MIS Quarterly

RESEARCH ARTICLE

TOP PERSUADER PREDICTION FOR SOCIAL NETWORKS[1]

Xiao Fang
Lerner College of Business and Economics, University of Delaware,
Newark, DE 19716-3130 U.S.A. {xfang@udel.edu}

Paul Jen-Hwa Hu
David Eccles School of Business, University of Utah,
Salt Lake City, UT 84112 U.S.A. {paul.hu@eccles.utah.edu}

Top persuaders in a social network are social entities whose adoption of a product or service will result in the largest numbers of other entities in the network adopting the same product or service. Predicting top persuaders is critical to an expanding array of important social network-centric applications, such as viral marketing, customer retention, and political message promotion. This study formulates the top persuader prediction problem and develops a novel method to predict top persuaders. Our method development is rooted in eminent social network theories that reveal several forces central to social persuasion, including social influence, entity similarity, and structural equivalence. Our method innovatively integrates these forces to predict top persuaders in a social network, in contrast to existing methods that primarily focus on social influence. Specifically, we introduce persuasion probability that denotes the likelihood of persuasion conditioned on these forces. We then propose how to estimate persuasion probabilities, develop an algorithm to predict top persuaders using the estimated persuasion probabilities, and analyze the theoretical property of the algorithm. We conduct an extensive evaluation with real-world social network data and show that our method substantially outperforms prevalent methods from representative previous research and salient industry practices.

Keywords: Top persuader, influential user, machine learning, social network, social persuasion, social influence, entity similarity, structural equivalence, data-driven method

Introduction

A social network consists of a set of social entities and the relationships among them (Wasserman and Faust 1994). Such networks range from traditional social networks, such as mail correspondence among people (Travers and Milgram 1969), to technology-enabled social networks, such as those supported by social media (e.g., Twitter) or online games (Hemp 2006; Kleinberg 2008). An important phenomenon within social networks is *social persuasion* that refers to the principles and processes by which a social entity's attitude, belief, or behavior is affected by other entities in a social network (Chaiken et al. 2000).[2] This phenomenon gives organizations the opportunity to predict essential social entities, namely top persuaders, whose adoption of a product, service, or opinion will result in the largest numbers of other entities in a social

[1]Kartik Hosanagar was the accepting senior editor for this paper. Zhengrui Jiang served as the associate editor.

The appendices for this paper are located in the "Online Supplements" section of the *MIS Quarterly*'s website (http://www.misq.org).

[2]Social persuasion closely relates to social contagion (Burt 1987) and network diffusion (Valente 1995), in that these concepts all center on the spread of attitudes, opinions, or behaviors within a social network, with a focus on potential changes in the attitude or behavior of a person who has not yet adopted, as a result of his or her exposure to and interaction with others who have adopted the focal product or service (Petty and Cacioppo 1986; Simon 1976).

network to adopt the same product, service, or opinion. Particularly, in a technology-enabled social network, massive and rich data reflecting individual social entities' behaviors are readily available for developing data-driven methods to predict top persuaders.

Predicting top persuaders is critical to an expanding array of important social network-centric applications. For example, seeding has emerged as a key strategy for firms to market products and services (e.g., viral marketing), for health agencies to foster healthy lifestyles, and for politicians to promote political views or public policies (Carr 2008; Watts and Peretti 2007). In general, seeding targets a group of social entities as "seeds" and exploits their persuasion effects to reach out and affect a broader population. The effectiveness of seeding apparently depends on the persuasion power of the selected social entities, which underscores the significance of top persuader prediction. Firms can also leverage these persuaders for customer retention. According to a Forbes Insights report (2011), customer retention has become a top marketing priority for firms and often accounts for the largest portion of their marketing budget. By retaining top persuaders, a firm not only keeps these customers (i.e., top persuaders) but also benefits from their persuasion effects to retain other customers they affect. On the other hand, when these persuaders churn, the firm loses them and other customers they affect. Therefore, a cost-effective customer retention strategy is to focus on retaining top persuaders, which requires accurate prediction of such persuaders.

Because of its importance, top persuader prediction has attracted attention from both academic researchers and industry practitioners. In industry, methods built on degree centrality have been developed to identify top persuaders. For example, according to Weng et al. (2010), Twitter considers its users with the largest numbers of followers as top persuaders. In academia, prior research has developed methods for identifying influential individual entities in a social network (e.g., Cha et al. 2010; Ghosh and Lerman 2010). A review of industry practices and academic research on top persuader prediction suggests a predominant focus on *social influence*, which refers to the process by which a social entity's attitude, belief, or behavior is affected by (direct) communications and interactions with his or her neighboring entities (Friedkin 1998; Rice et al. 1990). However, according to social network theories, social persuasion can arise from multiple distinct forces that include social influence, entity similarity, and structural equivalence (Burt 1987; Knoke 1990). *Entity similarity* denotes the degree to which two social entities are "similar" in their individual characteristics. One entity in a social network can use other similar entities as "a frame of reference" and adjust his or her attitude, belief, or behavior accordingly. Thus, a social entity could affect another's adoption of a product or service, partly due to the similarity in their individual characteristics. Furthermore, similarity also exists in structural position within a social network. Structurally equivalent people occupy the same position in a social structure; specifically, social entities are considered structurally equivalent if they relate in the same ways to other entities in a social network (Lorrain and White 1971). *Structural equivalence* matters because the similarity in structural position within a social network can affect individual entities' beliefs, attitudes, and behaviors (Burt 1987). In this vein, predicting top persuaders requires proper consideration of different essential forces underlying social persuasion, a fundamental challenge that motivates this research.

To address this challenge, we propose a novel method to predict top persuaders. Different from existing methods that primarily focus on social influence, our method predicts top persuaders by simultaneously considering all three important forces underlying social persuasion. In particular, we innovatively integrate these forces by introducing persuasion probability, or the probability that one social entity persuades another to adopt, given the impacts of social influence, entity similarity, and structural equivalence from the former to the latter. We then propose how to estimate persuasion probabilities and how to predict top persuaders with the estimated persuasion probabilities. The remainder of the paper proceeds as follows. We first review relevant previous research, highlight the research gaps to be addressed, and discuss the differences between our proposed method and existing methods. Next, we describe the foundation of our method development, anchored in social network theories. We then formulate the top persuader prediction problem and propose a data-driven method to predict top persuaders. To illustrate its practical value, we empirically evaluate the effectiveness of our proposed method with real-world social network data, using representative existing methods as benchmarks. Finally, we summarize our contributions, discuss important implications, and conclude with several future research directions.

Related Work

Existing methods for identifying top persuaders in a social network include centrality-based methods, PageRank-based methods, and TwitterRank. In general, centrality-based methods rank social entities according to a centrality measure (Wasserman and Faust 1994) and consider those entities with high centrality scores influential and thus top persuaders (Borgatti et al. 2009; Brass 1984; Cha et al. 2010; Ghosh and Lerman 2010). Since the initial conceptualization of centrality by Bavelas (1948), different centrality measures have been developed; among them, degree centrality (Albert et al. 2000; Walter et al. 1996), topological centrality (Borgatti et al. 2009; Brass 1984;), and eigencentrality (Ballester et al. 2006; Bonacich 1972) are widely applied. The degree of a social entity refers to the number of other entities to which it is connected directly (Freeman 1979). An entity that has a higher degree is likely to communicate and interact with more entities in a social network and therefore is more influential. Although commonly utilized, degree centrality is not particularly effective for identifying influential entities in a social network (Watts and Dodds 2007). Compared with degree centrality, topological centrality is more computationally expensive. A representative topological centrality is closeness centrality. The closeness of a social entity is calculated as the inverse of the sum of its distances to all other entities in a social network (Freeman 1979). Clearly, the closer a social entity is to other entities in a social network, the more efficiently that entity can communicate and interact with others and, as a result, the more influential it becomes. Another common topological centrality is betweenness centrality. In essence, the betweenness of a social entity indicates its frequency of falling on the shortest paths that link pairs of other entities in a social network (Freeman 1979; White and Borgatti 1994). Betweenness reflects a social entity's potential of controlling the communications and interactions between pairs of other entities; the higher an entity's betweenness, the more influential that entity becomes. In addition to degree and topological centrality, eigencentrality represents another popular centrality measure. Eigenvector centrality, an important eigencentrality, weights directly connected social entities according to their centrality scores (Bonacich 1972, 1987, 2007). Another important eigencentrality is intercentrality, which considers both a social entity's centrality and its contribution to every other entity's centrality (Ballester et al. 2006). Intercentrality can be applied to identify the key player in a network of delinquents whose removal would lead to the highest aggregate delinquency reduction (Ballester et al. 2010). Different from centrality metrics described above, percolation centrality measures the centrality of a social entity according to its percolation state, in addition to its topological connectivity (Piraveenan et al. 2013).

Previous research also analyzes weights of relationships from the lens of social influence and incorporates such weights to identify top persuaders, in general or in a particular application domain. One prevalent method uses the PageRank algorithm (Brin and Page 1998) to rank individual entities according to their social influences (Romero et al. 2010). Several other studies develop methods to identify top persuaders in a particular application domain. Taking blogging as an example, Agarwal et al. (2008) measure the influence of a blog, using such properties as number of inlinks, number of comments, number of outlinks, and length, and thereby identify top bloggers as those who write influential blogs. For Internet-enabled social networking, Trusov et al. (2010) examine the log-in activities in a social network and consider influential users as those whose activity level significantly affects others' activity levels. Romero et al. (2010) argue that the influence of a Twitter user depends on not only the number of his or her followers but also his or her ability to overcome the passivity of these followers. They therefore develop an algorithm to compute the influence and passivity score of each Twitter user, on the basis of the interaction between influence and passivity. Zhao et al. (2014) identify influential users in an online health community and define an influential response as one whose sentiment is aligned with the sentiment change of a community member; accordingly, influential users are those who post the greatest number of influential responses. Finally, Probst et al. (2013) provide a detailed survey of top persuader identification from the lens of social influence (i.e., identifying influential users).

While existing methods focus on social influence, entity similarity could contribute to social persuasion too. Several studies have shown empirically the important role of entity similarity in driving social persuasion (Anagnostopoulos et al. 2008; Crandall et al. 2008). For example, Crandall et al. (2008) examine the relative importance of social influence and entity similarity for predicting social persuasion, and report that, while both contribute to social persuasion, social influence seems more powerful for predicting social persuasion on Wikipedia whereas entity similarity appears more important for predicting social persuasion on LiveJournal. Although prior literature has recognized the significance of entity similarity, few studies consider both entity similarity and social influence to identify top persuaders in a social network. One exception is Weng et al. (2010), who develop a method, namely TwitterRank, to identify influential Twitter users on the basis of the tweet topic similarity (i.e., entity similarity) and the following relationship (i.e., social influence) between Twitter users.

According to our literature review, most existing methods identify top persuaders from the lens of social influence, except for TwitterRank, which combines social influence and entity similarity. However, social network theories suggest that multiple distinct forces—namely, social influence, entity similarity, and structural equivalence—jointly determine social persuasion (Burt 1987; Knoke 1990). Anchored in eminent social network theories, our proposed method differs from existing methods in that it simultaneously considers these distinct forces crucial to social persuasion. Furthermore, our method integrates these forces to predict top persuaders by introducing persuasion probability that denotes the

Table 1. Comparison of Our Method and Existing Methods				
	Our Method	**Centrality-Based Methods** (e.g., Ghosh and Lerman 2010)	**PageRank-Based Methods** (e.g., Romero et al. 2010)	**Twitter Rank** (Weng et al. 2010)
Considers social influence?	Yes	Yes	Yes	Yes
Considers entity similarity?	Yes	No	No	Yes
Considers structural equivalence?	Yes	No	No	No
Integrates forces underlying social persuasion through persuasion probability?	Yes	No	No	No

Notes: A method that considers structural equivalence would take into account the persuasion effect between structurally similar social entities in a social network.

likelihood of persuasion conditioned on these forces. Specifically, we propose how to estimate persuasion probabilities and then how to use the estimated persuasion probabilities to predict top persuaders. In contrast, Weng et al. (2010) combine social influence and entity similarity to identify top persuaders by multiplying them without proper theoretical justification. Finally, we demonstrate, through an extensive evaluation with real-world social network data, that our proposed method significantly outperforms representative existing methods. In Table 1, we compare our method and existing methods.

Theoretical Foundation

Social influence constitutes a common focus of previous top persuader research, and its effects can be explained by *social comparison theory* (Festinger 1954). According to this theory, persuasion occurs when a person influences others through his or her ability to reduce the uncertainty in a decision situation that those others face. Pfeffer et al. (1976) specifically refer to such social influences as *informational social influences*. When interacting with others in a social network, a person conveys a view that could influence those others' opinions or behaviors by simplifying their evaluation or decision task at hand, such as by reducing the range of viable alternatives or making a recommendation. To create social influence, the initiator usually sends a message, issues an opinion, or delivers a well-articulated view to recipients to evoke their response (Wood 2000; Zanden 1987).

Friedkin (1998) further explains the effects of social influence with *social influence network theory,* which suggests that each entity is endowed with an initial opinion or attitude, receives and responds to the information disseminated in the network, and then modifies his or her opinion or attitude accordingly. The magnitude of the resulting social persuasion depends on the social influence carried through the interentity communications and interactions. Such effects and influences echo the importance of the linkages that connect individual entities in a social network. People tend to compare themselves to those with whom they have close ties, then seek to evaluate and emulate the attitudes or actions of these intimates. Interpersonal communications and interactions, via network ties, are recurrent and often prevail in small, close-knit neighborhoods within the grand social network. In this vein, Agarwal et al. (2009) highlight the importance of the social influence of local reference groups and geographical peer effects in a social network. Similarly, Susarla et al. (2012) emphasize the effect of social influence for diffusing user-generated content in social networks.

Both social comparison theory and social influence network theory also shed light on the importance of individual characteristics. People in a social network may exhibit similar opinions or behaviors, simply because they share individual characteristics (Aral et al. 2009; Shalizi and Thomas 2011). All else being equal, individuals who are more similar in their demographics, cultural backgrounds, or personal preferences are more likely to have common views and behaviors. According to *attribution theory* (Weiner 1974, 1986), people gain cognitive control over their environment by explaining and understanding the causes of behaviors and environmental occurrences. In this light, people who share highly similar characteristics or backgrounds likely have comparable understandings of and reasoning for underlying causes and therefore exhibit similar opinions and behaviors.

Structural characteristics can also determine the extent to which an entity affects others' opinions and behaviors (Fang et al. 2013; Wejnert 2002). Burt (1987) analyzes social contagion for innovation diffusion in social structural circumstances and suggests that contagion arises when people who are proximate in the structure use one another to manage the uncertainty of an innovation. Thus both cohesion and structural equivalence models are crucial. The structural equivalence model emphasizes entities' structural positions in a network; the cohesion model focuses on socialization between an ego (an entity that has not yet adopted) and an alter (an entity that has adopted already) in a network. According to the cohesion model, the more frequent and empathic the interaction between an ego and an alter, the more likely an adoption by the latter is to trigger an adoption by the former. Burt et al. (1994) also use a contagion model to describe how a network structure draws individuals together, such that one's opinion or behavior affects those of the other. Structural equivalence can explain observable changes in people's opinions or behaviors in the presence of information dissemination in a social network.

Structural equivalence also represents a crucial characteristic of a social network. People occupying identical structural positions in a network tend to exhibit more similar opinions and behaviors than otherwise, partially because structural equivalence can reduce the uncertainty associated with an evaluation or decision task that people face and therefore can affect their attitudes and behaviors, even if no social ties connect them directly. As Burt explains, the spread of an opinion or behavior in a social network is contingent, to some degree, on the way the structure of the network brings entities together. If they connect to the same group of others through identical links, individuals likely exhibit similar opinions or behaviors, because they vicariously experience or even mimic each other, which reduces the uncertainty associated with their adoption of an opinion or behavior (Rice and Aydin 1991).

Burt emphasizes cohesion and structural equivalence for social contagion that closely relates to the social persuasion we study. According to Knoke (1990), persuasion and selection processes are fundamental to any comprehensive explanation of how individuals' social relationships shape their attitudes and activities. Therefore, persuasion and entity similarity are both important. Persuasion, as described by Knoke, and cohesion, as defined by Burt, intrinsically correspond to the social influence we examine. By anchoring in eminent social network theories and integrating the insights by Burt and Knoke, we consider three forces central to social persuasion: social influence, entity similarity, and structural equivalence.

Top Persuader Prediction Problem and Method

In this section, we formulate the top persuader prediction problem and then detail our method for predicting top persuaders.

Problem Formulation

Let V denote a set of n social entities v_i in a social network (i.e., $v_i \in V$, $i = 1,2,\ldots,n$). Social entities in a network are connected by pairwise relationships. A relationship can be directional or nondirectional (Wasserman and Faust 1994). For example, a social network of avatars in an online game consists of a set of avatars (i.e., social entities) connected by nondirectional communication relationships, whereas a social network on Twitter consists of a set of users (i.e., social entities) connected by directional relationships (e.g., one Twitter user following another user). For a social entity v_i, let vector $\boldsymbol{f}_i$ represent the profile of the entity (i.e., individual characteristics of the entity); taking an avatar as an example, its profile describes the avatar's characteristics such as gender and profession. Let s_{ij} denote the strength of social interactions from entity $v_i \in V$ to entity $v_j \in V$ in the relationship connecting them (e.g., the number of tweets published by v_i and retweeted by v_j). Specifically, s_{ij} is 0 if there is no relationship from v_i to v_j. We summarize important notations used in this paper in Table 2.

We formulate the top persuader prediction problem as follows: In a social network, we observe at current time a set of social entities that has adopted a focal item (e.g., a product, service, or opinion). The objective is to predict top K persuaders (i.e., K individual entities whose adoption of the item will result in the largest numbers of other entities in the network to adopt in a future time period).

To solve this problem, we need to properly address two challenges. First, for each ordered pair of social entities v_i and v_j, we need to estimate the probability that v_i persuades v_j to adopt (i.e., persuasion probability). Second, we have to predict top persuaders on the basis of the estimated persuasion probabilities. In the following, we detail how we address each challenge.

Estimating Persuasion Probabilities

The probability that v_i persuades v_j to adopt depends on the social influence from v_i to v_j as well as the entity similarity and structural equivalence of v_i to v_j. We use the strength of social interactions to measure social influence, because social influence results from and therefore is manifested by social interactions (Rice et al. 1990). Accordingly, the social influence I_{ij} from v_i to v_j is measured as

$$I_{ij} = \frac{s_{ij}}{\sum_{v_h \in V, h \neq j} s_{hj}} \tag{1}$$

Table 2. Notation

V = set of social entities in a social network	v_i = social entity, $v_i \in V$
$\boldsymbol{f}_i$ = profile of v_i	s_{ij} = strength of social interactions from v_i to v_j
I_{ij} = social influence from v_i to v_j	M_{ij} = entity similarity of v_i to v_j
R_{ij} = structural equivalence of v_i to v_j	d_{ij} = distance from v_i to v_j
z_{ij} = normalized distance, see Equation (5)	γ = attenuation factor, $0 < \gamma < 1$
D_j adoption state of v_j	p_{ij} = persuasion probability, see Equation (8)
c_i = persuasion score of entity v_i	$\boldsymbol{C}$ = vector of persuasion scores

Equation (1) measures social influence from v_i to v_j as the proportional social interaction strength from v_i to v_j, in relation to the total social interaction strength from all $v_h \in V$ to v_j, $h \neq j$. The way we measure social influence is appropriate and has been commonly used by previous research (Kempe et al. 2003; Romero et al. 2010). Social influence I_{ij} is within the range of [0,1]; the higher the value of I_{ij}, the more powerful is the influence of v_i on v_j, relative to the influence of any other social entity on v_j.

We measure entity similarity in a similar way; specifically, the entity similarity M_{ij} of v_i to v_j is

$$M_{ij} = \frac{sim(\boldsymbol{f}_i, \boldsymbol{f}_j)}{\sum_{v_h \in V, h \neq j} sim(\boldsymbol{f}_h, \boldsymbol{f}_j)} \tag{2}$$

where $sim()$ is a similarity function that takes the profiles of two social entities as input and returns their profile similarity. The choice of the similarity function depends on the types of attributes in $\boldsymbol{f}_i$ (Tan et al. 2006). For example, matching coefficient is an appropriate similarity function if $\boldsymbol{f}_i$ only contains nominal attributes, whereas a heterogeneous similarity function is proper if $\boldsymbol{f}_i$ contains both nominal and numeric attributes. The specific similarity function used in our evaluation study is discussed in the next section. Entity similarity M_{ij} is within the range of [0,1]; the higher the value of M_{ij}, the more similar v_i is to v_j, relative to the similarity of any other social entity to v_j.

Perfect structural equivalence is rare in real-world social networks; it is common to measure the structural equivalence (or inequivalence) between social entities as the extent to which they are structurally similar (or dissimilar) according to some distance function, such as the Euclidean distance function (Wasserman and Faust 1994). According to Burt (1987), the Euclidean distance between social entities v_i and v_j in a social network with directional relationships is defined as

$$ed(v_i, v_j) = \left[(z_{ij} - z_{ji})^2 + \sum_{v_k \in V, k \neq i,j} (z_{ik} - z_{jk})^2 + \sum_{v_k \in V, k \neq i,j} (z_{ki} - z_{kj})^2 \right]^{1/2} \tag{3}$$

where z_{ij} is the normalized path distance from v_i to v_j. Similarly, $ed(v_i, v_j)$ in a social network with nondirectional relationships is defined as

$$ed(v_i, v_j) = \left[\sum_{v_k \in V, k \neq i,j} (z_{ik} - z_{jk})^2 \right]^{1/2} \tag{4}$$

For a social network with nondirectional relationships, we have $z_{ij} = z_{ji}$, $z_{ik} = z_{kj}$, and $z_{jk} = z_{kj}$. Hence, the first and third terms in Equation (3) do not appear in Equation (4). Following Burt, we define normalized path distance z_{ij} as

$$z_{ij} = \begin{cases} \frac{d_{ij}}{n}, & \text{if there is a path from } v_i \text{ to } v_j \\ 1, & \text{if there is no path from } v_i \text{ to } v_j \end{cases} \tag{5}$$

where d_{ij} denotes the distance from v_i to v_j and n is the number of social entities in a social network. As its name implies, the Euclidean distance function $ed()$ measures structural inequivalence between social entities in a social network. By applying monotonically decreasing transformation, which is commonly used for transforming inequivalence to equivalence (Tan et al. 2006), we define the structural equivalence function $se()$ as

$$se(v_i, v_j) = \frac{1}{1 + ed(v_i, v_j)} \tag{6}$$

We thus measure the structural equivalence R_{ij} of v_i to v_j as

$$R_{ij} = \frac{se(v_i, v_j)}{\sum_{v_h \in V, h \neq j} se(v_h, v_j)} \tag{7}$$

Structural equivalence R_{ij} is within the range of [0,1]; the higher the value of R_{ij}, the greater the structural equivalence of v_i to v_j, relative to the structural equivalence of any other social entity to v_j.

Having introduced social influence I_{ij}, entity similarity M_{ij}, and structural equivalence R_{ij} from v_i to v_j, we now define persuasion probability p_{ij} that v_i persuades v_j to adopt. According to salient social network theories, social persuasion is driven by social influence, entity similarity, and structural equivalence. We therefore define, in Equation (8), persuasion probability as the probability that v_j adopts, given the impacts of social influence, entity similarity, and structural equivalence from v_i to v_j. That is,

$$p_{ij} = P\left(D_j = 1 | I_{ij}, \gamma^{d_{ij}-1} M_{ij}, \gamma^{d_{ij}-1} R_{ij}\right) \tag{8}$$

where $D_j = 1$ represents the adoption by v_j and $D_j = 0$ implies otherwise. Furthermore, γ denotes the attenuation factor, where $0 < \gamma < 1$, and $\gamma^{d_{ij}-1}$ models the scenario in which the impact of entity similarity (or structural equivalence) from v_i to v_j attenuates as the path distance d_{ji} from v_i to v_j increases (Jackson and Wolinsky 1996; Katz 1953; Watts 2001). Fang et al. (2013) develop an expectation–maximization (EM) based method to predict adoption probabilities. The major difference between adoption probability and persuasion probability is the consideration of hidden power by the former but not by the latter. According to Fang et al., adoption probability is affected by interentity impacts (including social influence, structural equivalence, and entity similarity) as well as unobserved factors (i.e., hidden power). Therefore, it is reasonable to include hidden power in the definition of adoption probability. In this study, persuasion probability is the probability that one social entity persuades another entity to adopt (i.e., interentity impacts). Thus, our definition of persuasion probability focuses on interentity impacts only and does not consider hidden power. The EM-based method by Fang et al. is developed to handle hidden power which, however, is not relevant to persuasion probability. Therefore, we propose a novel method to predict persuasion probabilities, instead of using the EM-based method.

To estimate p_{ij}, we need to construct training data from the currently observed adoption information and develop a method for estimating p_{ij} from the training data. At current time T, we observe the adoption decision of each social entity in a social network and the adoption time for those who have already adopted. If a social entity v_k adopted before current time, we set its adoption decision $D_k = 1$ and its adoption decision time as its time of adoption. If a social entity v_k has not adopted until current time, we set $D_k = 0$ and its adoption decision time as T.[3] According to prevalent social diffusion and persuasion models in social networks (Granovetter 1978; Kleinberg 2007), a social entity's adoption decision is affected by other entities that have already adopted, through their impacts of social influence, entity similarity and structural equivalence. Therefore, a social entity v_k's adoption decision D_k is affected by

$$I_k = \sum_{v_i \in V_D} I_{ik}$$
$$M_k = \sum_{v_i \in V_D} \gamma^{d_{ik}-1} M_{ik}$$
$$R_k = \sum_{v_i \in V_D} \gamma^{d_{ik}-1} R_{ik}$$

where V_D is the set of entities that adopted before v_k's adoption decision time, γ is the attenuation factor, and d_{ik} denotes the distance from v_i to v_k. We can construct training data by calculating I_k, M_k, and R_k and setting D_k for each social entity in a social network. The constructed training data thus have n records, where n is the number of social entities in the social network. Each record consists of predictor attributes I_k, M_k, and R_k and class label attribute D_k.

To estimate p_{ij} from the training data, a potential method is decision tree induction with Laplace estimation, which has been used to estimate class membership probabilities (Pazzani et al. 1994). This method first learns a decision tree from the training data. According to the value of I_{ij}, $\gamma^{d_{ij}-1} M_{ij}$, and $\gamma^{d_{ij}-1} R_{ij}$, a leaf node of the learned decision tree is reached next. The Laplace estimation of p_{ij} is then produced by

$$p_{ij} = \frac{n_D + 1}{n_L + 2} \tag{9}$$

where n_L is the number of training records in the reached leaf node and n_D is the number of training records with $D_k = 1$ in the node. Although commonly used, this method has a limitation. Laplace estimation assumes equal prior probabilities, which does not hold in most real-world scenarios (Zadrozny and Elkan 2001). In our study, Laplace estimation assumes that adoption and nonadoption cases are equiprobable (i.e., $P(D_k = 1) = P(D_k = 0) = 1/2$). This assumption

[3]For a social entity that has not adopted until current time T, its latest observable adoption decision was made at T.

is unrealistic, because nonadoption cases often substantially outnumber adoption ones (Aral et al. 2009; Fang et al. 2013). Furthermore, prior studies have shown, theoretically and empirically, the use of bagging to significantly improve probability estimation accuracy by decision trees (Breiman 1996). Bagging creates bootstrap samples from the training data,[4] learns a decision tree from each sample, and yields the average of probabilities, each of which is estimated by a decision tree. Hence, we can use bagging to enhance the accuracy of estimating p_{ij}.

We therefore propose to estimate p_{ij} using bagged decision tree induction with m-estimation. Instead of assuming equal prior probabilities, m-estimation effectively incorporates prior probabilities gauged from the training data for probability estimation (Cussens 1993). Previous research has shown that m-estimation produces more accurate probability estimations than does Laplace estimation (Sulzmann and Fürnkranz 2009). Different from Equation (9), the m-estimation of p_{ij} is given by

$$p_{ij} = \frac{n_D + mP(D_k = 1)}{n_L + m} \tag{10}$$

where parameter m can be set as $\sqrt{n_L}$ (Cussens 1993) and prior probability $P(D_k = 1)$ can be estimated as the ratio of records with $D_k = 1$ in the training data (Mitchell 1997). As we show in Figure 1, our proposed method consists of two phases: training and probability estimation. In the training phase, a set of decision trees is constructed, each of which is learned from a bootstrap sample of the training data using C4.5, a widely used decision tree learning algorithm (Quinlan 1993). Following Breiman (1996), we create 50 bootstrap samples and build 50 decision trees accordingly.[5] In the probability estimation phase, p_{ij} is estimated as the average of probability estimations across 50 decision trees.

Predicting Top Persuaders

We define the persuasion score of a social entity as the extent to which that entity can persuade other entities to adopt. Accordingly, top K persuaders refer to social entities with the highest K persuasion scores. For a social entity $v_i \in V$, let $c_i \geq 0$ be the persuasion score of that entity. Persuasion score c_i depends on the probabilities that v_i persuades other entities to adopt (i.e., p_{ij} where $v_j \in V$ and $j \neq i$). Understandably, the higher such probabilities, the greater v_i's persuasion score c_i. In addition, persuasion score c_i also depends on the persuasion scores of the social entities that v_i persuades. Intuitively, v_i could have a high persuasion score if the entities persuaded by v_i have high persuasion scores themselves. We therefore calculate c_i as

$$c_i = \sum_{v_j \in V, j \neq i} p_{ij} c_j \tag{11}$$

Let the vector of persuasion scores $c = \begin{pmatrix} c_1 \\ c_2 \\ \cdots \\ c_n \end{pmatrix}$, each element of which denotes a social entity's persuasion score. Let $\boldsymbol{P}$ be the $n \times n$ matrix of persuasion probabilities, where n is the number of social entities in a social network, a non-diagonal element of the matrix p_{ij} is the probability that v_i persuades v_j to adopt, and all diagonal elements are set to 0. By Equation (11), we have

$$\boldsymbol{C} = \boldsymbol{PC} \tag{12}$$

To predict top K persuaders, we need to compute $\boldsymbol{C}$ with Equation (12); top K persuaders are those with the highest K persuasion scores in the computed $\boldsymbol{C}$.

In response, we propose an algorithm to compute $\boldsymbol{C}$ and predict top K persuaders, based on the power iteration method (Golub and Van Loan 1996). We then show the convergence of the proposed algorithm and specify how to initialize $\boldsymbol{C}$ for the algorithm (see Figure 2).

The proposed algorithm, shown in Figure 2, first initializes $\boldsymbol{C}$, then iteratively updates $\boldsymbol{C}$ according to Equation (12) until convergence, and finally outputs social entities with the highest K persuasion scores in the converged $\boldsymbol{C}$. Two details of the algorithm warrant further explanation. First, $\|\boldsymbol{C}^{x+1}\|_\infty$ yields the maximum of the absolute values of $\boldsymbol{C}^{x+1}$'s elements: $\|\boldsymbol{C}^{x+1}\|_\infty = \max_{1 \leq i \leq n} |\boldsymbol{C}_i^{x+1}|$; and $\boldsymbol{C}^{x+1}/\|\boldsymbol{C}^{x+1}\|_\infty$ normalizes $\boldsymbol{C}^{x+1}$. This normalization is a commonly accepted practice for power iteration to avoid the risk of overflow or underflow (Heath 2002). Moreover, such normalization does not change persuasion score ranks of individual entities and therefore does not affect the final result of predicted top K persuaders. Second, one-norm $\|\boldsymbol{Z}\|_1$ of a vector $\boldsymbol{Z}$ is defined as the sum of the absolute values of its elements, that is, $\|\boldsymbol{Z}\|_1 = \sum_{i=1}^{n} |\boldsymbol{Z}_i|$

[4]A bootstrap sample contains the same number of records as the training data; each sample record is randomly drawn from the training data with replacement (Breiman 1996).

[5]There is no theory regarding how many bootstrap samples to create for bagging (Breiman 1996); we follow the rule of thumb of 50 bootstrap samples suggested by Breiman.

(a) Training	(b) Probability Estimation
Input: training data **Output**: 50 decision trees Create 50 bootstrap samples from the training data. Learn a decision tree from each sample using C4.5.	**Input**: $I_{ij}, \gamma^{d_{ij}-1} M_{ij}, \gamma^{d_{ij}-1} R_{ij}$ **Output**: p_{ij} **For** each decision tree learned Reach a leaf node according to $I_{ij}, \gamma^{d_{ij}-1} M_{ij}, \gamma^{d_{ij}-1} R_{ij}$. Estimate p_{ij} using Equation (10). **End for** Output average p_{ij} across 50 decision trees.

Figure 1. Bagged Decision Tree Induction with M-Estimation

Input: K, $\boldsymbol{P}$
Output: top K persuaders

Initialize $\boldsymbol{C} = \boldsymbol{C}^0$.
$x = 0$.
Do
    $\boldsymbol{C}^{x+1} = \boldsymbol{PC}^x$.
    $\boldsymbol{C}^{x+1} = \boldsymbol{C}^{x+1}/||\boldsymbol{C}^{x+1}||_\infty$.
    $\mu = ||\boldsymbol{C}^{x+1} - \boldsymbol{C}^x||_1/n$.
    $x = x + 1$.
Until $\mu < \epsilon$. //ϵ: predefined convergence threshold
Output social entities with the highest K persuasion scores in $\boldsymbol{C}^x$.

Figure 2. Predicting Top *K* Persuaders

(Heath 2002). Therefore, $||\boldsymbol{C}^{x+1} - \boldsymbol{C}^x||_1/n$ represents the average absolute difference of persuasion scores between two consecutive updates of $\boldsymbol{C}$. The iterative process of our algorithm terminates if the difference is sufficiently small. We now show the convergence of the algorithm.

Theorem 1: The algorithm for predicting top K persuaders converges and the converged $\boldsymbol{C}$ is unique, independent of the initialization of $\boldsymbol{C}$.

Proof: See Appendix A.

Theorem 1 guarantees the convergence of our algorithm and assures the uniqueness of the converged $\boldsymbol{C}$. In addition, Theorem 1 stipulates that both the convergence of the algorithm and the uniqueness of the converged $\boldsymbol{C}$ are independent of the initialization of $\boldsymbol{C}$. Hence, a random initialization of $\boldsymbol{C}$ is sufficient for the algorithm. Without loss of generality, we initialize $\boldsymbol{C}$ as $\boldsymbol{C}^0 = \begin{pmatrix} 1 \\ 1 \\ \dots \\ 1 \end{pmatrix}$; that is, all social entities have same persuasion scores initially.

Empirical Evaluation and Results

We evaluated our proposed method with real-world social network data to demonstrate its superior prediction performance over several prevalent methods, and then empirically analyzed its important properties. In this section, we describe the data and benchmark methods, detail our evaluation design, and report important evaluation results.

Data and Benchmark Methods

We conducted our evaluation in the context of a social network of avatars. Avatar social networks have been popularly used to analyze or evaluate social network theories and methods (Bainbridge 2007; Centola 2010). Specifically, our evaluation employed three related data sets collected from a major online game company. One data set contains 1.34 million records of complete message communications among all avatars of an online game over 20 weeks, starting from the

Table 3. Summary of Avatar Profile Attributes

Attribute Name	Attribute Type	Description
Gender	Nominal	Avatar gender: male or female
Profession	Nominal	Avatar profession, such as farmer, solider etc.
Level	Ordinal	Avatar level in the game, ranging from level 1 (lowest) to level 30 (highest)
Community	Nominal	Community to which an avatar belongs
Gold	Numeric	Amount of gold possessed by an avatar
Silver	Numeric	Amount of silver possessed by an avatar
Bonus	Numeric	Bonus points earned by an avatar
Experience	Numeric	Avatar score in experience
Intelligence	Numeric	Avatar score in intelligence
Strength	Numeric	Avatar score in strength
Agility	Numeric	Avatar score in agility
Spirit	Numeric	Avatar score in spirit
Attack	Numeric	Avatar score in attack capability
Defense	Numeric	Avatar score in defense capability

launch of the game. Each record consists of the timestamp of a message communication, and the respective identities of the two avatars participating in the communication. Using this data set, we constructed a social network of avatars, in which a social entity represents an avatar and a relationship between two entities exists if there are message communications between their corresponding avatars. We further measured the strength of social interactions between two social entities using the number of messages communicated between their corresponding avatars. For each of the 5,162 avatars, we collected data regarding whether the avatar adopted a particular virtual item during the study period and, if adopted, in which week. This particular virtual item is not a necessity for avatars. Following extant literature (e.g., Bass 1969; Aral et al. 2009), we considered the initial purchase of the item as the adoption of that item. Over the study period, a total of 1,021 avatars adopted the virtual item. The third data set contains individual profiles of the 5,162 avatars. As we summarize in Table 3, the profile of an avatar includes 14 attributes.

Because of the mixed attribute types in each profile, we applied a heterogeneous similarity function to compute profile similarity (Tan et al. 2006). For two avatar profiles $\boldsymbol{f}_i$ and $\boldsymbol{f}_j$, let $\boldsymbol{f}_{ik}$ and $\boldsymbol{f}_{jk}$ be their respective k^{th} attribute, $k = 1,2,\ldots,14$. According to Tan et al. (2006), if the k^{th} attribute is nominal, the attribute similarity $sim(\boldsymbol{f}_{ik}, \boldsymbol{f}_{jk})$ between $\boldsymbol{f}_{ik}$ and $\boldsymbol{f}_{jk}$ is given by

$$sim\left(\boldsymbol{f}_{ik}, \boldsymbol{f}_{jk}\right) = \begin{cases} 0 & if \;\; \boldsymbol{f}_{ik} \neq \boldsymbol{f}_{jk} \\ 1 & if \;\; \boldsymbol{f}_{ik} = \boldsymbol{f}_{jk} \end{cases}$$

and if the k^{th} attribute is ordinal or numeric, $sim(\boldsymbol{f}_{ik}, \boldsymbol{f}_{jk})$ is defined as

$$sim\left(\boldsymbol{f}_{ik}, \boldsymbol{f}_{jk}\right) = 1 - \frac{d - min_d}{max_d - min_d}$$

where difference $d = |\boldsymbol{f}_{ik} - \boldsymbol{f}_{jk}|$ and max_d and min_d denote the maximum and the minimum among all differences respectively. The profile similarity $sim(\boldsymbol{f}_i, \boldsymbol{f}_j)$ between $\boldsymbol{f}_i$ and $\boldsymbol{f}_j$ is then calculated as the average attribute similarity across all the profile attributes.

As we describe in our review of related research, salient prior methods focus on social influence and often employ a centrality measure to identify top persuaders as those that score high in the measure. We therefore benchmarked our proposed method against existing methods developed on the basis of major centrality measures, including degree centrality, topological centrality, percolation centrality, and eigencentrality. According to Freeman (1979), the degree centrality $deg(v_i)$ of a social entity v_i is measured as

$$deg\left(v_i\right) = \sum_{j=1}^{n} a_{ij}$$

where $a_{ij} = 1$ if there exists a relationship that connects entities v_i and v_j, $a_{ij} = 0$ otherwise, and n is the number of social entities in a social network. Unlike degree centrality, topological centrality, such as closeness or betweenness, incurs more computational cost. According to Freeman, the closeness centrality $cls(v_i)$ of entity v_i is evaluated as

$$cls\left(v_i\right) = \frac{1}{\sum_{j=1}^{n} d_{ij}}$$

where d_{ij} denotes the distance from entity v_i to entity v_j; and the betweenness centrality $btn(v_i)$ of entity v_i is given by

$$btn(v_i) = \sum_{j=1}^{n} \sum_{k=1,k>j}^{n} b_{jk}(v_i)$$

where $b_{jk}(v_i) = g_{jk}(v_i)/g_{jk}$, g_{jk} is the number of the shortest paths linking entity v_j and entity v_k, $g_{jk}(v_i)$ is the number of the shortest paths linking entity v_j and entity v_k that contain entity v_i, and $i \neq j \neq k$. Unlike betweenness centrality, percolation centrality considers percolation states of social entities, which are adoption states of entities in our study; and the percolation centrality $pc(v_i)$ of entity v_i is defined as (Piraveenan et al. 2013)

$$pc(v_i) = \frac{1}{n-2} \sum_{j-1}^{n} \sum_{k=1,k>j}^{n} b_{jk}(v_i) \frac{D_j}{\left[\sum_{h=1}^{n} D_h\right] - D_i}$$

where $D_j = 1$ if entity v_j adopts, $D_j = 0$ otherwise, and $i \neq j \neq k$. One important eigencentrality metric is eigenvector centrality. By representing a social network as an adjacency matrix $\boldsymbol{A} = [a_{ij}]_{n \times n}$, we can compute the principal eigenvector of the matrix, which contains eigenvector centrality scores for each social entity in the social network (Bonacich 1972, 1987, 2007). Another critical eigencentrality metric is intercentrality. Ballester et al. (2006) define the intercentrality of entity v_i as the square of the number of paths that start at v_i divided by the number of paths from v_i to v_i itself, where paths of length l are weighted by β^l, β is a decay factor, and $0 < \beta < 1$.

We also benchmarked our method against a method that employs the PageRank algorithm to rank social entities based on their social influences I_{ij} (Brin and Page 1998; Romero et al. 2010), namely INF-RANK in our evaluation. It is essential to compare with prior methods that consider both social influence and entity similarity for identifying top persuaders, though few such methods have been proposed. In particular, Weng et al. (2010) develop a transition matrix $\boldsymbol{B} = [b_{ij}]_{n \times n}$ for a social network and define b_{ij} as

$$b_{ij} = I_{ij} \times M_{ij}$$

where social influence I_{ij} and entity similarity M_{ij} are given by Equations (1) and (2) respectively. The principal eigenvector of the transition matrix contains rank scores of each social entity in the social network. In our evaluation, we refer to this benchmark method as INF-SIM. It is also important to include as benchmark a random selection method that randomly selects social entities as top persuaders. The inclusion of this method provides a comparison baseline; we anticipate our method and all other benchmark methods to outperform the random selection method. In Table 4, we summarize each benchmark method included in our evaluation.

In terms of time complexity, computing the degree, closeness, and betweenness centrality for each entity in a social network takes $O(r)$, $O(n^3)$, and $O(n^3)$ time respectively (Zhuge and Zhang 2010), where r is the number of relationships and n denotes the number of entities in a social network. Percolation centrality is a variation of betweenness centrality and it takes $O(n^3)$ time to calculate the percolation centrality for each entity in a social network. The time required to compute the intercentrality for each entity in a social network is dominated by the time of multiplying two $n \times n$ matrices, which takes $O(n^3)$ time. For eigenvector centrality, INF-RANK, INF-SIM, and our method (given persuasion probabilities), the respective running time is dominated by the time of multiplying an $n \times n$ matrix and an $n \times 1$ vector, which takes $O(n^2)$ time. It is worth noting that our method incurs additional time to estimate persuasion probabilities. According to our time complexity analysis, the running time of all the investigated methods, except for degree centrality, scales up superlinearly with the number of entities in a social network. Although degree centrality takes $O(r)$ time, r is close to $O(n^2)$ for a densely connected social network. Therefore, designing more efficient top-persuader prediction methods for large-scale social networks deserves future research attention.

Evaluation Design

We first introduce persuasion credit, a concept essential for evaluating different top persuader prediction methods. The persuasion credit of a social entity refers to the credit the entity receives by *persuading* other entities to adopt. It is generally accepted to measure the persuasion (or influence) power of a social entity based on other entities' adoption behaviors after the entity's adoption because these subsequent adoptions provide an observable indicator of an entity's persuasion (or influence) power (Bakshy et al. 2011; Trusov et al. 2010). We therefore follow this practice to measure persuasion credit. Specifically, persuasion credit is extended from influence score, a well-established concept proposed by Bakshy et al. (2011), which is defined as the score an entity obtains by *influencing* other entities to adopt. We use the following example to illustrate influence score and its calculation.

Consider a scenario in which social entity E adopts after the adoption by entities A, B, C, and D. According to Bakshy et al. (2011), the adoption by entity E can be attributed to the social influence of E's direct neighbors that have adopted earlier than E and each direct neighbor receives an equal credit.

Table 4. Summary of Benchmark Methods

Method	Force(s) Considered	Source(s)
Degree Centrality	Social Influence	Freeman 1979; Wasserman and Faust 1994; Borgatti et al. 2009
Closeness Centrality	Social Influence	
Betweenness Centrality	Social Influence	
Percolation Centrality	Social Influence	Piraveenan et al. 2013
Eigenvector Centrality	Social Influence	Bonacich 1972, 1987, 2007
Intercentrality	Social Influence	Ballester et al. 2006
INF-RANK	Social Influence	Brin and Page 1998; Romero et al. 2010
INF-SIM	Social Influence, Entity Similarity	Weng et al. 2010
Random Selection	/	Baseline

Table 5. An Illustration of Influence Score and Persuasion Credit

Social Entity	Distance to E	Influence Score	Persuasion Credit
A	1	0.50	0.38
B	1	0.50	0.38
C	2	0	0.19
D	4	0	0.05

As shown in Table 5, entities A and B are E's direct neighbors because their distance to E is 1. Furthermore, entities A and B adopted earlier than entity E. Therefore, the adoption by entity E is attributed to the social influence of A and B and, as a result, each receives half of a credit. As we show in Table 5, the respective influence score of A and B is 0.5, because of the contribution of E's adoption. Entities C and D are not E's direct neighbors and hence receive no credit from E's adoption. As illustrated in Table 5, the influence score of entity C or D is 0. Table 5 lists influence scores resulting from E's adoption only. As more entities become adopters, additional credits should be added to influence scores, according to the procedure described above.

Our persuasion credit extends influence score in two ways. First, influence score measures a social entity's power of social influence and therefore its consideration of only direct neighbors of the entity seems reasonable (Bakshy et al. 2011). Persuasion credit, on the other hand, measures a social entity's power of social persuasion, which results from the combined effects of social influence, entity similarity, and structural equivalence. The evaluation of a social entity's persuasion credit thus requires the proper consideration of both direct neighbors (i.e., those with distance to the entity equal 1) and indirect neighbors (i.e., those with distance to the entity greater than 1) of the entity because indirect neighbors of an entity could be persuaded by the entity through entity similarity or structural equivalence. Second, as discussed in the previous section, a social entity's power of persuading another attenuates as the distance between them increases (Jackson and Wolinsky 1996; Katz 1953; Watts 2001). It is therefore necessary to introduce the attenuation factor $0 < \gamma < 1$ in the evaluation of persuasion credit. We continue using the example in Table 5 to illustrate the calculation of persuasion credit. Let the persuasion credit of a social entity, before it is discounted by the attenuation factor, be x. Due to the adoption by E, we have

$$x\gamma^{1-1} + x\gamma^{1-1} + x\gamma^{2-1} + x\gamma^{4-1} = 1$$

The four terms in the left hand side of the above equation represent the respective persuasion credit of entities A, B, C, and D, after being discounted by the attenuation factor. By assuming $\gamma = 0.5$ and solving the equation, we note that $x = 0.38$ and obtain persuasion credits of A, B, C, and D as shown in Table 5. Table 5 again lists persuasion credits due to E's adoption only. As more entities become adopters, additional credits should be added to persuasion credits following the described procedure.

Next, we detail our evaluation design. We divided evaluation data into two parts. One part contained data over the first 10 weeks of the study period, which we used to train our method and each benchmark method for ranking social entities and predicting top persuaders for the second 10 weeks of the study period. The other part contained data over the second 10 weeks of the study period, which we used to calculate the persuasion credit of each social entity during this period. We then assigned ranks to social entities according to their persuasion credits and thereby identified top persuaders for the second 10-week period. The performance of each method was evaluated by comparing top persuaders and ranks predicted by the method against top persuaders and ranks identified with the second 10 weeks of data, using the following metrics.

We employed top-K precision (Manning et al. 2008) to compare predicted top persuaders and top persuaders identified with the second 10 weeks of data. In our evaluation, the metric is calculated as

$$\text{top-}K\text{ precision} = \frac{|\text{predicted top-}K\text{ persuaders} \cap \text{identified top-}K\text{ persuaders}|}{K}$$

We also evaluated the extent to which ranks predicted by each method agree or disagree with ranks computed with the second 10 weeks of data. A common measure is Spearman's rank correlation coefficient that ranges from –1 (i.e., complete disagreement) to 0 (i.e., independence) and then +1 (i.e., complete agreement) (Kendall and Gibbons 1990; Snedecor and Cochran 1989). The higher the coefficient, the better ranks predicted by a method correlate with ranks computed with the second 10 weeks of data (Kendall and Gibbons 1990; Snedecor and Cochran 1989). Finally, it is interesting to assess to what extent each method achieves the objective of the top persuader prediction problem (i.e., persuading as many social entities to adopt as possible). We therefore computed total persuasion credit for a method as

$$\text{total persuasion credit} = \sum_{v \in V_K} pc(v)$$

where V_K denotes the set of top-K persuaders predicted by the method and $pc(v)$ is the persuasion credit of entity v. For each method, its total persuasion credit reflects the number of social entities persuaded to adopt by the top persuaders predicted by that method.

Evaluation Results and Analyses

Following the described procedure, we conducted our evaluation, with $\gamma = 0.5$ and K varying from 50 (i.e., approximately 1% of the total number of avatars) to 500 (i.e., approximately 10% of the total number of avatars). In Table 6, we compared our method and benchmark methods in terms of top-K precision. The top-K precision of our method is 0.83, averaged across K; that is, on average, 83% of the top-K persuaders predicted by our method are among top-K persuaders identified for the second 10-week period. As shown, the top-K precision of our method is substantially higher than that of any benchmark method across K. For example, averaged across K, the top-K precision of our method is 52.69% higher than that of eigenvector centrality, the best performing benchmark method in terms of average top-K precision.[6] It is also worth noting that the top-K precision of random selection is the worst among all the investigated methods, far below that of our method or any other benchmark method. In addition to demonstrating the magnitude of our method's improvement over benchmark methods, we also applied the Wilcoxon signed-ranks test to the performance data in Table 6 (Snedecor and Cochran 1989) and noted that our method significantly outperformed each benchmark method in terms of top-K precision (p <0.001).

We also compared our method and benchmark methods in terms of the Spearman coefficient. As shown in Table 7, the average Spearman coefficient of our method is 0.77, which indicates a high, positive correlation between ranks predicted by our method and ranks computed with the second 10 weeks of data. On the other hand, the average Spearman coefficient of intercentrality, the best performing benchmark method according to average Spearman coefficient, is 0.21, far below that of our method. The average Spearman coefficient of random selection is 0, which reveals that ranks predicted by this method are generally not correlated with ranks computed with the second 10 weeks of data at all. We further applied the Wilcoxon signed-ranks test to the performance data in Table 7; the results showed that our method significantly outperformed each benchmark method in terms of Spearman coefficient ($p < 0.001$).

Table 8 summarizes total persuasion credits of our method and those of each benchmark method across K. Taking K = 50, for example, the total persuasion credit of our method is 24.55, implying that the top 50 persuaders predicted by our method take full credit for the virtual item purchase by approximate 24 other avatars. Averaged across K, the total persuasion credit of our method is 111.49. On average, the total persuasion credit of our method is 56.37% higher than that of eigenvector centrality, the best performing benchmark method according to average total persuasion credit. This improvement implies that top persuaders predicted by our method could lead to a substantially larger number of avatars to purchase than those predicted by the best performing benchmark method. Considering that the company from which we collected data gleans \$45 million annually from selling virtual items, such improvement could translate into

[6]The improvement in top-K precision (e.g., 52.69%) is calculated as the top-K precision difference between our method and a benchmark method divided by the top-K precision of the benchmark method.

Table 6. Performance Comparison on Top-*K* Precision (γ = 0.5)

K	Our Method	Degree	Closeness	Between-ness	Percolation	Eigenvector	Intercentrality (β = 0.1)	INF-RANK	INF-SIM	Random
50	0.72	0.54	0.60	0.46	0.50	0.64	0.64	0.40	0.32	0.01
100	0.81	0.56	0.60	0.48	0.52	0.61	0.58	0.42	0.40	0.02
150	0.78	0.55	0.57	0.53	0.54	0.58	0.57	0.47	0.41	0.03
200	0.82	0.55	0.55	0.54	0.54	0.58	0.57	0.45	0.44	0.04
250	0.83	0.56	0.56	0.54	0.54	0.55	0.55	0.47	0.47	0.05
300	0.85	0.54	0.53	0.52	0.53	0.52	0.52	0.48	0.48	0.06
350	0.85	0.53	0.52	0.51	0.50	0.53	0.51	0.49	0.47	0.07
400	0.87	0.50	0.51	0.50	0.49	0.51	0.50	0.49	0.46	0.08
450	0.88	0.49	0.50	0.48	0.48	0.49	0.49	0.49	0.46	0.09
500	0.89	0.48	0.48	0.48	0.48	0.49	0.48	0.49	0.45	0.10
AVG	0.83	0.53	0.54	0.50	0.51	0.55	0.54	0.46	0.44	0.06
SD	0.05	0.03	0.04	0.03	0.02	0.05	0.05	0.03	0.05	0.03

Notes: We evaluated the performance of intercentrality with decay factor β ranging from 0.1 to 0.9, in increments of 0.1. Our method substantially outperformed intercentrality in all three performance metrics, across all β values we investigated. For the interest of space, we only report the best performing intercentrality in this paper. For example, in this table, intercentrality achieves the best average top-*K* precision when β = 0.1. For the random selection method, the performances reported in Tables 6–8 as well as the tables in Appendixes B and D are the average performance of running the method 1,000 times.

Table 7. Performance Comparison on Spearman's Rank Correlation Coefficient (γ = 0.5)

K	Our Method	Degree	Closeness	Between-ness	Percolation	Eigenvector	Intercentrality (β = 0.1)	INF-RANK	INF-SIM	Random
50	0.68	0.19	0.17	0.14	0.01	0.02	0.21	0.19	0.19	0.00
100	0.65	0.19	0.13	0.15	0.05	0.12	0.17	0.12	0.16	0.00
150	0.73	0.20	0.19	0.10	0.13	0.21	0.15	0.18	0.14	0.01
200	0.73	0.12	0.20	0.09	0.10	0.14	0.16	0.20	0.11	0.01
250	0.79	0.08	0.12	0.08	0.07	0.18	0.19	0.14	0.09	0.00
300	0.80	0.14	0.17	0.14	0.15	0.19	0.19	0.16	0.11	0.00
350	0.82	0.17	0.18	0.18	0.23	0.16	0.28	0.16	0.16	0.00
400	0.83	0.25	0.20	0.24	0.26	0.19	0.26	0.20	0.21	0.00
450	0.85	0.25	0.23	0.26	0.24	0.23	0.26	0.23	0.21	0.00
500	0.86	0.25	0.25	0.26	0.24	0.23	0.28	0.24	0.22	0.00
AVG	0.77	0.19	0.18	0.16	0.15	0.17	0.21	0.18	0.16	0.00
SD	0.07	0.06	0.04	0.07	0.09	0.06	0.05	0.04	0.05	0.00

Table 8. Performance Comparison on Total Persuasion Credit (γ = 0.5)

K	Our Method	Degree	Closeness	Between-ness	Percolation	Eigenvector	Intercentrality (β = 0.1)	INF-RANK	INF-SIM	Random
50	24.55	16.81	17.34	17.40	18.40	19.18	18.80	16.91	15.68	2.89
100	47.00	31.33	32.58	30.46	32.29	33.41	32.31	31.97	28.59	5.73
150	67.80	43.49	44.78	44.45	44.94	45.08	45.73	44.61	39.84	8.48
200	87.78	56.95	55.30	57.23	57.67	58.00	57.93	54.89	51.84	11.51
250	106.42	69.11	68.03	68.70	69.31	66.84	65.89	66.90	63.26	14.23
300	124.16	78.10	76.73	77.73	78.85	76.11	75.60	77.43	73.72	17.02
350	140.97	86.87	86.01	85.88	85.44	86.89	84.25	87.88	81.69	19.88
400	157.30	92.17	94.03	92.61	92.45	93.93	92.35	95.38	88.41	22.91
450	172.30	100.67	101.37	99.25	101.02	101.14	101.11	102.78	95.65	25.63
500	186.62	108.30	107.70	106.51	108.92	109.22	108.00	110.10	101.80	28.42
AVG	111.49	68.38	68.39	68.02	68.93	68.98	68.20	68.89	64.05	15.67
SD	54.45	30.59	30.35	30.04	30.09	29.98	29.64	31.33	29.35	8.61

significant financial gains for the company. The total persuasion credit of random selection remains the worst among all the evaluated methods. By applying the Wilcoxon signed-ranks test to the performance data in Table 8, we noted that our method significantly outperformed each benchmark method in total persuasion credit (p <0.001).

To ensure the robustness of our evaluation results, we performed additional evaluations with different γ, ranging from 0.1 to 0.9, in increments of 0.1, except for $\gamma = 0.5$. By varying γ from 0.1 to 0.9, we examined our method and benchmark methods in different attenuation scenarios, from the most attenuation (γ = 0.1) to the least attenuation (γ = 0.9). We obtained evaluation results largely similar to those in Tables 6–8 and reported detailed evaluation results in Appendix B. Across the values we investigated, our method outperformed the best performing benchmark method by 39.15% to 60.11% in top-K precision, 256.53% to 280.33% in Spearman coefficient, and 36.99% to 62.78% in total persuasion credit. Together, the evaluation results in Tables 6–8 and the additional results in Appendix B consistently suggest the following. First, our method substantially outperforms each benchmark method in all three evaluation metrics (i.e., top-K precision, Spearman coefficient, and total persuasion credit), across a wide range of K and γ. Second, ranks predicted by random selection have no correlation with ranks computed with the second 10 weeks of data. This finding highlights the need to develop effective methods to rank social entities and predict top persuaders, instead of relying on random selection. The superiority of our method over benchmark methods can be attributed to two factors. First, our method is anchored in relevant social network theories and considers three distinct forces central to social persuasion suggested by these theories, whereas most benchmark methods focus on social influence exclusively. We affirm the importance of considering all three forces for top persuader prediction by demonstrating significant outperformance of our method over methods that take only two of the three forces into account in Appendix C. Second, our method properly integrates these forces to predict top persuaders, through the introduction of persuasion probability. The effect of proper integration is manifested by the substantial performance improvement of our method over INF-SIM, which multiplies social influence and entity similarity without due justification.

Finally, we empirically analyzed key properties of our proposed method, including its sensitivity to the number of bootstrap samples and its convergence. Figure 3 plots top-K precision of our method against the number of bootstrap samples created by the method, ranging from 0 to 100, for K = 100 and γ = 0.5. We note that zero bootstrap sample means no bagging, whereas greater than zero bootstrap sample indicates bagging is employed by our method. As shown in Figure 3, the top-K precision of our method without bagging is 0.77, which is lower than that of our method with bagging regardless of the number of bootstrap samples created but is substantially higher than that of the best performing benchmark method (i.e., 0.61 according to Table 6). Furthermore, the top-K precision of our method attains the maximum when the number of bootstrap samples is 50. We varied K and γ and observed largely similar results. We also studied the sensitivity of our method to the number of bootstrap samples, in terms of Spearman coefficient and total persuasion credit, and observed similar results. Overall, our empirical results suggest that bagging can improve the performance of our method, and that creating 50 bootstrap samples for bagging seems reasonable, consistent with extant literature (Breiman 1996). Moreover, according to our empirical results, our method substantially outperforms the best performing benchmark method, regardless of bagging or not bagging.

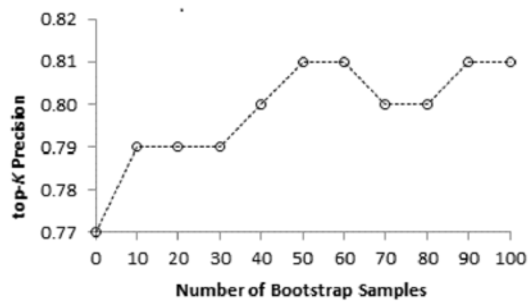


Figure 3. Top-*K* Precision of Our Method Against the Number of Bootstrap Samples Created (K = 100, γ = 0.5)

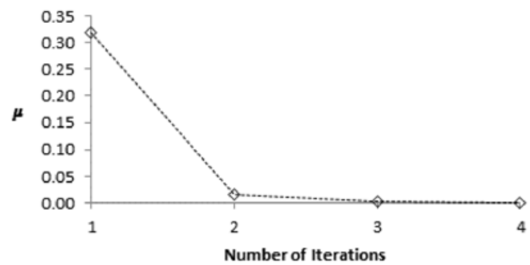


Figure 4. Convergence of Our Method (γ = 0.5)

To analyze the convergence of our method, we ran the algorithm for predicting top persuaders (Figure 2) and plotted μ after each iteration of updating $\boldsymbol{C}$ in Figure 4. As shown, the value of μ becomes trivial and the algorithm converges after four iterations for $\gamma = 0.5$. We varied γ from 0.1 to 0.9 and noted the convergence of the algorithm after three or four iterations for all γ. The observed convergence of the algorithm is consistent with Theorem 1.

Discussion and Conclusion

In the networked economy, organizations strive to learn and experiment with innovative ways of utilizing social networks for business purposes. Toward that end, predicting top persuaders in a social network is critical. This study formulates the top persuader prediction problem and develops a novel method to predict top persuaders. Our method is premised in eminent social network theories (Burt 1987; Festinger 1954; Knoke 1990; Weiner 1986), which serve as kernel theories to inform the design of our method by revealing several key forces central to social persuasion (Gregor and Hevner 2013). We consider these forces to develop a top persuader prediction method by drawing the knowledge and techniques from several disciplines that include machine learning and sociology. We empirically examine the effectiveness of our method with real-world social network data and demonstrate its superior performance over several prevalent methods.

Our study contributes to extant literature in several ways. First, we propose a novel method for predicting top persuaders in a social network. Our method simultaneously considers several important forces central to social persuasion (i.e., social influence, entity similarity, and structural equivalence), while most existing methods focus on social influence. Second, we design an innovative way to properly integrate these forces in our method. According to salient social network theories, social influence, entity similarity, and structural equivalence jointly determine social persuasion. We therefore introduce persuasion probability that denotes the likelihood of persuasion conditioned on these three essential forces. Persuasion probability thus functions as a nexus for integrating these forces in our method. We propose how to estimate persuasion probabilities, design a power iteration algorithm to predict top persuaders using the estimated persuasion probabilities, and analyze the theoretical property of the algorithm. Finally, we demonstrate the effectiveness of our proposed method by conducting an empirical evaluation involving real-world social network data. Overall, our evaluation results show that our proposed method substantially outperforms benchmark methods from representative prior research and salient industry practices.

Our study has implications for business and society as well. By analyzing massive data sets accumulated in organizations, our proposed method can effectively predict top persuaders and therefore offers great value for various social network-centric applications. For example, a firm could use our method to predict top persuaders among its potential customers, entice them to adopt its product or service, and then leverage their high persuasion power to foster such adoptions among many other customers. As shown in our evaluation results, top persuaders predicted by our method could lead to a much larger number of avatars purchasing a virtual item than those predicted by any benchmark method. Given the $14.8 billion annual global virtual goods market, which is expected to continue growing (Wohn 2014), the successful application of our method could generate significant and positive financial impacts. The application of our method is not restricted to online games; rather, this method could be used for various product or service marketing in other industries.

Another potential application of our method is the promotion of healthy lifestyles or practices, such as promoting a healthy diet in a population, or fostering the acceptance of modern medicine in a population of an underdeveloped area. In these situations, we could construct a social network for the population in which a healthy lifestyle or practice is promoted and use the proposed method to predict top persuaders in the social network. We would then target and convince these top persuaders to adopt the promoted healthy lifestyle or practice and exploit their high persuasion power to promote it to the whole population. In addition, the rapid growth of voluminous social network data encourages innovative applications of our method. An interesting case in point is to adapt our method for infectious disease containment. In this case, individuals, if infected, who would cause wide-spread contagion in a population, resemble the top persuaders we examined. Timely prediction and immunization of such contagious individuals is indispensable to effective containment efforts, which requires proper adaptation of our proposed method to the infectious disease containment scenario.

There are several directions to extend our study. In our evaluation, we follow a common practice to approximate true top persuaders with top persuaders identified according to persuasion credit. Future work is needed to identify true top persuaders for evaluating different top persuader prediction methods. It is also worthwhile to conduct field experiments to evaluate our proposed method. In such experiments, we can observe the number of followers after the adoption of top persuaders predicted by our method and measure social influence, entity similarity, and structural equivalence from a top persuader to a follower. As a result, we can directly observe the performance of our method and, at the same time, discover the mechanism leading to the observed performance. By combining our current evaluation results with archival data and field experimental outcomes, we are able to produce more comprehensive and compelling empirical evidence suggesting the effectiveness and practical value of our method. Another area worthy of future investigation is to explore other algorithms for estimating persuasion probabilities, such as Naïve Bayes or support vector machines. Understandably, the accuracy of estimated persuasion probabilities directly affects the performance of our method. By experimenting with other probability estimation algorithms, we can explore whether these algorithms could enhance the accuracy of estimated persuasion probabilities, in comparison to the decision tree-based algorithm currently used in our method, and subsequently improve the performance of our method. Furthermore, susceptibility to persuasion also seems to play an important role in the spread of adoptions (Aral and Walker 2012), in addition to persuasion. Therefore, it could be interesting to predict social entities that are most susceptible to persuasion. Finally, the avatar data used in our evaluation have several limitations that deserve future research attention. First, our evaluation results are obtained using data spanning over the first 20-week period of an online game. It would be interesting to further evaluate the performance of the proposed and benchmark methods with avatar data collected from other time periods or data sets from a different application. Toward that end, we conducted an additional evaluation with another data set and observed evaluation results suggesting the superiority of our method over the benchmark methods. We report detailed evaluation results in Appendix D. Second, we did not control for the impacts of marketing efforts in our evaluation. Such efforts could impact the adoption behaviors of Avatars and communications among them, which in turn might affect our evaluation results. Future research should find appropriate ways to control for marketing efforts. Third, we use the profile of an avatar as a proxy for the profile of a real life user who creates the avatar and implicitly assume that similarly profiled real life users will create similarly profiled avatars. Future work should collect profile data of real users to better measure entity similarity.

Acknowledgments

The authors thank the senior editor, the associate editor, and the three anonymous reviewers for their guidance and constructive comments that have tremendously improved the paper. Xiao Fang is the corresponding author of the paper.

References

Agarwal, N., Liu, H., Tang, L., and Yu, P. S. 2008. "Identifying the Influential Bloggers in a Community," in *Proceedings of the 1st ACM International Conference on Web Search and Data Mining*, Vancouver, October 26-31, pp. 207-217.

Agarwal, R., Animesh, A., and Prasad, K. 2009. "Research Note-Social Interactions and the 'Digital Divide': Explaining Variations in Internet Use," *Information Systems Research* (20:2), pp. 277-294.

Albert, R., Jeong, H., and Barabasi, A. 2000. "Error and Attack Tolerance of Complex Networks," *Nature* (406), pp. 378-382.

Anagnostopoulos, A., Kumar, R., and Mahdian, M. 2008. "Influence and Correlation in Social Networks," in *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Las Vegas, NV, August 24-27, pp. 7-15.

Aral, S., Muchnik, L., and Sundararajan, A. 2009. "Distinguishing Influence Based Contagion from Homophily Driven Diffusion in Dynamic Networks," *Proceedings of the National Academy of Sciences* (106:51), pp. 21544-21549.

Aral, S., and Walker, D. 2012. "Identifying Influential and Susceptible Members of Social Networks," *Science* (337:6092), pp. 337-341.

Bainbridge, W. S. 2007. "The Scientific Research Potential of Virtual Worlds," *Science* (317:5837), pp. 472-476.

Bakshy, E., Hofman, J. M., Mason, W. A., and Watts, D. J. 2011. "Everyone's an Influencer: Quantifying Influence on Twitter," in *Proceedings of the 4th ACM International Conference on Web Search and Data Mining*, Hong Kong, February 9-12, pp. 65-74.

Ballester, C., Calvó-Armengol, A., and Zenou, Y. 2006. "Who's Who in Networks. Wanted: The Key Player," *Econometrica* (74:5), pp. 1403-1417.

Ballester, C., Calvó-Armengol, A., and Zenou, Y. 2010. "Delinquent Networks," *Journal of the European Economic Association* (8:1), pp. 34-61.

Bass, F. 1969. "A New Product Growth for Model Consumer Durables," *Management Science* (15:5), pp. 215-227.

Bavelas, A. 1948. "A Mathematical Model for Group Structures," *Human Organization* (7:3), pp. 16-30.

Bonacich, P. 1972. "Factoring and Weighing Approaches to Status Score and Clique Identification," *Journal of Mathematical Sociology* (2:1), pp. 113-120.

Bonacich, P. 1987. "Power and Centrality: A Family of Measures," *American Journal of Sociology* (92:5), pp. 1170-1182.

Bonacich, P. 2007. "Some Unique Properties of Eigenvector Centrality," *Social Networks* (29:4), pp. 555-564.

Borgatti, S. P., Mehra, A., Brass, D. J., and Labianca, G. 2009. "Network Analysis in the Social Sciences," *Science* (323:5916), pp. 892-895.

Brass, D. J. 1984. "Being in the Right Place: A Structural Analysis of Individual Influence in an Organization," *Administrative Science Quarterly* (29:4), pp. 518-539.

Breiman, L. 1996. "Bagging Predictors," *Machine Learning* (24:2) pp. 123-140.

Brin, S., and Page, L. 1998. "The Anatomy of a Large-Scale Hypertextual Web Search Engines," *Computer Networks and ISDN Systems* (30:1-7), pp. 107-117.

Burt, R. S. 1987. "Social Contagion and Innovation: Cohesion Versus Structural Equivalence," *American Journal of Sociology* (92:6), pp. 1287-1335.

Burt, R. S., Gabbay, S. M., Holt, G., and Moran, P. 1994. "Contingent Organization as a Network Theory: The Culture-Performance Contingency Function," *Acta Sociologica* (37:4), pp. 345-370.

Carr, D. 2008. "How Obama Tapped into Social Networks' Power," *The New York Times*, November 9.

Centola, D. 2010. "The Spread of Behavior in an Online Social Network Environment," *Science* (329:5996), pp. 1194-1197.

Cha, M., Haddadi, H., Benevenuto, F., and Gummadi, K. P. 2010. "Measuring User Influence in Twitter: The Million Follower Fallacy," in *Proceedings of the 4th International AAAI Conference on Weblogs and Social Media*, Washington, DC, May 23-26, pp. 10-17.

Chaiken, S. L., Gruenfeld, D. H., and Judd, C. M. 2000. "Persuasion in Negotiations and Conflict Situations," *The Handbook of Conflict Resolution: Theory and Practice*, M. Deutsch and P. T. Coleman (Eds.), San Francisco: Jossey-Bass, pp. 144-165.

Crandall, D., Cosley, D., Huttenlocher, D., Kleinberg, J., and Suri, S. 2008. "Feedback Effects between Similarity and Social Influence in Online Communities," in *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Las Vegas, NV, August 24-27, pp. 160-168.

Cussens, J. 1993. "Bayes and Pseudo-Bayes Estimates of Conditional Probabilities and Their Reliability, " in *Proceedings of the European Conference on Machine Learning*, San Mateo, CA: Morgan Kaufmann, pp. 136-152.

Fang, X., Hu, P. J., Li, Z., and Tsai, W. 2013. "Predicting Adoption Probabilities in Social Networks," *Information Systems Research* (24:1), pp.128-145.

Festinger, L. 1954. "A Theory of Social Comparison Processes," *Human Relations* (7:2), pp.117-140.

Forbes Insights. 2011. "Bringing 20/20 Foresight to Marketing," May (http://www.forbes.com/forbesinsights/bringing_foresight_to_marketing/).

Freeman, L. C. 1979. "Centrality in Social Networks: Conceptual Clarification," *Social Networks* (1:3), pp. 215-239.

Friedkin, N. E. 1998. *A Structural Theory of Social Influence*, New York: Cambridge University Press.

Ghosh, R., and Lerman, K. 2010. "Predicting Influential Users in Online Social Networks," in *Proceedings of the 4th KDD Workshop on Social Network Analysis*, Washington, DC, July 25-28.

Golub, G. H., and Van Loan, C. F. 1996. *Matrix Computations* (3rd ed.), Baltimore, MD: Johns Hopkins University Press.

Granovetter, M. 1978. "Threshold Models of Collective Behavior," *American Journal of Sociology* (83:6), pp. 1420-1443.

Gregor, S., and Hevner, A. 2013. "Positioning and Presenting Design Science Research for Maximum Impact," *MIS Quarterly* (37:2), pp. 337-355.

Heath, M. T. 2002. *Scientific Computing: An Introductory Survey* (2nd ed.), New York: McGraw-Hill.

Hemp, P. 2006. "Avatar-Based Marketing," *Harvard Business Review* (84:6), pp. 48-57.

Jackson, M. O., and Wolinsky, A. 1996. "A Strategic Model of Social and Economic Networks," *Journal of Economic Theory* (71:1), pp. 44-74.

Katz, L. 1953. "A New Status Index Derived from Sociometric Analysis," *Psychometrika* (18:1), pp. 39-43.

Kempe, D., Kleinberg, J., and Tardos, E. 2003. "Maximizing the Spread of Influence through a Social Network," in *Proceedings of the 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Washington, DC, August 24-27, pp.137-146.

Kendall, M. G., and Gibbons, J. D. 1990. *Rank Correlation Methods*, Oxford, UK: Oxford University Press.

Kleinberg, J. 2007. "Cascading Behavior in Networks: Algorithmic and Economic Issues," in *Algorithmic Game Theory*, Cambridge, UK: Cambridge University Press.

Kleinberg, J. 2008. "The Convergence of Social and Technological Networks," *Communications of the ACM* (51:11), pp. 66-72.

Knoke, D. 1990. "Networks of Political Action: Toward Theory Construction," *Social Forces* (68:4), pp. 1041-1063.

Lorrain, F., and White, H. C. 1971. "Structural Equivalence of Individuals in Social Networks," *Journal of Mathematical Sociology* (1:1), pp. 49-80.

Manning, C. D., Raghavan, P., and Schütze, H. 2008. *Introduction to Information Retrieval*, Cambridge, UK: Cambridge University Press.

Mitchell, T. M. 1997. *Machine Learning*, New York: McGraw-Hill. (New chapters can be accessed at http://www.cs.cmu.edu/~tom/NewChapters.html.)

Pazzani, M., Merz, C., Murphy, P., Ali, K., Hume, T., and Brunk, C. 1994. "Reducing Misclassification Costs," in *Proceedings of the 11th International Conference of Machine Learning*, Boca Raton, FL, December 12-15, pp. 217-225.

Petty, R. E., and Cacioppo, J. T. 1986. *Communication and Persuasion: Central and Peripheral Routes to Attitude Change*, New York: Springer-Verlag.

Pfeffer, J., Salancik, G. R., and Leblebici, H. 1976. "The Effect of Uncertainty on the Use of Social Influence in Organizational Decision Making," *Administrative Science Quarterly* (21:2), pp. 227-245.

Piraveenan, M., Prokopenko, M., and Hossain, L. 2013. "Percolation Centrality: Quantifying Graph-Theoretic Impact of Nodes During Percolation in Networks," *PLoS One* (8:1), pp. 1-14.

Probst, F., Grosswiele, L., and Pfleger, R. 2013. "Who Will Lead and Who Will Follow: Identifying Influential Users in Online Social Networks—A Critical Review and Future Research Directions," *Business and Information Systems Engineering* (5:3), pp. 179-193.

Quinlan, J. R. 1993. *C4.5: Programs for Machine Learning*, San Francisco: Morgan Kaufmann Publishers.

Rice, R. E., and Aydin, C. 1991. "Attitudes toward New Organizational Technology: Network Proximity as a Mechanism for Social Information Processing," *Administrative Science Quarterly* (36:2), pp. 219-244.

Rice, R. E., Grand, A. E., Schmitz, J., and Torobin, J. 1990. "Individual and Network Influences on the Adoption and Perceived Outcomes of Electronic Messaging," *Social Networks* (12:1), pp. 27-55.

Romero, D. M., Galuba, W., Asur, S., and Huberman, B. A. 2010. "Influence and Passivity in Social Media," Working Paper, HP Labs.

Shalizi C. R., and Thomas, A. C. 2011. "Homophily and Contagion are Generically Confounded in Observational Social Network Studies," *Social Methods Research* (40:2), pp. 211-239.

Simon, H. A. 1976. *Administrative Behavior*, New York: Free Press.

Snedecor, G. W., and Cochran, W. G. 1989. *Statistical Methods*, Ames, IA: Blackwell Publishing Professional.

Sulzmann, J. N., and Furnkranz, J. 2009. "An Empirical Comparison of Probability Estimation Techniques for Probabilistic Rules," in *Proceedings of the 12th International Conference on Discovery Science*, Porto, Portugal, October 3-5, pp. 317-331.

Susarla, A., Oh, J. H., and Tan, Y. 2012. "Social Networks and the Diffusion of User-generated Content: Evidence from YouTube," *Information Systems Research* (23:1), pp. 23-41.

Tan, P. N., Steinbach, M., and Kumar, V. 2006. *Introduction to Data Mining*, Boston: Pearson-Addison Wesley.

Travers, J., and Milgram, S. 1969. "An Experimental Study of the Small World Problem," *Sociometry* (32:4), pp. 425-443.

Trusov, M., Bodapati, A. V., and Bucklin, R. E. 2010. "Determining Influential Users in Internet Social Networks," *Journal of Marketing Research* (47:4), pp. 643-658.

Valente, T. W. 1995. *Network Models of the Diffusion of Innovations*, Cresskill, NJ: Hampton Press.

Walter, W. P., Koput, K. W., and Smith-Doerr, L. 1996. "Interorganizational Collaboration and the Locus of Innovation: Networks of Learning in Biotechnology," *Administrative Science Quarterly* (41:1), pp. 116-145.

Wasserman, S., and Faust, K. 1994. *Social Network Analysis: Methods and Applications*, Cambridge, UK: Cambridge University Press.

Watts, A. 2001. "A Dynamic Model of Network Formation," *Games and Economic Behavior* (34:1), pp. 331-341.

Watts, D. J., and Dodds, P. S. 2007. "Influentials, Networks, and Public Opinion Formation," *Journal of Consumer Research* (34:4), pp. 441-458.

Watts, D. J., and Peretti, J. 2007. "Viral Marketing for the Real World," *Harvard Business Review* (85:5), pp. 22-23.

Weiner, B. 1974. *Achievement Motivation and Attribution Theory*, Morristown, NJ: General Learning Press.

Weiner, B. 1986. *An Attributional Theory of Motivation and Emotion*, New York: Springer Verlag.

Wejnert, B. 2002. "Integrating Models of Di usion of Innovations: A Conceptual Framework," *Annual Review of Sociology* (28), pp. 297-326.

Weng, J., Lim, E., Jiang, J., and He, Q. 2010. "TwitterRank: Finding Topic-sensitive Influential Twitterers," in *Proceedings* *of the 3rd ACM International Conference on Web Search and Data Mining*, New York, February 3-6, pp. 261-270.

White, D. R., and Borgatti, S. P. 1994. "Betweenness Centrality Measures for Directed Graphs," *Social Networks* (16:4), pp. 335-346.

Wohn, D. 2014. "Spending Real Money: Purchasing Patterns of Virtual Goods in an Online Social Game," in *Proceedings of the 2014 ACM SIGCHI Conference on Human Factors in Computing Systems*, Toronto, April 26-May 1, pp. 3359-3368.

Wood, W. 2000. "Attitude Change: Persuasion and Social Influence," *Annual Review of Psychology* (51), pp. 539-570.

Zadrozny, B., and Elkan, C. 2001. "Learning and Making Decisions When Costs and Probabilities Are Both Unknown," in *Proceedings of the 7th International Conference on Knowledge Discovery and Data Mining*, San Francisco, August 26-29, pp. 204-213.

Zanden, J. W. 1987. *Social Psychology* (4th ed.), New York: Random House.

Zhao, K., Yen, J., Greer, G., Qiu, B., Mitra, P., and Portier, K. 2014. "Finding Influential Users of Online Health Communities: A New Metric Based on Sentiment Influence," *Journal of the American Medical Informatics Association* (21:e2), pp. e212-218.

Zhuge, H., and Zhang, J. 2010. "Topological Centrality and its e-Science Applications," *Journal of the American Society for Informatics Science and Technology* (61:9), pp. 1824-1841.

About the Authors

Xiao Fang is an associate professor of Management Information Systems at the Lerner College of Business and Economics, University of Delaware. He received his B.S. and M.S. degrees from Fudan University, China, and his Ph.D. in Management Information Systems from the University of Arizona. He studies business and social network analytics with research methods and tools drawn from reference disciplines including Management Science (e.g., Optimization) and Computer Science (e.g., Data Mining and Machine Learning). He has published in *MIS Quarterly*, *Management Science*, *Information Systems Research*, *Operations Research*, *INFORMS Journal on Computing*, *Journal of Management Information Systems*, *ACM Transactions on Information Systems*, *IEEE Transactions on Knowledge and Data Engineering*, among others.

Paul Jen-Hwa Hu is David Eccles Chair Professor at the David Eccles School of Business, the University of Utah. He received his Ph.D. in Nanagement Information Systems from the University of Arizona. His current research interests include information technology for health care, technology implementation management, digital transformation of business and government, technology impacts and evaluations, and knowledge management. Paul has published in *MIS Quarterly, Information Systems Research, Journal of Management Information Systems, Journal of the AIS, Decision Sciences*, and various IEEE and ACM transactions.

MIS Quarterly

RESEARCH ARTICLE

TOP PERSUADER PREDICTION FOR SOCIAL NETWORKS

Xiao Fang
Lerner College of Business and Economics, University of Delaware,
Newark, DE 19716-3130 U.S.A. {xfang@udel.edu}

Paul Jen-Hwa Hu
David Eccles School of Business, University of Utah,
Salt Lake City, UT 84112 U.S.A. {paul.hu@eccles.utah.edu}

Appendix A

Proof of Theorem 1

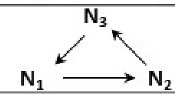
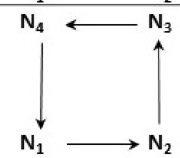
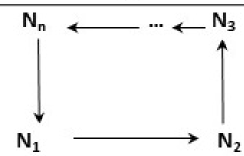
Our proof is built on Perron-Frobenius theorem, a seminal work in matrix theory (Meyer 2000). By Perron-Frobenius theorem, the power iteration algorithm for predicting top K persuaders converges to a unique $\boldsymbol{C}$ and this convergence is independent of the initialization of $\boldsymbol{C}$ if the persuasion probability matrix $\boldsymbol{P}$ is nonnegative, irreducible, and aperiodic (Heath 2002). We first show that $\boldsymbol{P}$ is nonnegative. Each component of the right hand side of Equation (10) is positive except $n_D \geq 0$; thus, persuasion probability p_{ij} estimated with Equation (10) is positive, for all $i, j = 1, 2, \ldots, n$ and $i \neq j$. Because all diagonal elements of $\boldsymbol{P}$ are equal to zero and all non-diagonal elements of $\boldsymbol{P}$ are positive persuasion probabilities, $\boldsymbol{P}$ is nonnegative.

We now prove that $\boldsymbol{P}$ is irreducible and aperiodic. Let $G(\boldsymbol{P})$ be the directed graph associated with $\boldsymbol{P}$. According to Meyer (2000), $G(\boldsymbol{P})$ is defined as a directed graph with n nodes $\{N_1, N_2, \ldots, N_n\}$, and there exists an edge from N_i to N_j if element p_{ij} of $\boldsymbol{P}$ is positive, where $i, j = 1, 2, \ldots, n$ and $i \neq j$. To prove $\boldsymbol{P}$ is irreducible and aperiodic is equivalent to show that $G(\boldsymbol{P})$ is strongly connected and aperiodic (Meyer 2000). We first show $G(\boldsymbol{P})$ is strongly connected. A directed graph is strongly connected if for every ordered pair N_i, N_j of its nodes there exists a path from N_i to N_j (West 2001). Since $p_{ij} > 0$ for all $i, j = 1, 2, \ldots, n$ and $i \neq j$, by the definition of $G(\boldsymbol{P})$, there is an edge from N_i to N_j for any ordered pair N_i, N_j of $G(\boldsymbol{P})$. That is, for any ordered pair N_i, N_j of $G(\boldsymbol{P})$, there is a path of length 1 from N_i to N_j. Therefore, $G(\boldsymbol{P})$ is strongly connected.

The period of a directed graph is the greatest common divisor of the lengths of its cycles and a directed graph is aperiodic if its period is 1 (Denardo 1977; Jarvis and Shier 1999). As discussed above, there exists an edge from N_i to N_j for any ordered pair N_i, N_j of $G(\boldsymbol{P})$. We thus list cycles that start and end at N_1 and their respective lengths in Table A1.

Apparently, the greatest common divisor of the cycle lengths in Table A1 is 1. Table A1 only lists part of the cycles in $G(\boldsymbol{P})$. Hence, the greatest common divisor of the lengths of all the cycles in $G(\boldsymbol{P})$ must be 1, and by definition, $G(\boldsymbol{P})$ is aperiodic. This completes the proof.

Table A1. Some Cycles in *G*(*P*) and Their Lengths

Cycle	Cycle Length
$N_1 \rightleftarrows N_2$	2
	3
	4
...	...
	n

References

Denardo, E. V. 1977. "Periods of Connected Networks and Powers of Nonnegative Matrices," *Mathematics of Operations Research* (2:1), pp. 20-24.

Heath, M. T. 2002. *Scientific Computing: An Introductory Survey* (2nd ed.), New York: McGraw-Hill.

Jarvis, J. P., and Shier, D. R. 1999. "Graph-Theoretic Analysis of Finite Markov Chains," in *Applied Mathematical Modeling: A Multidisciplinary Approach*, Boca Raton, FL: CRC Press.

Meyer, C. D. 2000. *Matrix Analysis and Applied Linear Algebra*, Philadelphia, PA: Society for Industrial and Applied Mathematics.

West, D. B. 2001. *Introduction to Graph Theory* (2nd ed.), Upper Saddle River, NJ: Prentice Hall.

Appendix B

Additional Evaluation Results

We conducted additional evaluations with γ ranging from 0.1 to 0.9 in increments of 0.1, except for $\gamma = 0.5$. As summarized in Tables B1 to B24, our method substantially outperformed each benchmark method in all three evaluation metrics, across γ and K.

Table B1. Performance Comparison on Top-*K* Precision ($\gamma = 0.1$)

K	Our Method	Degree	Closeness	Between-ness	Percolation	Eigenvector	Intercentrality ($\beta = 0.1$)	INF-RANK	INF-SIM	Random
50	0.70	0.56	0.58	0.50	0.52	0.60	0.58	0.42	0.36	0.01
100	0.78	0.57	0.58	0.49	0.53	0.58	0.55	0.45	0.41	0.02
150	0.77	0.56	0.57	0.53	0.55	0.59	0.58	0.49	0.42	0.03
200	0.76	0.57	0.56	0.55	0.55	0.57	0.56	0.47	0.46	0.04
250	0.75	0.57	0.55	0.55	0.54	0.53	0.52	0.48	0.49	0.05
300	0.75	0.55	0.55	0.54	0.54	0.54	0.54	0.51	0.49	0.06
350	0.78	0.55	0.54	0.52	0.51	0.55	0.53	0.52	0.48	0.07
400	0.80	0.52	0.54	0.53	0.50	0.53	0.53	0.51	0.48	0.08
450	0.79	0.52	0.52	0.52	0.51	0.52	0.52	0.51	0.48	0.09
500	0.80	0.51	0.51	0.51	0.50	0.52	0.51	0.50	0.48	0.10
AVG	0.77	0.55	0.55	0.52	0.52	0.55	0.54	0.49	0.45	0.06
SD	0.03	0.02	0.02	0.02	0.02	0.03	0.02	0.03	0.04	0.03

Table B2. Performance Comparison on Spearman's Rank Correlation Coefficient ($\gamma = 0.1$)

K	Our Method	Degree	Closeness	Between-ness	Percolation	Eigenvector	Intercentrality ($\beta = 0.1$)	INF-RANK	INF-SIM	Random
50	0.54	0.18	0.15	0.12	0.00	0.01	0.21	0.16	0.18	-0.01
100	0.60	0.19	0.13	0.15	0.04	0.12	0.16	0.12	0.16	0.00
150	0.67	0.18	0.19	0.10	0.14	0.20	0.14	0.18	0.14	0.00
200	0.71	0.12	0.20	0.09	0.11	0.14	0.15	0.19	0.11	0.00
250	0.76	0.08	0.12	0.08	0.08	0.17	0.19	0.14	0.09	0.00
300	0.78	0.14	0.17	0.14	0.15	0.19	0.18	0.16	0.11	0.00
350	0.77	0.17	0.17	0.18	0.23	0.16	0.28	0.16	0.16	0.00
400	0.78	0.25	0.20	0.24	0.26	0.19	0.26	0.20	0.21	0.00
450	0.82	0.25	0.23	0.26	0.24	0.23	0.26	0.23	0.21	0.00
500	0.84	0.26	0.25	0.26	0.24	0.23	0.28	0.24	0.22	0.00
AVG	0.73	0.18	0.18	0.16	0.15	0.16	0.21	0.18	0.16	0.00
SD	0.10	0.06	0.04	0.07	0.09	0.07	0.05	0.04	0.05	0.00

Table B3. Performance Comparison on Total Persuasion Credit ($\gamma = 0.1$)										
K	**Our Method**	**Degree**	**Closeness**	**Between-ness**	**Percolation**	**Eigenvector**	**Intercentrality ($\beta = 0.1$)**	**INF-RANK**	**INF-SIM**	**Random**
50	43.92	31.24	32.05	31.93	33.31	35.05	34.37	29.39	27.45	2.87
100	77.81	54.96	56.55	52.03	55.39	57.62	55.54	51.63	47.11	5.69
150	104.84	74.15	75.30	74.01	73.75	75.99	75.84	70.52	62.78	8.60
200	127.70	92.10	89.93	91.35	91.29	93.11	92.11	83.48	79.64	11.49
250	147.22	107.75	106.42	105.84	105.21	105.02	102.98	97.88	95.19	14.13
300	163.48	119.04	117.42	116.70	117.39	116.85	115.97	111.76	108.96	17.22
350	179.63	130.61	128.23	126.67	124.86	129.27	125.82	125.39	117.92	19.96
400	192.71	136.62	138.02	134.54	133.14	137.40	135.15	134.18	126.94	22.76
450	202.45	146.50	146.56	142.56	142.62	145.85	145.52	142.63	135.41	25.83
500	210.62	155.08	153.40	150.29	151.34	154.46	153.05	150.10	141.22	28.45
AVG	145.04	104.81	104.39	102.59	102.83	105.06	103.64	99.70	94.26	15.70
SD	55.85	41.03	40.32	39.69	38.90	39.46	39.23	40.48	38.92	8.63

Table B4. Performance Comparison on Top-*K* Precision ($\gamma = 0.2$)										
K	**Our Method**	**Degree**	**Closeness**	**Between-ness**	**Percolation**	**Eigenvector**	**Intercentrality ($\beta = 0.6$)**	**INF-RANK**	**INF-SIM**	**Random**
50	0.72	0.54	0.56	0.48	0.50	0.62	0.60	0.40	0.34	0.01
100	0.79	0.57	0.59	0.49	0.53	0.59	0.57	0.44	0.41	0.02
150	0.77	0.55	0.57	0.53	0.55	0.59	0.58	0.49	0.42	0.03
200	0.80	0.57	0.57	0.56	0.55	0.58	0.58	0.47	0.46	0.04
250	0.79	0.57	0.56	0.55	0.55	0.54	0.54	0.48	0.48	0.05
300	0.80	0.56	0.55	0.53	0.53	0.54	0.54	0.51	0.49	0.06
350	0.83	0.55	0.54	0.53	0.51	0.55	0.53	0.52	0.49	0.07
400	0.83	0.51	0.52	0.51	0.49	0.52	0.51	0.50	0.47	0.08
450	0.84	0.50	0.50	0.50	0.49	0.50	0.50	0.49	0.46	0.09
500	0.84	0.50	0.49	0.49	0.49	0.50	0.50	0.50	0.46	0.10
AVG	0.80	0.54	0.55	0.52	0.52	0.55	0.54	0.48	0.45	0.06
SD	0.04	0.03	0.03	0.03	0.03	0.04	0.04	0.04	0.05	0.03

Table B5. Performance Comparison on Spearman's Rank Correlation Coefficient ($\gamma = 0.2$)										
K	**Our Method**	**Degree**	**Closeness**	**Between-ness**	**Percolation**	**Eigenvector**	**Intercentrality ($\beta = 0.1$)**	**INF-RANK**	**INF-SIM**	**Random**
50	0.59	0.19	0.16	0.13	0.00	0.01	0.21	0.18	0.18	0.00
100	0.65	0.19	0.13	0.15	0.04	0.12	0.16	0.12	0.16	0.00
150	0.70	0.17	0.19	0.10	0.14	0.20	0.14	0.18	0.14	0.00
200	0.72	0.12	0.20	0.09	0.10	0.14	0.16	0.19	0.11	0.00
250	0.78	0.08	0.12	0.08	0.08	0.17	0.19	0.14	0.09	0.00
300	0.78	0.14	0.17	0.14	0.15	0.19	0.19	0.16	0.11	0.00
350	0.80	0.17	0.17	0.18	0.23	0.16	0.28	0.16	0.16	0.00
400	0.81	0.25	0.20	0.24	0.26	0.19	0.26	0.20	0.21	0.00
450	0.84	0.25	0.23	0.26	0.24	0.23	0.26	0.23	0.21	0.00
500	0.86	0.26	0.25	0.26	0.24	0.23	0.28	0.24	0.22	0.00
AVG	0.75	0.18	0.18	0.16	0.15	0.16	0.21	0.18	0.16	0.00
SD	0.09	0.06	0.04	0.07	0.09	0.06	0.05	0.04	0.05	0.00

Table B6. Performance Comparison on Total Persuasion Credit (γ = 0.2)

K	Our Method	Degree	Closeness	Between-ness	Percolation	Eigenvector	Intercentrality (β = 0.1)	INF-RANK	INF-SIM	Random
50	36.02	25.15	25.93	25.74	27.06	28.49	27.97	24.25	22.48	2.84
100	65.59	45.27	46.86	43.17	45.94	47.89	46.29	43.72	39.54	5.74
150	90.76	61.92	63.20	62.00	62.27	63.75	64.09	60.15	53.34	8.43
200	113.39	78.29	76.47	77.92	78.07	79.59	79.10	72.23	68.39	11.38
250	132.90	92.97	91.86	91.39	91.48	90.47	88.96	85.78	82.52	14.27
300	150.89	103.63	102.14	101.73	102.78	101.57	100.86	98.60	95.23	17.08
350	167.05	114.11	112.68	111.14	109.91	113.77	110.58	111.12	103.96	19.87
400	181.46	120.02	122.03	118.73	117.80	121.68	119.70	119.52	111.98	22.88
450	193.51	129.55	130.31	126.35	127.08	129.87	129.64	127.64	120.08	25.82
500	204.28	138.02	137.21	134.02	135.65	138.55	137.19	135.18	126.20	28.45
AVG	133.59	90.89	90.87	89.22	89.81	91.56	90.44	87.82	82.37	15.68
SD	56.46	37.52	37.13	36.39	35.91	36.48	36.16	37.25	35.45	8.65

Table B7. Performance Comparison on Top-*K* Precision (γ = 0.3)

K	Our Method	Degree	Closeness	Between-ness	Percolation	Eigenvector	Intercentrality (β = 0.6)	INF-RANK	INF-SIM	Random
50	0.72	0.56	0.58	0.48	0.52	0.64	0.62	0.42	0.34	0.01
100	0.84	0.58	0.61	0.50	0.53	0.61	0.58	0.44	0.42	0.02
150	0.79	0.55	0.57	0.53	0.54	0.58	0.57	0.47	0.42	0.03
200	0.83	0.56	0.56	0.55	0.55	0.58	0.58	0.46	0.46	0.04
250	0.81	0.56	0.56	0.55	0.54	0.54	0.54	0.48	0.48	0.05
300	0.83	0.55	0.54	0.52	0.53	0.54	0.53	0.49	0.48	0.06
350	0.84	0.54	0.53	0.52	0.51	0.54	0.52	0.51	0.48	0.07
400	0.85	0.51	0.52	0.51	0.50	0.52	0.51	0.50	0.47	0.08
450	0.86	0.50	0.50	0.49	0.49	0.50	0.50	0.49	0.46	0.09
500	0.86	0.49	0.48	0.48	0.48	0.49	0.49	0.49	0.45	0.10
AVG	0.82	0.54	0.55	0.51	0.52	0.55	0.54	0.47	0.45	0.06
SD	0.04	0.03	0.04	0.02	0.02	0.05	0.04	0.03	0.04	0.03

Table B8. Performance Comparison on Spearman's Rank Corrrelation Coefficient (γ = 0.3)

K	Our Method	Degree	Closeness	Between-ness	Percolation	Eigenvector	Intercentrality (β = 0.3)	INF-RANK	INF-SIM	Random
50	0.57	0.19	0.16	0.13	0.01	0.01	0.21	0.18	0.18	0.00
100	0.64	0.19	0.13	0.15	0.04	0.12	0.16	0.12	0.16	0.00
150	0.75	0.17	0.19	0.10	0.13	0.21	0.14	0.18	0.14	0.00
200	0.75	0.12	0.20	0.09	0.10	0.14	0.16	0.20	0.11	0.00
250	0.78	0.08	0.12	0.08	0.07	0.17	0.19	0.14	0.09	0.00
300	0.81	0.14	0.17	0.14	0.15	0.19	0.19	0.16	0.11	0.00
350	0.85	0.17	0.17	0.18	0.23	0.16	0.28	0.16	0.16	0.00
400	0.86	0.25	0.20	0.24	0.26	0.19	0.26	0.20	0.21	0.00
450	0.88	0.25	0.23	0.26	0.24	0.23	0.26	0.23	0.21	0.00
500	0.89	0.26	0.25	0.26	0.24	0.23	0.28	0.24	0.22	0.00
AVG	0.78	0.18	0.18	0.16	0.15	0.16	0.21	0.18	0.16	0.00
SD	0.10	0.06	0.04	0.07	0.09	0.06	0.05	0.04	0.05	0.00

Table B9. Performance Comparison on Total Persuasion Credit (γ = 0.3)

K	Our Method	Degree	Closeness	Between-ness	Percolation	Eigenvector	Intercentrality (β = 0.1)	INF-RANK	INF-SIM	Random
50	30.93	21.43	22.12	22.01	23.21	24.37	23.92	21.00	19.46	2.86
100	57.84	39.14	40.62	37.60	39.96	41.58	40.21	38.64	34.76	5.72
150	81.23	54.08	55.23	54.35	54.78	55.67	56.20	53.45	47.47	8.57
200	103.27	69.12	67.40	69.04	69.32	70.38	70.11	64.82	61.25	11.31
250	123.06	82.84	81.78	81.73	82.09	80.45	79.21	77.77	74.27	14.30
300	141.13	92.89	91.46	91.59	92.70	90.87	90.26	89.69	86.11	17.15
350	157.21	102.67	101.57	100.52	99.62	102.60	99.63	101.37	94.57	19.94
400	172.49	108.38	110.45	107.81	107.17	110.22	108.41	109.43	102.02	22.78
450	186.30	117.56	118.40	115.06	116.21	118.07	117.93	117.28	109.79	25.53
500	198.69	125.76	125.15	122.62	124.57	126.60	125.29	124.78	116.00	28.36
AVG	125.22	81.39	81.42	80.23	80.96	82.08	81.12	79.82	74.57	15.65
SD	56.27	34.75	34.47	33.85	33.59	33.95	33.61	34.87	32.96	8.59

Table 10. Performance Comparison on Top-*K* Precision (γ = 0.4)

K	Our Method	Degree	Closeness	Between-ness	Percolation	Eigenvector	Intercentrality (β = 0.1)	INF-RANK	INF-SIM	Random
50	0.74	0.54	0.60	0.46	0.50	0.64	0.64	0.40	0.32	0.01
100	0.82	0.57	0.60	0.49	0.52	0.61	0.58	0.43	0.41	0.02
150	0.78	0.55	0.57	0.52	0.53	0.58	0.57	0.47	0.41	0.03
200	0.82	0.56	0.56	0.55	0.55	0.58	0.58	0.46	0.45	0.04
250	0.81	0.56	0.56	0.54	0.54	0.55	0.54	0.48	0.47	0.05
300	0.85	0.55	0.54	0.52	0.53	0.53	0.53	0.49	0.48	0.06
350	0.85	0.54	0.53	0.52	0.51	0.54	0.52	0.50	0.48	0.07
400	0.87	0.50	0.51	0.51	0.49	0.51	0.50	0.49	0.46	0.08
450	0.87	0.50	0.50	0.49	0.48	0.50	0.50	0.50	0.46	0.09
500	0.87	0.49	0.48	0.48	0.48	0.49	0.49	0.49	0.45	0.10
AVG	0.83	0.53	0.54	0.51	0.51	0.55	0.54	0.47	0.44	0.06
SD	0.04	0.03	0.04	0.03	0.02	0.05	0.05	0.03	0.05	0.03

Table B11. Performance Comparison on Spearman's Rank Correlation Coefficient (γ = 0.4)

K	Our Method	Degree	Closeness	Between-ness	Percolation	Eigenvector	Intercentrality (β = 0.6)	INF-RANK	INF-SIM	Random
50	0.61	0.19	0.17	0.14	0.01	0.02	0.21	0.19	0.19	0.00
100	0.63	0.19	0.13	0.15	0.04	0.12	0.17	0.12	0.16	-0.01
150	0.73	0.18	0.19	0.10	0.13	0.21	0.14	0.18	0.14	0.00
200	0.72	0.12	0.20	0.09	0.10	0.14	0.16	0.20	0.11	0.00
250	0.79	0.08	0.12	0.08	0.07	0.17	0.19	0.14	0.09	-0.01
300	0.80	0.14	0.17	0.14	0.15	0.19	0.19	0.16	0.11	0.00
350	0.83	0.17	0.17	0.18	0.23	0.16	0.28	0.16	0.16	0.00
400	0.83	0.25	0.20	0.24	0.26	0.19	0.26	0.20	0.21	0.00
450	0.86	0.25	0.23	0.26	0.24	0.23	0.26	0.23	0.21	0.00
500	0.87	0.25	0.25	0.26	0.24	0.23	0.28	0.24	0.22	0.00
AVG	0.77	0.18	0.18	0.16	0.15	0.17	0.21	0.18	0.16	0.00
SD	0.09	0.06	0.04	0.07	0.09	0.06	0.05	0.04	0.05	0.00

Table B12. Performance Comparison on Total Persuasion Credit ($\gamma = 0.4$)

K	Our Method	Degree	Closeness	Between-ness	Percolation	Eigenvector	Intercentrality ($\beta = 0.1$)	INF-RANK	INF-SIM	Random
50	27.34	18.80	19.41	19.39	20.48	21.43	21.02	18.70	17.32	2.89
100	51.67	34.73	36.09	33.57	35.64	36.98	35.77	34.90	31.29	5.73
150	73.52	48.36	49.37	48.79	49.27	49.73	50.34	48.50	43.19	8.53
200	94.58	62.32	60.64	62.45	62.81	63.47	63.33	59.29	55.99	11.47
250	113.56	75.20	74.14	74.47	74.98	72.88	71.81	71.74	68.14	14.32
300	132.04	84.69	83.29	83.90	85.02	82.68	82.14	82.90	79.23	17.13
350	148.50	93.92	92.98	92.42	91.77	93.93	91.14	93.91	87.44	19.94
400	164.64	99.42	101.41	99.41	99.02	101.26	99.57	101.68	94.48	22.97
450	178.92	108.26	109.04	106.35	107.83	108.78	108.69	109.30	101.97	25.62
500	192.41	116.17	115.58	113.76	115.95	117.09	115.83	116.71	108.17	28.59
AVG	117.72	74.19	74.19	73.45	74.28	74.82	73.96	73.76	68.72	15.72
SD	55.49	32.48	32.24	31.79	31.69	31.82	31.47	32.95	30.99	8.64

Table B13. Performance Comparison on Top-*K* Precision (γ - 0.6)

K	Our Method	Degree	Closeness	Between-ness	Percolation	Eigenvector	Intercentrality ($\beta = 0.1$)	INF-RANK	INF-SIM	Random
50	0.72	0.56	0.60	0.46	0.52	0.66	0.64	0.40	0.32	0.01
100	0.80	0.56	0.60	0.48	0.52	0.61	0.58	0.42	0.40	0.02
150	0.78	0.55	0.57	0.53	0.54	0.58	0.57	0.47	0.41	0.03
200	0.84	0.55	0.56	0.54	0.54	0.58	0.58	0.44	0.44	0.04
250	0.83	0.55	0.54	0.54	0.53	0.54	0.54	0.46	0.46	0.05
300	0.86	0.54	0.53	0.52	0.53	0.52	0.52	0.48	0.48	0.06
350	0.86	0.52	0.51	0.50	0.50	0.52	0.50	0.49	0.46	0.07
400	0.88	0.49	0.50	0.49	0.48	0.50	0.49	0.48	0.46	0.08
450	0.89	0.48	0.48	0.48	0.47	0.48	0.48	0.48	0.45	0.09
500	0.89	0.48	0.47	0.48	0.48	0.48	0.47	0.48	0.45	0.10
AVG	0.83	0.53	0.54	0.50	0.51	0.55	0.54	0.46	0.43	0.06
SD	0.05	0.03	0.04	0.03	0.03	0.06	0.05	0.03	0.05	0.03

Table B14. Performance Comparison on Spearman's Rank Correlation Coeffiient ($\gamma = 0.6$)

K	Our Method	Degree	Closeness	Between-ness	Percolation	Eigenvector	Intercentrality ($\beta = 0.6$)	INF-RANK	INF-SIM	Random
50	0.59	0.19	0.17	0.15	0.01	0.02	0.21	0.20	0.19	0.00
100	0.63	0.19	0.13	0.15	0.05	0.12	0.17	0.13	0.16	0.01
150	0.71	0.20	0.20	0.10	0.13	0.21	0.15	0.18	0.14	-0.01
200	0.73	0.12	0.20	0.09	0.10	0.14	0.16	0.20	0.11	0.00
250	0.78	0.08	0.12	0.08	0.08	0.18	0.19	0.14	0.09	0.00
300	0.79	0.14	0.17	0.14	0.15	0.20	0.19	0.16	0.11	0.00
350	0.81	0.17	0.18	0.18	0.23	0.17	0.28	0.16	0.16	0.00
400	0.83	0.25	0.20	0.24	0.26	0.19	0.26	0.20	0.21	0.00
450	0.85	0.25	0.23	0.26	0.24	0.23	0.26	0.23	0.21	0.00
500	0.86	0.25	0.25	0.26	0.24	0.23	0.28	0.24	0.22	0.00
AVG	0.76	0.19	0.19	0.16	0.15	0.17	0.22	0.18	0.16	0.00
SD	0.09	0.06	0.04	0.07	0.09	0.06	0.05	0.04	0.05	0.00

Table B15. Performance Comparison on Total Persuasion Credit ($\gamma = 0.6$)

K	Our Method	Degree	Closeness	Between-ness	Percolation	Eigenvector	Intercentrality ($\beta = 0.1$)	INF-RANK	INF-SIM	Random
50	22.29	15.22	15.70	15.82	16.73	17.39	17.03	15.47	14.36	2.84
100	43.20	28.59	29.75	27.95	29.58	30.52	29.51	29.58	26.39	5.74
150	62.92	39.90	41.06	40.93	41.42	41.31	41.98	41.41	37.09	8.61
200	82.07	52.56	50.93	52.97	53.45	53.50	53.48	51.25	48.43	11.40
250	100.29	64.07	62.99	63.92	64.61	61.84	61.00	62.88	59.22	14.23
300	117.73	72.62	71.27	72.60	73.71	70.64	70.17	72.86	69.13	17.09
350	134.63	80.98	80.18	80.42	80.16	80.99	78.49	82.81	76.88	20.05
400	150.94	86.10	87.84	86.90	86.95	87.77	86.27	90.08	83.34	22.72
450	166.56	94.30	94.89	93.27	95.31	94.70	94.69	97.28	90.34	25.61
500	181.40	101.65	101.01	100.38	102.99	102.54	101.36	104.51	96.43	28.66
AVG	106.20	63.60	63.56	63.52	64.49	64.12	63.40	64.81	60.16	15.70
SD	53.48	28.92	28.69	28.52	28.71	28.38	28.04	29.94	27.95	8.64

Table B16. Performance Comparison on Top-*K* Precision (γ - 0.7)

K	Our Method	Degree	Closeness	Between-ness	Percolation	Eigenvector	Intercentrality ($\beta = 0.1$)	INF-RANK	INF-SIM	Random
50	0.70	0.54	0.60	0.44	0.50	0.64	0.64	0.38	0.30	0.01
100	0.78	0.56	0.61	0.48	0.53	0.61	0.59	0.42	0.40	0.02
150	0.78	0.55	0.57	0.53	0.54	0.58	0.57	0.47	0.41	0.03
200	0.84	0.55	0.56	0.54	0.54	0.58	0.58	0.44	0.44	0.04
250	0.84	0.55	0.54	0.54	0.54	0.54	0.54	0.46	0.46	0.05
300	0.86	0.53	0.52	0.51	0.52	0.51	0.51	0.47	0.47	0.06
350	0.88	0.51	0.51	0.49	0.49	0.51	0.49	0.48	0.45	0.07
400	0.88	0.48	0.49	0.49	0.47	0.49	0.48	0.48	0.45	0.08
450	0.89	0.48	0.48	0.47	0.47	0.48	0.48	0.48	0.45	0.09
500	0.90	0.47	0.46	0.47	0.47	0.47	0.46	0.47	0.44	0.10
AVG	0.84	0.52	0.53	0.50	0.51	0.54	0.53	0.46	0.43	0.06
SD	0.06	0.03	0.05	0.03	0.03	0.06	0.06	0.03	0.05	0.03

Table B17. Performance Comparison on Spearman's Rank Correlation Coefficient ($\gamma = 0.7$)

K	Our Method	Degree	Closeness	Between-ness	Percolation	Eigenvector	Intercentrality ($\beta = 0.6$)	INF-RANK	INF-SIM	Random
50	0.60	0.19	0.17	0.15	0.01	0.02	0.21	0.20	0.19	-0.01
100	0.67	0.20	0.13	0.15	0.05	0.13	0.17	0.13	0.16	0.00
150	0.69	0.20	0.20	0.10	0.13	0.21	0.15	0.18	0.14	0.00
200	0.71	0.12	0.20	0.09	0.10	0.14	0.16	0.20	0.11	0.00
250	0.77	0.08	0.13	0.08	0.08	0.18	0.19	0.14	0.09	0.00
300	0.81	0.14	0.17	0.14	0.15	0.20	0.19	0.16	0.11	0.00
350	0.80	0.17	0.18	0.18	0.23	0.17	0.28	0.16	0.16	0.00
400	0.82	0.25	0.20	0.24	0.26	0.19	0.26	0.20	0.21	0.00
450	0.83	0.25	0.23	0.26	0.24	0.23	0.26	0.23	0.21	0.00
500	0.85	0.25	0.25	0.26	0.24	0.23	0.28	0.23	0.22	0.00
AVG	0.76	0.19	0.19	0.16	0.15	0.17	0.22	0.18	0.16	0.00
SD	0.08	0.06	0.04	0.07	0.09	0.06	0.05	0.04	0.05	0.00

Table B18. Performance Companion on Total Persuasion Credit (γ = 0.7)										
K	**Our Method**	**Degree**	**Closeness**	**Between-ness**	**Percolation**	**Eigenvector**	**Intercentrality (β = 0.1)**	**INF-RANK**	**INF-SIM**	**Random**
50	20.47	13.92	14.34	14.52	15.36	15.91	15.58	14.27	13.26	2.91
100	39.93	26.32	27.40	25.85	27.33	28.12	27.18	27.57	24.55	5.73
150	58.80	36.90	37.94	37.98	38.46	38.15	38.83	38.71	34.77	8.63
200	77.19	48.86	47.25	49.37	49.88	49.70	49.72	48.16	45.53	11.43
250	94.98	59.81	58.72	59.86	60.61	57.61	56.85	59.44	55.78	14.23
300	112.05	67.95	66.63	68.22	69.33	65.99	65.55	68.93	65.20	17.01
350	129.04	75.95	75.18	75.75	75.64	75.94	73.55	78.45	72.74	19.94
400	145.35	80.89	82.51	81.99	82.24	82.47	81.04	85.50	78.98	22.84
450	161.39	88.80	89.31	88.12	90.41	89.13	89.15	92.52	85.78	25.61
500	176.74	95.90	95.22	95.08	97.87	96.75	95.60	99.65	91.79	28.43
AVG	101.59	59.53	59.45	59.67	60.71	59.98	59.30	61.32	56.84	15.67
SD	52.55	27.46	27.23	27.19	65.25	26.96	63.68	28.71	26.73	8.59

Table B19. Performance Comparison on Top-*K* Precision (γ = 0.8)										
K	**Our Method**	**Degree**	**Closeness**	**Betweenness**	**Percolation**	**Eigenvector**	**Intercentrality (β = 0.1)**	**INF-RANK**	**INF-SIM**	**Random**
50	0.72	0.54	0.60	0.44	0.50	0.64	0.64	0.40	0.30	0.01
100	0.79	0.56	0.60	0.48	0.52	0.61	0.58	0.43	0.40	0.02
150	0.78	0.54	0.57	0.53	0.53	0.58	0.57	0.47	0.40	0.03
200	0.84	0.55	0.56	0.54	0.54	0.58	0.58	0.44	0.44	0.04
250	0.84	0.55	0.54	0.54	0.53	0.54	0.54	0.46	0.46	0.05
300	0.87	0.53	0.52	0.51	0.52	0.51	0.51	0.47	0.47	0.06
350	0.88	0.51	0.50	0.49	0.49	0.51	0.49	0.47	0.45	0.07
400	0.90	0.48	0.49	0.48	0.47	0.48	0.47	0.47	0.45	0.08
450	0.91	0.46	0.47	0.46	0.46	0.47	0.46	0.47	0.44	0.09
500	0.92	0.46	0.45	0.46	0.46	0.46	0.45	0.47	0.43	0.10
AVG	0.84	0.52	0.53	0.49	0.50	0.54	0.53	0.45	0.42	0.06
SD	0.06	0.04	0.05	0.04	0.03	0.06	0.06	0.02	0.05	0.03

Table B20. Performance Comparison on Spearman's Rank Corellation Coefficient (γ = 0.8)										
K	**Our Method**	**Degree**	**Closeness**	**Between-ness**	**Percolation**	**Eigenvector**	**Intercentrality (β = 0.6)**	**INF-RANK**	**INF-SIM**	**Random**
50	0.60	0.19	0.17	0.15	0.01	0.02	0.21	0.20	0.19	0.00
100	0.65	0.20	0.13	0.15	0.06	0.13	0.17	0.13	0.16	0.00
150	0.69	0.21	0.20	0.10	0.13	0.21	0.15	0.18	0.14	0.00
200	0.72	0.13	0.20	0.09	0.10	0.14	0.16	0.20	0.11	0.00
250	0.77	0.09	0.13	0.08	0.07	0.18	0.19	0.14	0.09	0.00
300	0.80	0.15	0.17	0.14	0.15	0.20	0.19	0.16	0.11	0.00
350	0.82	0.17	0.18	0.18	0.23	0.17	0.28	0.15	0.16	0.00
400	0.83	0.25	0.20	0.24	0.26	0.19	0.27	0.20	0.21	0.00
450	0.85	0.25	0.23	0.26	0.24	0.23	0.26	0.23	0.21	0.00
500	0.86	0.25	0.25	0.26	0.24	0.23	0.28	0.23	0.22	0.00
AVG	0.76	0.19	0.19	0.17	0.15	0.17	0.22	0.18	0.16	0.00
SD	0.09	0.06	0.04	0.07	0.09	0.06	0.05	0.04	0.05	0.00

Table B21. Performance Comparison on Total Persuasion Credit (γ = 0.8)

K	Our Method	Degree	Closeness	Between-ness	Percolation	Eigenvector	Intercentrality (β = 0.1)	INF-RANK	INF-SIM	Random
50	18.92	12.83	13.21	13.42	14.20	14.66	14.35	13.25	12.33	2.90
100	37.29	24.39	25.40	24.07	25.41	26.07	25.20	25.84	22.98	5.73
150	55.24	34.35	35.28	35.46	35.92	35.46	36.13	36.38	32.76	8.57
200	72.92	45.68	44.09	46.26	46.81	46.44	46.47	45.47	43.02	11.46
250	90.26	56.12	55.02	56.34	57.15	53.96	53.26	56.43	52.79	14.27
300	107.16	63.89	62.59	64.40	65.53	61.95	61.54	65.48	61.77	17.16
350	123.93	71.57	70.82	71.66	71.71	71.53	69.25	74.63	69.12	19.88
400	140.40	76.34	77.85	77.69	78.13	77.84	76.47	81.47	75.17	22.78
450	156.53	83.99	84.41	83.59	86.11	84.25	84.29	88.32	81.76	25.77
500	172.34	90.84	90.13	90.40	93.37	91.64	90.53	95.35	87.70	28.50
AVG	97.50	56.00	55.88	56.33	57.43	56.38	55.75	58.26	53.94	15.70
SD	51.60	26.15	25.93	25.99	26.42	25.69	25.37	27.61	25.64	8.62

Table B22. Performance Comparison on Top-*K* Precision (γ = 0.9)

K	Our Method	Degree	Closeness	Between-ness	Percolation	Eigenvector	Intercentrality (β = 0.1)	INF-RANK	INF-SIM	Random
50	0.72	0.54	0.60	0.44	0.50	0.64	0.64	0.40	0.30	0.01
100	0.79	0.56	0.61	0.48	0.53	0.61	0.59	0.43	0.40	0.02
150	0.78	0.53	0.57	0.52	0.53	0.58	0.58	0.47	0.40	0.03
200	0.83	0.55	0.56	0.54	0.54	0.58	0.58	0.44	0.44	0.04
250	0.84	0.55	0.54	0.54	0.53	0.54	0.54	0.46	0.46	0.05
300	0.87	0.53	0.52	0.51	0.52	0.51	0.51	0.47	0.47	0.06
350	0.88	0.51	0.50	0.49	0.49	0.51	0.49	0.47	0.45	0.07
400	0.89	0.48	0.49	0.48	0.47	0.48	0.47	0.47	0.44	0.08
450	0.91	0.46	0.47	0.45	0.46	0.47	0.46	0.46	0.44	0.09
500	0.92	0.45	0.45	0.45	0.46	0.46	0.45	0.46	0.43	0.10
AVG	0.84	0.52	0.53	0.49	0.50	0.54	0.53	0.45	0.42	0.06
SD	0.06	0.04	0.06	0.04	0.03	0.06	0.06	0.02	0.05	0.03

Table B23. Performance Comparison on Spearman's Rank Correlation Coefficient (γ = 0.9)

K	Our Method	Degree	Closeness	Between-ness	Percolation	Eigenvector	Intercentrality (β = 0.6)	INF-RANK	INF-SIM	Random
50	0.59	0.19	0.17	0.15	0.01	0.02	0.22	0.21	0.19	0.01
100	0.63	0.20	0.14	0.15	0.06	0.13	0.17	0.14	0.16	0.00
150	0.68	0.21	0.21	0.10	0.13	0.22	0.15	0.18	0.15	0.00
200	0.72	0.13	0.21	0.09	0.10	0.14	0.16	0.20	0.11	0.00
250	0.76	0.09	0.13	0.08	0.07	0.18	0.19	0.14	0.09	0.00
300	0.80	0.15	0.17	0.14	0.15	0.20	0.19	0.16	0.11	0.00
350	0.81	0.18	0.18	0.18	0.23	0.17	0.28	0.15	0.16	0.00
400	0.82	0.25	0.20	0.24	0.26	0.19	0.27	0.19	0.21	0.00
450	0.84	0.25	0.23	0.26	0.24	0.23	0.26	0.23	0.20	0.00
500	0.86	0.26	0.25	0.26	0.24	0.23	0.28	0.23	0.22	0.00
AVG	0.75	0.19	0.19	0.17	0.15	0.17	0.22	0.18	0.16	0.00
SD	0.09	0.06	0.04	0.07	0.09	0.06	0.05	0.04	0.05	0.00

Table B24. Performance Comparison on Total Persuasion Credit (γ = 0.9)

K	Our Method	Degree	Closeness	Between-ness	Percolation	Eigenvector	Intercentrality (β = 0.1)	INF-RANK	INF-SIM	Random
50	17.59	11.89	12.24	12.48	13.20	13.59	13.30	12.37	11.53	2.85
100	34.93	22.73	23.68	22.52	23.75	24.31	23.48	24.33	21.60	5.73
150	52.10	32.13	32.97	33.27	33.71	33.12	33.79	34.33	31.00	8.51
200	69.11	42.91	41.33	43.55	44.12	43.58	43.64	43.10	40.81	11.40
250	86.00	52.87	51.78	53.24	54.10	50.76	50.10	53.76	50.15	14.33
300	102.67	60.32	59.04	61.03	62.17	58.40	58.01	62.42	58.74	16.94
350	119.29	67.69	66.97	68.04	68.23	67.62	65.45	71.22	65.90	19.99
400	135.75	72.32	73.72	73.87	74.48	73.72	72.41	77.87	71.77	22.73
450	152.06	79.71	80.05	79.56	82.30	79.90	79.96	84.56	78.18	25.60
500	168.22	86.33	85.59	86.22	89.37	87.08	86.00	91.49	84.04	28.56
AVG	93.77	52.89	52.74	53.38	54.54	53.21	52.61	55.55	51.37	15.67
SD	50.67	24.97	24.75	24.91	25.44	24.54	24.23	26.61	24.65	8.62

Appendix C

Analysis of Three Versus Two Forces for Top Persuader Prediction

Our method integrates three important forces (i.e., social influence, entity similarity, and structural equivalence) to predict top persuaders. To demonstrate the importance and value of considering three forces for top persuader prediction, we conducted additional analyses using the data and evaluation design detailed in the "Empirical Evaluation and Results" section. Specifically, we removed structural equivalence from our method and persuasion probability originally defined in Equation (8) became

$$p_{ij} = P\left(D_j = 1 \middle| I_{ij}, \gamma^{d_{ij}-1} M_{ij}\right) \tag{C1}$$

We labeled this new method without structural equivalence as R-removed. Similarly, we removed entity similarity only and labeled the method without entity similarity as M-removed; we dropped social influence only and labeled the method without social influence as I-removed. The performance differences between our method and these new methods, each of which considers only two forces, reveal the need to consider three forces for top persuader prediction; they also shed light on the contribution of each removed force to the performance of our method.

In Tables C1–C3, we report the performance comparisons between our method and each method that only considers two forces, with $\gamma = 0.5$. Our method substantially outperforms R-removed, M-removed, and I-removed in each evaluation metric, across the *K* values we investigated. Averaged across *K*, our method outperforms R-removed, M-removed, and I-removed by 20.28%, 17.66%, and 68.85%, respectively, in terms of top-*K* precision and by 76.60%, 99.68%, and 483.31%, respectively, in terms of Spearman coefficient. On average, our method is 22.72% higher in total persuasion credit than R-removed, 20.21% higher than M-removed and 32.34% higher than I-removed. We further analyzed different values ranging from 0.1 to 0.9, in increments of 0.1, except for $\gamma = 0.5$, and observed results largely similar to those reported in Tables C1–C3. Across , our method outperforms R-removed by 11.95% to 27.20% in top-*K* precision, 71.14% to 258.14% in Spearman coefficient, and 13.60% to 27.25% in total persuasion credit. Similarly, our method outperforms M-removed by 11.80% to 19.88% in top-*K* precision, 66.38% to 204.92% in Spearman coefficient, and 13.41% to 23.54% in total persuasion credit; it also outperforms I-removed by 66.22% to 80.35% in top-*K* precision, 128.09% to 537.94% in Spearman coefficient, and 31.82% to 40.31% in total persuasion credit. Overall, our method significantly outperforms the methods that consider only two forces, suggesting the necessity of considering three forces for top persuader prediction. Our results also indicate that each of the three forces contributes to the performance of our method; in particular, social influence seems to contribute the most among the three forces.

Table C1. Performance Comparison on Top=*K* Precision ($\gamma = 0.5$)

K	Our Method	R-removed	M-removed	I-removed
50	0.72	0.60	0.64	0.30
100	0.81	0.69	0.69	0.43
150	0.78	0.71	0.72	0.50
200	0.82	0.71	0.73	0.52
250	0.83	0.72	0.73	0.47
300	0.85	0.71	0.73	0.55
350	0.85	0.70	0.72	0.56
400	0.87	0.70	0.71	0.58
450	0.88	0.68	0.69	0.58
500	0.89	0.68	0.69	0.55
AVG	0.83	0.69	0.71	0.50
SD	0.05	0.03	0.03	0.09

Table C2. Performance Comparison on Spearman's Rank Correlation Coefficient ($\gamma = 0.5$)

K	Our Method	R-removed	M-removed	I-removed
50	0.68	0.26	0.19	0.02
100	0.65	0.24	0.25	0.19
150	0.73	0.38	0.37	0.20
200	0.73	0.41	0.41	0.17
250	0.79	0.45	0.46	0.37
300	0.80	0.50	0.43	0.29
350	0.82	0.56	0.48	0.35
400	0.83	0.63	0.53	0.39
450	0.85	0.68	0.53	0.38
500	0.86	0.70	0.55	0.42
AVG	0.77	0.48	0.42	0.28
SD	0.07	0.16	0.12	0.13

Table C3. Performance Comparison on Total Persuasion Credit ($\gamma = 0.5$)

K	Our Method	R-removed	M-removed	I-removed
50	24.55	21.61	23.56	21.24
100	47.00	41.44	41.48	38.50
150	67.80	58.51	59.81	54.96
200	87.78	74.02	76.10	70.19
250	106.42	87.99	88.60	78.13
300	124.16	99.70	101.28	94.01
350	140.97	109.73	112.48	104.24
400	157.30	122.08	122.90	112.39
450	172.30	131.72	132.96	120.10
500	186.62	141.30	143.28	124.08
AVG	111.49	88.81	90.25	81.78
SD	54.45	39.66	39.82	35.36

Appendix D

Empirical Evaluation with Another Data Set

We conducted an additional evaluation with another data set. The data set contains 6.01 million records of phone communications among 28,440 mobile phone users over 20 weeks. Each record corresponds to a phone communication and consists of the timestamp and duration of the communication as well as the respective identities of the two users participating in the communication. We can construct a social network from these data. A social entity of the network represents a mobile phone user, a relationship between two entities exists if there are phone communications between their corresponding users, and the strength of social interactions between two entities is measured as the communication time between their corresponding users. The data set also contains adoption information: whether a user adopted a particular mobile phone service during the study period (i.e., initial purchase of the service) and, if adopted, in which week. Over the study period, a total of 3,129 users adopted the service. We also have data about each user's profile, including gender, age, and membership levels in the two most recent years.

Following the same procedure described in the "Evaluation Design" subsection, we used the data over the first 10 weeks of the study period to train our method and each benchmark method for predicting top-K persuaders, where K varies from 280 (i.e., approximately 1% of the total number of users) to 2,800 (i.e., approximately 10% of the total number of users), in increments of 280. We then employed data over the second 10 weeks of the study period to evaluate the prediction performance of each method. In Tables D1–D3, we report the performance of our method and each benchmark method in terms of top-K precision, Spearman's rank correlation coefficient, and total persuasion credit respectively, with attenuation factor $\gamma = 0.5$. As shown, our method substantially outperforms all the benchmark methods in each performance metric, across the investigated K values. Across K, our method, on average, is 312.01% higher in top-K precision than eigenvector centrality (the best performing benchmark method in terms of average top-K precision), 82.04% higher in the Spearman coefficient than eigenvector centrality (the best performing benchmark method according to average Spearman coefficient), and 207.68% higher in total persuasion credit than intercentrality (the best performing benchmark method in terms of average total persuasion credit). In addition, we applied the Wilcoxon signed-ranks test to the performance data in these tables and noted that our method significantly outperformed each benchmark method in any performance metric ($p < 0.001$). To ensure the robustness of our evaluation results, we conducted more evaluations with different γ, ranging from 0.1 to 0.9, in increments of 0.1, except for $\gamma = 0.5$. We obtained evaluation results largely similar to those in Tables D1–D3. Overall, across the values we investigated, our method outperforms the best performing benchmark method by a range of 208.59% to 379.07% in top-K precision, 57.48% to 110.65% in Spearman coefficient, and 115.01% to 263.26% in total persuasion credit.

Table D1. Performance Comparison on Top-*K* Precision ($\gamma = 0.5$)

K	Our Method	Degree	Closeness	Between-ness	Percolation	Eigenvector	Intercentrality ($\beta = 0.1$)	INF-RANK	INF-SIM	Random
280	0.80	0.16	0.05	0.14	0.16	0.21	0.17	0.09	0.06	0.01
560	0.80	0.17	0.08	0.14	0.14	0.19	0.18	0.12	0.09	0.02
840	0.78	0.15	0.10	0.14	0.14	0.16	0.17	0.12	0.09	0.03
1120	0.76	0.15	0.10	0.14	0.14	0.15	0.16	0.13	0.11	0.04
1400	0.73	0.15	0.11	0.13	0.14	0.15	0.15	0.12	0.11	0.05
1680	0.70	0.16	0.12	0.14	0.15	0.15	0.16	0.13	0.12	0.06
1960	0.66	0.17	0.13	0.14	0.16	0.17	0.17	0.13	0.12	0.07
2240	0.63	0.16	0.13	0.14	0.16	0.17	0.17	0.14	0.13	0.08
2520	0.59	0.16	0.14	0.15	0.16	0.17	0.18	0.15	0.13	0.09
2800	0.58	0.18	0.16	0.16	0.17	0.20	0.20	0.16	0.14	0.10
AVG	0.70	0.16	0.11	0.14	0.15	0.17	0.17	0.13	0.11	0.05
SD	0.08	0.01	0.03	0.01	0.01	0.02	0.01	0.02	0.03	0.03

Table D2. Performance Comparison on Spearman's Rank Correlation Coefficient ($\gamma = 0.5$)

K	Our Method	Degree	Closeness	Between-ness	Percolation	Eigenvector	Intercentrality ($\beta = 0.9$)	INF-RANK	INF-SIM	Random
280	0.52	0.10	0.04	0.15	0.01	0.22	0.36	0.13	0.12	0.00
560	0.52	0.23	-0.06	0.04	0.09	0.48	0.40	0.03	0.08	0.00
840	0.55	0.22	-0.03	0.06	0.07	0.49	0.35	0.08	0.01	0.00
1120	0.55	0.19	0.04	0.06	0.12	0.43	0.32	0.07	-0.03	0.00
1400	0.58	0.16	-0.02	0.09	0.09	0.37	0.31	0.10	0.00	0.00
1680	0.59	0.18	0.11	0.07	0.11	0.35	0.28	0.08	0.03	0.00
1960	0.61	0.19	0.01	0.10	0.11	0.31	0.27	0.09	0.05	0.00
2240	0.63	0.21	-0.02	0.09	0.09	0.30	0.24	0.07	0.03	0.00
2520	0.63	0.20	0.00	0.08	0.10	0.28	0.22	0.05	0.04	0.00
2800	0.63	0.20	-0.01	0.09	0.14	0.24	0.22	0.07	0.04	0.00
AVG	0.58	0.19	0.01	0.08	0.09	0.35	0.30	0.08	0.04	0.00
SD	0.04	0.04	0.05	0.03	0.04	0.10	0.06	0.03	0.04	0.00

Table D3. Performance Comparison on Total Persuasion Credit ($\gamma = 0.5$)

K	Our Method	Degree	Closeness	Between-ness	Percolation	Eigenvector	Intercentrality ($\beta = 0.1$)	INF-RANK	INF-SIM	Random
280	195.17	58.00	20.30	37.24	43.30	84.94	67.65	35.38	25.29	11.56
560	341.35	94.54	43.79	71.46	74.63	115.48	110.64	66.81	46.14	22.99
840	458.04	121.89	69.19	100.37	101.51	134.01	140.84	88.65	68.43	34.27
1120	550.64	150.69	92.49	124.65	128.15	157.92	165.08	111.75	95.96	46.11
1400	617.37	183.13	112.53	141.36	155.97	180.09	185.52	129.74	116.02	56.80
1680	677.77	202.90	140.09	165.47	181.58	199.38	207.60	150.31	131.69	67.07
1960	715.98	220.63	166.56	185.38	206.50	223.65	230.23	168.08	147.83	79.90
2240	750.42	234.20	184.71	205.66	227.57	241.07	250.54	191.25	171.27	91.15
2520	780.36	250.02	207.77	229.83	252.31	258.07	276.12	215.75	188.05	101.87
2800	804.90	269.20	232.11	253.92	262.54	282.53	299.61	229.00	204.89	113.87
AVG	589.20	178.52	126.95	151.53	163.41	187.71	193.38	138.67	119.56	62.56
SD	202.53	70.23	71.25	69.76	75.62	64.83	73.96	64.12	60.60	34.30