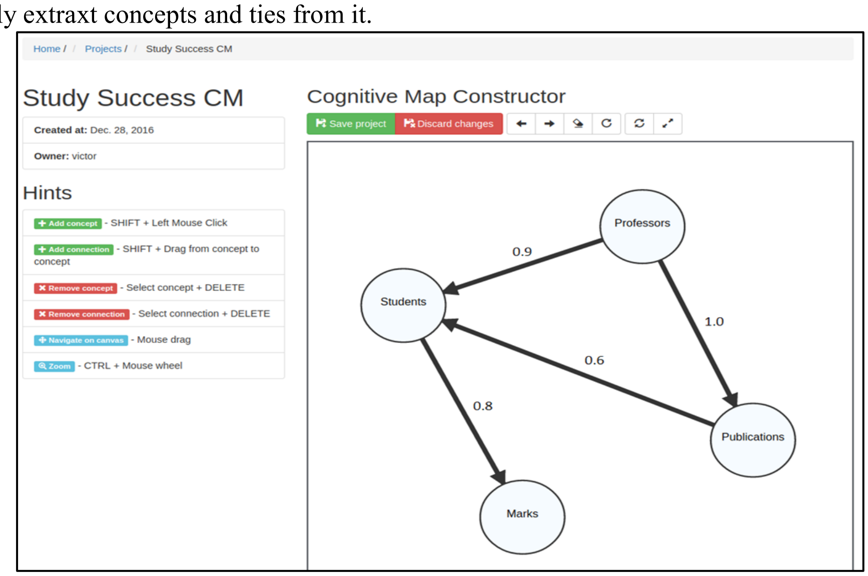

Fig. 2 Project page of OCAM

4 Conclusions

Fuzzy cognitive maps building methods constitute one of the topical areas of research in modelling. This article features the descriptions of algorithms and architecture developed by the authors of the OCAM system to automatically build fuzzy cognitive maps based on website logs data. Further research will be focused on extending the system functioning with additional algorithms to build cognitive maps and extended compatibility with scenario modelling applications.

References

1. Kulinich, A. A. Computer Cognitive Map Modeling Systems. Control Science, 2010, 3, pp. 2–16.
2. Samsonovich, A.V. and Ascoli, G.A. Cognitive map dimensions of the human value system extracted from the natural language. In Advances in Artificial General Intelligence (Proc. 2006 AGIRI Workshop), ed. by Goertzel B. (IOS Press, Amsterdam), 2007, pp. 111-124.
3. Gray S. A. et al. Mental modeler: a fuzzy-logic cognitive mapping modeling tool for adaptive environmental management //System Sciences (HICSS), 2013 46th Hawaii International Conference on. – IEEE, 2013. – pp. 965-973.
4. Kosko B. Fuzzy cognitive maps //International Journal of man-machine studies. – 1986. – T. 24. – №. 1. – pp. 65-75.
5. Carvalho J. P., Tomè J. A. B. Rule based fuzzy cognitive maps-fuzzy causal relations //Computational Intelligence for Modelling, Control and Automation, Edited by M. Mohammadian. – 1999.
6. Mitra S., Hayashi Y. Neuro-fuzzy rule generation: survey in soft computing framework //IEEE transactions on neural networks. – 2000. – T. 11. – №. 3. – pp. 748-768.
7. Chen G., Wei Q. Fuzzy association rules and the extended mining algorithms //Information Sciences. – 2002. – T. 147. – №. 1. – pp. 201-228.
8. Au W. H., Chan K. C. C. FARM: A data mining system for discovering fuzzy association rules //Fuzzy Systems Conference Proceedings, 1999. FUZZ-IEEE'99. 1999 IEEE International. – IEEE, 1999. – T. 3. – pp. 1217-1222.
9. Sison L. G., Chong E. K. P. Fuzzy modeling by induction and pruning of decision trees //Intelligent Control, 1994., Proceedings of the 1994 IEEE International Symposium on. – IEEE, 1994. – pp. 166-171.
10. Umanol M. et al. Fuzzy decision trees by fuzzy ID3 algorithm and its application to diagnosis systems //Fuzzy Systems, 1994. IEEE World Congress on Computational Intelligence., Proceedings of the Third IEEE Conference on. – IEEE, 1994. – pp. 2113-2118.
11. Janikow C. Z. Fuzzy decision trees: issues and methods //IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics). – 1998. – T. 28. – №. 1. – pp. 1-14.
12. Abonyi J. et al. Modified Gath–Geva clustering for fuzzy segmentation of multivariate time-series //Fuzzy Sets and Systems. – 2005. – T. 149. – №. 1. – pp. 39-56.
13. Gath I., Geva A. B. Unsupervised optimal fuzzy clustering //IEEE Transactions on pattern analysis and machine intelligence. – 1989. – T. 11. – №. 7. – pp. 773-780.
14. Jang J. S. R. ANFIS: adaptive-network-based fuzzy inference system //IEEE transactions on systems, man, and cybernetics. – 1993. – T. 23. – №. 3. – pp. 665-685.
15. Tahmasebi P., Hezarkhani A. A hybrid neural networks-fuzzy logic-genetic algorithm for grade estimation //Computers & geosciences. – 2012. – T. 42. – pp. 18-27.
16. Pach F. P., Abonyi J. Association rule and decision tree based methods for fuzzy rule base generation //World Academy of Science, Engineering and Technology. – 2006. – T. 13. – p. 45-50.
17. Agrawal R., Imieliński T., Swami A. Mining association rules between sets of items in large databases //Acm sigmod record. – ACM, 1993. – T. 22. – №. 2. – pp. 207-216.
18. Brin S., Motwani R., Silverstein C. Beyond market baskets: Generalizing association rules to correlations //ACM SIGMOD Record. – ACM, 1997. – T. 26. – №. 2. – pp. 265-276.