Fuzzy Cognitive Maps: a Business Intelligence Discussion

Gonzalo Nápoles, Greg Van Houdt, Manal Laghmouch, Wouter Goossens, Quinten Moesen, and Benoît Depaire

Hasselt University, Belgium

Abstract. Modeling complex systems by means of computational models has enabled experts to understand the problem domain without the need of waiting for the real events to happen. In that regard, Fuzzy Cognitive Maps (FCMs) have become an important modeling tool in the neural computing field because of their flexibility and transparency. However, obtaining an FCM-based model able to align its dynamical behavior with the system semantics is not always trivial. It is of importance to align a business' needs to the insights gained from FCM simulations. In this paper, we discuss some aspects to be considered when designing FCM-based simulation models from the perspective of a business intelligence case study. In a nutshell, when the fixed point is unique, we recommend to focus on the number of iterations to converge instead of focusing on the reached system equilibrium, and stress the importance of the transfer function chosen in the model.

Keywords: Fuzzy Cognitive Maps, Recurrent Neural Networks, Simulation, Neural Cognitive Modeling, Interpretability.

1 Introduction

Fuzzy Cognitive Maps (FCMs) are neural systems used to perform causal reasoning among several elements or concepts by the use of fuzzy graph structures [13]. The strength of causal dependencies are modeled using approximations instead of crisp numerical values. Estimates are obtained using either expert-driven, data-driven or hybrid approaches [5, 19].

Many applications of FCMs have proven their value in the areas of political and social sciences [2, 3], medicine [7, 18, 4], engineering [9, 22], business [23, 24, 10], information technology [20, 16, 21] and pattern recognition [6, 17], but none of these applications question other aspects than the reached equilibrium as valuable. When it comes to the FCM design, there is a lack of general guidelines to model the system due to FCMs are problem-dependent, so each model has its own particularities to be taken into account.

Jetter and Kok [12] proposed an elaborate framework for modeling FCMs which consists of six steps, whereas Alizadeh and Jetter [1] proposed some guidelines to extract causal relationships from qualitative data, specifically from secondary data sources, as an expansion to this framework. While these papers mainly focus on what needs to be taken into account in terms of the static analysis, there is less discussion devoted to the issues that impact the FCM's dynamic behavior. To what extent estimates of a specific reasoning rule are realistic causal estimates in the case of data-driven approaches is also subject to discussion. Furthermore, the consequences of selecting a specific continuous transfer function that uses heuristic learning methods and other FCM aspects than the equilibrium cannot be neglected.

The contribution of this paper is two-fold. Firstly, we present an FCM-based model concerning the Business & Information Systems Engineering program of Hasselt University, which will be used to illustrate the discussion. Secondly, besides the practical insights we can draw from the simulation results, we go over several aspects that are deemed pivotal when designing FCM-based systems used in scenario analysis. As such, we investigate the consequences of selecting a specific transfer function and how to interpret the results of the model from the business intelligence perspective.

The remainder of this paper is structured as follows. Section 2 goes over the theoretical background of fuzzy cognitive mapping, while Section 3 introduces the case study. In Section 4, we discuss relevant aspects to be taken into account when performing the simulations. Finally, Section 5 presents some concluding remarks and further research issues to be explored.

2 Fuzzy Cognitive Maps

Within the neural network family, FCMs are considered recurrent neural systems that allow the modeling and simulation of complex systems. In such knowledge-based structures [8], the problem domain is represented as a weighted, directed graph comprised of well-defined neural concepts and causal relationships.

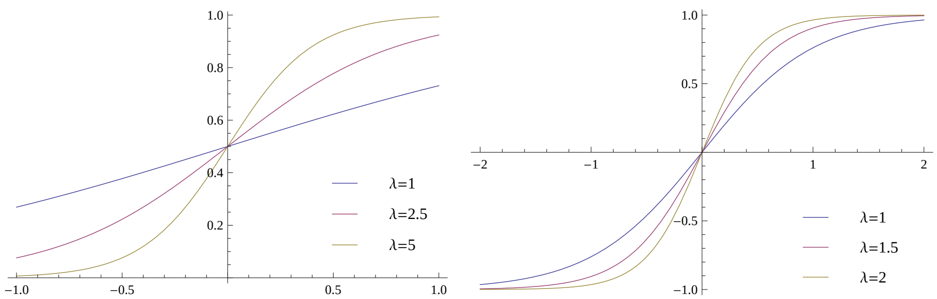
Mathematically speaking, the FCM reasoning model can be roughly defined by a 4-tuple $\langle \mathcal{C}, \mathcal{W}, \mathcal{A}, f \rangle$ where $\mathcal{C} = \{C_1, C_2, \ldots, C_M\}$ denotes the set of neural concepts, $\mathcal{W} : \mathcal{C} \times \mathcal{C} \rightarrow [-1, 1]$ is the causal weight matrix such that $w_{ij} \in [-1, 1]$ defines the causal relation between C_i and C_j. The value of w_{ij} determines the sign and intensity (magnitude) of the edge connecting C_i with C_j. The function $\mathcal{A} : \mathcal{C} \times \mathbb{N} \rightarrow \mathbb{R}$ associates, at iteration $t \in \mathbb{N}$, the neural concept C_i with an activation value by using a neural reasoning rule. In the formulation proposed by Kosko in [13], both the causal weights and the activation of neurons involve the quantification of a fuzzy linguistic variable. Equation 1 formalizes the reasoning rule used in most FCMs reported in the literature,

$$A_i^{(t+1)} = f\left(\sum_{\substack{j=1 \\ i \neq j}}^{M} w_{ji} A_j^{(t)}\right) \tag{1}$$

where $f : \mathbb{R} \rightarrow I$ is the transfer function which aggregates the impact of multiple causal events over the target concept and clamps the result to the predefined activation interval I over $[0,1]$ in case of the sigmoid function (Equation 2a) or $[-1,1]$ for the hyperbolic tangent function (Equation 2b).

$$f(x) = \frac{1}{1+e^{-\lambda x}} \quad \text{(2a)} \qquad\qquad f(x) = \frac{e^{2\lambda x}-1}{e^{2\lambda x}+1} \quad \text{(2b)}$$

Figures 1a and 1b display the shape of these continuous transfer functions for different slope λ values.



(a) Sigmoid function. (b) Hyperbolic tangent.

Fig. 1: Sigmoid and hyperbolic tangent functions for different slope values.

The reader can notice that the constraint $i \neq j$ prevents a concept to be caused by itself. While this helps preserve the coherence in the model, there are real-world situations in which the current state of a concept is conditioned by the concept's previous state.

This FCM reasoning is repeated until either the system stabilizes or meets a maximal number of iterations [15]. The former implies that a hidden pattern was discovered [14] while the latter suggests that the FCM is either cyclic or chaotic. If the FCM reaches an equilibrium point, then the variations between two consecutive state vectors will be infinitesimal.

Although FCMs are indeed neural networks, their power does not necessarily lie in their prediction capabilities but in their interpretability. In that regard, the activation values and the understandable meaning of causal relations play a pivotal role. Consequently, also the relations between concepts can be interpreted. Overall, there are two possible types of causal relations:

- $w_{ij} > 0$: Higher (lower) activation values of C_i in the t-th iteration will lead to higher (lower) activation values of C_j in the (t+1)-th iteration
- $w_{ij} < 0$: Lower (higher) activation values of C_i in the t-th iteration will lead to higher (lower) activation values of C_j in the (t+1)-th iteration.

After obtaining this clear view of the FCM theory, we will apply this in a business intelligence context. Consequently, we provide some guidelines on how to use FCMs in decision making scenarios.

3 Problem Description and Modeling

In this section, we describe an FCM-based model concerning a business intelligence case study. More specifically, we are interested in gaining insights into the academic Business & Information Systems Engineering program at Hasselt University, Belgium. Aiming at doing so, we involved 12 experts (i.e., professors and students) in the knowledge engineering process. Therefore, the concepts and connections among them rely on expert judgment. Each expert is associated with a confidence weight ranging from 0 to 1, which represents the expert's confidence in his or her own estimations. While such a weight will not entirely remove the uncertainty in the model, it allows to reduce the impact of poor judgments in the cognitive network. Equation 3 shows how to compute the aggregated weights $w_{ij}^{(+)}$ with α_k denoting the confidence attached to expert k.

$$w_{ij}^{(+)} = \sum_{k=1}^{K} \alpha_k w_{ji}^{(k)} / K \tag{3}$$

The concepts in the FCM involve key factors in the academic Business & Information Systems Engineering program that in some way have an influence on the program. The following concepts are represented: (C1) New students, (C2) Student potential, (C3) Strong Bachelor and Master's program and the Business Informatics Community, (C4) Successful student projects, (C5) Business relations, (C6) Understanding business needs, (C7) Applied research projects, (C8) Successful professional program, (C9) Revenue, (C10) Research capacity, (C11) Research output, (C12) External partners and (C13) Public image. The network is comprised of these 13 neural concepts and 27 causal relationships, which result in a 13x13 weight matrix. Figure 2 shows the resultant model. Notice that all weights are non-negative in this system.

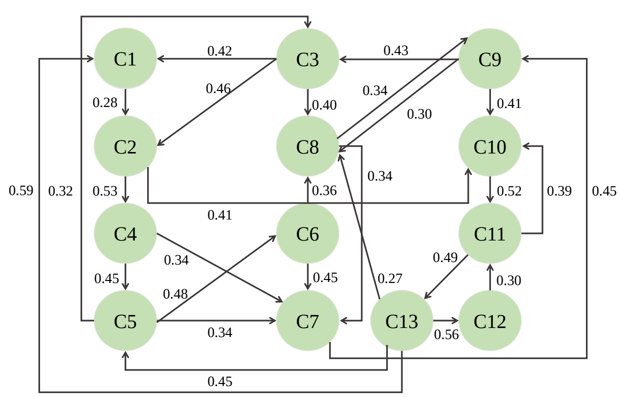


Fig. 2: The FCM-based simulation model concerning the Business & Information Systems Engineering program at Hasselt University, Belgium.

It is worth mentioning that our goal is to extract business knowledge from the FCM-based model by performing WHAT-IF simulations. In what follows, we will focus on the configuration issues such as the impact of selecting a specific transfer function and how to interpret the simulation results.

4 Simulations and Discussion

Academic programs in universities are similar to a business context since they need enough assets while producing enough revenue. In this section, we will analyze what influence C1 or C13 have on C9, according to different settings. We used the standard Kosko's activation rule to run our scenarios, so we assume that the activation value of a concept in each iteration does not depend on its previous state. In addition, we adopt either the sigmoid function or the hyperbolic tangent to bound the concepts' activation values.

Before moving forward, we need to give attention to the semantics of iterations in a simulation. We propose to define this according to the expected frequency of the change of key variables in the model. For example, it is expected for the concept "new students" (C1) to change at the beginning of each academic semester. Of course, we have to take into account that Equation 1 assumes that all concepts change with the same frequency, which might not always be realistic. In what follows, we will assume that each iteration step is equal to an academic semester.

First, we use the sigmoid function with $\lambda = 1$ and $A_1^{(0)} \in \{0.0, 0.2\}$. The model converges to the same fixed-point attractor, even when there is no initial activation. This is illustrated in Fig. 3, on which the x-axis represents the number of iterations and the y-axis denotes the activation value of C9. Shifting the function can alleviate this problem, yet the function will continue to produce only positive values, also in a system characterized by negative weights. Therefore, we will not add simulations using the sigmoid function.

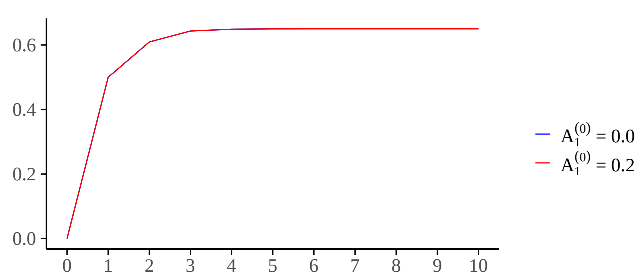


Fig. 3: Behavior of concept C9 given $A_1^{(0)} \in \{0.0, 0.2\}$.

Using the hyperbolic tangent function allows us to enhance the model, since this function is not limited to positive values. This means that the neural cognitive system will be able to capture a decrease or increase of concepts more effectively, even when the system converges to a unique fixed point.

As a first scenario, we investigate the effect of C1 on C9 using a low and a high activation value. Specifically, we choose $A_1^{(0)} = 0.1$ and $A_1^{(0)} = 0.9$ as extreme positive activation values. Moreover, we investigated the effect of the slope parameter λ, which regulates the degree of non-linearity of neural concepts. In the extreme case, with a very high λ, the concepts become binary: it is either fully activated positively, or fully activated negatively. During the simulations we will work with $\lambda \in \{1, 1.5, 2\}$ using the hyperbolic tangent function.

We start the simulations with $\lambda = 1$ and $A_1^{(0)} = 0.1$. Figure 4 displays the simulation results for this scenario, where the x-axis represents the iterations while the y-axis denotes the activation value for each concept in a given iteration. What immediately catches our attention, is that with $\lambda = 1$ all concepts move towards zero. We obtain the same results with $A_1^{(0)} = 0.9$. This is not valuable for the decision making context since an increase of C1, which has a positive causal relationship with other concepts, has barely an effect on the system in the long run. It should be noticed that C1 has only one outgoing connection, which might be an indicator that we need to increase the excitability of neurons to capture small changes in the system.

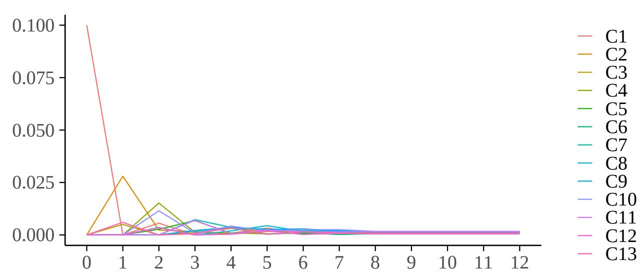
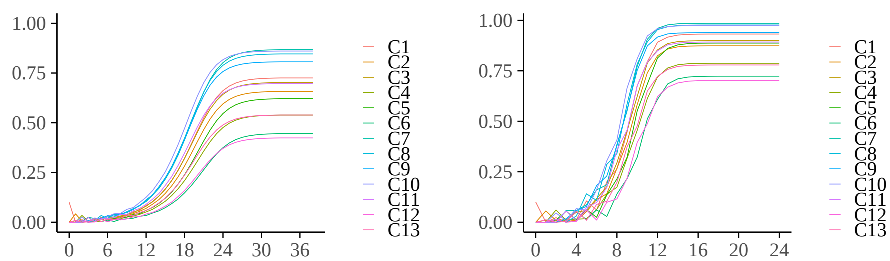


Fig. 4: Simulation results for $\lambda = 1$ and $A_1^{(0)} = 0.1$.

One aspect that needs to be discussed is the convergence issue. The authors in [11] proved that a unique fixed point always exists for log-sigmoid and hyperbolic tangent FCM models. This result holds not only for $\lambda = 1$, but every λ value. In contrast with prediction scenarios, in our problem this unique fixed point on itself is not an issue, but the results are interesting. Given that the system moves towards zero as a fixed point, it takes longer to move there from $A_1^{(0)} = 0.9$ than from $A_1^{(0)} = 0.1$. The same conclusion can be drawn for every other concept we activate. However, returning to the business aspect, saying that whatever the activation values are, there will barely be any effect on the system in the long run advocates against this configuration.

We can alter the configuration by altering the excitability degree of neurons through the λ parameter. This will result in a different fixed-point attractor. The simulations with $\lambda = 1.5$ and $\lambda = 2$, both with $A_1^{(0)} = 0.1$, are shown in Figures 5a and 5b. Again, the x-axes represent the number of iterations, while the y-axes show the activation value for each concept in a given iteration. With these slopes, we do get interesting results. If $\lambda = 1.5$, then it would take 33 iterations to reach the fixed-point attractor when $A_1^{(0)} = 0.1$, or 24 semesters when $A_1^{(0)} = 0.9$. If $\lambda = 2$, then it takes the system 18 and 14 semesters with $A_1^{(0)} = 0.1$ and $A_1^{(0)} = 0.9$, respectively, to reach the equilibrium.



(a) $\lambda = 1.5$ and $A_1^{(0)} = 0.1$.

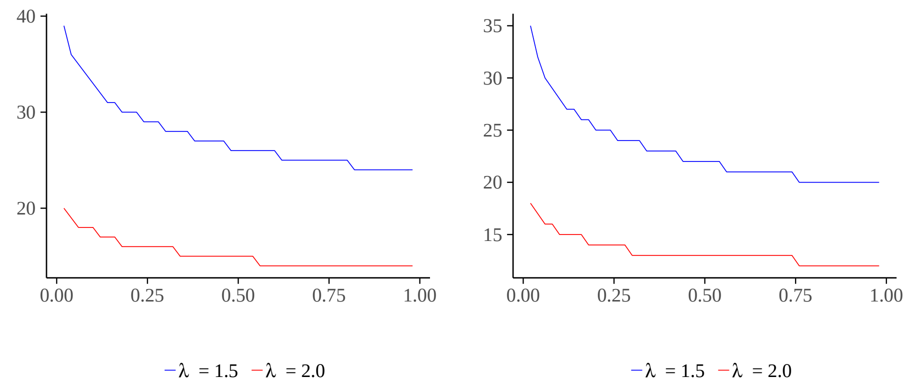
(b) $\lambda = 2$ and $A_1^{(0)} = 0.1$.

Fig. 5: Simulation results using the hyperbolic tangent function.

It gets really interesting if we repeat the same simulations for the effect of C13 on C9. Given the insights we obtained earlier, the sigmoid function is neglected in this case. Additionally, we will not perform a simulation with $\lambda = 1$, due to all values ranging to zero. We proceed with $\lambda = 1.5$ and $\lambda = 2$ instead. With respect to the activation values, we will use $A_{13}^{(0)} = 0.1$ and $A_{13}^{(0)} = 0.9$. The activation of C13 with $\lambda = 1.5$ results in an equilibrium sooner. The FCM stabilizes in only 28 semesters if $A_{13}^{(0)} = 0.1$ and in 20 semesters if $A_{13}^{(0)} = 0.9$. In comparison with the activation values of C1, the system stabilizes respectively 5 semesters and 4 semesters earlier. If $\lambda = 2$, then it takes the system 15 and 12 semesters with respectively $A_{13}^{(0)} = 0.1$ and $A_{13}^{(0)} = 0.9$ to reach the fixed point. Again, this fixed point is reached sooner than with the activation of C1.

These numbers do not necessarily mean it is more interesting to invest in a better public image (C13) rather than obtaining new students (C1). A generalization of our findings from these simulations is hard to make, but we can say that the higher $A_i^{(0)}$, the faster the system stabilizes. We cannot forget about the costs related to increasing either concept. In the end, it is up to the decision maker to decide which concept to stimulate.

On the other hand, the more excited the neuron, the more impact it has on the system even when using the same activation values. Figures 6a and 6b show the relationship between the excitement of neurons and the time to reach the equilibrium point, for both $\lambda = 1.5$ and $\lambda = 2$. The x-axis denotes the initial activation value of (a) C1 and (b) C13, the y-axis shows the number of iterations it took the system to reach the fixed-point attractor.



(a) New students

(b) Public image

Fig. 6: Relationship between $A_i^{(0)}$ and the number of iterations to stabilize.

Before concluding our paper, it seems convenient to briefly discuss the role of the λ parameter on the simulation results. As illustrated, we can obtain different dynamics if we alter the neurons' excitation, so this cannot be done arbitrarily! One option would be to conduct a comparative analysis of scenarios by using the same excitation degree, so we can draw relative conclusions. The second option would be to determine the precise non-linearity degree during the knowledge engineering phase or using a supervised learning process.

5 Conclusion

In this paper, we have presented a Business Intelligence application of Fuzzy Cognitive Maps. Besides the application itself, we have discussed issues related to the interpretation of simulation results for different configurations. The fact that our problem is described by non-negative weights ensures that the fixed-point attractor exists and that it is unique, regardless of the neurons' excitation degree. However, business decisions should be based on both the final values and on how quick the modeled system converged to the unique fixed-point attractor. In that sense, we have introduced a new way to look at the simulation so that decision makers could gain meaningful insights from the duration of reaching a certain equilibrium. This enables experts to activate the concepts to reach an equilibrium faster or slower, depending on the desired outcome. It is worth mentioning that the costs of the concepts is not included in an FCM-based model, thus a decision maker should consider the trade-off between the number of iterations and the costs linked to activating the concepts.

Another aspect discussed in this research refers to the non-linearity of neural concepts, which is implemented via the transfer function used to clamp the activation value of each concept to the desired interval. However, the experiments have illustrated how arbitrary changes on the slope parameter leads to different dynamic behaviors, which might be difficult to justify. Therefore, properly configuring the mathematical FCM model to realistically represent the physical system under analysis is a major research challenge.

References

[1] Alizadeh, Y., Jetter, A.: Content analysis using fuzzy cognitive map (fcm): A guide to capturing causal relationships from secondary sources of data. In: Portland International Conference on Management of Engineering and Technology (PICMET). pp. 1–11. IEEE (2017)

[2] Andreou, A.S., Mateou, N.H., Zombanakis, G.A.: Soft computing for crisis management and political decision making: the use of genetically evolved fuzzy cognitive maps. Soft Computing **9**(3), 194–210 (2005)

[3] Andreou, A., Mateou, N., Zombanakis, G.A.: Evolutionary fuzzy cognitive maps: A hybrid system for crisis management and political decision making. In: Conference Proceedings on Computational Intelligence for Modelling Control and Automation. vol. 1, pp. 1–12 (2003)

[4] Deja, R., Froelich, W., Deja, G., Wakulicz-Deja, A.: Hybrid approach to the generation of medical guidelines for insulin therapy for children. Information Sciences **384**, 157–173 (2017)

[5] Felix, G., Nápoles, G., Falcon, R., Froelich, W., Vanhoof, K., Bello, R.: A review on methods and software for fuzzy cognitive maps. Artificial Intelligence Review pp. 1–31 (2017)

[6] Froelich, W.: Towards improving the efficiency of the fuzzy cognitive map classifier. Neurocomputing **232**, 83–93 (2017)

[7] Froelich, W., Papageorgiou, E.I., Samarinas, M., Skriapas, K.: Application of evolutionary fuzzy cognitive maps to the long-term prediction of prostate cancer. Applied Soft Computing **12**(12), 3810–3817 (2012)

[8] Froelich, W., Salmeron, J.L.: Advances in fuzzy cognitive maps theory. Neurocomputing **232**, 1 – 2 (2017), advances in Fuzzy Cognitive Maps Theory

[9] Gonzalez, J.L., Aguilar, L.T., Castillo, O.: A cognitive map and fuzzy inference engine model for online design and self fine-tuning of fuzzy logic controllers. International Journal of Intelligent Systems **24**(11), 1134–1173 (2009)

[10] Hajek, P., Prochazka, O., Froelich, W.: Interval-valued intuitionistic fuzzy cognitive maps for stock index forecasting. In: 2018 IEEE Conference on Evolving and Adaptive Intelligent Systems (EAIS). pp. 1–7. IEEE (2018)

[11] Harmati, I.Á., Hatwágner, M.F., Kóczy, L.T.: On the existence and uniqueness of fixed points of fuzzy cognitive maps. In: International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems. pp. 490–500. Springer (2018)
[12] Jetter, A.J., Kok, K.: Fuzzy cognitive maps for futures studies - a methodological assessment of concepts and methods. Futures **61**, 45 – 57 (2014)
[13] Kosko, B.: Fuzzy cognitive maps. International journal of man-machine studies **24**(1), 65–75 (1986)
[14] Kosko, B.: Hidden patterns in combined and adaptive knowledge networks. International Journal of Approximate Reasoning **2**(4), 377–393 (1988)
[15] Kosko, B.: Neural networks and fuzzy systems: a dynamical systems approach to machine intelligence. Prentice Hall, Upper Saddle River (1992)
[16] Lopez, C., Salmeron, J.L.: Dynamic risks modelling in erp maintenance projects with fcm. Information Sciences **256**, 25–45 (2014)
[17] Nápoles, G., Espinosa, M.L., Grau, I., Vanhoof, K., Bello, R.: Fuzzy cognitive maps based models for pattern classification: Advances and challenges. In: Soft Computing Based Optimization and Decision Models, pp. 83–98. Springer (2018)
[18] Papageorgiou, E.I.: A new methodology for decisions in medical informatics using fuzzy cognitive maps based on fuzzy rule-extraction techniques. Applied Soft Computing **11**(1), 500–513 (2011)
[19] Papageorgiou, E.I.: Review study on fuzzy cognitive maps and their applications during the last decade. In: Fuzzy Systems (FUZZ), 2011 IEEE International Conference on. pp. 828–835. IEEE (2011)
[20] Papageorgiou, E.I., Froelich, W.: Application of evolutionary fuzzy cognitive maps for prediction of pulmonary infections. IEEE Transactions on Information Technology in Biomedicine **16**(1), 143–149 (2012)
[21] Salmeron, J.L.: Fuzzy cognitive maps-based it projects risks scenarios. In: Fuzzy Cognitive Maps, pp. 201–215. Springer (2010)
[22] Salmeron, J.L., Ruiz-Celma, A., Mena, A.: Learning fcms with multi-local and balanced memetic algorithms for forecasting industrial drying processes. Neurocomputing **232**, 52–57 (2017)
[23] Trappey, A.J., Trappey, C.V., Wu, C.R.: Genetic algorithm dynamic performance evaluation for rfid reverse logistic management. Expert Systems with Applications **37**(11), 7329–7335 (2010)
[24] Wei, Z., Lu, L., Yanchun, Z.: Using fuzzy cognitive time maps for modeling and evaluating trust dynamics in the virtual enterprises. Expert Systems with Applications **35**(4), 1583–1592 (2008)