This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

The Affective Evolution of Social Norms in Social Networks

Seyyed Hadi Sajadi, MohammadAmin Fazli, and Jafar Habibi

***Abstract*—Social norms are a core concept in social sciences and play a critical role in regulating a society's behavior. Organizations and even governmental bodies use this social component to tackle varying challenges in the society, as it is a less costly alternative to establishing new laws and regulations. Social networks are an important and effective infrastructure in which social norms can evolve. Therefore, there is a need for theoretical models for studying the spread of social norms in social networks. In this paper, by using the intrinsic properties of norms, we redefine and tune the Rescorla–Wagner conditioning model in order to obtain an effective model for the spread of social norms. We extend this model for a network of people as a Markov chain. The potential structures of steady states in this process are studied. Then, we formulate the problem of maximizing the adoption of social norms in a social network by finding the best set of initial norm adopters. Finally, we propose an algorithm for solving this problem that runs in polynomial time and experiments it on different networks. Our experiments show that our algorithm has superior performance over other methods.**

***Index Terms*—Conditioning, Markov chains, norm, Rescorla–Wagner model, social networks, spread.**

I. Introduction

SOCIAL networks have become one of the most important social organizations. In this space, many connected users participate in innovation, social development, and the creation of new ideas. It is through such collaboration that social networks have a large amount of impact on the arts, culture, science, education, and so on [1]. These interactions can also help in evolving social norms based on the needs of a society.

Social norms, also called the social grammar [2], are standards of behavior which are accepted in a group or community, where violating them results in some sort of punishment. Norms play an important role in achieving the goals of a society, especially to regulate (order) various behaviors [3].

Social norms specify what is acceptable in a group, so they can even be used to assure adherence to society's laws [4]. In some cases in societies, social norms are the main tool for maintaining order [5]. Different claims are made by scholars about the interaction between informal rules (social norms) and formal rules (laws), some think that they are complementary and some think they are substitutes, but there is a serious belief in the context that formal rules are costly [6]. Thus, researchers in various fields have focused on creating and evolving social norms (see [7]–[9]).

Manuscript received October 27, 2017; revised April 4, 2018; accepted July 8, 2018. *(Corresponding author: MohammadAmin Fazli.)*

S. H. Sajadi is with the School of Science and Technology, Sharif University of Technology International Campus, Kish 79417-76655, Iran.

M. Fazli and J. Habibi are with the Sharif University of Technology, Tehran 11155-8639, Iran (e-mail: fazli@sharif.edu).

Digital Object Identifier 10.1109/TCSS.2018.2855417

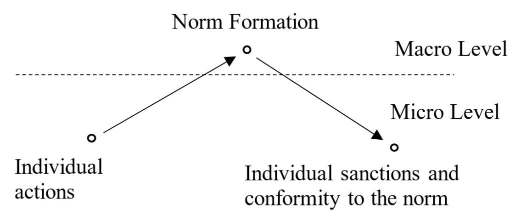

Fig. 1. Micro and macro levels in the emergence of a norm [11]. According to this model, the formation of norms starts at the micro level and continues to the macro level and ends again at the micro level, where conformity to the norms is guaranteed using sanctions.

To create a social norm, we need to devise methods for promulgating the norm and also sanctions that can prevent their violation [10]. This is based on James Coleman's model for social norms [11]. This model, as can be seen in Fig. 1, starts at the micro level, where individual interact, continues to the macro level, where the norm is formed, and finally ends again at the micro level, where conformity to the norms is guaranteed using sanctions [9].

Any model for social-norm evolution must provide solutions for steps 1 and 3 of Coleman's model at the micro level. In this paper, we emphasize on step 3, i.e., promulgation, where for analyzing the nature of social norms, we want to define a new method to model the process in which an actor adopts a social norm. In our model, we focus only on the effective aspects of this process and discard other factors like cognition (see [12]). We make use of methods based on classical conditioning and associative learning. These methods are the main tool for understanding many aspects and have many practical uses in different fields such as psychotherapy and law (see [13]–[19]).

Classical conditioning, also called the Pavlov conditioning and associative learning, is a learning procedure in which different stimuli are used on an organism for learning a habit or behavior. Classical conditioning is composed of the following components [13].

1) *Unconditional Stimulus (US):* Food.
2) *Unconditional Response (UR):* Food-induced salivation.
3) *Conditional Stimulus (CS):* Bell.
4) *Conditional Response (CR):* Bell-induced salivation.

The examples given for each above-mentioned component come from the classic experiment done using classical conditioning in which US food alongside CS bell was used for conditioning dogs (the organism) so that they would salivate when hearing the sound of a bell. Using associative learning, we can see that when the two types of stimulus happen concurrently many times, CS can alone to induce the CR (here salivation) [12]. In our model, promulgation of norms is done using a classical conditioning procedure. We assume that people are conditioned through their interactions with their friends, family, and other people with whom they have social relations.

One of the most famous models in the context of conditioning and associative learning is the Rescorla–Wagner model that has been found very useful for describing animals as well as human conditioning in a variety of contexts [20]. One of the main such works is the work of Epstein [12], which uses this model to analyze social behavior. In this model, the following formula is used to correlate learning to the amount of effort (the number of tries required to condition the organism) [21]:

$$\upsilon_{t+1} = \upsilon_t + \alpha(\lambda_t - \upsilon_t) \tag{1}$$

where υ_t is the amount of learning after tth trial (the associative strength between the CS and US), α is the learning parameter, and λ_t is the intensity of the US on that trial.

In this paper, we use the Rescorla–Wagner model as a basis for modeling the conformance of people to a social norm. Then, we propose an extension of it to model the spread of social norms in a social network. We show that our model can be considered as a standard Markov chain process. Then, we study the steady states of this process and extract their form. We show that the structure of the network and the initial condition of nodes affect the steady state. Finally, we propose an efficient algorithm to find the initial setting for which the maximum amount of conformance to a social norm happens and compare our algorithm with different heuristics and methods. Our experiments on different networks show that the amount of conformance to a social norm which can be achieved by our algorithm is much larger than those methods.

This paper is structured as follows. In the continuation of this section, we present the related works. In Section II, we propose our extension to the Rescorla–Wagner model and show that how it can be represented as a Markov chain process. In Section III, we analyze the steady states of our model and find their potential forms and show how the structure of the network and the initial state of nodes can affect these states. In Section IV, the problem of maximizing the spread of social norms in a given social network is defined and an efficient algorithm for solving this problem is proposed. Finally, in Section V, we establish an experiment to check the performance of our algorithm in comparison to other algorithms which can be used to solve the target problem.

A. Related Works

Much research has been done on utilizing social norms to tackle various social problems. In this regard, a large body of work on the topic of social norms has been conducted by Deitch-Stackhouse *et al.* [22]. This line of research includes ways to tackle social challenges like sexual violence, alcoholism, and smoking in schools and universities. This is in line with an approach called the social norm approach where positive aspects of adopting a norm are emphasized instead of prioritizing negative aspects such as fear. For example, one of the results shows that only 5% of men conduct sexual assault, and an effective way to control this 5% is to make use of the other 95%. In another example, the work of Rahwan [23] about using artificial intelligence for regulating societies and social contract is significant. This research explains how we need tools to program an algorithmic social contract between various human stakeholders, mediated by machines. Social norms can be used in this approach to avoid violation of social contracts.

Some other studies have focused on specific norms inside social networks. McLaughlin and Vitak [24] use an expectation violation theory [25] to analyze and extract social norms related to the ethics of social interactions and activities on Facebook. Social networks are also an optimum medium for promulgating social norms. References [8] and [9] use this medium to promote social norms related to controlling the spread of the HIV virus. In another example, Dickie *et al.* [26] showed that social norm can be used for promulgating hand wash and is an effective method of preventing the spread of infectious illness.

What we emphasize in this paper is the evolution of social norms by utilizing the social networks structure. Much research in this context has been done on developing behaviors in social networks [7], [27], [28]. For example, in [7] from MIT's Media Laboratory utilizes behavior analysis methods in networks. Setting social networks to decrease the speed of dissemination of ideas to prevent the herd behavior in choosing strategies and approaches is one of the most important technology-based methods used by him for changing behaviors.

Another related research areas are social contagion [29], [30], spread of influence [31]–[33], diffusion of innovation [34], and cascading behavior [35]. In all of these research areas, the spread of different types of information between various nodes in networks is analyzed, and efficient algorithms to facilitate such flow that can be social norms are devised and proposed.

To the best of our knowledge, [21] and [31] are the nearest work to the findings of our paper. Kempe *et al.* [31] propose an algorithm for finding the best set of initial adopters for spreading an influence in a social network. This paper proposes two simple models for spreading influences through social networks that are called the threshold model and the cascade model. Then, it proves that the NP-hardness of finding the best set of initial adopters for both of these models and designs is an approximation algorithm for this problem. Finally, it conducts experiments for comparing their algorithm with simple heuristics. In our paper, we use the same approach for the spread of social norms.

Wei *et al.* [21] propose a generalized version of the Rescorla–Wagner model based on Markov chains to enable the study of connections among agents and their effects on the dynamic of the learning process. To this end, the author replaces all scalars with matrices. Then, the stability and convergence issues of the proposed model are studied. Our paper, on the other hand, proposes a specialized extension of the Rescorla–Wagner model for the spread of social norms, and the initial set selection problem is studied.

II. Proposed Model

The disposition to social norms and adherence to them is a gradual process [36], where accepting a social norm by a node is related to the number of trials that his neighboring nodes apply [12].

As we previously noted in Colman's model for social norm individual actions play a vital role in increasing a node's disposition to accept a social norm. To model such actions, different methods have been proposed, from the marginal utility theory [11] to the game theory [3], [37]. In our model, we use Epstein's intuition in utilizing the Rescorla–Wagner conditioning theory. Similarly, other studies have also used the theory of Rescorla–Wagner to propose neurocomputational models and their applications in social norm creation [38].

Before formally expressing the model, consider the following example. Consider an enterprise in which the collaborations among employees form a network. In this enterprise, we want to promulgate the social norm of "having formal interactions in the work environment." There is an initial set of nodes (employees) whose interactions are currently in the formal form. When these employees interact with their friends in the work environment, they may be influenced by their behavior and thus become inclined to the norm. We can model this situation in a classical conditioning setting. Assume that a man who does not adhere to the norm (person A) has linked with a person who adheres to the norm (person B).

1) *US:* B interacts with A according to the norm.
2) *UR:* The disposition toward the norm induced by US.
3) *CS:* The social position of B and her influence on A.
4) *CR:* The disposition toward the norm induced by CS.

In accordance to [12], when the US and the CS are repeated simultaneously for a number of times for an employee, he becomes conditioned, and therefore, the US is not needed and the desired effect is achieved only by CS. Therefore, in each interaction in the work environment, even with one with a low impact on him, he will act formally. After this phase, these nodes' adherence to the social norm may influence their colleagues and make them inclined to the social norm too.

This scenario happens for many other social norms. Generally, assume that we have a network $\mathcal{N}$ with n nodes for which some of its nodes (the set S) are initially adhered to one of such social norms. In our model, we hire the Rescorla–Wagner conditioning formula (1) to model the spread of these norms in a social network. To do this, let us first define the intensity of US for a node as the average amount of adherence of his neighboring nodes to the social norm. Therefore, we have

$$\lambda_t^i = \frac{\sum_{j \in N^i} v_t^j}{\deg^i(\mathcal{N})} \tag{2}$$

where

1) λ_t^i is the intensity of the US in the tth simultaneous occurence of the CS and the US.
2) $v_t^i \in [0, 1]$ is the amount of node i's adherence to the social norm after tth simultaneous occurence of the CS and the US. $v_t^i = 1$ means that all of his actions are adapted to the social norm and $v_t^i = 0$ means the opposing case.
3) N^i is the set of i's neighboring nodes and $\deg^i(\mathcal{N}) = |N^i|$ is the degree of node i in the network $\mathcal{N}$.

Now, we can reformulate the Rescorla–Wagner model for each node i according to the above-mentioned definition of λ_t^i

$$v_t^i = v_{t-1}^i + \alpha^i\left(\lambda_{t-1}^i - v_{t-1}^i\right) \tag{3}$$

where $\alpha^i \in [0, 1]$ is the learning parameter of node i.

The above-mentioned formula repeats for each node. Their simple form enables us to write them as one matrix formula

$$V_t = V_{t-1} + A(\Lambda_{t-1} - V_{t-1}). \tag{4}$$

where V_t is equal to an $n \times 1$ matrix whose ith entriy is v_t^i. Λ_t is the $n \times 1$ vector of λ_i^ts. Finally, A is an $n \times n$ diagonal matrix for which $A_{i,i} = \alpha^i$.

By (2), we can write Λ_n as a linear function of V_t. To this aim, consider the adjacency matrix of the network (a symmetric (0, 1) matrix for which the (i, j) entry is 1 if there is an edge between nodes i and j and 0 if there is no edge and normalize it so that the sum of the entries in each row is 1. We define this matrix as P where its (i, j) entry is

$$P_{i,j} = \begin{cases} 1/\deg^i(\mathcal{N}), & \text{If } i \text{ is connected to } j \text{ in } \mathcal{N} \\ 0 & \text{Otherwise.} \end{cases} \tag{5}$$

By reconsidering (2), one can easily see that

$$\Lambda_t = PV_t. \tag{6}$$

According to (6), we can rewrite (4) as follows:

$$\begin{aligned} V_t &= V_{t-1} + A(PV_{t-1} - V_{t-1}) \\ &= V_{t-1} + A(P - I)V_{t-1} \\ &= (A(P - I) + I)V_{t-1} \\ &= MV_{t-1} \end{aligned} \tag{7}$$

where I is the $n \times n$ identity matrix and M is equal to $A(P - I) + I$. M is the row stochastic, and thus, (7) defines a standard Markov chain.

Observation 1: M is row stochastic.

Proof: Consider the ith row of matrix $M = A(P - I) + I$. $M_{i,i} = 1 - \alpha^i$ and for each $j \neq i$, we have

$$M_{i,j} = \begin{cases} \alpha^i/\deg^i(\mathcal{N}), & \text{If } i \text{ is connected to } j \text{ in } \mathcal{N} \\ 0 & \text{Otherwise.} \end{cases} \tag{8}$$

Thus, trivially each $M_{i,j} \geq 0$ and

$$\sum_{j=1}^{n} M_{i,j} = 1 - \alpha^i + \frac{\deg^i(\mathcal{N})\alpha^i}{\deg^i(\mathcal{N})} = 1. \tag{9}$$

□

Sections III and IV, we need to know more about the eigenvalues and eigenvectors of M. Perron's famous theorem can be helpful here. Many different proofs are proposed for this theorem [39].

Theorem 1 (Perron [40]): The eigenvalue of the largest absolute value of a positive (square) matrix A is both simple and positive and belongs to a positive eigenvector. All other eigenvalues are smaller in absolute value.

The following observation specifies the maximum eigenvalue of M and its associated eigenvector.

Observation 2: M's maximum eigenvalue is $\lambda = 1$ and its associate eigenvector is $\vec{1} = [1\ 1\ 1 \cdots 1]^T$. All Ms other than eigenvalues are smaller in absolute value.

Proof: M is row stochastic, therefore, it can be simply shown that $\lambda = 1$ and $v = [1\ 1\ 1 \cdots 1]^T$ are M's eigenvalue and eigenvector, respectively. To see this, consider that the ith row of the $n \times 1$ matrix Mv is the sum of the entries in M's ith row, which has been shown to be equal to 1. Thus, we have $Mv = v$ which proves our claim.

Now, suppose that for a $n \times 1$ vector x and $\lambda' > 1$, $Mx = \lambda' x$. The rows of M are nonnegative and sum to 1. Therefore, each element of vector Mx is a convex combination of the x's entries. This can be no greater than $x_{\max}$ (the largest entry of x). On the other hand, at least one element of $\lambda' x$ is greater than $x_{\max}$, which proves that the eigenvalue $\lambda' > 1$ is impossible.

Therefore, the largest eigenvalue of M is equal to 1 and from Perron's theorem, all its other eigenvalues are smaller in absolute value. □

In Section III, we use this observation to find the potential forms of the system's steady state.

III. Steady-State Analysis

In this section, we study the steady state of our social norm evolution process. In Section II, we showed that this process can be represented with a standard Markov chain process model. We call our process is in a steady state if variables v_t^is which define the amount of nodes' adherence to the social norm are unchanging in time. Therefore, returning back to our model

$$V_t = MV_{t-1}. \tag{10}$$

The steady state is V_p for some $p > 0$ when $V_{p+1} = V_p$. For simplicity, we represent the steady state with notation V_∞. We show that the steady state is sensitive to the start state (V_0). Theorem 2 shows this fact.

Theorem 2: The steady state of the social norm evolution process $V_t = MV_{t-1}$ on a connected graph is in the form of $V_\infty = c_1\vec{1}$, where

$$V_0 = c_1\ u_1 + c_2\ u_2 + \cdots + c_n u_n$$

and $u_1, u_2, \ldots,$ and u_n are the eigenvectors of M with $|u_1| > |u_2| > \cdots > |u_n|$.

Proof: Assume that $u_1, u_2, \ldots, u_n$ are M's eigenvectors associated with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. From Observation 2, we know that $\lambda_1 = 1$ and $u_1 = [1\ 1\ 1 \cdots 1]^T$. These vectors define a basis for $\mathcal{R}^n$, so we can write V_0 as a linear combination of these vectors

$$V_0 = c_1 u_1 + c_2 u_2 + \cdots + c_n u_n \tag{11}$$

where $(c_1, c_2, \ldots, c_n)$ describes the coordinates of V_0 in terms of the ordered basis $\mathcal{B} = \{u_1, u_2, \ldots, u_n\}$. In this setting, we have

$$\begin{aligned} V_1 &= MV_0 = c_1 M u_1 + c_2 M u_2 + \cdots + c_n M u_n \\ &= c_1 \lambda_1 u_1 + c_2 \lambda_2 u_2 + \cdots + c_n \lambda_n u_n. \end{aligned} \tag{12}$$

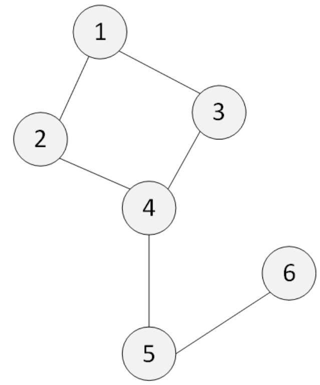

Fig. 2. Sample graph used for understanding the fact that the steady state of our social norm evolution process is determined by V_0.

TABLE I

Different Values for V_0 and Their Corresponding Steady State V_∞—The Maximum Amount of Adherence to the Social Norm Is Achieved by Choosing Nodes 3 and 4 as Initial Adopters

V_0	V_∞
[1 1 0 0 0 0]	[0.333 0.333 0.333 0.333 0.333 0.333]
[0 0 1 1 0 0]	[0.416 0.416 0.416 0.416 0.416 0.416]
[0 0 0 0 1 1]	[0.250 0.250 0.250 0.250 0.250 0.250]

With the same justification, we have

$$V_t = c_1 \lambda_1^t u_1 + c_2 \lambda_2^t u_2 + \cdots + c_n \lambda_n^t u_n. \tag{13}$$

From Observation 2, $\lambda_2, \lambda_3, \ldots, \lambda_n$ are all less than 1, therefore all terms in right-hand sideof (13) except $c_1 \lambda_1^t u_1$ vanishes for $t \to \infty$. Thus, with $\lambda_1 = 1$ and $u_1 = [1\ 1\ 1 \cdots 1]^T$, we have

$$V_\infty = \lim_{t \to \infty} V_t = c_1 u_1 = c_1 \vec{1} = [c_1\ c_1\ c_1 \cdots c_1]^T. \tag{14}$$

□

Therefore, the steady state of our Markov process is uniquely determined by V_0. We check this fact by simulating the process for different values of V_0 in the graph shown in Fig. 2. In this simulation, we assume that the learning factor of all nodes is equal to 0.5. The values of V_0 and the corresponding steady state (V_∞) are presented in Table I.

As it can be seen in this table, three different scenarios are simulated, each with a different initial state. In each of these initial states, the amount of adherence to social norm is set to zero for all nodes except two of them which we consider as an initial adopters of the norm. Our simulation shows that among these scenarios, the scenario in which nodes 3 and 4 are chosen as initial adopters of the social norm cause the maximum amount of adherence to the social norm in the steady state. In Section IV, we devise algorithms for optimizing the amount of adherence to the social norm by carefully setting the initial state of the process.

IV. Maximizing the Spread of Social Norms

As we saw in Section III, the social norm evolution process and its steady state are sensitive to its initial state. Therefore, we can control the steady state of this process by controlling the initial state values. In this section, we consider the problem of maximizing the adherence to the social norm in the steady state by setting the initial state. To set the initial state, we must find k champions (initial adopters) among the nodes of the network and set the values of v_0^i to one for them. This problem can be defined formally as follows:

$$\begin{aligned} &\max \sum_{i=1}^{n} v_\infty^i \\ &\text{s.t.} \sum_{i=1}^{n} v_0^i = k \\ &\quad v_0^i \in \{0,1\}, \quad i = 1,\ldots,n \\ &\quad v_t^i = v_{t-1}^i + \alpha^i\left(\lambda_i^{t-1} - v_{t-1}^i\right), \quad i = 1,\ldots,n \\ &\quad \text{and } t = 1,2,3,\ldots \end{aligned} \tag{15}$$

From here, we call this problem the social norm spread maximization problem or MaxSNSP. In Theorem 3, we show that MaxSNSP is solvable in polynomial time and propose an algorithm for this aim.

Theorem 3: There exists an optimal algorithm that solves the MaxSNSP problem in polynomial time.

Proof: We show that Algorithm 1 returns the optimal solution for the MaxSNSP problem. This algorithm can trivially can in polynomial time $O(n^3)$.

By Theorem 2, we showed that the steady state of the social norm evolution process $V_t = MV_{t-1}$ is in the form of $v_\infty^i = c_1$ for each $1 \le i \le n$. We showed that c_1 is the coordinate associated with the largest eigenvector of M in absolute value when we express V_0 in the basis containing all M's eigenvectors. Therefore, if we have

$$V_0 = c_1\ u_1 + c_2\ u_2 + \cdots + c_n u_n$$

where u_1, $u_2, \ldots,$ and u_n are the eigenvectors of M with $|u_1| > |u_2| > \cdots > |u_n|$, we can conclude that $V_\infty = c_1\vec{1}$. In this section, our aim is to maximize $\sum_{i=1}^{n} v_\infty^i = n \times c_1$. Thus, the MaxSNSP problem reduces to maximizing c_1.

For this aim, first in lines 1 and 2 of the algorithm, we calculate the eigenvectors of M and sort them according to the absolute values of their corresponding eigenvalues. Therefore, we assume that $|\lambda_1| > |\lambda_2| > \cdots > |\lambda_n|$. From Observation 2, we know that $u_1 = \vec{1}$ and its associate eigenvalue is $\lambda_1 = 1$. Then, in line 3, we calculate the change of basis matrix Q to transform coordinatewise representation of vectors to their equivalent representation with respect to basis $\mathcal{B} = \{u_1, u_2, \ldots, u_n\}$

$$Q = \begin{bmatrix} u_{1,1} & u_{2,1} & \cdot & \cdot & u_{n,1} \\ u_{1,2} & u_{2,2} & \cdot & \cdot & u_{n,2} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ u_{1,n} & u_{2,n} & \cdot & \cdot & u_{n,n} \end{bmatrix}^{-1} \tag{16}$$

where $u_{i,j}$ represents the jth component of u_i. Thus,

$$[c_1\ c_2\ \cdots\ c_n]^T = QV_0. \tag{17}$$

Algorithm 1 MaxSNSP Problem

1: Calculate the eigenvectors of M and call them u_1, $u_2, \ldots, u_n$.
2: Sort the eigenvectors according to the absolute values of their corresponding eigenvalues. $\{|\lambda_1| > |\lambda_2| \ge \cdots \ge |\lambda_n|\}$
3: Set $Q = [u_1^T\ u_2^T\ \cdots\ u_n^T]^{-1}$
4: Set S equal to the set of k indices $i_1, i_2, \ldots, i_k$ for which q_{1,i_j}s $(1 \le j \le k)$ are the largest k values among all $q_{1,i}$s $(1 \le i \le n)$.
5: **for all** $i \in \{1, 2, \cdots, n\}$ **do**
6:   **if** $i \in S$ **then**
7:     $v_0^i = 1$
8:   **else**
9:     $v_0^i = 0$
10:   **end if**
11: **end for**

Replacing V_0 with $[v_0^1\ v_0^2\ \cdots\ v_0^n]^T$ and Q with $[q_{i,j}]_{1\le i,j\le n}$, we have

$$\begin{bmatrix} c_1 \\ c_2 \\ \cdot \\ \cdot \\ c_n \end{bmatrix} = \begin{bmatrix} q_{1,1} & q_{1,2} & \cdot & \cdot & q_{1,n} \\ q_{2,1} & q_{2,2} & \cdot & \cdot & q_{2,n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ q_{n,1} & q_{n,2} & \cdot & \cdot & q_{n,n} \end{bmatrix} \begin{bmatrix} v_0^1 \\ v_0^2 \\ \cdot \\ \cdot \\ v_0^n \end{bmatrix}. \tag{18}$$

Therefore, for c_1, we have

$$c_1 = q_{1,1}v_0^1 + q_{1,2}v_0^2 + q_{1,3}v_0^3 + \cdots + q_{1,n}v_0^n. \tag{19}$$

Thus, the MaxSNSP problem can be redefined as follows:

$$\begin{aligned} &\max \sum_{i=1}^{n} q_{1,i}v_0^i \\ &\text{s.t.} \sum_{i=1}^{n} v_0^i = k \\ &\quad v_0^i \in \{0,1\}, \quad i = 1,\ldots,n. \end{aligned} \tag{20}$$

Therefore, a simple greedy algorithm that picks k largest values from $q_{1,i}$s (line 4 of the algorithm) and set their corresponding coefficient v_0^i to 1 (lines 5–11) can maximize c_1 and solve the MaxSNSP problem. □

V. Experiments

In addition to obtaining an optimization algorithm for the MaxSNSP problem, we are interested in understanding its behavior in practical environments and comparing its performance with other heuristics for identifying influential nodes in social networks. These heuristics benefit from widely used structural measures in a social network analysis. Different heuristics (measures) and networks have been taken into consideration in our experiments. Our results show the superior performance of our algorithm over other heuristics.

A. Networks

We are interested in studying the effect of different network topologies on the spread of social norms in social networks.

Particularly, we experiment with scale-free networks, small-world networks, and random networks, each captures different aspects of social networks' structural properties. These models are considered as the core for studying the propagation of norms over networks [41].

Scale-free networks represent the logical connectivity of people in social networks [42], which is guided by the "richers get richer" rule. This rule decribes the network has the structural property that the connectivity of the network follows a power-law degree distribution [43]. This property tells that the network has a small number of nodes (hubs), which have a very high connectivity while the most of other nodes are sparsely connected. In this paper, we use the Barabasi–Albert model for generating scale-free networks [44]. This method starts generating the network with a complete initial network with m_0 nodes. Then, new nodes are added to the network one at a time. Each new node draws m edges to the previous nodes with a probability proportional to the number of links that nodes already have.

A small-world network is a network, in which most nodes are not connected directly, but the neighbors of each node are likely to be directly connected and the average path length of the network (the average number of hops needed to reach from one node to another node) is small. Such networks represent situations where the nodes are more likely to interact with other nodes in their physical proximity [45]. In this paper, the Watts–Strogatz model is used for modeling such properties of networks [46]. This method puts nodes around a ring and connects each node to its 2 m nearest neighbors on the ring (m nodes on its left side and m on its right side). Then, some of the edges are rewired randomly with probability p independently of the other edges.

Random networks are used as a model to answer questions about how networks' properties form in a random environment. In fact, most of the network models have some randomness in nature, but the one which is almost inclusively called the random network is the Erdos–Renyi network model [47]. In this paper, we choose this model to generate random networks. In this model, each edge has a fixed probability (parameter p) of being present or absent, independently of the other edges.

We also include two real-world networks to see how our algorithm works in real scenarios. The first real work that we call BuisNet is a network extracted from interactions in a business company. This network is drawn from the enterprise social network that is installed and used in this company. This network has 297 nodes and 2053 edges. The second network is called Zachary's karate club network [48] with 34 nodes and 78 edges. This network shows the ties between the members of a karate club.

B. Heuristics

We compare our algorithms with heuristics based on different network structural properties. The main line of all these heuristics is to first calculate one of the structural properties of the network's nodes. Then, k nodes with the highest value of the mentioned property are chosen as initial adopters.

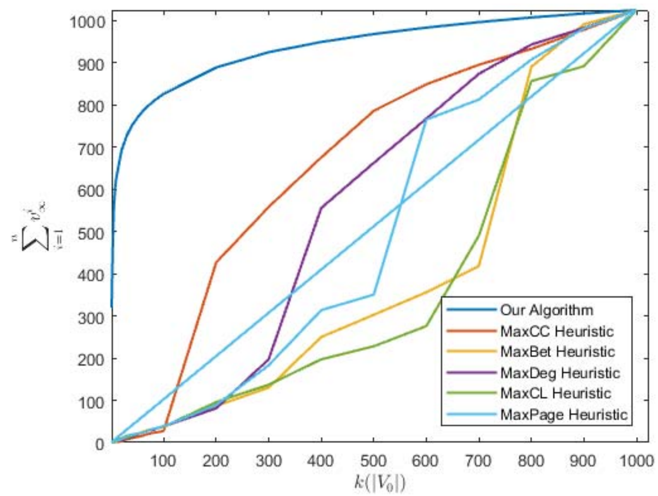

Fig. 3. Comparison of Algorithm 1 and the other heuristics over a Watts–Strogatz network with $n = 1000$, $m = 10$, and $p = 0.1$.

Therefore, for each of the following heuristics, we only explain the base structural property [49].

1) *MaxDegree:* This heuristic tries to find k nodes with the highest degree (the number of edges connected to the node).
2) *MaxCC:* In this heuristic, the local clustering coefficient (cc_i) is used. This property looks all pairs of node i's neighboring nodes and counts how many of them are linked together. That is,

$$\mathrm{cc}_i = \frac{\#\{jk \in E \mid j \neq k,\, j, k \in N_i\}}{\deg_i \cdot (\deg_i - 1)/2}$$

where N_i is the set of i's neighboring nodes and E is the set of network's edges.

3) *MaxBet:* This heuristic sort nodes by their betweenness. Betweenness (bet_i) counts the number of shortest paths which goes through node i. If we define $P_i(jk)$ as the number of shortest paths between nodes j and k that i lies on and $P(jk)$ as the number of all shortest paths between j and k, we have

$$\mathrm{bet}_i = \frac{P_i(jk)/P(jk)}{n(n-1)/2}.$$

4) *MaxPage:* This heuristic considers the pagerank (pr_i) of nodes. Nodes' pagerank is computed from a set of n equations as follows:

$$\forall_i: \ \mathrm{pr}_i = \frac{1-d}{n} + d \sum_{j \in N_i} \frac{\mathrm{pr}_j}{\deg_j}$$

where d is a constant between 0 and 1.

5) *MaxCL:* This heuristic sort nodes by nodes' closeness (cl_i). Closeness tracks down how close a given node is to any other node. If we define $l(ij)$ as the distance between the nodes i and j, then we have

$$\mathrm{cl}_i = \frac{n-1}{\sum_{j \neq i} l(ij)}.$$

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

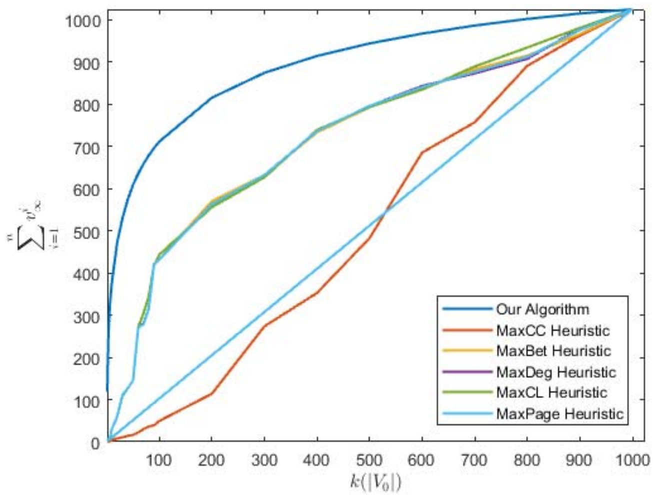

Fig. 4. Comparison of Algorithm 1 and the other heuristics over a Barabasi–Albert network with $n = 1000$, $m_0 = 20$, and $m = 20$.

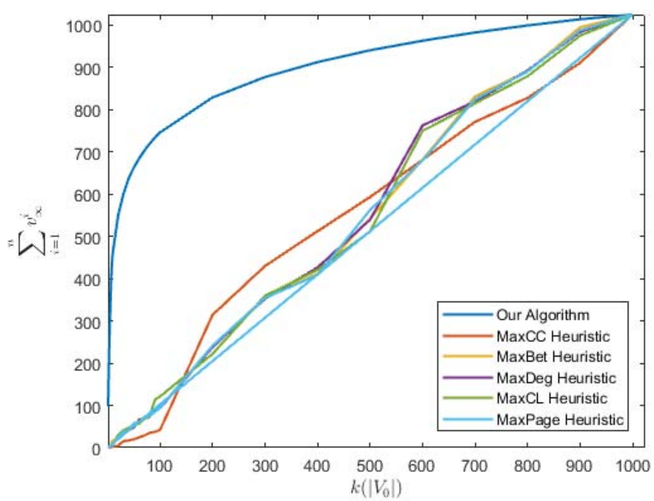

Fig. 5. Comparison of Algorithm 1 and the other heuristics over a Erdos–Renyi network with $n = 1000$ and $p = 0.02$.

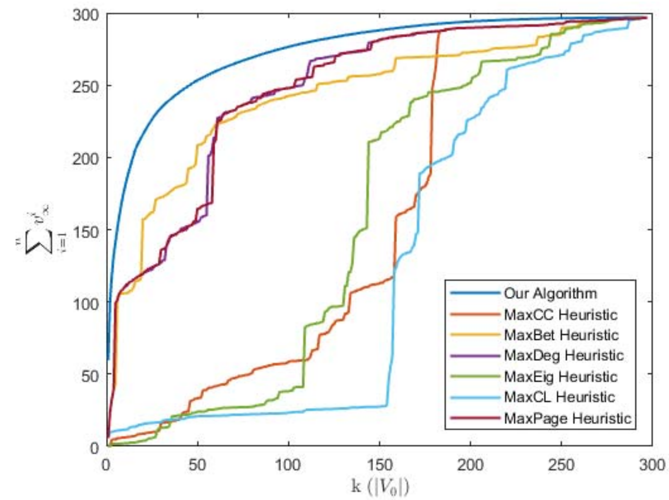

Fig. 6. Comparison of Algorithm 1 and the other heuristics over the BuisNet network.

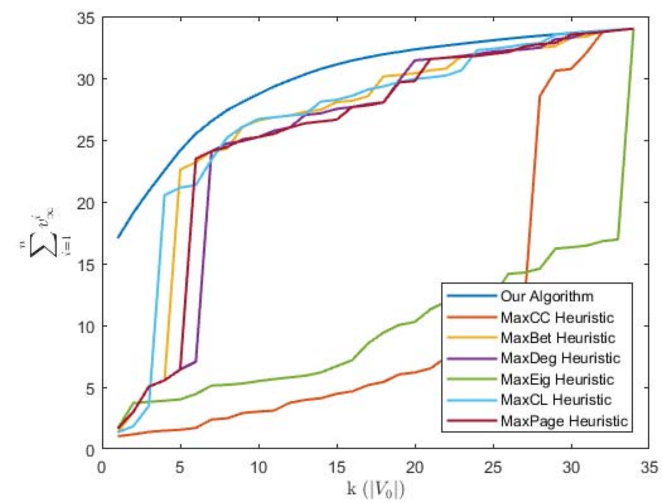

Fig. 7. Comparison of Algorithm 1 and the other heuristics over Zachary's karate club network.

C. Results

We compare the performance of our proposed algorithm with the mentioned heuristics on different networks. The main purpose of these experiments is to show that the complexity in Algorithm 1 is needed and no simple heuristic (as one may claim) can beat this algorithm. The results are presented in Figs. 3–7. All of these figures show the outperformance of Algorithm 1 over the heuristics. In each of these figures, we used the following simulation scenario.

1) Using each network model, we generate a network. The parameters used for generating each network are depicted in Figs. 3–7 captions. For each of its nodes, we set the learning factor parameter (α_i) to a uniform random number between 0 and 1.
2) For each $1 \leq k \leq n$, we run our algorithm and other heuristics over the network generated by the network model. Each of these algorithms returns a subset of k nodes as the initial adopters. We construct the vector V_0 from this returned set.
3) The spread of social norm process is simulated over the network starting from each of the returned sets of initial adopters. For simulation, an iterative process is started at V_0 and in each iteration of this process V_{i+1} is set to MV_i. This process continues until it reaches to a steady state, where $V_{i+1} = V_i$.
4) The amount of social norm adherence in the steady state of each process is saved and finally plotted in Figs. 3–7.

VI. Conclusion

In this paper, by using the very famous Rescorla–Wagner conditioning model, we obtained an effective model for the spread of social norms. We created an extension to this model for a social network of people and showed that this new model is a standard Markov chain process. The steady states of this process were studied and the potential structures of steady states of it were extracted.

Then, by using our mathematical model, we formulated the problem of maximizing the adherence to a social norm in a social network by finding the best set of initial norm adopters. We proposed a linear programming algorithm for solving this problem that runs in polynomial time. Finally, the superior performance of this algorithm in comparison to other typical algorithms that are typically used in social network contexts was shown by simulating the process over different network models.

This research can have many opportunities for followup works, which we have listed some of them in the following.

1) One can use the algorithm for the MaxSNSP problem in a real-world scenario in which a social norm must be promulgated among a set of people. This environment can be, for example, a business company with an established enterprise social network by which one can extract the friendship network among people. The learning factors of each person can be estimated by mining their content over the social network. Then by running Algorithm 1, the best set of initial adopters can be calculated and tested for spreading the social norm.
2) In this paper, we considered the static control in which the steady state of the process is examined. The dynamics of this process is also important and must be controlled which can be modeled by tools such as dynamical systems [50].
3) In our model, we focused on affective aspects of adopting a behavior. There are some other interesting factors such as cognition, mass media, and so on. Considering each of these factors needs serious modifications in the model that can be interesting.
4) While our algorithm works well, the intuition behind it is not clear, because our algorithm and formulations are directly extracted from mathematics of the model. More than that our experiments in which our algorithm is compared with heuristics (which are based on different structural properties) show that no networks structural property can estimate the chosen initial set by our algorithm. Thus, it could be an interesting research direction to find how the measure used by our algorithm can be interpreted with the network's structure.

REFERENCES

[1] P. T. Mitchell, *Network Centric Warfare and Coalition Operations: The New Military Operating System*. Evanston, IL, USA: Routledge, 2009.

[2] C. Bicchieri, *The Grammar of Society: The Nature and Dynamics of Social Norms*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[3] H. P. Young, "The evolution of social norms," *Economics*, vol. 7, no. 1, pp. 359–387, 2015.

[4] D. Acemoglu and M. O. Jackson, "Social norms and the enforcement of laws," *J. Eur. Econ. Assoc.*, vol. 15, no. 2, pp. 245–295, 2017.

[5] E. A. Posner, *Law and Social Norms*. Cambridge, MA, USA: Harvard Univ. Press, 2009.

[6] Y. Zasu, "Sanctions by social norms and the law: Substitutes or complements?" *J. Legal Stud.*, vol. 36, no. 2, pp. 379–396, 2007.

[7] A. Pentland, *Social Physics: How Good Ideas Spread-the Lessons From a New Science*. Baltimore, MD, USA: Penguin, 2014.

[8] C. A. Latkin, V. Forman, A. Knowlton, and S. Sherman, "Norms, social networks, and HIV-related risk behaviors among urban disadvantaged drug users," *Social Sci. Med.*, vol. 56, no. 3, pp. 465–476, 2003.

[9] C. Latkin *et al.*, "Relationships between social norms, social network characteristics, and HIV risk behaviors in Thailand and the United States," *Health Psychol.*, vol. 28, no. 3, pp. 323–329, 2009.

[10] R. A. Posner and E. B. Rasmusen, "Creating and enforcing norms, with special reference to sanctions," *Int. Rev. Law Econ.*, vol. 19, no. 3, pp. 369–382, 1999.

[11] J. Coleman, *Foundations of Social Theory*. Cambridge, MA, USA: Harvard Univ. Press, 1990.

[12] J. M. Epstein, *Agent_Zero: Toward Neurocognitive Foundations for Generative Social Science*. Princeton, NJ, USA: Princeton Univ. Press, 2014.

[13] M. H. Olson, *An Introduction to Theories of Learning*. Abingdon, U.K.: Routledge, 2015.

[14] A. P. Field, "Is conditioning a useful framework for understanding the development and treatment of phobias?" *Clin. Psychol. Rev.*, vol. 26, no. 7, pp. 857–875, 2006.

[15] D. Hermans, M. G. Craske, S. Mineka, and P. F. Lovibond, "Extinction in human fear conditioning," *Biol. Psychiatry*, vol. 60, no. 4, pp. 361–368, 2006.

[16] S. Mineka and K. Oehlberg, "The relevance of recent developments in classical conditioning to understanding the etiology and maintenance of anxiety disorders," *Acta Psychol.*, vol. 127, no. 3, pp. 567–580, 2008.

[17] S. Mineka and A. Öhman, "Phobias and preparedness: The selective, automatic, and encapsulated nature of fear," *Biol. Psychiatry*, vol. 52, no. 10, pp. 927–937, 2002.

[18] A. L. Strauss, *Psychological Modeling: Conflicting Theories*. Evanston, IL, USA: Routledge, 2017.

[19] A. J. Budney and J. T. Borodovsky, "The potential impact of cannabis legalization on the development of cannabis use disorders," *Preventive Med.*, vol. 104, pp. 31–36, Nov. 2017.

[20] S. Siegel and L. G. Allan, "The widespread influence of the rescorla-wagner model," *Psychonomic Bull. Rev.*, vol. 3, no. 3, pp. 314–321, 1996.

[21] J. Wei, J. Wu, M. Molinari, V. Cvetkovic, and K. H. Johansson. (2017). "On the modeling of neural cognition for social network applications." [Online]. Available: https://arxiv.org/abs/1704.03192

[22] J. Deitch-Stackhouse *et al.*, "The influence of social norms on advancement through bystander stages for preventing interpersonal violence," *Violence Against Women*, vol. 21, no. 10, pp. 1284–1307, 2015.

[23] I. Rahwan, "Society-in-the-loop: Programming the algorithmic social contract," *Ethics Inf. Technol.*, vol. 20, no. 1, pp. 5–14, 2018.

[24] C. McLaughlin and J. Vitak, "Norm evolution and violation on Facebook," *New Media Soc.*, vol. 14, no. 2, pp. 299–315, 2012.

[25] J. K. Burgoon, "Expectancy violations theory," in *International Encyclopedia of Interpersonal Communication*. Hoboken, NJ, USA: Wiley, 2009.

[26] R. Dickie, S. Rasmussen, R. Cain, L. Williams, and W. MacKay, "The effects of perceived social norms on handwashing behaviour in students," *Psychol., Health Med.*, vol. 23, no. 2, pp. 154–159, 2018.

[27] M. O. Jackson, B. W. Rogers, and Y. Zenou, "The economic consequences of social-network structure," *J. Econ. Literature*, vol. 55, no. 1, pp. 49–95, 2017.

[28] R. S. Burt, "Structural holes versus network closure as social capital," in *Social Capital*. Evanston, IL, USA: Routledge, 2017, pp. 31–56.

[29] N. O. Hodas and K. Lerman, "The simple rules of social contagion," *Sci. Rep.*, vol. 4, Mar. 2014, Art. no. 4343.

[30] K. Lerman, X. Yan, and X.-Z. Wu, "The 'Majority illusion' in social networks," *PLoS ONE*, vol. 11, no. 2, p. e0147617, 2016.

[31] D. Kempe, J. Kleinberg, and É. Tardos, "Maximizing the spread of influence through a social network," in *Proc. 9th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2003, pp. 137–146.

[32] A. Sathanur, M. Halappanavar, Y. Shi, and W. Sagduyu. (2017). "Exploring the role of intrinsic nodal activation on the spread of influence in complex networks." [Online]. Available: https://arxiv.org/abs/1707.05287

[33] A. J. Morales, J. Borondo, J. C. Losada, and R. M. Benito, "Efficiency of human activity on information spreading on Twitter," *Social Netw.*, vol. 39, pp. 1–11, Oct. 2014.

[34] M. W. Peng and C. O. Vlas, "Diffusion of a twentieth-century innovation," *Acad. Strategic Manage. J.*, vol. 16, no. 1, pp. 172–174, 2017.

[35] L. Liu, B. Chen, W. Jiang, X. Qiu, and L. He, "Patterns of cascading behavior in WeChat moments," in *Proc. 4th Int. Conf. Inf., Cybern. Comput. Social Syst. (ICCSS)*, Jul. 2017, pp. 400–405.

[36] A. Lindbeck, "Welfare state disincentives with endogenous habits and norms," *Scand. J. Econ.*, vol. 97, no. 4, pp. 477–494, 1995.

[37] I. Kolstad, *The Evolution of Social Norms*. Bergen, Norway: Chr. Michelsen Institute, 2003.

[38] H. Wang, C. Coble, and P. Bello, "Cognitive-affective interactions in human decision-making: A neurocomputational approach," in *Proc. 28th Annu. Conf. Cogn. Sci. Soc.*, R. Sun and N. Miyake, Eds. 2006, p. 234146.

[39] C. R. MacCluer, "The many proofs and applications of perron's theorem," *SIAM Rev.*, vol. 42, no. 3, pp. 487–498, 2000.

[40] O. Perron, "Grundlagen für eine theorie des jacobischen kettenbruchalgorithmus," *Math. Ann.*, vol. 64, no. 1, pp. 1–76, 1907.
[41] M. A. Mahmoud, M. S. Ahmad, M. Z. M. Yusoff, and A. Mustapha, "A review of norms and normative multiagent systems," *Sci. World J.*, vol. 2014, Jul. 2014, Art. no. 684587.
[42] L. A. N. Amaral, A. Scala, M. Barthélémy, and H. E. Stanley, "Classes of small-world networks," *Proc. Nat. Acad. Sci. USA*, vol. 97, no. 21, pp. 11149–11152, 2000.
[43] M. E. J. Newman, "The structure and function of complex networks," *SIAM Rev.*, vol. 45, no. 2, pp. 167–256, 2003.
[44] R. Albert and A.-L. Barabási, "Statistical mechanics of complex networks," *Rev. Mod. Phys.*, vol. 74, no. 1, p. 47, 2002.
[45] O. Sen and S. Sen, "Effects of social network topology and options on norm emergence," in *Coordination, Organizations, Institutions and Norms in Agent Systems V*. Springer, 2010, pp. 211–222.
[46] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks," *Nature*, vol. 393, no. 6684, pp. 440–442, 1998.
[47] B. Bollobás, "The evolution of random graphs," *Trans. Amer. Math. Soc.*, vol. 286, no. 1, pp. 257–274, 1984.
[48] W. W. Zachary, "An information flow model for conflict and fission in small groups," *J. Anthropol. Res.*, vol. 33, no. 4, pp. 452–473, 1977.
[49] M. O. Jackson, *Social and Economic Networks*. Princeton, NJ, USA: Princeton Univ. Press, 2010.
[50] G. Teschl, *Ordinary Differential Equations and Dynamical Systems*, vol. 140. Providence, RI, USA: AMS, 2012.

Seyyed Hadi Sajadi received the B.Sc. degree from the Department of Electrical and Computer Engineering, Shahid Beheshti University, Tehran, Iran, and the M.Sc. degree in economic and social systems engineering, Mazandaran University of Science and Technology, Babol, Iran. He is currently pursuing the Ph.D. degree in software engineering, Sharif University of Technology International Campus, Kish, Iran.

His current research interests include ethics in cyberspace, child online protection, and also leveraging ICT in social applications, especially the evolution of social norms using online social networks.

MohammadAmin Fazli received the B.Sc. degree in hardware engineering and the M.Sc. and Ph.D. degrees in software engineering from the Sharif University of Technology, Tehran, Iran, in 2009, 2011, and 2015, respectively.

He is currently an Assistant Professor with the Sharif University of Technology, where he is also a Research and Development Supervisor with the Intelligent Information Solutions Center. His current research interests include game theory, combinatorial optimization, computational business and economics, graphs and combinatorics, complex networks, and dynamical systems.

Jafar Habibi is an Associate Professor with the Department of Computer Engineering, Sharif University of Technology, Tehran, Iran, where he is also the Managing Director of the Intelligent Information Solutions Center and a Supervisor with the RoboCup Simulation Group and the Software Engineering Laboratory. His current research interests include software engineering, simulation systems, MIS, DSS, and evaluation of computer systems performance.