A Multidisciplinary Survey of Social Network Diffusion Models

Paulo Shakarian

***Abstract*—Various models for the diffusion of information and behavior in a social network have been introduced in various disciplines. This paper (a companion to tutorials presented at IJCAI-2015 and AAAI-2016) provides an overview of several major families of models. In particular, we describe deterministic tipping, linear threshold/independent cascade, logic programming diffusion models, and evolutionary graph theory.**

***Index Terms*—social network analysis, social network diffusion, social influence**

I. Introduction

In recent years, research on diffusion process in social networks has grown in a variety of fields including computer science, physics, and biology. Recently, we have reviewed some of the major models in each of these disciplines in a tutorial we presented at IJCAI-2015 - which will also be presented at AAAI-2016. This paper gives an overview of these paradigms. Please refer to our recent book [1] for more detailed technical descriptions. [1] Specifically, we will review the following:

- *Deterministic Models.* The classic deterministic model first introduced for social networks by Mark Granovetter [2] is sometimes referred to as "opinion dynamics." Under this paradigm, each individual in a social network adopts a new behavior once the number of influencing friends previously adopting that behavior exceeds a certain threshold.
- *Independent Cascade and Linear Threshold Models.* Introduced in the seminal work of [3], these probabilistic models were designed to capture the intuition of various previously-introduced paradigms such as the susceptible-infected-recovered (SIR) model. They have become established as the standard models to study information diffusion in computer science.
- *Evolutionary Graph Theory.* Originally introduced to model the spread of a mutant gene in a structured population in the classic work of [4], these models are much-studied in theoretical biology and statistical physics. They are also used in research on game theory - primarily to study the conditions that can lead to the emergence of cooperation in a social network.
- *Logic Programming Models.* Leveraging years of established research from artificial intelligence, these frameworks allow for more fine-grain modeling of the conditions upon which influence among individuals occurs by allowing for the consideration of attributes of both individuals and their relationships.

We believe that by understanding the various models from a variety of disciplines, researchers can better understand which model is appropriate for a given application - or which model can be most easily modified to address a new research concern. For example, the traditional deterministic tipping, linear threshold, and independent cascade models make the "progressive assumption" - meaning that the number of adopters of a new behavior is increasing with time while this assumption is not made in evolutionary graph theory. Likewise, logic programming models allow for diffusion to also depend on the attributes of nodes and edges - which is generally not the case for the other paradigms. We do not argue for one model over the rest as a "one size fits all" solution but rather that one must consider various aspects of the models involved while considering them in a given application.

Throughout this paper, we will assume that there is an underlying population of n individuals amongst which there are m directed relationships - allowing us to represent the population as a graph $G = (V, E)$ where each $(i, j) \in E$ in interpreted as individual i having the ability to influence individual j. We use the notation $\eta_i^{(in)}$ and $\eta_i^{(out)}$ to denote the incoming and outgoing neighbors of individual i respectively. In some of the probabilistic models - such as independent cascade and evolutionary graph theory, we will use the notation p_{ij} associated with edge (i, j) to denote the probability of j being infected by i conditioned on i being infected previously. In other models (such as linear threshold and logic programming approaches) there is a weight associated with the edges - denoted w_{ij} which specifies a strength on the influence relationship but does not necessarily have a probabilistic interpretation. In many models, unweighted graphs are considered – which can often be treated as a special case of a weighted or probabilistic version of the model.

II. Deterministic Tipping Models

The *deterministic tipping model* sometimes referred to as *opinion dynamics* was initially studied in both sociology [2] and economics [5]. In this framework, individuals can be thought to be in one of two states - active (those who adopted the behavior) or inactive. In most work under this paradigm, individuals can only move from inactive to active. Each individual i in the population is associated with a threshold (κ_i). When κ_i individuals in the set $\eta_i^{(in)}$ are active, then individual i also becomes active (i.e. adopts that behavior). When an initial group if individuals adopts a new behavior (often called a *seed set*) they initiate a deterministic cascading process that must terminate in n steps or fewer.

Ariozna State University, Tempe, AZ, USA; e-mail shak@asu.edu.

[1] Slides for the IJCAI and AAAI tutorials, along with a preprint of the book can be found at http://lab.engineering.asu.edu/cysis/diffusion/.

Hence, while it is relatively simple to simulate a cascading process under deterministic tipping dynamics, a natural problem to study is can we identify a seed set of size k such that at least x number of individuals in the population are active. This is often referred to as the *target set selection* problem or when k is sought to be minimized the *min-seed* problem. Dryer and Roberts [6] introduce this problem and prove it to be NP-hard - even in the case of certain threshold settings (i.e. when the threshold for all individuals in the network is 2). The hardness of approximation for this problem is described in [7]. The work of [8] presents an algorithm for target-set selection whose complexity is determined by the tree-width of the graph. The work of [9] proves a non-trivial upper bound on the smallest seed set. Despite the intractability of this problem and associated difficulty of approximation, scalable heuristics are available that can find small seed sets in practice [10][11]. However, there are drawback with deterministic tipping dynamics - specifically that it makes the *progressive* or *monotonic* assumption - in that the number of active individuals increases with time. Further, as it is deterministic, it does not represent uncertainty. However, real world uses are possible - for instance in [12] it was used as a way to create effective features in a graph-based machine learning problem.

III. The Linear Threshold and Independent Cascade Models

One way to address the issue of determinism in the tipping model is to have all nodes draw their thresholds from a uniform random distribution - the intuition being that actual thresholds will be difficult to observe in practice. Such a model was introduced in [3] and is known as the *linear threshold* (LT) model. A related model, the *independent cascade* (IC) model was also introduced in the same paper. In the IC model, each edge is associated with a probability (as described in the introduction). So, when node i is infected in a given time step, it has a single chance to infect each outgoing neighbor j with a probability p_{ij}. This model can be considered a variant of the popular susceptible-infected-recovered (SIR) model that is well-studied in epidemiology and physics [13][14]. In a similar manner, non-negative real-valued weights are assigned to edges of the graph in the LT model such that for each node j the quantity $\sum_{i \in \eta_j^{in}} w_{ij}$ is less than or equal to one. Hence, the threshold for each node is selected uniformly at random from the interval $[0, 1]$ and the node is active when the sum of incoming active weights exceeds the threshold.

As the LT and IC models are stochastic, the quantity often studied is the expected number of active nodes upon completion of the diffusion process. For a given seed set $S \subseteq V$, the expected number of active nodes is often denoted $\sigma(S)$. It turns out that evaluation of σ is #P-hard for both models [15][16] and often simulation runs are used to approximate this value - though several heuristics are available - notably MIIA for IC [15] and SIMPATH-SPREAD for LT [16].

The reduction used to show the #P-hardness of calculating $\sigma(S)$ used a proof technique called the *live edge model*. This technique often used in the formal analysis of the IC and LT models. With this technique, the stochastic process is mapped to a set of deterministic processes that each occur in a subgraph of G - each of which is considered as a possible (and disjoint) world and can be associated with a probability based on the model. For example, in the IC model, the probability associated with subgraph $G' = (V, E')$ is $\prod_{(i,j) \in E'} p_{ij} \times \prod_{(i,j) \in E \setminus E'} (1 - p_{ij})$. Note that within a given subgraph G' (often referred to as a *realization* of the diffusion process), the expected number of infected nodes given seed set S is simply all nodes in G' for which there exists a path from S in that graph (often termed *reachability* and denoted $R_{G'}(S)$).

One important result shown under both IC and LT models shown using the live edge model is the *submodularity* of the σ function. The intuition behind this mathematical property is that there are diminishing returns. Formally, for $S' \subseteq S \subseteq V$ and $i \in V \setminus S$ we have:

$$\sigma(S \cup \{i\}) - \sigma(S) \leq \sigma(S' \cup \{i\}) - \sigma(S')$$

Hence adding node i provides a larger increase to the expected number of active nodes when added to a subset. Submodularity of σ follows from the submodularity of reachability and that, using the live-edge model, $\sigma(S)$ is equal to a positive linear combination of submodular functions (which is also submodular).

The property of submodularity plays an important role in the *influence maximization* problem - the stochastic analogue to the target set selection problem. In this problem, one seeks to find a set $S \subseteq V$ of size k or less such that $\sigma(S)$ is maximized. Even with access to an oracle that can efficiently compute σ, the influence maximization problem for both IC and LT is NP-hard by reductions from well-known combinatorial problems [3]. However, as σ is submodular, monotonically increasing (for $S' \subseteq S, \sigma(S') \leq \sigma(S)$), and normalized ($\sigma(\emptyset) = 0$), then by the result of [17], the standard greedy algorithm provides a $1 - 1/e$ approximation (where e is the base of the natural logarithm) under the assumption that there is access to an oracle for σ.

Another model known as the *generalized threshold* model is shown to capture both LT and IC as special cases. In this model, each node i is associated with a function $f_i : 2^{\eta_i^{(in)}} \rightarrow [0, 1]$ which maps subsets of active incoming neighboring nodes to a normalized non-negative real number. In this model, each node again selects a threshold (i.e. θ_i) uniformly at random and the node is activated when for a set of active in-neighbors (η') the function f_i exceeds the threshold ($f_i(\eta') \geq \theta_i$). In a very interesting result, when the associated f_i function is submodular for each node i, then computing the the expected number of infectees under this model is also submodular - allowing for the greedy approximation even in this more general case.

IV. Evolutionary Graph Theory

Another important class of stochastic diffusion models that has received much attention is known as evolutionary graph theory (EGT). Originally introduced by [4], EGT studies the ability of a mutant gene to overtake a finite structured

population. Here the population's structure is a directed graph and the progression of the mutant gene through the population is the diffusion process. Since its introduction, numerous results on EGT, both analytical and experimental, have been produced - see the survey [18] for an overview. Additionally, several extensions to the model have been proposed, including game-theoretic ones. The application of EGT to game theory has provided researchers new insight about the evolution of cooperation and other game-theoretic concepts in structured populations.

The dynamics of EGT is an extension of an earlier model of the spread of a mutant gene in a population of n individuals where there is no specified graph-structure relating them to each other (this is known as a *well-mixed* population). The *Moran Process* of [19] is a stochastic process used to model evolution in such a population. It is defined as follows. At each time-step a randomly selected individual is chosen to reproduce. Then, a second individual is chosen at random to die – replaced by a duplicate of the first individual. Individuals are selected for reproduction based on *fitness*. Typically, each individual is assigned one of two labels - *resident* and *mutant* - and often residents are assigned a fitness of 1 and mutants are assigned a fitness of r - a positive real value. The mutant is *advantageous* if $r > 1$ and disadvantageous when $r < 1$. The case where $r = 1$ is known as *neutral drift.* An often-studied problem is determining the probability that a single mutant will eventually overtake the population. This is known as the *fixation probability* (the opposite event - that all mutants die out - is called *extinction* and a population with a lower fixation probability is deemed more *evolutionarily stable* as it is resistant to invasion by a mutant). This probability, ρ_1, arising from this n original Moran Process, is often termed the *Moran probability* and can be shown to be equal to the quantity $\frac{1-1/r}{1-1/r^n}$.

In the original work that introduced EGT [4], Lieberman et al. generalize the model of the Moran Process by specifying relationships between the n individuals of the population in the form of a directed, weighted graph (again, we will use the notation $G = (V, E)$). We also assume a probability associated with each edge - just as with the IC model, except here $\forall i$, $\sum_j p_{ij} = 1$. The dynamics proceed as follows. At each step, first an individual is selected from the population proportional to its fitness (just as with the standard Moran process, this is $r/(n_{mutant}r + n - n_{mutant})$ for mutants and $1/(n_{mutant}r + n - n_{mutant})$ for residents – where n_{mutant} is the number of individuals in the population with a mutant label). This individual is selected for "birth." Then, a single outgoing neighbor j of node i is chosen with a probability p_{ij}. Individual j then "dies" and is replaced with a clone of node i. In other words, j adopts i's label for the next iteration. Again, a key problem explored in the literature on EGT is to determine the fixation probability - the probability that all members in the population adopt a mutant label given an initial invasion of mutants.

There has been much research on the computation of fixation probability in EGT. To compute this value for an initial, single, randomly-placed mutant, [4] shows that the network structure plays a significant role in this computation as this is only equal to the Moran probability for a special class of graphs referred to as *isothermal* that is for all nodes (i), the quantity $\sum_j p_{ji}$ is the same. This quantity is often called the *temperature* as nodes will change label more often if it is higher (hence in *isothermal* graphs the temperature is the same for all nodes). Many researchers [20][21][24][22][23] have studied the problem of computing the probability of fixation given that a certain subset of nodes are mutants. If the mutants inhabit set $C \subseteq V$, then this probability is written P_C. Hence the fixation probability for a randomly selected mutant (ρ) is simply the average of the P_C for all singleton sets. In [25] the authors provide a set of linear constraints for solving for P_C - though there are an intractable number of these constraints. As with LT and IC, simulation is often used to estimate fixation probabilities. However, analytical results are available in many special cases of graphs and algorithms such as that of [22] can provide faster approximations for certain cases.

One of the most popular applications of EGT is game theory. In the game theoretic context, nodes of a graph represent agents and edges represent potential for interaction between them. Interactions between agents are games played that can be described using a normal game theoretic payoff matrix. EGT thus provides a structural component for interactions in populations of agents. Evolutionary game theory, which is concerned with the population-dependent success of game theoretic strategies, has initially mostly focused on well-mixed populations in which interactions between all agents are equally likely. Combining EGT with evolutionary game theory can take into account the effect of population structure, which has the capacity to crucially impact evolutionary trajectories, outcomes, and strategy success. Thus EGT is a welcome tool to explore how many of the results for well-mixed populations are affected by population structure. In game-theoretic applications of EGT, the evolutionary fitness (f_i) of individual i is often related to their game theoretic payoff ($\boldsymbol{p}$) (based on game-play with neighbors) with the following relationship: $f_i = 1 - \boldsymbol{w} + \boldsymbol{w} \cdot \boldsymbol{p}$. Where the parameter $\boldsymbol{w}$ relates the payoff acquired from games played to fitness. If $\boldsymbol{w} = 1$, the payoff acquired is equal to the fitness. If $\boldsymbol{w} = 0$, the game is irrelevant and we are at neutral drift. An often explored special case is *weak selection*, where $\boldsymbol{w} << 1$, which reflects the assumption that the game of interest plays only a partial role in the overall fitness of individuals. Using this paradigm, researchers have reached a variety of important conclusions on the effects of population structure on game-theoretic concepts. For instance, Santos et al. [26] investigate the effects of single-scale and scale-free networks on cooperation in the Prisoner's Dilemma, Snow-Drift, and Stag-Hunt games through simulations. The authors find that in degree-heterogeneous graphs cooperation is easier to sustain than in well-mixed populations and thus identify heterogeneity as a "powerful mechanism for the emergence of cooperation." Additionally, the authors find that the sustainability of cooperation also depends on"detailed and intricate ties" between agents. As evidence of this, scale free networks which exhibit properties like those that emerge from models of growth from preferential attachment (Albert-Barbarasi topology) are shown to produce higher cooperation than random scale-free networks.

V. Logic Programming Based Frameworks

Attributes about individuals within a social network, along with characteristics about the relationships among them, can play a significant role in diffusion. For instance, a close friend may have a stronger influence relationship than an office co-worker. Likewise, individuals of different ages, genders, and education levels may respond to various social contagion in different ways. While models such as deterministic tipping, IC, and LT can capture the structure of a population, they do not inherently capture attributes of the individuals, their relationships, and the social contagion itself. Logic programming brings a natural representation of these additional factors - along with a suite of long-established results. The intuition is that a graph with multiple labels on nodes and edges is embedded into a logic program - along with additional rules that specify complex diffusion relationships.

The logic-programming approach to social network diffusion first introduced in [27] and later extended in [28]. Since its introduction, there have been other variants of the logic-based approach that have leveraged formalisms such as probabilistic soft logic (PSL) [29] and modal logic [30] in addition to tackling problems such as non-monotonic diffusion reasoning [31] and informing the creation of diffusion-specific centrality measures [32]. A key advantage is with these frameworks is that they do not specify a single diffusion model, but rather provide a language for reasoning about a whole class of diffusion models. These approaches even allow for the composition of models - enabling reasoning about multiple diffusion processes that occur at the same time and potentially interact.

The well-known annotated logic - Generalized Annotated Programs (GAP) was the first to be adapted for social network diffusion [27][28]. In this case, a social network was defined as a 5-tuple: $(V, E, \ell_{node}, \ell_{edge}, w)$ where V is the set of nodes, E is a *multi-set* of relationships, ℓ_{node}, ℓ_{edge} are functions that label the nodes and edges respectively, and w is a weighting function that assigns weights to multi-edges. This structure can be easily embedded into a logic program along with associated diffusion rules.

To provide more concrete intuition for how a social network and associated diffusion processes can be embedded in a logic program, consider Figure 1 which shows a toy social network the cell phone company might use. Here, we might have a set of node labels $\{male, female, adopter, temp\_adopter, non\_adopter\}$ denoting the sex and past adoption behavior of each vertex; and a set of edge labels $\{phone, email, IM\}$ denoting the types of interactions between nodes (phone call, email, and instant messaging respectively). The function ℓ_{node} is shown in Figure 1 by the shape (denoting past adoption status) and shading (male/female). The type of edges (bold for phone, dashed for email, dotted for IM) is used to depict ℓ_{edge}. $w(v_1, v_2)$ denotes the percentage of communications of type $\ell_{edge}(v_1, v_2)$ initiated by v_1 that were with v_2 (measured either w.r.t. time or bytes).

We can easily embed this social network into a GAP. For instance, we would include the rule $female(v_1) : 1 \leftarrow$ meaning that node v_1 is assigned an annotation of 1 (signifying truth). Likewise, $IM(v_3, v_1 3) : 0.4 \leftarrow$ specifies that there is an instant messaging relationship between v_3 and $v_1 3$ with a weight of 0.4. We can specify the diffusion process through GAP rules as follows (note we use capital letters to denote nodes as these refer to arbitrary nodes rather than specific ones).

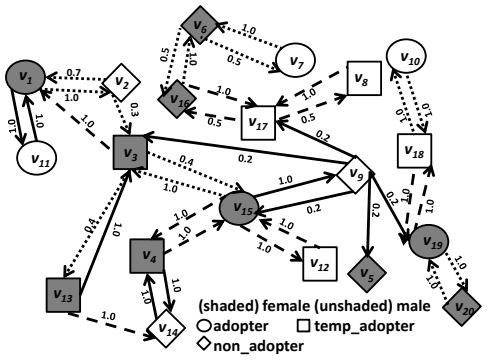

Fig. 1. Example cellular social network.

1) $will\_adopt(V_0) : 0.8 \times X + 0.2 \leftarrow adopter(V_0) : 1 \wedge male(V_0) : 1 \wedge IM(V_0, V_1) : 0.3 \wedge female(V_1) : 1 \wedge will\_adopt(V_1) : X.$
2) $will\_adopt(V_0) : 0.9 \times X + 0.1 \leftarrow adopter(V_0) : 1 \wedge male(V_0) : 1 \wedge IM(V_0, V_1) : 0.3 \wedge male(V_1) : 1 \wedge will\_adopt(V_1) : X.$
3) $will\_adopt(V_0) : 1 \leftarrow temp\_adopter(V_0) : 1 \wedge male(V_0) : 1 \wedge email(V_1, V_0) : 1 \wedge female(V_1) : 1 \wedge will\_adopt(V_1) : 1.$

Rule 1 says that if V_0 is a male adopter and V_1 is female and the weight of V_0's instant messages to V_1 is 0.3 or more, and we previously thought that V_1 would be an adopter with confidence X, then we can infer that V_0 will adopt the new plan with confidence $0.8 \times X + 0.2$. The other rules may be similarly read.

Due to the results of [33], determining the outcome of a diffusion process under this model can be computed efficiently under some natural assumptions. However, a solving a *social network diffusion optimization problem* (SDNOP) in such a framework (the analogue to influence maximization or target set selection) remains NP-hard as the tipping model can be easily embedded into this framework. There are also special cases of GAPs where the diffusion process exhibits submodularity (known as *linear GAPs*) and allow for the greedy approximation as in the case of IC and LT. However, it should be noted that, in general, the annotations associated with the atomic propositions in this framework are not necessarily probabilistic - and the efficiency of the progression of diffusion in this framework precludes exact embeddings of probabilistic models such as IC, LT, and EGT.

VI. Conclusion

This paper surveyed some of the major social network diffusion models from a variety of disciplines and described some key results. However, this area of study will continue to evolve. Lately, network diffusion research where historical traces of diffusion processes are available are becoming more prevalent – and empirical studies examining influence and diffusion in such datasets will lead to further refinements of these models - and perhaps result in new paradigms.

Acknowledgment

The author is supported through the AFOSR Young Investigator Program (YIP) grant FA9550-15-1-0159, ARO grant W911NF-15-1-0282, and the DoD Minerva program.

References

[1] P. Shakarian, A. Aleali, A. Bhatnagar, R. Guo, and E. Shaabani, *Diffusion in Social Networks*, Springer, 2015.

[2] M. Granovetter, "Threshold models of collective behavior." *The American Journal of Sociology* (6), 1420-1443, 1978.

[3] D. Kempe, J. Kleinberg, and . Tardos, "Maximizing the spread of influence through a social network." *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining.* ACM, 2003.

[4] Lieberman, E., Hauert, C., Nowak, M. A., "Evolutionary dynamics on graphs." *Nature* 433 (7023), 312-316, 2005.

[5] T. Schelling, *Micromotives and Macrobehavior*, W.W. Norton and Co., 1978.

[6] P. Dreyer, F. Roberts, "Irreversible threshold processes: Graph-theoretical threshold models of the spread of disease and of opinion," *Discrete Applied Mathematics*, 157 (7), 1615-1627, 2009.

[7] N. Chen, "On the approximability of influence in social networks," *SIAM J. Discrete Math,* 23, 1400-1415, 2009.

[8] O. Ben-Zwi, D. Hermelin, D. Lokshtanov, I. Newman, I., "Treewidth governs the complexity of target set selection," *Discrete Optimization,* 8 (1), 87-96, 2011.

[9] D. Reichman, "New bounds for contagious sets," *Journal, Discrete Mathematics,* Volume 312 Issue 10, May, 2012.

[10] P. Shakarian, D. Paulo, "Large Social Networks can be Targeted for Viral Marketing with Small Seed Sets," *IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining,* 2012.

[11] P. Shakarian, S. Eyre, and D. Paulo, "A Scalable Heuristic for Viral Marketing Under the Tipping Model," *Social Network Analysis and Mining,* Springer 3(4), 2013.

[12] E. Shaabani, A. Aleali, P. Shakarian, and J. Bertetto, "Early Identification of Violent Criminal Gang Members," *21st ACM SIGKDD Conference on Knowledge, Discovery, and Data Mining*, 2015.

[13] R. Anderson, R. May, "Population biology of infectious diseases: Part I," *Nature,* 280 (5721), 361, 1979.

[14] M. Kitsak, L. Gallos, S. Havlin, F. Liljeros, L. Muchnik, E. Stanley, H. Eugene, and H. Makse, "Identification of influential spreaders in complex networks," *Nat Phys,* 6 (11), 888–893, 2010.

[15] W. Chen, C. Wang, and Y. Wang, "Scalable influence maximization for prevalent viral marketing in large-scale social networks," *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining,* 1029–1038, 2010.

[16] A. Goyal, W. Lu, L.V. Lakshmanan, "SIMPATH: An efficient algorithm for influence maximization under the linear threshold model," *11th International Conference Data Mining (ICDM)*, IEEE, 211–220, 2011.

[17] G. Nemhauser, L. Wolsey, and M. Fisher, "An analysis of approximations for maximizing submodular set functions," *Mathematical Programming,* 14(1) 265–294, 1978.

[18] P. Shakarian, P. Roos, and A. Johnson, "A Review of Evolutionary Graph Theory with Applications to Game Theory," *BioSystems*, 107(2), 2012.

[19] P. Moran, "Random processes in genetics," *Mathematical Proceedings of the Cambridge Philosophical Society,* 54 (01), 60-71, 1958.

[20] M. Broom, J. Rychtar, "An analysis of the fixation probability of a mutant on special classes of non-directed graphs," *Proc. of the Royal Society A,* 464, 2609-2627, 2008.

[21] N. Masuda, H. Ohtsuki, "Evolutionary dynamics and fixation probabilities in directed networks," *New Journal of Physics,* 11, 033012, 2009.

[22] P. Shakarian, P. Roos, "Fast and deterministic computation of fixation probability in evolutionary graphs," *The Sixth IASTED Conference on Computational Intelligence and Bioinformatics,* 2011.

[23] P. Shakarian, P. Roos, and G. Moores. "A Novel Analytical Method for Evolutionary Graph Theory Problems," *BioSystems,* 111(2), 2013.

[24] V. Barbosa, R. Donangelo, and S. Souza, "Early appraisal of the fixation probability in directed networks," *Phys. Rev. E,* 82 (4), 046114, 2010.

[25] J. Rychtar, B. Stadler, "Evolutionary dynamics on small-world networks," *International Journal of Computational and Mathematical Sciences,* 2 (1), 2008.

[26] E. Santos, J. Pacheco, and T. Lenaerts, "Evolutionary dynamics of social dilemmas in structured heterogeneous populations," *PNAS* 103 (9), 3490-3494, 2006.

[27] P. Shakarian, V.S. Subrahmanian, and M.L. Sapino, "Using Generalized Annotated Programs to Solve Social Network Optimization Problems," *26th Intl. Conference on Logic Programming,* 2010.

[28] P. Shakarian, M. Broecheler, V.S. Subrahmanian, and C. Molinaro, "Using generalized annotated programs to solve social network diffusion optimization problems," *ACM Transactions on Computational Logic,* 14(2), 2013.

[29] M. Broecheler, P. Shakarian, and V. S. Subrahmanian. "A scalable framework for modeling competitive diffusion in social networks," *IEEE Conference on Social Computing,* IEEE, 2010.

[30] Z. Christoff, J. Hansen, "A logic for diffusion in social networks," *Journal of Applied Logic,* 13(1), 2015.

[31] P. Shakarian, G.I. Simari, and D. Callahan, "Reasoning about Complex Networks: A Logic Programming Approach," *29th Intl. Conference on Logic Programming,* 2013.

[32] C. Kang, C. Molinaro, S. Kraus, Y. Shavitt, and V.S. Subrahmanian, "Diffusion Centrality in Social Networks," *IEEE ASONAM,* 2012.

[33] M. Kifer, V.S. Subrahmanian, "Theory of generalized annotated logic programming and its applications," *The Journal of Logic Programming,* 12(4), 1992.