Comprehensive Influence Propagation Modelling for Hybrid Social Network

Weihua Li[1], Quan Bai[1] and Minjie Zhang[2]

[1] Auckland University of Technology, New Zealand,
[2] University of Wollongong, Australia
weihua.li@aut.ac.nz; quan.bai@aut.ac.nz; minjie@uow.edu.au

Abstract. The evolution of influencer marketing relies on a social phenomenon, i.e., influence diffusion. The modelling and analysis of influence propagation in social networks has been extensively investigated by both researchers and practitioners. Nearly all of the works in this field assume influence is driven by a single factor, e.g., friendship affiliation. However, influence spread through many other pathways, such as face-to-face interactions, phone calls, emails, or even through the reviews posted on web-pages. In this paper, we modelled the influence-diffusion space as a hybrid social network, where both direct and indirect influence are considered. Furthermore, a concrete implementation of hybrid social network, i.e., Comprehensive Influence Propagation model is articulated. The proposed model can be applied as an effective approach to tackle the multi-faceted influence diffusion problems in social networks. We also evaluated the proposed model in the influence maximization problem in different scenarios. Experimental results reveal that the proposed model can perform better than those considering a single aspect of influence.

Keywords: Hybrid Social Network; Indirect Influence; Influence Propagation; Influence Maximization

1 Introduction

With the prevalence of social networks and spectacular growth of social media, people tend to spend more time on-line conducting social activities. Prior studies show that consumers perceive peers' influence from social networks as more trustworthy and persuasive than traditional media, such as radio and TV advertising [1]. Motivated by this background, how to analyse and model the influence diffusion in social networks has drawn great attention in the contemporary research field. Decision making applications, such as maximization of product adoption [2], are developed based on the influence propagation models. Specifically, the models are capable of estimating the spread of influence through the network topology, which can assist the business owners to make decisions on how to promote new products.

Most research works investigate influence diffusion in social networks based on the existing network topological structure, where friendship-affiliation links represent influence propagation channels, and the strength of links is considered as the only factor affecting the influence propagation probability. Therefore, the assumption is friendship affiliation links are equivalent to influential links. However, this cannot hold in general, as they are naturally two different types of links coexisting in a social network.

Influence is a hybrid effect, which can be decomposed into multiple components focusing on different activities of human-beings [3]. Specifically, influence is presented as a mixed types of communications and interactions, such as perceiving information posted by the friends of on-line social networks, delivering messages or emails, conducting face-to-face discussions, reviewing the comments from web-pages, etc. Hence, any of these behaviours are capable of exerting influence and impacting individuals' decisions. However, most researchers ignore the multiple possible interactive diffusion channels. On the other side, in many situations, individuals are more likely getting influenced by the 'stimuli' left by others, especially in E-commerce domain. Feedbacks of a particular product, such as reviews, ratings, comments from previous buyers influence the purchasing decisions of others, even they are not adjacent neighbours and without any immediate interactions. Therefore, the influence is still capable of propagating through the users in the same context without explicit links, since they are affiliated implicitly via other features, such as similar preference and criteria for items. Obviously, this is an important feature to be considered in influence propagation modelling, but unfortunately, is ignored in most existing works.

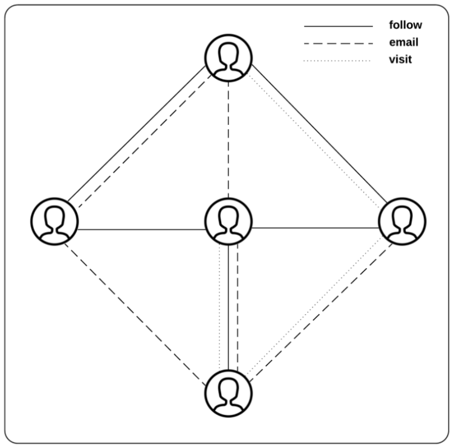
In this paper, we model influence diffusion space as a Hybrid Social Network (HSN), which is formed by merging a number of homogeneous or heterogeneous networks representing possible influence propagation channels. Comprehensive Influence Propagation (CIP) model has been proposed to capture the multiple pathways of influence propagation. Furthermore, experiments have been conducted by making use of influence maximization [4][5] as a typical application. We evaluate the CIP in different scenarios, i.e., in the same network with divergent scale of initial negative influencers. Experimental results reveal that the proposed model can perform better than those considering a single aspect of influence.

The reminder of this paper is organized as follows. Section 2 introduces the general ideas of influence propagation in HSN. In Section 3, the CIP model has been elaborated. In Section 4, the CIP model has been evaluated in influence maximization problem, meanwhile, the experimental results are demonstrated. Finally, the paper concluded in Section 5.

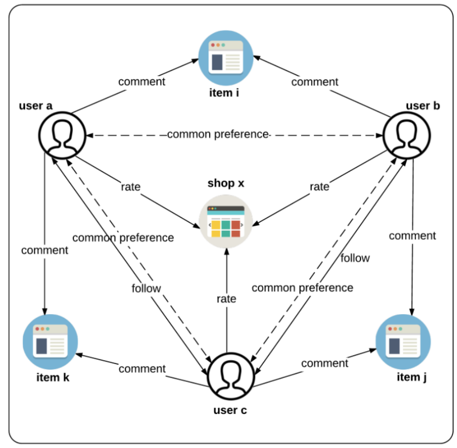
2 Influence Propagation in Hybrid Social Networks

2.1 Influence Diffusion Models

Independent Cascade Model (ICM) and Linear Threshold Model (LTM) are two fundamental influence diffusion models which have been widely applied in many research works [5][6][7]. In both models, each node has two possible states, i.e.,



(a) Direct Influence Propagation



(b) Indirect Influence Propagation

Fig. 1. Influence Diffusion in Social Networks

active and inactive. At the beginning, a limited set of influencers (nodes), i.e., **seed set**, are supposed to be selected as the initial active nodes, which attempt to propagate influence and affect the inactive neighbours at a certain probability. If any neighbour is activated, the state will be converted to active and it starts to propagate influence to its neighbours.

ICM is a non-deterministic diffusion model, where the receiver's state is not deterministically decided by the itself, but is affected and influenced with a predefined probability by the senders [8]. By contrast, LTM is a deterministic diffusion model, where each node is assigned a fixed threshold affecting the activation. In this paper, ICM is employed by extending its features, since individual's level of influence acceptance is not considered in this research work.

2.2 Direct and Indirect Influence

Influence diffusion is a sort of communication which concerns the spread of messages perceived as new ideas or innovations [9]. Thus, direct and indirect influence in this paper lay emphasis on the types of communicational channels. To be more specific, direct influence refers to the immediate interactions or message reciprocations among any users with explicit links, such as friendship affiliation.

The concept of indirect communicational influence stems from the ant and stigmergy algorithms, where ants interact each other and conduct group activities by leaving and sensing pheromones [10]. By tailoring this idea, in this paper, indirect influence describes a form of indirect communications among the users mediated by modifications of the environment. Specifically, some users are not connected explicitly, but they diffuse influences by leaving the messages, such as ratings, comments, reviews, beliefs, etc. Meanwhile, individuals are getting influenced by reading the information produced by their counterparts, since they locate in the same environment. Sometimes, the strength of indirect influence is even more prominent than direct influence, especially in sparse networks.

Figure 1 demonstrates two typical examples focusing on direct and indirect communicational influence. In Figure 1a, three possible direct influence-diffusion channels exist. Whereas, other types of nodes, i.e., item and shop, are involved in Figure 1b, where user a and b potentially influence each other via the messages delivered to the item and shop, though they are not connected explicitly in this social context. As for the other two pairs of users, i.e., user a and c, user c and b, they share both implicit and explicit influential links.

2.3 Hybrid Social Network

In this paper, HSN refers to an implicit heterogeneous user-centric network comprised of a number of social networks concerning possible direct and indirect influence propagation channels. It aims to model various influential relations among the individuals. Meanwhile, it implies the decomposition of influence effects, which gives high extensibility and flexibility. Specifically, when other available influential factors are added or the existing factors are changed, the model can be adapted easily by updating a particular facet.

Figure 2 demonstrates the generic model of HSN. Social networks inside the rectangle container represent different influence diffusion channels, and they are supposed to be extracted from the original heterogeneous social network. There are two types of network, the direct and indirect influence propagation network. The former is a homogeneous network, while the latter is a heterogeneous network, where the indirect influential relationships are established by via its corresponding object layer. Specifically, in this figure, $G_I^{a\prime}$ is object/item layer of network G_I^a. HSN G_{HSN} is constructed by merging all the social networks.

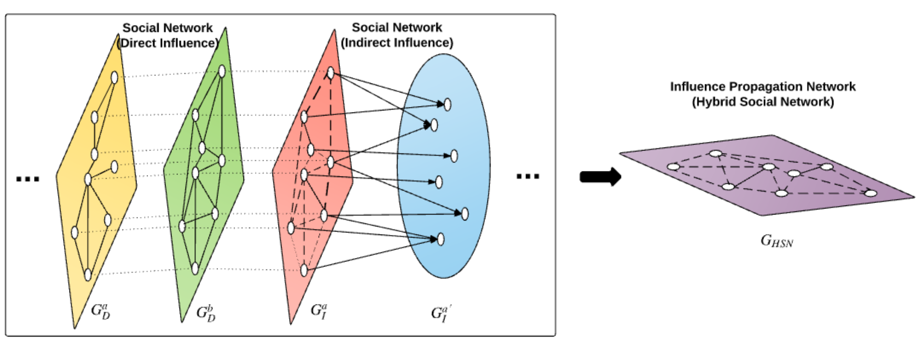


Fig. 2. Hybrid Social Network Composition

3 Comprehensive Influence Propagation (CIP)

Comprehensive Influence Propagation (CIP) model is a concrete and simplified representations of influence propagation using HSN concept. CIP constructs HSN by extracting two types of networks, i.e., **direct influence-diffusion network** and **indirect influence-diffusion network**.

3.1 Formal Definitions

Definition 1: A Social Network in general refers to a graph $G = (V, E)$ containing a set of vertices $V = \{v_1, v_2, v_3, ..., v_n\}$ interconnected by edge set $E = \{e_1, e_2, e_3, ..., e_n\}$. An edge is represented by $e_{ij} = \{(v_i, v_j)|v_i, v_j \in V \wedge v_i \neq v_j\}$, and $w(e_{ij})$ denotes the weight of the corresponding edge e_{ij}. Recall that, HSN tends to model the users' direct and indirect influential relations, thus, two types of social network present in the current setting, i.e., direct influence-diffusion network G_D^n and indirect influence-diffusion network G_I^n, where n denotes the network index, $n \in \mathbb{N}$.

Definition 2: User is defined as a vertex $v_i, v_i \in V$ in any network $G_X^n = (V, E_X^n)$. Each user has a different set of neighbours in divergent networks. User v_i's neighbourhood in a particular graph G_X^n is represented as $\Gamma(v_i|G_X^n)$, where $(e_{ij}|G_X^n) \in E_X^n,\ v_j \in \Gamma(v_i|G_X^n)$. We regard v_i and v_j as adjacent neighbours in G_X^n if their tie strength $w(e_{ij}|G_X^n)$ is larger than threshold σ.

Another type of node, i.e., item also exists in the social context. **Item** is defined as a type of entity, such as product, shop, etc., that is visited by users. $I = \{i_1, ..., i_n\}$ represents the item set, where i_x denotes the x^{th} item, $i_x \in I, x \in \mathbb{N}$. v_j has a **preference state** towards item i_x, i.e., $s_{jx}, s_{jx} \in \{PA, NA, IA\}$. $s_{jx} = PA$ implies that v_j shows a favour towards item i_x and tends to diffuse positive influence to its neighbours $\Gamma(v_j)$. Similarly, $s_{jx} = NA$ indicates that v_j expresses disfavour towards item i_x. While, IA refers to neutral opinion. The initial NA users are regarded as the **negative influencers**.

Definition 3: Direct Influence-Diffusion Network $G_D^n = (V, E_D^n)$ is a homogeneous network representing users V and their direct influential relations of n^{th} layer, where the edges E_D^n presenting direct interactions are explicitly formed. Edge weight $w(e_{ij}|G_D^n)$ denotes the interaction frequency between users v_i and v_j in G_D^n. Equation 1 aims to calculate the edge weight in direct influence-diffusion network, where $f_i|G_D^n$ denotes user v_i's interactions with neighbours in G_D^n, and $max(f_i|G_D^n)$ refers to v_i's maximum interaction frequency G_D^n.

$$w(e_{ij}|G_D^n) = \frac{w(e_{ij}|G_D^n) - min(f_i|G_D^n)}{max(f_i|G_D^n) - min(f_i|G_D^n)} \cdot \frac{w(e_{ij}|G_D^n) - min(f_j|G_D^n)}{max(f_j|G_D^n) - min(f_j|G_D^n)} \qquad (1)$$

Definition 4: Indirect Influence-Diffusion Network $G_I^n = (V, I, E_I^n, E_I^{n\prime})$ represents one of the indirect influence propagation spaces, where the users' relations E_I^n can be established implicitly based on the user-item interactions $E_I^{n\prime} = \{(v_m, i_n)|v_m \in V, i_n \in I\}$. Meanwhile, $\varphi(v_i, G_I^n)$ represents the interacted items of v_i in G_I^n. The edge weight $w(e_{ij}|G_I^n)$ denotes the implicit influential relations, such as similar preference, attention or criteria for items.

In this context, we assume $w(e_{ij}|G_I^n)$ is derived from the ratings to items. Specifically, r_{jx} indicates the rating value of i_x given by v_j, and $w(e_{ij}|G_I^n)$ is formulated using Jaccard index [11] by considering rating differences in Equation 2.

$$w(e_{ij}|G_I^n) = (1 - \sqrt{\sum\nolimits_{i_x \in \varphi(v_i, G_I^n) \cap \varphi(v_j, G_I^n)} (r_{ix} - r_{jx})^2}) \cdot \frac{|\varphi(v_i, G_I^n) \cap \varphi(v_j, G_I^n)|}{|\varphi(v_i, G_I^n) \cup \varphi(v_j, G_I^n)|} \tag{2}$$

Definition 5: Hybrid Social Network $G_{HSN} = (V, E_H^n)$ is an influence-diffusion network constructed by merging a collection of social networks in $G_D = \{G_D^1, G_D^2, G_D^3, ..., G_D^p\}$ and $G_I = \{G_I^1, G_I^2, G_I^3, ..., G_I^q\}$. Edge $e_{ij}|G_{HSN}$ denotes a comprehensive influence-diffusion channel between v_i and v_j, while the corresponding edge weight represents the possibility that influence propagates from one node to the other, which can be formulated in Equation 3. $|G_X|$ denotes the cardinality of graph collection G_X.

$$w(e_{ij}|G_{HSN}) = 1 - \prod_{G_X^n \in \{G_D, G_I\}} \prod_{n=1}^{|G_X|} 1 - w(e_{ij}|G_X^n) \tag{3}$$

3.2 Influence Propagation Mechanism of CIP

The influence diffusion approach leveraged in the proposed model inherits and extends the key features of classic ICM, i.e., propagation and attenuation. The influence initiates from the **seed set**, i.e., activated nodes (both PA and NA). They transfer their influence through the correlation graph, whereas the power of this effect decreases when hopping further and further away from the activated nodes. Mathematically, $A = \{v_{a_1}, v_{a_2}, v_{a_3}, ..., v_{a_n}\}$ denotes an initial set of activated users, where $A \subset V, s_{v_{a_n}} \in \{PA, NA\}$. Furthermore, The classic IC model has been extended by accommodating the influence strength. In order to balance the influential capacity of both PA and NA nodes, we utilise the breadth-first influence diffusion algorithm, which is described in Algorithm 1.

In Algorithm 1, the input A_n is a set of activated vertices of level n and influence diffusion threshold σ_p. While, the output is global activated set A that is accumulated by each level of A_n. In each iteration, all the inactive neighbours of vertices in A_n are supposed to be activating candidates; the activated users including both PA and NA are added to S_n of current level. Lines 1-3 shows the termination criteria of this recursive algorithm. Line 5 aims to balance the positive and negative influence by randomizing the sequence of activated nodes set. Line 15 checks whether the propagation criteria meets. Similarly, Line 17 aims to determine if the neighbour v_j to be activated by considering **Influence Propagation Attenuation** (IPA). Line 19 updates the IPA after a successful propagation. Lines 24-26 demonstrate the increment of global activated vertices A. The algorithm is recursive, thus, S_n will be taken as input for the next level by invoking itself in Line 27.

3.3 CIP-based Influence Maximization

A typical application of CIP is to tackle the influence maximization problem [4][5], where social network plays the medium for promoting a particular item. It aims to select a set of influential users with limited budgets to maximize the positive influence of a particular item in social networks [12]. The selection process is named as **seed selection**, and the selected set of users is **seed set**.

Algorithm 1 Breadth-First Influence Propagation Algorithm

Input: A_n, G_{HSN}, σ_p
Output: A
1: **if** $A_n = \emptyset$ **then**
2:     Return
3: **else**
4:     Initialize $S_n := \emptyset$
5:     Shuffle the sequence of A
6: **end if**
7: **for** $\forall v_i \in A_n$ **do**
8:     **if** $NotRecursiveInoke$ **then**
9:         Initialize $v_i.IPA := 1$
10:     **end if**
11:     **for** $\forall v_j \in \Gamma(v_i|G_{HSN}) \wedge w(e_{ij}|G_{HSN}) \neq 0$ **do**
12:         **if** $s_j \neq IA$ **then**
13:             Next
14:         **end if**
15:         **if** $w(e_{ij}|G_{HSN}) \cdot v_i.IPA \geq \sigma_p$ **then**
16:             Generate a random decimal $d_r, 0 \leq d_r \leq 1$
17:             **if** $d_r \leq w(e_{ij}|G_{HSN}) \cdot v_i.IPA$ **then**
18:                 $s_{jx} := s_{ix}$
19:                 $v_j.IPA := v_i.IPA \cdot w(e_{ij}|G_{HSN})$
20:                 $S_n := S_n \cup \{v_j\}$
21:             **end if**
22:         **end if**
23:     **end for**
24:     **if** $S_n \neq \emptyset$ **then**
25:         $A := A \cup S_n$
26:     **end if**
27:     Invoke self and input S_n as variable
28: **end for**

The CIP-based influence maximization problem tends to select seeds from a heterogeneous network having different types of nodes and links, where both positive and negative influence towards a particular item coexist, and they are capable of propagating through various channels. It aims to maximize the positive impact with different scale of negative influencers. Therefore, the evaluation metric of **influence effectiveness** is formulated in Equation 4, where $|PA|$, $|NA|$ and $|V|$ represent the size of positive, negative and all the users respectively. The reward from positive nodes and penalization from the negative ones present in an asymmetric way by using a trade-off factor $\beta, \beta \in [0, 1]$, which is determined based on the specific business needs.

$$\xi(PA, NA) = \beta \cdot \frac{|PA|}{|V|} + (1 - \beta) \cdot (1 - \frac{|NA|}{|V|}) \tag{4}$$

There are a number of classic seed selection algorithms for influence maximization problem, such as greedy selection, rank-based selection and random selection [7][5]. In the current setting, by considering multiple influence-diffusion channels, we utilise the following seed selection algorithms to evaluate the model performance in different scales of negative influencers.

- *Greedy Selection:* Obtain the maximum influence marginal gain in selecting each seed.
- *Rank-based Selection:* Rank users based on the node degree in a particular social network.
- *Random Selection:* Select users randomly.

4 Experiment and Analysis

In this section, experiment has been conducted to evaluate the proposed model in CIP-based influence maximization problem. Due to the current technical limitations and great privacy concerns, it is impossible to capture all the possible behaviours and relations of each individual. Therefore, we select two social networks for this experiment, i.e., the Trust Network (TNT) and Preference Network (PNT), presenting direct influence-diffusion network and indirect influence-diffusion network respectively.

4.1 Experiment Scenario

An organization intends to market a particular new product i_x via social networks. As the product has been introduced via other approaches, some users already have a prospective attitude towards i_x. Due to the limited budget, the organization plans to select a finite set of users as initial positive influencers, hoping that they can recommend i_x to their friend circles and spread positive influence in the network. We assume the possible negatives influencers are known. If any of them are selected as seeds, the preference state is not revised and they are not supposed to exert any positive influence on the neighbours.

4.2 Dataset

Movielens[3] dataset [13] has been used in our experiment. It is a stable benchmark dataset, which contains 1,000,209 anonymous ratings of approximately 3,900 movies made by 6,040 MovieLens users who joined MovieLens in 2000.

In order to reduce the computational cost, 500 users are selected randomly. Three social networks have been constructed based on the dataset, and the basic properties show in Table 1.

Table 1. Properties of Social Networks

Networks	**#Nodes**	**#Edges**	**Average Degree**	**Average Clustering Coefficient**
Hybrid Social Network (HSN)	500	3443	15.068	0.263
Preference Network (PNT)	500	4046	17.707	0.122
Trust Network (TNT)	500	5737	24.141	0.822

4.3 Experimental Results

Performance of the CIP is evaluated in influence maximization problem by using various traditional seed selection algorithms, where three rank-based approaches are based on the node degree of the corresponding social network. The evaluation metric, i.e., influence effectiveness, has been defined in Equation 4. By considering both positive and negative influence, we assign $\beta = 0.5$ in our experiment. As influence diffusion and infection is a stochastic model, the values of influence effectiveness are averaged over 100 trials.

Selecting seed set from HSN implies identifying influencers by considering multiple influence-diffusion channels, i.e., the trust connectivity and user preference in the current setting. Whereas, selecting seed set from PNT or TNT indicates the consideration only covers a particular aspect of influence. Among all the seed selection algorithms, the greedy selection normally performs better than others though it is not scalable [6]. Therefore, we regard the seed set and its corresponding influence effectiveness produced by greedy algorithm as the optimal solution. In this experiment, seeds are selected from HSN using greedy selection, subsequently, they have been input into the TNT and PNT to evaluate the influence effectiveness against other algorithms. Next, we select seeds from TNT and PNT using greedy algorithm and apply them in HSN for evaluation.

Figures 3 and 4 demonstrate the influence effectiveness comparison in TNT and PNT respectively, where only 15 negative influencers present in this context. As we can observe from both figures, the seeds from HSN give more considerable performance, since the effectiveness is pretty close to the seeds produced by greedy algorithm, especially in TNT.

[3] https://grouplens.org/datasets/movielens

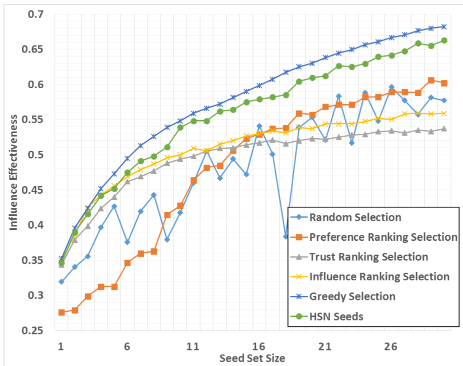


Fig. 3. Evaluation in TNT

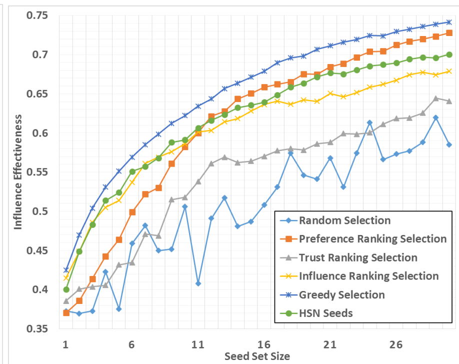


Fig. 4. Evaluation in PNT

Moreover, we evaluate the proposed model further by comparing the influence effectiveness difference in various scenarios, where the "difference" refers to the influence effectiveness gap between the seed set selected by a particular algorithm and that produced by the greedy selection in the same network and scenario. Table 2 compares the effectiveness difference in different networks and scenarios. For example, TNT (random):0.130 refers to the influence effectiveness difference between random selection and greedy selection in TNT, when 15 negative influencers exist and the size of the seed set is 10. We also find that the optimal solution from HSN still performs well and even outperforms some classic heuristic-based algorithms in other social networks.

Table 2. Influence Effectiveness Difference Comparison

Neg Influencers Size	15	15	15	30	30	30	50	50	50
Seed Set Size	10	20	30	10	20	30	10	20	30
TNT (random)	0.130	0.077	0.105	0.103	0.113	0.103	0.069	0.106	0.096
TNT (pref-rank)	0.120	0.073	0.080	0.105	0.076	0.081	0.074	0.074	0.087
TNT (trust-rank)	0.054	0.107	0.145	0.053	0.100	0.137	0.034	0.077	0.118
TNT (seeds from HSN)	**0.009**	**0.021**	**0.020**	**0.019**	**0.022**	**0.015**	**0.026**	**0.031**	**0.026**
PNT (random)	0.116	0.166	0.156	0.174	0.137	0.172	0.113	0.148	0.189
PNT (pref-rank)	0.040	**0.032**	**0.014**	0.040	**0.014**	**0.003**	0.028	**0.022**	**0.013**
PNT (trust-rank)	0.104	0.121	0.101	0.113	0.135	0.110	0.106	0.139	0.141
PNT (seeds from HSN)	**0.031**	0.035	0.041	0.021	0.024	0.021	**0.023**	0.026	0.030
HSN (random)	0.144	0.116	0.140	0.099	0.150	0.101	0.150	0.139	0.138
HSN (pref-rank)	0.105	0.049	0.042	0.097	**0.053**	0.057	0.082	0.057	0.053
HSN (trust-rank)	0.060	0.087	0.086	0.063	0.098	0.104	0.062	0.096	0.104
HSN (seeds from PNT)	0.048	0.049	0.052	0.058	0.062	0.067	**0.051**	**0.052**	**0.047**
HSN (seeds from TNT)	**0.031**	**0.044**	**0.038**	**0.051**	0.066	**0.048**	0.057	0.072	0.058

Whereas, the optimal solutions from TNT and PNT present higher influence effectiveness difference, though some are better than other algorithms under certain scenarios. In Figures 5 and 6, by comparing the difference of exchanging the solutions, it is obvious that the seeds from HSN are closer to the optimal solutions in other social networks.

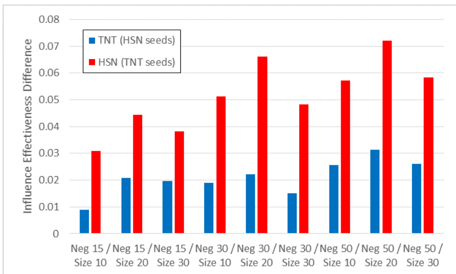


Fig. 5. Solution Exchange: TNT and HSN

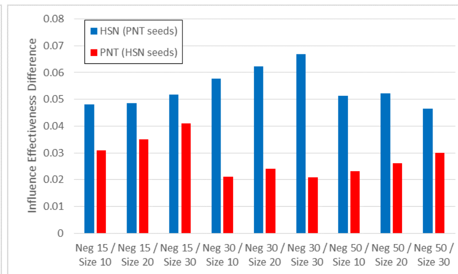


Fig. 6. Solution Exchange: PNT and HSN

Based on the experiment and the above discussion, we can conclude that the proposed model can deliver more considerable and stable seed set than those considering a single aspect of influence, such as trust connectivity or user's preference.

5 Conclusion and Future Work

In this paper, we proposed a novel approach, i.e., hybrid social network, to model the influence propagation in hybrid networks. We articulated the multi-faceted nature of influence, introduced the decomposition of influence effects, and defined direct/indirect influence. Furthermore, a concrete implementation of hybrid social network, i.e., CIP has been introduced, and validated in influence maximization problem. The experimental results reveal that seed set selected from HSN using greedy algorithm gives considerable and stable performance in other social networks, but not vice versa.

As we claimed that CIP is just one of the applications of hybrid social network. There are various potential directions to investigate influence diffusion by leveraging this generic approach, thus, the future work is set as follows.

- **Capture the dynamics of influence diffusion using HSN.** Hybrid social network implies the decomposition of influence effects, which gives high extensibility and flexibility. Specifically, when other available influential factors are added or the existing factors are changed, the model can be extended by granting adaptation capabilities, which is able to update a particular influence facet with the evolution of social networks.

– **Model the influence diffusion in a decentralised manner.** Influence agents can be allocated to each social network to monitor and manage the influential relations, while a central agent takes charge of the hybrid social network by communicating with the other agents.
– **Analyse major channels of influence diffusion.** The hybrid social network model can be extended by considering the impact factor in each influence facet. A particular influence can be diffused through various channels with different chances/possibilities. Further research works can be set to analyse the major channels for influences.

References

1. Christy MK Cheung and Dimple R Thadani. The impact of electronic word-of-mouth communication: A literature analysis and integrative model. *Decision Support Systems*, 54(1):461–470, 2012.
2. Smriti Bhagat, Amit Goyal, and Laks VS Lakshmanan. Maximizing product adoption in social networks. In *Proceedings of the fifth ACM international conference on Web search and data mining*, pages 603–612. ACM, 2012.
3. Anthony. The nature of influence, March 2009. [Online; posted 02-March-2009].
4. Pedro Domingos and Matt Richardson. Mining the network value of customers. In *Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 57–66. ACM, 2001.
5. David Kempe, Jon Kleinberg, and Éva Tardos. Maximizing the spread of influence through a social network. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 137–146. ACM, 2003.
6. Wei Chen, Yajun Wang, and Siyu Yang. Efficient influence maximization in social networks. In *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 199–208. ACM, 2009.
7. Wei Chen, Yifei Yuan, and Li Zhang. Scalable influence maximization in social networks under the linear threshold model. In *Data Mining (ICDM), 2010 IEEE 10th International Conference on*, pages 88–97. IEEE, 2010.
8. Yichuan Jiang and JC Jiang. Diffusion in social networks: A multiagent perspective. *Systems, Man, and Cybernetics: Systems, IEEE Transactions on*, 45(2):198–213, 2015.
9. Yi-Cheng Chen, Wen-Yuan Zhu, Wen-Chih Peng, Wang-Chien Lee, and Suh-Yin Lee. Cim: community-based influence maximization in social networks. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 5(2):25, 2014.
10. Marco Dorigo, Eric Bonabeau, and Guy Theraulaz. Ant algorithms and stigmergy. *Future Generation Computer Systems*, 16(8):851–871, 2000.
11. Seung-Seok Choi, Sung-Hyuk Cha, and Charles C Tappert. A survey of binary similarity and distance measures. *Journal of Systemics, Cybernetics and Informatics*, 8(1):43–48, 2010.
12. Matthew Richardson and Pedro Domingos. Mining knowledge-sharing sites for viral marketing. In *Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 61–70. ACM, 2002.
13. F Maxwell Harper and Joseph A Konstan. The movielens datasets: History and context. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 5(4):1–19, 2015.