

OPIU: Opinion Propagation in Online Social Networks Using Influential Users Impact

Amir Mohammadinejad
Institut Mines-Télécom
Télécom SudParis
CNRS UMR 5157 SAMOVAR
France
amir.mohammadinejad@it-sudparis.eu

Reza Farahbakhsh
Institut Mines-Télécom
Télécom SudParis
CNRS UMR 5157 SAMOVAR
France
reza.farahbakhsh@it-sudparis.eu

Noël Crespi
Institut Mines-Télécom
Télécom SudParis
CNRS UMR 5157 SAMOVAR
France
noel.crespi@it-sudparis.eu

Abstract—The current research in opinion propagation is largely based on content analysis of social interactions of users on a network. However, the social power of users in propagating the opinion is not considered. In this paper, we study the impact of influential users on propagating the opinion on signed networks and propose an opinion propagation model considering influential users (OPIU). In particular, we identified two major factors involved in a user opinion propagation: (i) the influential users' effect, induced by the presence of a highly confident individual in the network, and (ii) the neighbors' effect, caused by the presence of the users who have a link with the current user. To assess the performance of the proposed method, we applied it to two large datasets of Epinions signed and Etsy unsigned networks. The results are compared with similar opinion propagation algorithm which indicates the influential users have a significant impact on propagation and considering them can effectively improve the propagation extend.

Index Terms—Influential Users, Link Analysis, Opinion Propagation, Voter Model, Social Networks

I. INTRODUCTION

Today, millions of users participate in different social networks and make a lot of social links with others. However, it is still difficult to determine the extent to which such users affect the opinions of others. For example, in online shopping centers such as Amazon, eBay and etc. the most important issue is satisfying the customers in order to have a successful business. These websites have products and the network of their users in which each user can review them. One of the most situation is when a user wants to buy a product and she has an inquiry about it. The websites usually put the other users' reviews to help her to decide. Thus, these websites can spread the effective opinions through their users. However, they should know which opinion has a better effect in order to assure the users' decision in shopping. In this case, these websites can recommend specific users' opinions who have a constructive impact. In other words, the online shopping centers need to propagate the effective opinions through their users to have more benefit. Moreover, with propagating the opinion, these websites can realize the current and future opinions of their customers regarding each product and adapt their products based on that.

In recent years, opinion propagation in online social networks has become a widespread phenomenon that indicates

its importance. It can be even said that almost all social interactions are shaped by users beliefs and opinions [1]. Thus it is of high value to study opinion dynamics, and up to now many researchers have proposed various models to analyze the evolution of the opinion dynamics and propagation from various aspects [2]–[4]. Current studies on opinion propagation are based on the links a user has in the network which means the users' opinion is affected by the opinion of her neighbors. However, the impact of neighbors are different, i.e. some of them have a greater impact. We consider such neighbors as influential users who are more popular or trusted among others.

In this paper, we present an opinion propagation model based on the impact of influential users for signed and unsigned networks. Consider a shopping website that has products and users. These users are partially connected with each other (the network of users) and rated some of the products as their opinions. Thus, there are two kinds of links: 1) the link between users and 2) the rates or opinions of users toward the products. Two users are neighbors if there is a link between them. In signed networks, the links have positive and negative values and in directed networks the links have direction. In reality, the user decision is influenced by the opinion of neighbors, influential users and current users' knowledge. The first two can be considered through the link analysis and the third one can be determined by analyzing user' profiles. Here, we utilize the link analysis and considering the neighbors, we propose a model for opinion propagation based on the impact of influential users. The main contributions of this study are: (1) Introducing a ranking method to distinguish the users' importance in signed and unsigned networks as well as their opinion propagation impact. (2) Introducing influential user as a new feature for opinion propagation and propose an opinion propagation model based on the impact of this feature. (3) Introducing a new property of opinion propagation namely Fuzzy Majority Opinion as a new measure to analyze the performance of it.

The remaining of the paper is organized as follows. In section II we overview the existing methods used for opinion propagation followed by our proposed model in section III. Section IV presents the evaluation of our work. Finally, section

V concludes the study.

II. RELATED WORKS

Social network analysis [5] studies the relationships between social entities like members of a group. Thus, it enables us to inspect their structural properties such as links, neighbors, centrality and etc.

Kou et al. [6] studied the opinion dynamics with multilevel confidences in Hegselman and Krause (HK) model by defining three clusters for the users namely, close-mind, moderate-mind and open-mind based on social differentiation theory. They divided the network to three sub-group but they did not consider the impact of each user. In another work, Liang et al. [7] considered the impact of both the bounded confidence and influence radius of agents on the opinion dynamics and they found that heterogeneity did not always promote consensus and there is an optimal heterogeneity under which the relative size of the largest opinion cluster reaches its peak point. Zhang et al. in [8] focused on mining features namely double propagation. They used two improvements based on part-whole and 'no' patterns to increase the recall. They applied feature ranking to the extracted feature candidates to improve the precision of the top-ranked candidates. Yang et al. in [9] developed a linear influence model where rather than requiring the knowledge of the social network and then modeling the diffusion by predicting which user will influence other users of the network, they focused on modeling the global influence of a user based on the rate of diffusion through the network. Cha et al. in [10] analyzed the information diffusion in the Flickr social network. they found that even popular photos do not spread widely throughout the network. Also, they implied that the information exchanged between friends is likely counted for some of the favorite markings but with a remarkable delay at each hop. Shang et al. in [11] proposed an opinion formation model under bounded confidence over multiplex networks, consisting of edges at different topological and temporal scales. They found that the existence of multiplexity prevents the convergence and that working with the aggregated or summarized simplex network is inaccurate since it misses vital information.

These studies focused on different opinion propagation models in the network regardless of the impact of the users who make the opinions and propagation. For users' opinion, the above studies, as well as HK model, used the current users' links to her neighbors. However, the neighbors may have a different impression. As mentioned above in [6] the users divided into three clusters and the proposed propagation model based on each cluster manipulated. Nevertheless, each of the users of these groups may have a different impact on propagation. In this work, we propose an opinion propagation regarding the influence and impact of each user in the network.

III. OPINION PROPAGATION FORMULATION

This section consists of two distinct parts. In the first part, we review the base propagation model which leads us to a

model appropriate for our case and in the second part we describe the proposed model.

A. Opinion Propagation Method

Opinion propagation is consist of several methods. There are several models and among them Voter model [12], [13] is one of the promising ones which attracted a lot of attention.

The Voter model is a stochastic process and assumes that there is an interaction between a pair of voters (users). The opinions of any given user on the same issue changes at random times under the influence of the opinions of her neighbors. The model starts with an initial set of active users and for each user at time step t , one of her neighbors will be chosen at random and the user will assume the opinion of that neighbor. There are three opinion formation methods for Voter model:

Sznajd (S): It is used for discrete opinions, e.g. +1 and -1. In each time step t , two randomly selected users transfer their opinion to their neighbors if and only if they share the same opinion.

Deffuant (D): It is used for continuous opinions, e.g. in the range of $[0, 1]$. In each time step, one neighbor of the current user will be met and these two interact. The interaction will update the two users' opinions if the differences of their opinions are near to each other (confidence bound).

Karause and Hegselman (KH): The KH is used for continuous opinions. In each time step, one user is chose randomly and changes her opinion into the arithmetic average of her neighbor's opinions who are within her confidence bound. The user U_i will update her opinion x_i as follows

$$x_i(n+1) = x_i(n) + \frac{\mu}{N_i} \sum [x_j(n) - x_i(n)] \quad (1)$$

where μ is convergence parameter and is in the interval of $[0, 1]$ and N_i is the set of user U_i 's neighbors.

Algorithm 1 is the pseudo code of the whole process of proposed opinion propagation. It has three main processes. First, we find the influential user's scores (ranks). Then, we propagate the opinions of users considering the impact and rank of users and at the end we analyze the propagation with the Fuzzy Majority Opinion.

In this section, we will propose our opinion propagation model inspired by Voter model. The proposed model is based on link analysis which considers the connection of users in a network. Furthermore, the networks of interest are both signed (with positive and negative links) and unsigned ones. These properties convinced us to use Voter model as our baseline for opinion propagation since it uses the links between users and their neighbors. Moreover, we will use the results of Voter model in order to compare the performance of the proposed model.

B. Opinion propagation using influential users model (OPIU)

Usually the users' opinion changes by her prior opinion (initial opinion), the opinion of experts and friends (neighbors in the network). In our case, because the initial opinion is hard to access, we just consider the links and connections

Algorithm 1 Opinion propagation model based on influential users

```

1: Finding influential users in the network
2: Personality definition
3: Rank the users using SPRank
4: Using Credibility to rank and find the influential users
5: Propagating the opinion
6: for each user  $i$  [ $U_i$ ] do
7:   Consider all the  $neighbors_i$  and their ranks
8:   Set the neighbors whose opinions are in confidence bound
    $U_i$ 
9:   Updating the  $U_i$  opinion based on these neighbors
10:   $p_i^{t+1} = p_i^t + \frac{1}{|N_i^t|} \sum \mu \frac{SP_j}{SP_i + SP_j} [p_j^t - p_i^t]$ 
11: end for
12: Analyzing the propagation using Fuzzy Majority Opinion
13: for each product  $P_i$  do
14:   Let  $A = a_1, \dots, a_n$  as the users' opinions toward  $P_i$ 
15:   Let  $E$  as all the subsets of  $A$ 
16:   for each subset  $E_i$  do
17:     Compute the Fuzzy Majority Opinion (FMO)
18:   end for
19:   Assign the dominant opinion as the opinion of  $P_i$ 
20: end for
21: for each user  $U_i$  do
22:   for each opinion of  $U_i$  that changed do
23:     Consider the product  $P_j$  that  $U_i$  has opinion
24:     Consider the FMO of  $P_j$ 
25:     Compare the new opinions with the FMO
26:   end for
27: end for

```

of the user to propagate the opinion. The problem of current propagation methods is that they don't consider the impact of influential users and the impact of neighbors (some of them have more and some have less impact on users' opinion formation). In OPIU, we consider the influential users and update the users' opinion with the fact that each neighbor has a different influence on current user final opinion in the network. First, we discuss about the social influence then we introduce the method of finding influential users in the network and at the end we present the proposed model (OPIU).

1) *Social influence opinion propagation*: Social influence is the process that users adjust/modify their opinions or change them because of their social interactions with others [14]. A simple process model acquired from the observations implied how opinions in a group of interacting users can shape or spread over repeated interactions. In particular, the studies in this domain identified two major attractors of opinion: (i) the influential user's effect, derived from the presence of highly confident users in the network, and (ii) the majority effect, induced by the presence of a group of users sharing homological opinions.

Indeed, it is difficult to measure how opinions alter under experimental situations, as it depends on many social factors such as the personality of the users, their bounded confidence level, their social status, their credibility, or their social power [15]. The present work draws upon experimental methods motivated by the concepts of opinion propagation in sociology and psychology. In the proposed model, we utilize the first major opinion attractor namely influential users as an effective factor on propagation. Also, we use the second one as a criterion for evaluating the opinion propagation (see section

III-B4).

2) *Influence impact on opinion propagation*: Based on social influence studies, the people who are connected can change each other's opinion if their opinions are close enough. For example, the studies showed that the people sharing similar opinions have a strong tendency to amplify their confidence after interacting with each other [16]. Therefore, in our proposed model, for each user U_i we considered a set of neighbors whose opinions are not more than a certain confidence and then update the current users' opinion based on these neighbors. In reality, not all the users have the same influence on each other and some have social power e.g. have greater influence. One way to calculate the users' social power is their rank [17]. The OPIU model comes from our previous study [17] in which we employed the ranking of each user as their social power. This means that the users with more ranks have more social power and effect on others (we call them as influential users). Here, for a signed network we employ that ranking algorithm (which is based on PageRank) to compute the ranking of the users based on their links:

$$Rank^+(U_i) = (1 - \alpha) \frac{1}{N} + \alpha \sum_{U_j \in M(U_i)} \frac{PR^+(U_j)}{L^+(U_j)} \times Per_j \quad (2)$$

where $L^+(U_j)$ is the number of user j 's positive outgoing links (similar equation is used for $Rank^-(U_i)$ [17]). Similarly, the ranking of the users in an unsigned network is:

$$Rank(U_i) = (1 - \alpha) \frac{1}{N} + \alpha \sum_{U_j \in M(U_i)} \frac{PR(U_j)}{L(U_j)} \times Per_j \quad (3)$$

The ranking algorithm starts with some initial conditions and it converges to the final rank vectors after enough iterations. In the formula, the Per_j is the personality of user u_j based on optimist and pessimist scores of the users defined in [17], [18]. The final social power rank vector (SPRank) for the signed network is computed as $SPRank(P_i) = Rank^+(U_i) - Rank^-(U_i)$ and for unsigned networks is as $SPRank(P_i) = Rank(U_i)$. These formulas compute the social power score (rank) of each user of the network. The users who have higher scores are more influential in the network [17].

3) *OPIU model*: We gave a score to each user of the network using their links (which is used to detect the influential users). Now we formulate the OPIU as follows:

Given a directed network G , we observe the decision of users toward a particular product over it. The user U 's decision toward the product P is $Decision_{U \rightarrow P} = Function\{PK, C, R\}$ which PK is U 's prior knowledge, C is the U 's connection in the network and R is the review of others toward the product. There are two approaches to formulate a propagation model through a network: 1) information effects [19], 2) direct-benefit effects [20]. Network models based on direct-benefit effects involve the following significant consideration: The user has certain social network neighbors and her benefits in adopting a new opinion increase when more and more of these neighbors pursue it. We consider this on the users' decision which consist the connection of the user. The

connections consist of two kinds of impacts: 1) the impact of neighbors and 2) the impact of influential users.

Consider a weighted graph $G = (V, E, A)$ where V is the set of vertices with n users, E is the set of directed links, and A is the adjacency matrix. A neighborhood matrix G^t is used to represent the social relationships on A at time t . For all $i, j \in A$, $G_{ij}^t \in \{0, 1\}$ shows if there is a directed link from user i to j at time t . So the $n \times n$ matrix G^t is specified as:

$$G_{ij}^t = \begin{cases} 1 & \text{if } i \text{ pays its attention to } j \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

where $G_{ij}^t = 1$ denotes user i can receive an opinion from a supplier user j . In fact, we assume that each user is always connected with itself, i.e. $G_{ii}^t = 1$, for all $i \in A$, all t . G^t is asymmetric to describe a directed network, so that $G_{ij}^t \neq G_{ji}^t$, for some i, j . A user $i \in A$ only observes herself and her neighbors, including the users in the set of $j | G_{ij}^t = 1$ for all j , at time t . The opinions of n users at time t are appeared by an $1 \times n$ vector $P^t = (p_1^t, p_2^t, \dots, p_n^t)$, where p_i^t is the user i 's opinion at time t , $p_i^t \in (0, 1)$, $i \in A$. We define $diff_{ij}^t$ as the difference between opinion p_i^t and p_j^t : $diff_{ij}^t = |p_i^t - p_j^t|$ where $|p_i^t - p_j^t|$ is the absolute value of $p_i^t - p_j^t$. Obviously, we have $diff_{ii}^t = 0$ and $diff_{ij}^t = diff_{ji}^t$. Furthermore, we define w_{ij}^t as the weight of the influence of j on i .

$$W_{ij}^t = \begin{cases} 1 & \text{if } diff_{ij}^t \leq \epsilon \text{ and } G_{ij}^t = 1 \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

where ϵ is the confidence level (CL) and $w_{ii}^t = 1$ for $\epsilon \geq 0$ for all t and i . Each user will update her opinion by taking the average of all opinions which lie in her CL including her opinion at each time step t . The element p_i^{t+1} of new opinion vector P^{t+1} is calculated as:

$$P_i^{t+1} = \sum_{j=1}^n \frac{w_{ij}^t}{\sum_{a \in A} w_{ia}^t} p_j^t \quad (6)$$

The P vector will keep updating until it converges. The convergence criteria is $\sum_{i=1}^n (p_i^{t+1} - p_i^t)^2 \leq \xi$ where ξ is a very small positive number (e.g. 10^{-4}). Also, the influential users have great influence on other individuals in the society but, their opinions are hardly influenced. Let us suppose that there are M users and K of them are influential ones. We consider social power scores so the update rules for user i with p_i opinion will be:

$$p_i^{t+1} = \begin{cases} p_i^t + \frac{1}{|N_i^t|} \sum \mu \frac{SP_j}{SP_i + SP_j} [p_j^t - p_i^t], & N_i^t \neq 0 \\ p_i^t, & \text{Otherwise} \end{cases} \quad (7)$$

Where $N_i^t = \{j | |p_i^t - p_j^t| \leq \epsilon_i\}$ is the opinion neighbor set of user i at time t and $|N_i^t|$ is the cardinality of N_i^t .

4) *Fuzzy Majority Opinion*: In order to evaluate our experiments, we used the concept of the majority opinion (section IV-C). First, we indicate to some extend the Fuzzy Majority Opinion can be computed and then we present its usage to evaluate the opinion propagation.

There are two common ways to compute majority opinion [21], namely aggregation operators and fuzzy method. Here

we used the fuzzy method which provides in addition to a value for the majority opinion a sign of the strength of that value as a delegate of the majority opinion. To do so, consider $A = a_1, \dots, a_n$ be a set of values which establish the opinions of the users. Let E be a crisp subset of A . The first step is to specify the degree to which this is a subset carrying a majority opinion. A subset E holds a majority opinion if all the elements in E are similar and the cardinality of E satisfies the idea of being a majority of elements from A . Let $MOP(E)$ implies the degree to which the elements in E , form a majority opinion, are a majority of elements from A with similar values. Thus, $MOP(E) = Q(\frac{|E|}{n}) \wedge Sim(E)$ where \wedge shows the min operator and $Sim(E)$ is equal to $Min_{a_i, a_j \in E} [Sim(a_i, a_j)]$. Then, $Opi(E) = Average(E) = \frac{\sum_{a_i \in E} a_i}{|E|}$ is the opinion of the elements in E which is the mean value of the elements involved in E . Using above concepts, the fuzzy majority opinion FMO indicating the majority opinion of the set of elements in A is defined as:

$$FMO = \bigcup_{E \subseteq A} \left\{ \frac{MOP(E)}{Opi(E)} \right\} \quad (8)$$

So for each subset E , the value $MOP(E)$ indicates the degree to which the quantity $Opi(E)$ is a majority opinion. Also, following similarity relation is assumed:

$$Sim(a_i, a_j) = \begin{cases} 1 & \text{if } |a_i - a_j| < \sigma \\ \frac{2\sigma - |a_i - a_j|}{\sigma} & \text{if } \sigma < |a_i - a_j| < 2\sigma \\ 0 & \text{otherwise} \end{cases} \quad (9)$$

where σ is the standard deviation of a_1, \dots, a_n . Furthermore, for the formal definition of the quantity (Q), a definition of a majority in terms of a fuzzy subset Q is defined on the unit interval. In particular, $Q : [0, 1] \rightarrow [0, 1]$ such that $Q(0) = 0$, $Q(1) = 1$ and $Q(x) \geq Q(y)$ if $x > y$. $Q(x)$ is defined as below:

$$Q(x) = \begin{cases} 0 & \text{if } x \leq 0.4 \\ 5(x - 0.4) & \text{if } 0.4 < x \leq 0.6 \\ 1 & \text{otherwise} \end{cases} \quad (10)$$

IV. EXPERIMENTAL RESULTS

In this section, we evaluate the OPIU using real-world networks within opinion propagation. The evaluation consists of two main parts. First, we present the details of the datasets and discuss our observation on OPIU and then we evaluate the OPIU performance using FMO.

A. Dataset

To evaluate our work we used two directed datasets namely Epinions (signed) and ETSY (unsigned). The first one is the same dataset used in Stanford collection (SNAP, <https://snap.stanford.edu/data/>) and the second dataset is crawled by our crawler.

Etsy Dataset: Etsy is a peer-to-peer e-commerce website covering a wide range of products on handmade or vintage items and supplies, as well as unique factory-manufactured items. Etsy's top three competitors according to Hoovers

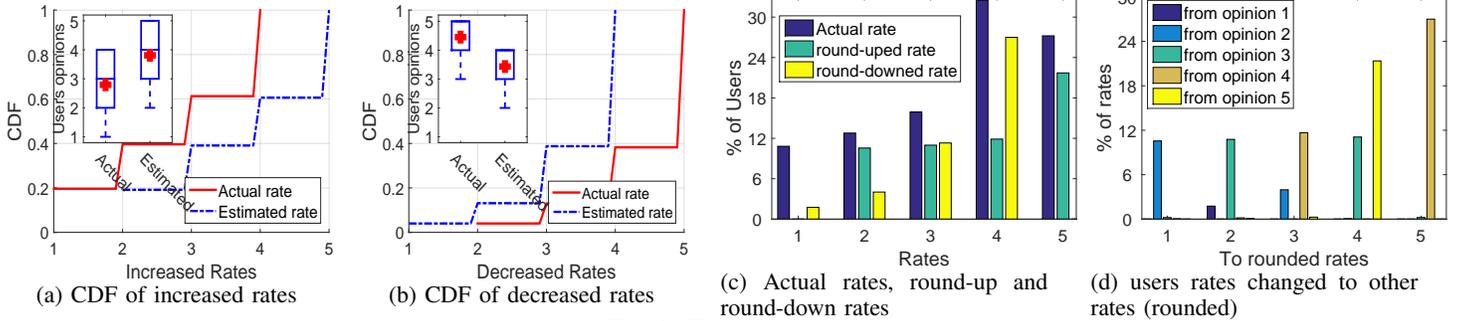


Fig. 1: Epinions dataset

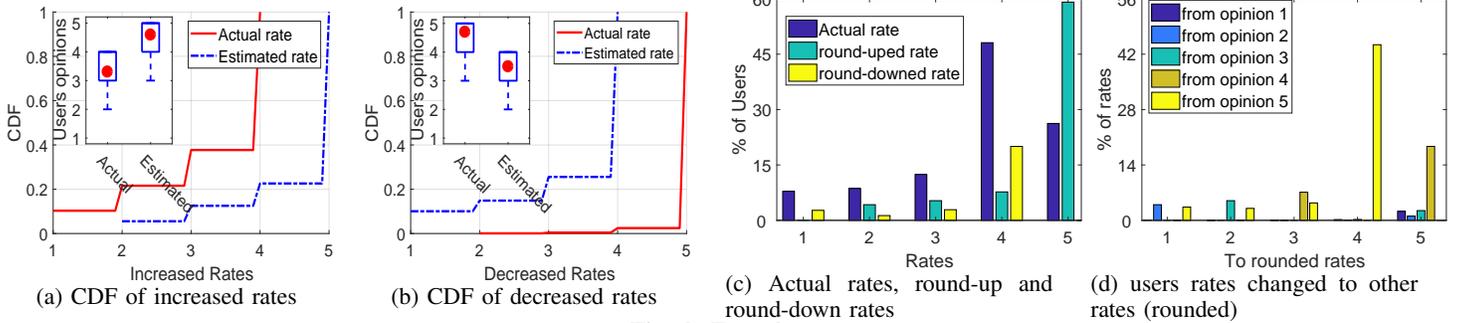


Fig. 2: Etsy dataset

Online are Amazon Handmade, Craigslist, and eBay. This website includes the network of users (unsigned links between users) and the user’s opinions toward the products. Same as Epinions, the users’ opinions toward the products contains the integer values between 1 and 5. This dataset is crawled from “https://www.Etsy.com”. In general, Etsy has six main categories and due to the huge number of users we selected one category (namely Home) for our experiments. The crawler is programmed in C# which goes to the product pages one by one and collect the user’s opinions for them. Then, for each user of the product, it collects the followers (who have a link to the current user) and following (who the current user has a link to them) in order to establish the network of the users. The most challenging part was the time consumed for crawling the users due to the fact that there is a huge number of links which took a month with a core i7 CPU and 16GB RAM computer. We crawled 239,237 users with 4,618,783 links. Same as Epinions, we did a filtering step to omit the users who had few number of links. The main characteristic of Etsy and Epinions datasets are presented in table I.

TABLE I: Epinions [22] and Etsy Datasets Characteristics

| Characteristics | Epinions | Etsy |
|---------------------------------|----------|-----------|
| Total number of users (crawled) | 131,828 | 239,237 |
| Total links (edges) | 841,372 | 4,618,783 |
| Number of filtered users | 49,289 | 72,528 |
| Filtered links | 507,592 | 1,914,852 |
| Positive links | 434,694 | - |
| Negative links | 72,898 | - |
| Number of Products | 139,738 | 24,362 |
| Number of Products’ ratings | 664,824 | 200,148 |

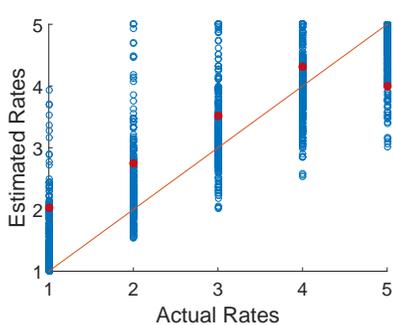
B. Experiments and Observations

We evaluated the proposed method from three different levels, namely opinions, users and products. In case of the opinions, we consider each opinion from users toward the products and analyze the differences between actual and estimated ones. In case of the users, we analyze the users and their opinion changes. And, in case of the products, we analyze the rates of products which changed significant and also analyze the users who made this changes. These levels provide us different visions and understanding of the impact of influential users to the opinion.

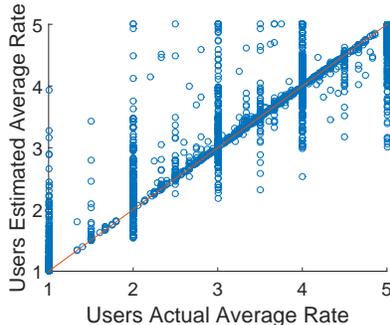
In the formula, when $\mu = 0$, it means user i will never consider other user’s opinion (we can treat it as leader). Without loss of generality, we assume $\mu = 0.5$ in our experiments and in order to find the neighbors of a user, we considered the outgoing links of her (those users who have a link from the current user are the neighbors of her).

1) *In the level of rates:* The rates of users are the users’ opinion toward the products. The number of rates from users to products has the mean value of 13.49 in Epinions and 19.53 in Etsy. We observed from Epinions that 2% of users have no rate, 3.5% have 1 rate, 65% have 1 to 10 rates and 97% have 1 to 100 rates toward the products. Also, for Etsy dataset 5% of users have no rate, 11% have 1 rate, 45% have 1 to 10 rates and 91% have 1 to 100 rates.

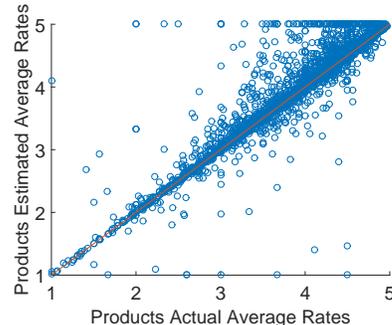
In our experiments, some of the opinions changed with OPIU while others remained unchanged (OPIU succeed to change 25.57% of the Epinions and 21.43% of Etsy rates). The rates whether increased or decreased. Hence, we separated the rates to two subgroups i.e. increased and decreased rates to analyze them. Figures 1a, 1b, 2a and 2b show the increased



(a) The spread of different rates of users. The red circle is the average of estimated rates

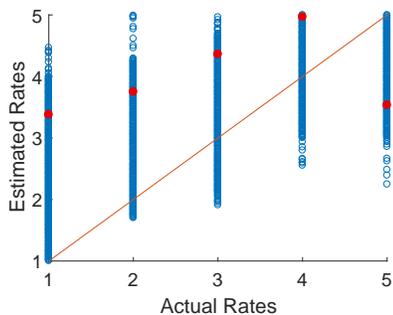


(b) The spread of different average rates of each user.

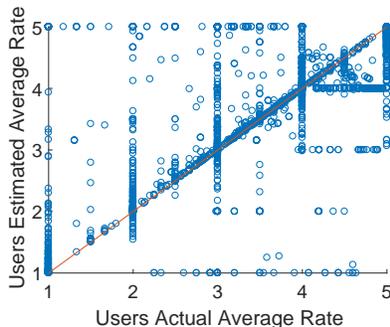


(c) The spread of different average rates of products.

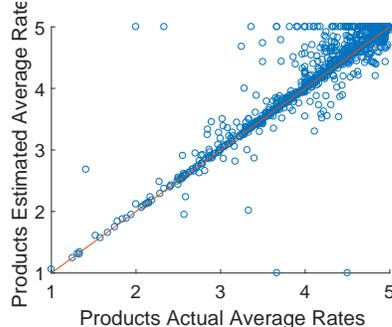
Fig. 3: Epinions dataset



(a) The spread of different rates of users. The red circle is the average of estimated rates



(b) The spread of different average rates of each user.



(c) The spread of different average rates of products.

Fig. 4: Etsy dataset

and decreased rates to compare the actual and estimated rates for datasets. These figures indicate that the users who gave low and high rates to the product tend to make their rates lower and higher respectively.

Figures 1c and 2c illustrate how much percentage of users have different rates toward the products. Note that the rates are in the range of 1 to 5 and we compared the actual rates with estimated rounded ones. We observed that the changes are mostly ascending i.e. from lower rates to upper ones and it means that users often tend to be positive rather than negative.

Figures 1d and 2d illustrate how many rates of users are changed to other rates (considering that we examined this with rounded estimated rates). These figures show the changes of rates are normally smooth (and not a big jump). Generally, these figures convey that it is hard for users to change their opinions to other ones which are very far from theirs.

Figures 3a and 4a show the spread of estimated rates of users in comparison to actual ones. Note that the red circles are the average rate of each estimated column. These figures show that our model significantly changed the user's opinion.

2) *In the level of Users:* In order to count the neighbors of a specific user, we considered the links from the current user to her neighbors (the user out-going links). In case of the Epinions dataset, if the link is negative, we consider the neighbor as a negative neighbor and otherwise positive neighbor. The average number of neighbors for each user is 10.29 (with 8.81 positive neighbors and 1.48 negative neighbors) in Epinions and 27.87 for Etsy. Figures 3b and 4b show the spread of average rates of users in their networks. These figures indicate how average rates of users changed. As

we can see, most of the changes are near to the actual rates which means that the average rates of user change around the actual rates.

3) *In the level of Products:* Figures 3c and 4c show the spread of average rates of products in the network and indicate how average rates of products changed. The changes are mostly near to the actual rates with a small fraction of users who made big changes. One interesting observation is analyzing the products which their average opinions changed significantly. Among the Epinions products, we found 22.68% of them have big jumps (significant difference between estimated and actual average opinions). This percentage is 26.59% for Etsy. Moreover, we investigate the products which have the most significant changes (top 2% of them) in order to evaluate their opinions and the users who rated them. We observed that the users rated these products have some neighbors who have high scores in the network which we tagged them as influential users in section III-B2 (the scores of the neighbors are in the range of top 5% ranks of users). This observation implies to the fact that influential users are involved in changing the average opinion of the products and indicates the impact of them in propagating the opinions.

C. FMO Evaluation

According to the third main part of the algorithm, we use Fuzzy Majority Opinion characterized in section III-B4 to evaluate our methodology. The opinion of group members will lead to the majority opinion [14]. In this study, they showed that the opinion of users will change to the opinion which is the majority opinion of the network users. We use this as a

TABLE II: MSE of OPIU and Voter opinions with normal and Fuzzy Majority Opinion

| Dataset | Model | Maj_Op1 | Maj_Op2 | Mode |
|----------|-------------|---------|---------|------|
| Epinions | OPIU | 0.77 | 0.71 | 1.23 |
| | Voter Model | 1.19 | 1.13 | 1.66 |
| Etsy | OPIU | 0.89 | 0.86 | 1.19 |
| | Voter Model | 1.36 | 1.29 | 1.44 |

criterion for evaluating the estimated opinion propagated by OPIU model. In other words, if the estimated opinions lead to the majority opinion (getting near to it), the model is working properly. We compare the Voter model and the proposed OPIU model considering the Fuzzy Majority Opinion. To do so, we first assign the dominant opinion for each product and then compare the OPIU computed opinions of users toward products with the FMO of each of them as algorithm 1. Note that here we used two similarity relation function (Sim). The first function (Maj_Op1) is described in equation 9 [23] and the second one (Maj_Op2) is defined as follows [21]:

$$Sim(a_i, a_j) = \begin{cases} 1 & \text{if } |a_i - a_j| < 2 \\ \frac{1}{2}(4 - |a_i - a_j|) & \text{if } 2 < |a_i - a_j| < 4 \\ 0 & \text{otherwise} \end{cases} \quad (11)$$

Table II shows the mean square error (MSE) of OPIU and voter model with Fuzzy Majority Opinion for both datasets. Note that this computation is done towards the users' opinions which are changed (there are some opinions which remain unchanged after applying the method). This result shows that the estimated users' opinions are leading to the Majority Opinions and the MSE of OPIU is better than Voter model for both datasets. This confirms that the performance of OPIU is better than baseline propagation Voter model.

V. CONCLUSION

Today people take lots of decisions in their lives such as shopping, where to go for a trip, renting hotel and etc. Some of these decisions are made online and as the statistic shows, the tendency of people for online shopping is growing day by day. Users of a website usually take the decision about its products based on the current information they have, the opinion of their neighbors and influential users of that website. There are a lot of studies paid attention to the dynamics of opinion which shows the importance of the subject. In this paper, we have proposed an opinion propagation model (OPIU), where the impact of influential users was considered as a crucial factor on propagation in both signed and unsigned networks. The OPIU is based on the link analysis approach inspired by baseline propagation method (Voter). For each user, OPIU considers her neighbors and the degree of their expertise in the network and based on that propagate their opinion toward the current user. In this case, we consider the impact of influential users of the network. Furthermore, in order to analyze the performance of the proposed model, we introduced

a method namely Fuzzy Majority Opinion. We found that users usually tend to improve their opinions rather than decreasing it e.g. diminish the rates. In addition, users rarely make a lot of changes in their opinions. Furthermore, the empirical experiments with the Epinions and Etsy datasets show that our approach outperforms baseline method significantly and the fact that identifying the expertise of neighbors (influential users) have a crucial impact on opinion propagation.

REFERENCES

- [1] D. Acemoglu and A. Ozdaglar, "Opinion dynamics and learning in social networks," *Dynamic Games and Applications*, vol. 1, no. 1, pp. 3–49, 2011.
- [2] C. Altafini, "Dynamics of opinion forming in structurally balanced social networks," *PLoS one*, vol. 7, no. 6, p. e38135, 2012.
- [3] T. V. Martins, M. Pineda, and R. Toral, "Mass media and repulsive interactions in continuous-opinion dynamics," *EPL (Europhysics Letters)*, vol. 91, no. 4, p. 48003, 2010.
- [4] G. Qiu, B. Liu, and C. Chen, "Opinion word expansion and target extraction through double propagation," *Computational linguistics*, vol. 37, no. 1, pp. 9–27, 2011.
- [5] S. Wasserman and K. Faust, *Social network analysis: Methods and applications*. Cambridge press, 1994.
- [6] G. Kou, Y. Zhao, Y. Peng, and Y. Shi, "Multi-level opinion dynamics under bounded confidence," *PLoS one*, vol. 7, no. 9, p. e43507, 2012.
- [7] H. Liang, Y. Yang, and X. Wang, "Opinion dynamics in networks with heterogeneous confidence and influence," *Physica A: Statistical Mechanics and its Applications*, vol. 392, no. 9, pp. 2248–2256, 2013.
- [8] L. Zhang, B. Liu, S. H. Lim, and E. O'Brien-Strain, "Extracting and ranking product features in opinion documents," in *Proceedings of the 23rd international conference on computational linguistics: Posters*, 2010, pp. 1462–1470.
- [9] J. Yang and J. Leskovec, "Modeling information diffusion in implicit networks," in *ICDM*. IEEE, 2010, pp. 599–608.
- [10] M. Cha, A. Mislove, and K. P. Gummadi, "A measurement-driven analysis of information propagation in the flickr social network," in *ACM WWW*, 2009, pp. 721–730.
- [11] Y. Shang, "Deffuant model of opinion formation in one-dimensional multiplex networks," *Journal of Physics A: Mathematical and Theoretical*, vol. 48, no. 39, p. 395101, 2015.
- [12] P. Clifford and A. Sudbury, "A model for spatial conflict," *Biometrika*, vol. 60, no. 3, pp. 581–588, 1973.
- [13] V. Sood, T. Antal, and S. Redner, "Voter models on heterogeneous networks," *Physical Review E*, vol. 77, no. 4, 2008.
- [14] M. Moussaïd, J. E. Kämmer, P. P. Analytis, and H. Neth, "Social influence and the collective dynamics of opinion formation," *PLoS one*, vol. 8, no. 11, p. e78433, 2013.
- [15] B. Latane *et al.*, "The psychology of social impact," *American psychologist*, vol. 36, no. 4, pp. 343–356, 1981.
- [16] D. G. Myers and G. D. Bishop, "Discussion effects on racial attitudes," *Science*, vol. 169, no. 3947, pp. 778–779, 1970.
- [17] A. Mohammadinejad, R. Farahbakhsh, and N. Crespi, "Employing personality feature to rank the influential users in signed networks," in *IEEE SocialCom*, 2016, pp. 346–353.
- [18] G. Beigi, J. Tang, and H. Liu, "Signed link analysis in social media networks," *arXiv preprint arXiv:1603.06878*, 2016.
- [19] B. Golub and M. O. Jackson, "Naive learning in social networks and the wisdom of crowds," *American Economic Journal: Microeconomics*, vol. 2, no. 1, pp. 112–149, 2010.
- [20] L. E. Blume, "The statistical mechanics of strategic interaction," *Games and economic behavior*, vol. 5, no. 3, 1993.
- [21] G. Pasi and R. R. Yager, "Modeling the concept of majority opinion in group decision making," *Information Sciences*, vol. 176, no. 4, pp. 390–414, 2006.
- [22] J. Leskovec, D. Huttenlocher, and J. Kleinberg, "Signed networks in social media," in *ACM SIGCHI*, 2010, pp. 1361–1370.
- [23] H. Aghaie, S. Shafieezadeh, and B. Moshiri, "A new modified fuzzy topsis for group decision making using fuzzy majority opinion based aggregation," in *ICEE*. IEEE, 2011, pp. 1–6.