

Tipo de artículo: Artículo original
Temática: Inteligencia artificial
Recibido: 14/05/2014 | Aceptado: 15/10/2014

Compensatory fuzzy logic for intelligent social network analysis

Lógica difusa compensatoria para el análisis inteligente de redes sociales

Maikel Y. Leyva-Vázquez^{1*}, Rafael Bello-Lara¹, Rafael Alejandro Espín-Andrade²

¹ Centro de Consultoría y Desarrollo de Arquitecturas Empresariales. Departamento de Soluciones SOA. Universidad de las Ciencias Informáticas, Carretera a San Antonio de los Baños, km 2 ½, Torrens, Boyeros, La Habana, Cuba. CP.: 19370.

² Universidad de Occidente, Mazatlán, México.

* Autor para correspondencia: mleyvaz@uci.cu

Abstract

Fuzzy graph theory has gained in visibility for social network analysis. In this work fuzzy logic and their role in modeling social relational networks is discussed. We present a proposal for extending the fuzzy logic framework to intelligent social network analysis using the good properties of robustness and interpretability of compensatory fuzzy logic. We apply this approach to the concept path importance taking into account the length and strength of the connection. Results obtained with our model are more consistent with the way human make decisions. Additionally a case study to illustrate the applicability of the proposal on a coauthorship network is developed. Our main outcome is a new model for social network analysis based on compensatory fuzzy logic that gives more robust results and allows compensation. Moreover this approach makes emphasis in using language for social network analysis.

Keywords: *coauthorship network, compensatory fuzzy logic, fuzzy graph, social network analysis.*

Resumen

La teoría de los grafos difusos ha ganado en visibilidad para el análisis de redes sociales. En este trabajo se discute el rol de las relaciones difusas y su papel en el modelado de redes sociales. En el artículo se presenta una propuesta para extender el marco de trabajo de la lógica difusa al análisis inteligente de las redes sociales usando las propiedades de robustez e interpretabilidad asociadas a la lógica difusa compensatoria. Mediante este enfoque es analizada la importancia de los caminos teniendo en cuenta la longitud y la fortaleza de la conexión entre nodos de la red. Los

resultados obtenidos resultan más consistentes con la forma de tomar decisiones en los humanos. Adicionalmente se presenta un estudio de caso basado en el análisis de una red de coautoría mostrando la aplicabilidad de la propuesta. El principal resultado obtenido radica en un nuevo modelo para el análisis de redes sociales basado en la lógica difusa compensatoria brindando resultados más robustos y permitiendo la compensación. Adicionalmente el modelo contribuye al uso del lenguaje en el proceso de análisis de redes sociales.

Palabras clave: análisis de redes sociales, grafos difusos, lógica difusa compensatoria, redes de coautoría.

Introduction

Recently the ideas of a fuzzy relationship and fuzzy graph have gained in visibility. Fuzzy cognitive maps (Leyva *et al.*, 2014) and the paradigm for intelligent social network analysis (PISNA) (Yager, 2008; Yager, 2014) are two examples. Social networks have become an important technology with examples like Facebook and LinkedIn or ResearchGate, a network dedicated to researchers.

However the current approaches for social network analysis (SNA) based on fuzzy graphs have some limitations, especially for dealing with sensitivity to changes in the values of truth and compensation when calculating the accuracy of compound predicates (Ortega *et al.*, 2013). Rickard and Yager (2013) propose a method for improving interpretability in SNA using interval type-2 (IT2) fuzzy sets (John and Coupland, 2006). The main drawback with this approach is the high computational cost and the need of preprocessing the information. Another proposal (Brunelli *et al.*, 2014) is based on the construction of higher dimensional fuzzy m-ary adjacency relations from the binary relations by means of ordered weighted averaging (OWA) functions (Yager, 1988). This approach introduces a flexible consensus measure but is not directly applicable to general tasks of social network analysis.

In this paper we propose a model for SNA combining fuzzy graph and compensatory fuzzy logic (CFL) (Andrade *et al.*, 2014a). This combination allows overcoming the limitation in traditional fuzzy logic for SNA to estimate the truth or falsity of observations about a network.

The outline of this paper is as follows: Section 1 is dedicated to fuzzy graphs and Section 2 to compensatory fuzzy logic. The intelligent fuzzy social network analysis based on CFL proposal is presented in Section 3. A case study is discussed in Section 4. The paper closes with concluding remarks, and discussion of future work in Section 5.

Computational methodology

Fuzzy Graph

While crisp relations are appropriate for describing relations like *parent of*, fuzzy set are better at capturing relationships of degree like *friendship* (Wierman, 2010). A fuzzy relationship on X is a mapping $R : X \times X \rightarrow [0, 1]$ where $R(x, y)$ indicates the degree of relationship between x and y (Yager, 2010). This allows extending the connections in a network from connected or not to fuzzy connections. Here we denote a fuzzy graph as $G = \langle V, E, R \rangle$ here V is set of vertices, E is the set of edges and R is a relation $R: V \times V \rightarrow [0,1]$.

Fuzzy set theory has been applied to social network analysis (SNA) modeling fuzzy relations that exist between entities as graph (Nair and Sarasamma, 2007). A social network can easily be represented by a fuzzy relation or a fuzzy graph extending the analyst's capabilities of examining networks (Yager, 2008).

Using fuzzy sets it is possible to formalize the idea of vocabulary. For any element $y \in Y$, its membership grade, $W(y) \in [0,1]$ indicate the compatibility of the value y with the linguistic concept W (Yager, 2008). For studying fuzzy graphs, there are a number of attributes about which it will be useful to have vocabularies. Strength of connection and the number of links in a path (link-length) are two of them.

It is possible to define a fuzzy adjacency matrix to represent a social network starting from real world information. If a value adjacency matrix $(V = v_{ij})_{n \times n}$ exist with $v_{ij} \in \mathbb{R}_{\geq 0}$, and v^* playing the role of the upper bound, we can rescale each v_{ij} into r_{ij} with a function $r_{ij} = h(v_{ij})$, $h: [0, v^*] \rightarrow [0,1]$ (Brunelli and Fedrizzi, 2009).

Compensatory Fuzzy Logic

Fuzzy logic is a multivalent logic system introduced by Zadeh (1965) in 1965 at the University of Berkeley (California). Although the many advantages of fuzzy logic to model ambiguous or vague knowledge it have certain drawbacks. The mains limitations in the modeling of knowledge can be summarized as (Alonso *et al.*, 2014):

1. The associative property of conjunction and disjunction operators.
2. The lack of sensitivity to changes in the values of truth of the basic predicates when calculating the truth value of compound predicates.
3. The lack of compensation for the truth values of basic predicates when calculating the accuracy of compound predicates.

CFL is s a variant of fuzzy logic that overcomes the preceding limitations. A CFL system is a quartet of operators: a conjunction, a disjunction, a negation and a strict fuzzy order that satisfies the axioms of compensation, commutativity,

strict growth, veto, fuzzy reciprocity, fuzzy transitivity and the Morgan's laws (Andrade, Fernández and González, 2014a; Andrade *et al.*, 2014c). In this work Geometric Mean Based Compensatory Logic (GBCFL) is used due to the robustness and relative simplicity of its operators (Ortega, Andrade and Gómez, 2013). In GBCFL conjunction is defined as follows:

$$c(x_1, x_2, \dots, x_n) = (x_1 \cdot x_2 \cdot \dots \cdot x_n)^{\frac{1}{n}} \quad (1)$$

The disjunction is defined as the dual of the conjunction:

$$d(x_1, x_2, \dots, x_n) = 1 - [(1 - x_1)(1 - x_2) \dots (1 - x_n)]^{\frac{1}{n}} \quad (2)$$

The fuzzy negation is:

$$n(x) = 1 - x \quad (3)$$

and the fuzzy strict order is:

$$o(x, y) = 0.5[c(x) - c(y)] + 0.5 \quad (4)$$

At any fuzzy predicate p over the universe U , universal and existential propositions are defined respectively as (Andrade *et al.*, 2014b):

$$\forall_{x \in U} p(x) = \bigwedge_{x \in U} p(x) \quad (5)$$

$$\exists_{x \in U} p(x) = \bigvee_{x \in U} p(x) \quad (6)$$

The min-max approach of the Zadeh family of operator (Wierman, 2010) used in the PISMA approach is not sensible to the changes in the truth values of predicates. With CFL we can express an “appealing” sensibility and attain more reliable operators according to the way that human take decisions. Combining the modeling capability of fuzzy graphs and CFL we can provide more realistic formulation of some concepts available in SNA as well as some new concepts.

Intelligent social network analysis based on CFL

The centrality measures have been used for the analysis in fuzzy graphs (e.g. (Samarasinghea and Strickert, 2011), (Yager, 2010)). The most used centrality measures to identifying a central node are: degree centrality, betweenness centrality, closeness centrality (Samarasinghea and Strickert, 2011). Another way of measuring the centrality of a vertex by the number of other vertices to which it directly connected by at most k steps (Yager, 2008):

$$C^k(x_i) = \sum_{\substack{j=1 \\ j \neq i}}^n R^k(v_i, v_j) \quad (7)$$

where $R^k(v_i, v_j)$ is the strength of the strongest path from v_i to v_j containing at most k links.

Strength (ST) of the path is usually defined as:

$$ST(\rho) = \text{Min}_{i=1 \text{ to } n} [R(v_{i-1}, v_i)] \quad (8)$$

One of the drawbacks of the *Min* operator is that it leads to a degradation of outcomes by compressing strength, for example on a given node, a *very strong* relation and a *moderate* one would have the same partial effect if the previous path strength of influence is *weak* (Montibeller and Belton, 2009).

In this paper instead we define degree of connection of a path (how strong is a path) taking into account the strength in all its edges:

$$S(\rho) = \forall_{i=1 \text{ to } n} [R(v_{i-1}, v_i)] \quad (9)$$

where \forall is the universal proposition in CFL.

Moreover the length of a path $L(\rho)$ can be defined as the number of edges the path contains. The word Far (F) can be defined a fuzzy subset $F: R \rightarrow [0,1]$ such that $F(k)$ is the degree to which a path of k links is *Far*. (Figure 1).

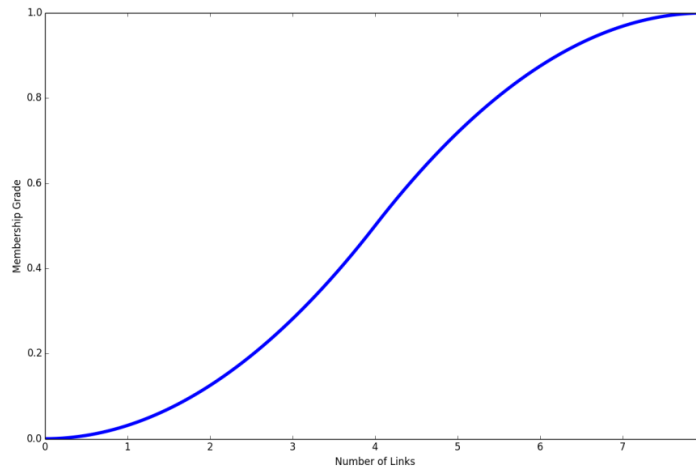


Figure 1. Example fuzzy set representation of far path.

We propose to find the better path between two vertices in terms of the negation of its farness and the degree of connection:

$$NFS(v_i, v_j) = \max_{\text{all paths from } i \text{ to } j} [\neg F(k) \wedge S(R^k(v_i, v_j))] \quad (10)$$

Where \neg and \wedge are the negation and conjunctive operator in CFL respectively.

Results and discussion

Coauthorship networks (Newman, 2004) are collaboration graph where nodes are scientific. Here we consider that two scientists are connected if they have appeared as authors in the same paper. Figure 2 shows the network formed by of research staff at the ISPAJE University inside the Eureka network.

Eureka network is a multinational scientific network. Its objective is the contribution to increase capabilities of useful knowledge discovering for the management of organizations in Iberian American region. For constructing coauthorship network three books published by Eureka have been taking into account (Andrade *et al.*, 2013; Espín *et al.*, 2011; Leyva *et al.*, 2013).

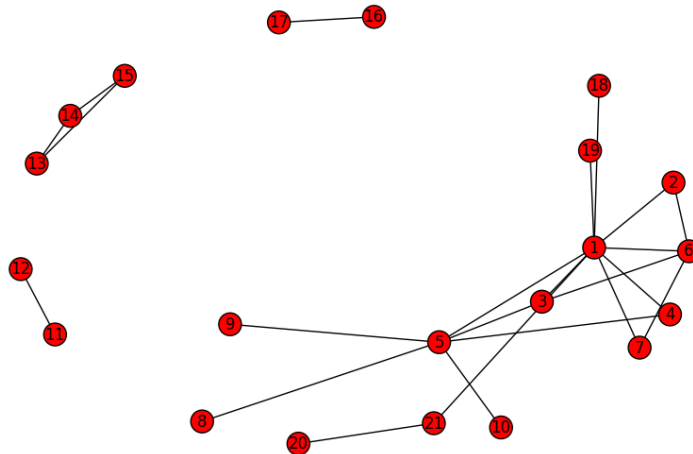


Figure 2. ISPAJE University coauthorship network.

In this case study we analyze the principal component of this network. The strength of connection among adjacency vertices is obtained using the sigmoid membership function with parameters a and b (fig 3):

$$\mu_a(x) = \begin{cases} 0, & \text{if } x \leq a \\ 2 \left[\frac{x-a}{b-a} \right]^2, & \text{if } a < x \leq m \\ 1 - 2 \left[\frac{x-b}{b-a} \right]^2, & \text{if } m < x < b \\ 1, & \text{if } x \geq b \end{cases} \quad (12)$$

where $m = \frac{a+b}{2}$.

In this case $a=0$ and $b = 5 = v^*$ (fig 3).

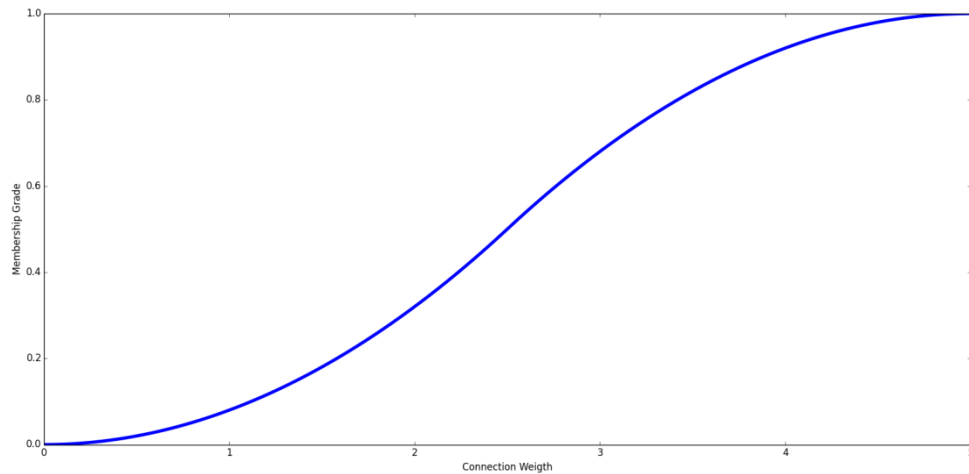


Figure 3. Fuzzy subset representation of term strength.

$F(k)$ is calculate with the sigmoid function (12) with parameter $a=0$, $b=8$ and $m = 4$ (fig 1).

The NFS values of paths among vertices 5 y 7 in shown in table I.

Table I. Value of *NFS* among v_5 and v_7

	<i>Path</i>	$\neg F(k) \wedge S(R^k(v_5, v_7))$
P_1	[5, 1, 2, 6, 7]	0.3188
P_2	[5, 1, 3, 6, 7]	0.3188
P_3	[5, 1, 6, 7]	0.4238
P_4	[5, 1, 7]	0.4677
P_5	[5, 3, 1, 2, 6, 7]	0.3361
P_6	[5, 3, 1, 6, 7]	0.5273
P_7	[5, 3, 1, 7]	0.7224
P_8	[5, 3, 6, 1, 7]	0.4917
P_9	[5, 3, 6, 2, 1, 7]	0.31785
P_{10}	[5, 3, 6, 7]	0.4145
P_{11}	[5, 4, 1, 2, 6, 7]	0.2277
P_{12}	[5, 4, 1, 3, 6, 7]	0.2277
P_{13}	[5, 4, 1, 6, 7]	0.32420
P_{14}	[5, 4, 1, 7]	0.3776

If a ranking were required, this is as follows:

$$P_7 \succ P_6 \succ P_8 \succ P_4 \succ P_3 \succ P_{10} \succ P_{14} \succ P_5 \succ P_{13} \succ P_1 \sim P_2 \succ P_9 \succ P_{11} \sim P_{12}$$

The better path in term of length and strength among researcher 5 and 7 ($NFS(v_5, v_7) = 0.7224$) is shown in conjunction with all paths among them in figure 4. The path found is neither shortest (P_4) nor strongest (P_6), different to results with the min-max approach of Zadeh operators (Wierman 2010) and more consistent with the way that human make decisions.

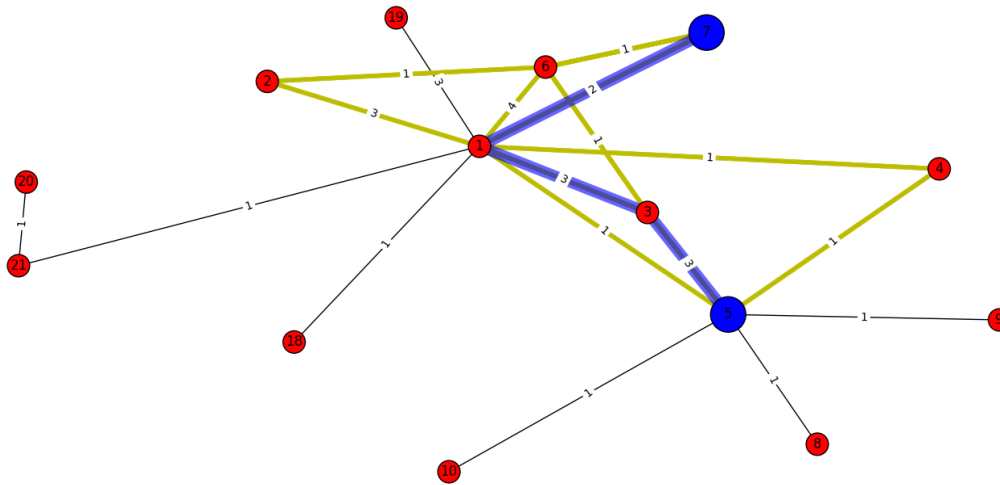


Figure 4. Better path in term NFS among researcher 5 and 7.

Comparing our proposal with the PISNA approach as proposed by Yager (2008), we notice that the former model will only give the most suitable or the shortest path not allowing to express compensation. Additionally our proposal will give more robust results because among fuzzy operators the those belonging to CFL are the most robust (Ortega, Andrade and Gómez, 2013). This results make our model appealing for combining decision making models and social networks analysis specially in recommendation tasks (Al Falahi, *et al.*, 2012) and consensus reaching (Brunelli and Fedrizzi, 2014). Our approach is an opportunity to use the language as key element of communication in the construction of semantic models in SNA (Yager, 2008). It is closer to the objective to make reasoning processes in environments of uncertainty and imprecision with words (Zadeh, 1999).

Conclusions

In this work we discussed the idea of fuzzy graph and their role in modeling social relational networks. Compensatory fuzzy logic was introduced and we discussed how these technologies can provide the analyst further flexibility for social network analysis. We applied this approach to the concept path importance taking into account the length and strength of the connection.

The new model for intelligent social network analysis based on CFL gives more robust outcomes and allow expressing compensation. The sensibility of the operators in our proposal gives more consistent results with the way that human make decisions. Moreover it bring closer the opportunity to use the language as crucial component in the construction of semantic models in SNA.

As future research we intend to develop models for combining decision making and social networks analysis. Discovering knowledge with linguistic summarization of social networks based on CFL is a promising area. The mining of social relation networks, the development of more flexible ways for querying social networks, and the development of a software tool are other areas of future work.

Acknowledgement

The authors are grateful to the anonymous reviewers for their valuable suggestions.

Author Maikel Leyva wants to thank to Dra. Macarena Espinilla from Jaén University and to MSc. Yasel Couce for their help.

References

- AL FALAHI, K., N. MAVRIDIS AND Y. ATIF. Social Networks and Recommender Systems: A World of Current and Future Synergies. In *Computational Social Networks*. Springer, 2012, p. 445-465.
- ALONSO, M. M., R. A. E. ANDRADE, V. L. BATISTA AND A. R. SUÁREZ. Discovering Knowledge by Fuzzy Predicates in Compensatory Fuzzy Logic Using Metaheuristic Algorithms. In *Soft Computing for Business Intelligence*. Springer, 2014, p. 161-174.
- ANDRADE, R. A. E., E. FERNÁNDEZ AND E. GONZÁLEZ. Compensatory Fuzzy Logic: A Frame for Reasoning and Modeling Preference Knowledge in Intelligent Systems In *Soft Computing for Business Intelligence*. Springer, 2014a.
- ANDRADE, R. A. E., E. FERNÁNDEZ AND E. GONZÁLEZ. Compensatory Fuzzy Logic: A Frame for Reasoning and Modeling Preference Knowledge in Intelligent Systems. In *Soft Computing for Business Intelligence*. Springer, 2014b, p. 3-23.
- ANDRADE, R. A. E., E. GONZÁLEZ, E. FERNÁNDEZ AND M. M. ALONSO. Compensatory Fuzzy Logic Inference. In *Soft Computing for Business Intelligence*. Springer, 2014c, p. 25-43.
- ANDRADE, R. A. E., R. B. PÉREZ, A. C. ORTEGA, J. M. GÓMEZ, *et al.*, *Soft Computing for Business Intelligence*. In.: Springer Berlin Heidelberg, 2013.
- BRUNELLI, M. AND M. FEDRIZZI. A fuzzy approach to social network analysis. In *Social Network Analysis and Mining, 2009. ASONAM'09. International Conference on Advances in*. IEEE, 2009, p. 225-230.

- BRUNELLI, M., M. FEDRIZZI AND M. FEDRIZZI Fuzzy m-ary adjacency relations in social network analysis: Optimization and consensus evaluation. *Information Fusion*, 2014, 17, 36-45.
- ESPÍN ANDRADE, R., MARX GÓMEZ J. AND R. V. A. Towards a transdisciplinary technology for Business Intelligence: Gathering Knowledge Discovery, Knowledge Management and Decision Making. In *Schriftenreihe der Oldenburger Wirtschaftsinformatik*. Aachen, 2011.
- JOHN, R. AND S. COUPLAND Extensions to type-1 fuzzy logic: Type-2 fuzzy logic and uncertainty. *Computational Intelligence: Principles and Practice*, 2006, 89-101.
- LEYVA-VAZQUEZ, M., K. PEREZ-TERUEL AND R. I. JOHN. A model for enterprise architecture scenario analysis based on fuzzy cognitive maps and OWA operators. In *Electronics, Communications and Computers (CONIELECOMP), 2014 International Conference on*. IEEE, 2014, p. 243-247.
- LEYVA LÓPEZ J. C., E. A. R., B. P. R. AND Á. C. P. Studies on Knowledge Discovery, Knowledge Management and Decision Support. In.: Atlantis Press, 2013.
- MONTIBELLER, G. AND V. BELTON. Qualitative operators for reasoning maps: Evaluating multi-criteria options with networks of reasons. In.: Elsevier, 2009, vol. 195, p. 829-840.
- NAIR, P. S. AND S. T. SARASAMMA. Data mining through fuzzy social network analysis. In *Fuzzy Information Processing Society, 2007. NAFIPS'07. Annual Meeting of the North American*. IEEE, 2007, p. 251-255.
- NEWMAN, M. E. Coauthorship networks and patterns of scientific collaboration. *Proceedings of the National Academy of Sciences*, 2004, 101(suppl 1), 5200-5205.
- ORTEGA, M., P. MICHEL, R. E. ANDRADE AND J. MARX GÓMEZ. Multivalued Fuzzy Logics: A Sensitive Analysis. In *Fourth International Workshop on Knowledge Discovery, Knowledge Management and Decision Support*. Atlantis Press, 2013.
- RICKARD, J. T. AND R. R. YAGER. Perceptual computing in social networks. In *IFSA World Congress and NAFIPS Annual Meeting (IFSA/NAFIPS), 2013 Joint*. IEEE, 2013, p. 691-696.
- SAMARASINGHEA, S. AND G. STRICKERT 2011. A New Method for Identifying the Central Nodes in Fuzzy Cognitive Maps using Consensus Centrality Measure. In *Proceedings of the 19th International Congress on Modelling and Simulation*, Perth, Australia, 12–16 December 2011 2011.
- WIERMAN, M. J. *An Introduction to the Mathematics of Uncertainty*. Edition ed. Omaha, Nebraska: Center for Mathematics of Uncertainty, Inc., 2010.

- YAGER, R. R. On ordered weighted averaging aggregation operators in multicriteria decisionmaking. *Systems, Man and Cybernetics, IEEE Transactions on*, 1988, 18(1), 183-190.
- YAGER, R. R. Intelligent social network analysis using granular computing. *International Journal of Intelligent Systems*, 2008, 23(11), 1197-1219.
- YAGER, R. R. Concept representation and database structures in fuzzy social relational networks. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, 2010, 40(2), 413-419.
- YAGER, R. R. Social Network Database Querying Based on Computing with Words. In *Flexible Approaches in Data, Information and Knowledge Management*. Springer, 2014, p. 241-257.
- ZADEH, L. A. Fuzzy sets. *Information and Control*, 1965, 8(3), 338-353.
- ZADEH, L. A. From Computing with Numbers to Computing with Words—From Manipulation of Measurements to Manipulation of Perceptions. *IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS—I: FUNDAMENTAL THEORY AND APPLICATIONS*, 1999, 45(1), 105.