# Information Diffusion

# 9

## 9.1 Overview

In the real world, social information can widely spread among people, and information exchange has become one of the most important social activities. The creation of the Internet and online social networks has rapidly facilitated the communication among people. Via the interactions among users in online social networks, information can easily be propagated from one user to other users. For instance, in recent years, online social networks have become the most important social occasion for news acquisition, and many outbreaking social events can get widely spread in the online social networks at a very fast speed. People as the multi-functional "sensors" can detect different kinds of event signals happening in the real world, and write posts to report their discoveries via the online social networks.

In this chapter, we will study the *information diffusion* problem [13, 42] in the online social networks. Formally, *diffusion* denotes the spreading process of certain entities (like information, idea, innovation, even heat in physics and disease in bio-medical science) through certain channels among the target group of objects inside a system. The *entity* to be spread, the propagation *channels*, the target group of *objects*, and the overall *system* can all affect the information diffusion process and lead to different diffusion observations. Therefore, different types of diffusion models have been proposed already, which will be introduced in this chapter.

Depending on the system where the diffusion process is studied, the diffusion models can be divided into (1) information diffusion models in social networks [13], (2) viral spreading models in the bio-medical system [15], and (3) heat diffusion models in physical system [21]. We will focus on the information diffusion in online social networks in this chapter. The channels for information diffusion actually belong to certain sources (or platforms), like the diffusion channels in the *online platform* [13], diffusion channels in the *offline world* [47], as well as the diffusion channels across *multiple online platforms* [38–40, 49]. Meanwhile, depending on the number of diffusion channels as well as the information sources available, the diffusion models include (1) single-channel diffusion model [13], (2) multi-source single-channel diffusion model [41], and (3) multi-source multi-channel diffusion model [38,47]. On the other hand, based on the number of topics to be spread in the diffusion process, the information diffusion models can be categorized into (1) single topic diffusion models [13, 38, 39], (2) multiple intertwined topics concurrent diffusion models [46, 47].


J. Zhang, P. S. Yu, *Broad Learning Through Fusions*,
https://doi.org/10.1007/978-3-030-12528-8_9

In the following part of this chapter, we will introduce different kinds of diffusion models proposed to depict how information propagates among users in online social networks. We will first talk about the classic diffusion models proposed for the single-network single-channel scenario, including the *threshold based models* [10, 13], *cascades based models* [13], *heat diffusion based models* [21], and *viral diffusion based models* [15]. After that, several diffusion models proposed for much more complicated scenarios will be introduced, including the *intertwined information diffusion model* [46], *signed network diffusion model* [50], *network coupling based diffusion model* [49], *cross-network random walk based diffusion model* [41], and *multi-source multi-channel diffusion model* [47].

## 9.2 Traditional Information Diffusion Models

The "*diffusion*" phenomenon has been observed in different disciplines, like social science, physics, and bio-medical science. Various diffusion models have been proposed in these areas already. We summarize the traditional information diffusion models in Fig. 9.1. In this section, we will provide a brief introduction to these models, and talk about how to apply or adapt them to describe information diffusion process in online social networks, which covers (1) how the information diffusion process starts, (2) how the information spreads, and (3) how the information diffusion process ends.

Let $G = (\mathcal{V}, \mathcal{E})$ represent the target network structure, based on which we want to study the information diffusion problem. Formally, given a user node $u \in \mathcal{V}$, we can represent the set of neighbors of $u$ as $\Gamma(u) = \{v | v \in \mathcal{V} \wedge (u, v) \in \mathcal{E}\}$. Each user node in the network $G$ will have an indicator denoting whether the user has been activated by certain information or not. We will use notation $s(u) = 1$ to denote that user $u$ has been activated, and $s(u) = 0$ to represent that $u$ is still inactive. Initially, all the users are inactive to a certain information. Information can be propagated from an initial influence seed user set $\mathcal{S} \subset \mathcal{V}$ who are exposed to and activated by the information at the very beginning. At a timestamp in the diffusion process, given user $u$'s neighbor, we can represent

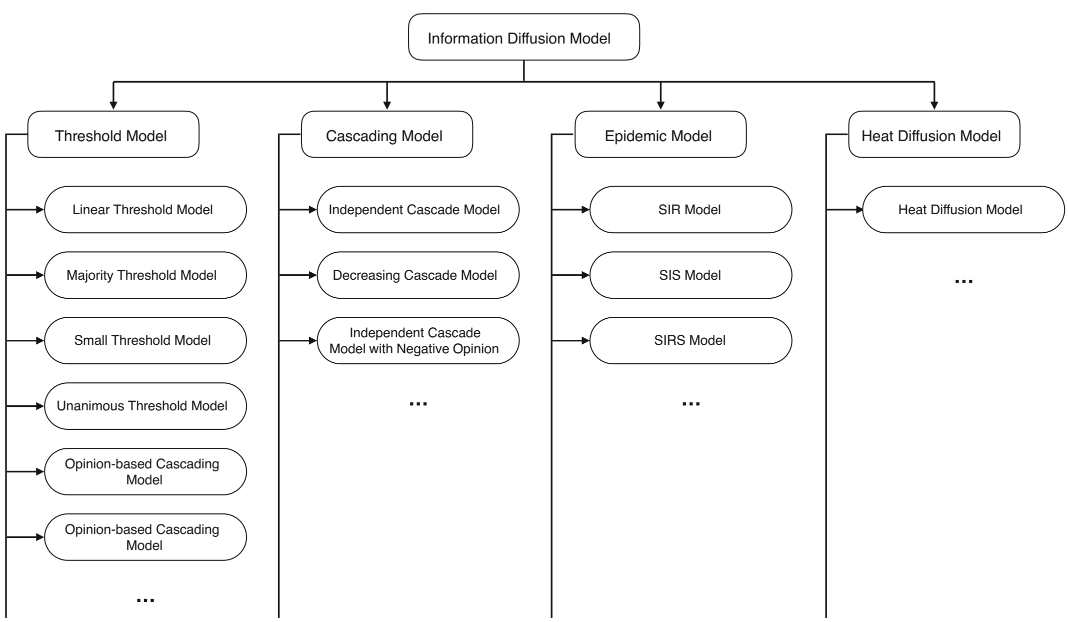


**Fig. 9.1** A summary of traditional information diffusion models

the subset of the $u$'s active neighbors as $\Gamma^a(u) = \{v | v \in \Gamma(u), s(v) = 1\} \subseteq \Gamma(u)$. The set of inactive neighbors can be represented as $\Gamma^i(u) = \Gamma(u) \setminus \Gamma^a(u)$. Generally, the information diffusion process will stop if no new activation is possible.

## 9.2.1 Threshold Based Diffusion Model

In this subsection, we will introduce the *threshold* based models [10], and will use *linear threshold model* [13] as an example to illustrate such a kind of models. Several different variants of the *linear threshold models* will be briefly introduced in this part as well.

Generally, the *threshold models* assume that individuals have a unique threshold indicating the minimum amount of required information for them to be activated by certain information. Information can propagate among the users, and the propagated information amount is determined by the closeness of the users. Close friends can influence each other much more than regular friends and strangers. If the information propagated from other users in the network surpasses the threshold of a certain user, the user will turn to an activated status and also start to influence other users. Therefore, the threshold values can determine both the status and actions of users in the online social networks. Depending on the setting of the thresholds as well as the amount of information propagated among the users, the *threshold models* have different variants.

### 9.2.1.1 LT Model

In the *linear threshold* (LT) model [13], each user has a unique threshold denoting the minimum required information to active the user. Formally, the threshold of user $u$ can be represented as $\theta_u \in [0, 1]$. In the simulation experiments, the threshold values are normally selected from the uniform distribution $U(0, 1)$. Meanwhile, for each user pair, like $u, v \in \mathcal{V}$, information can be propagated between them. As mentioned before, close friends will have larger influence on each other compared with regular friends and strangers. Formally, the amount of information users $u$ can send to $v$ is denoted as weight $w_{u,v} \in [0, 1]$. Generally, the total amount of information can send out is bounded. For instance, in the LT model, the total amount of information user $u$ can receive is bounded by 1, i.e., $\sum_{v \in \Gamma(u)} w_{v,u} \leq 1$. Different methods have been proposed to define the specific values of the weight $w_{u,v}$. In many of the cases, between the same user pair $u$ and $v$, weight $w_{u,v}$ can be different from weight $w_{v,u}$, since the information propagation between users is asymmetrical. However, in many simulation experiments, to simplify the setting, for the same user pair, $w_{u,v}$ and $w_{v,u}$ are usually assigned with the same value. For instance, in the LT models, *Jaccard's coefficient* can be used to calculate the closeness between the user pairs, where $w_{u,v}$ will be equal to $w_{v,u}$.

In the LT model, the information sent from the neighbors to user $u$ can be aggregated with linear summation. For instance, the total amount of information user $u$ can receive from his/her neighbors can be denoted as $\sum_{v \in \Gamma(u)} w(v, u)s(v)$ or $\sum_{v \in \Gamma^a(u)} w(v, u)$. To check whether a user can be activated or not, the LT model will confirm whether the following inequality holds or not:

$$\sum_{v \in \Gamma^a(u)} w(v, u) \geq \theta_u. \tag{9.1}$$

The above inequality denotes whether the received information can surpass the activation threshold of user $u$ or not. Here, we also need to notice that inactive neighbors will not send out information, and only the active neighbors can send out information. The information provided so far shows the critical details of the LT model. Next, we will provide the general framework of the LT model to illustrate how it works.

In the LT model, the initial activated seed user set can be represented as $\mathcal{S} \subset \mathcal{V}$, users in which can start the propagation of information to their neighbors. Generally, in the LT model, information propagates among users within the network step by step.

- *Diffusion Starts*: At step 0, only the seed users in $\mathcal{S}$ are active, and all the remaining users have inactive status.
- *Diffusion Spreads*: At step $t (t > 0)$, for each user $u$, if the information propagated from $u$'s active neighbors is greater than his threshold, i.e., $\sum_{v \in \Gamma^a(u)} w(v, u) \geq \theta_u$, $u$ will be activated with status $s(u) = 1$. All the activated users will remain active in the coming rounds, and can send out information to the neighbors in the next rounds. Active users cannot be activated again.
- *Diffusion Ends*: If no new activation happens in step $t$, the diffusion process will stop and all the activated users will be returned as the infection result.

Specifically, in the diffusion process, at step $t$, we don't need to check all the users to see whether they will be activated or not. The reason is that, in the diffusion process, for most of the inactive users, if the status of their neighbors is not changed in the previous step, i.e., step $t-1$, the influence they can receive in step $t$ will still be the same as in step $t-1$. And they will remain the same status as they are in the previous step, i.e., "*inactive*." Let $\mathcal{V}^a(t-1)$ denote the set of users who are recently activated in step $t-1$, we can represent the set of users they can influence as $\bigcup_{u \in \mathcal{V}^a(t-1)} \Gamma^i(u)$. In step $t$, these recently activated users will make changes to the information their neighbors can receive. Therefore, we only need to check whether the status of inactive users in the set $\bigcup_{u \in \mathcal{V}^a(t-1)} \Gamma^i(u)$ will meet the activation criterion or not.

After the diffusion process stops, a group of users with the active status will indicate the influence these seed users spread to, which can be represented as set $\mathcal{V}^a$. Generally, to denote the impact of certain information, we can introduce a mapping: $\sigma : \mathcal{S} \to |\mathcal{V}^a|$ to project the seed users to their expected number of activated users, which is formally called the influence function. Given the *influence function*, with different seed user sets as the input, the influence they can achieve is usually different. Choosing the optimal seed user who can lead to the maximum influence is named as the *influence maximization* problem [13], which will be introduced in the following Chap. 10 in detail.

#### 9.2.1.2 Other Threshold Models

The LT model assumes the cumulative effects of information propagated from the neighbors, and can illustrate the basic information diffusion process among users in the online social networks. The LT model has been well analyzed, and many other variant models have been proposed as well. Depending on the assignment of the threshold and weight values, many other different diffusion models can all be reduced to a special case of the LT model.

**Majority Threshold Model** Different from the LT mode, in the *majority threshold model* [27], an inactive user $u$ can be activated if majority of his/her neighbors are activated. The *majority threshold model* can be reduced to the LT model in the case that: (1) the influence weight between any friends $(u, v)$ in the network is assigned with value 1; (2) the threshold of any user $u$ is set as $\frac{1}{2}D(u)$, where $D(u)$ denotes the degree of node $u$ in the network. For the nodes with large degrees, like the central node in the star-structured diagram, their activation will lead to the activation of lots of surrounding nodes in the network.

***k*-Threshold Model** Another diffusion model similar to the LT model is called the *k-threshold diffusion model*, in which users can be activated of at least $k$ if his/her neighbors are active. The *k-threshold model* is equivalent to the LT model with the setting: (1) the influence weight between any friend pairs $(u, v)$ in the network is assigned with value 1; and (2) the activation thresholds of all the users are assigned with a shared value $k$. For each user $u$, if $k$ of his/her neighbors have been activated, $u$ will be activated.

Depending on the values of $k$, the *k-threshold model* will have different performance. When $k = 1$, a user will be activated of at least one of his/her neighbor is active. In such a case, all the users in the same connected components with the initial seed users will be activated finally. When $k$ is a very large value and even greater than the large node degree, e.g., $k > \max_{u \in \mathcal{V}} D(u)$, no nodes can be activated. When $k$ is a medium value, some of the users will be activated as the information propagates, but the other users with less than $k$ neighbors will never be activated.

### 9.2.2 Cascade Based Diffusion Model

An information cascade [9] occurs when people observe the actions of others and then engage in the same actions. Cascade clearly illustrates the information propagation routes, and the activating actions performed by users on their neighbors. In the cascade model, the information propagation dynamics is carried out in a step-by-step fashion. At each step, users can have trials to activate their neighbors to change their opinions with certain probabilities. If they succeed, the neighbors will change their status to follow the initiators. In the case that multiple users can all have the chance to activate a certain target user, the activation trials will be performed sequentially in an arbitrary order.

Depending on the activation trials and users' reactions to the activation trials, different cascade models have been proposed already. In this section, we will use the *independent cascade* (IC) model [13] as an example of the cascade based models to illustrate the model architecture.

#### 9.2.2.1 IC Model

In the diffusion process, about a certain target user, multiple activation trials can be performed by his/her neighbors. In the *independent cascade* model [13], each activation trial is performed independently regardless of the historical unsuccessful trials. The activation trials are performed step by step. When user $u$ who has been activated in the previous step and tries to activate his/her neighbor $v$ in the current step, the success probability is denoted as $p_{u,v} \in [0, 1]$. Generally, if users $u$ and $v$ are close friends, the activation probability will be larger compared with regular friends and strangers. The specific activation probability value is usually correlated with the social closeness between users $u$ and $v$, which can also be defined based on the Jaccard's coefficient in the simulation. The activation trials will only happen among the users who are friends. If $u$ succeeds in activating $v$, then user $v$ will change his/her status to "*active*" and will remain in the status in the following steps. However, if $u$ fails to activate $v$, $u$ will lose the chance and cannot perform the activation trials any more.

In the IC mode, we can represent the initial seed user as set $\mathcal{S} \subset \mathcal{V}$, who will spread the information to the remaining users. We illustrate the general information propagation procedure in the IC model as follows:

- *Diffusion Starts*: In the initial, the seed users will send out the information and start to activate their neighbors. For the users in set $\mathcal{S}$, the activation trials will start in a random order. For instance, if we pick user $u \in \mathcal{S}$ as the first user, $u$ will activate his/her inactive friends in $\Gamma(u)$ in a random order as well.

- *Diffusion Spreads*: In step $t$, only the users who have just been activated in the previous step can activate the other users. We can denote the users who have just been activated in the previous as set $\mathcal{V}^a(t-1) \subset \mathcal{V}$. Users in set $\mathcal{V}^a(t-1)$ will start to perform activation trials. For the users who are activated by these users, they will remain active in the following steps and will be added to the set $\mathcal{V}^a(t)$, who will start the activation trials in the next step.
- *Diffusion Ends*: If no new activation happens in a step, the diffusion process will stop and the activated users will be returned as the infection result.

In the IC model, the activation trials are performed by flipping a coin with certain probabilities, whose result is uncertain. Even with the same provided initial seed user set $\mathcal{S}$, the number of users who will be activated by the seed users can be different if we run the IC model twice. Formally, we can represent the set of activated users by the seed users as $\mathcal{V}^a \subset \mathcal{V}$. Therefore, in the experimental simulations, we usually run the diffusion model multiple times and calculate the average number of activated users, i.e., $|\mathcal{V}^a|$, to denote the expected influence achieved by IC on the provided seed user set $\mathcal{S}$.

#### 9.2.2.2 Other Cascade Models

Generally, the independent activation assumption renders the IC model the simplest cascade based diffusion models. In the real world, the diffusion process will be more complicated. For the users, who have failed to be activated by many other users, it probably indicates that the user is not interested in the information. Viewed in such a perspective, the probability for the user to be activated will decrease as more activation trials have been performed. In this part, we will introduce another cascade based diffusion model, *decreasing cascade model* [14].

**Decreasing Cascade (DC) Model** To illustrate the DC model more clearly and show its difference compared with the IC model, we use notation $P(u \to v|\mathcal{T})$ to represent the probability for user $u$ to activate $v$ given a set of users $\mathcal{T}$ who have performed but failed the activation trials to $v$ in the previous rounds. Let $\mathcal{T}$, $\mathcal{T}'$ denote two different historical activation user set, where $\mathcal{T} \subseteq \mathcal{T}'$. In the IC model, we have

$$P(u \to v|\mathcal{T}) = P(u \to v|\mathcal{T}'). \tag{9.2}$$

In other words, every activation trial is independent with each other, and user's activation probability will not be changed as more activation trials have been performed.

As introduced at the beginning of this subsection, the fact that users in set $\mathcal{T}$ fail to activate $v$ indicates that $v$ probably is not interested in the information, and the chance for $v$ to be activated afterwards will be lower. Furthermore, as more activation trials are performed by users in $\mathcal{T}'$, the probability for $u$ to active $v$ will be decreased steadily, i.e.,

$$P(u \to v|\mathcal{T}) \geq P(u \to v|\mathcal{T}'). \tag{9.3}$$

Intuitively, this restriction states that a contagious node's probability of activating some $v$ decreases if more nodes have already attempted to activate $v$, and $v$ is hence more "marketing-saturated." The DC model incorporates the IC model as a special case, and it is a much more general information diffusion model.

### 9.2.3 Epidemic Diffusion Model

The threshold and cascade based diffusion models introduced in the previous subsections mostly assume that "once a user is activated, he/she will remain in the active status forever." However, in the real world, these activated users can change their minds and they can still have the chance to recover to their originally inactive status. In the bio-medical science, the diffusion problem has also been studied for many years to model the spread of disease, and several *epidemic diffusion models* [15] have been introduced already. In the disease propagation, people who are susceptible to the disease can be get infected by other people. After some time, many of these infected people can get recovered and become immune to the disease, while many other users can recover but may get susceptible to the disease again. Depending on the people's reactions to the disease after recovery, several different *epidemic diffusion models* have been proposed.

In this subsection, we will introduce the *epidemic diffusion models*, and try to use them to model the information diffusion process in online social networks.

#### 9.2.3.1 Susceptible-Infected-Recovered (SIR) Diffusion Model

The SIR model [15] was proposed by W.O. Kermack and A.G. McKendrick in 1927 to model the spread of infectious diseases, which considers a fixed population in three main categories: *susceptible* (S), *infected* (I), and *recovered* (R). As the disease propagates, the individual status can change among $\{S, I, R\}$ with the following flow:

$$S \rightarrow I \rightarrow R. \tag{9.4}$$

In other words, the individuals who are susceptible to the disease can get infected, while those infected individuals also have the chance to recover from the disease. In this part, we will use the SIR model to describe the information cascading process in online social networks. Let $\mathcal{V}$ denote the set of users in the network. We introduce the following notations to represent the number of users in different categories:

- $S(t)$: the number of users who are *susceptible* to the information at time $t$, but have not been *infected* yet.
- $I(t)$: the number of users who are currently *infected* by the information, and can spread the information to others in the *susceptible* category.
- $R(t)$: the number of users who have been infected and already recovered from the information infection. After recovery, the users will become immune to the information and cannot be infected again in the future.

Based on the above notations, we have the following equations hold in the SIR model.

$$\begin{cases} (1) \quad S(t) + I(t) + R(t) = |\mathcal{V}|, \\ (2) \quad \frac{\mathrm{d}S(t)}{\mathrm{d}t} + \frac{\mathrm{d}I(t)}{\mathrm{d}t} + \frac{\mathrm{d}R(t)}{\mathrm{d}t} = 0, \\ (3) \quad \frac{\mathrm{d}S(t)}{\mathrm{d}t} = -\beta S(t)I(t), \\ (4) \quad \frac{\mathrm{d}I(t)}{\mathrm{d}t} = \beta S(t)I(t) - \gamma I(t), \\ (5) \quad \frac{\mathrm{d}R(t)}{\mathrm{d}t} = \gamma I(t). \end{cases} \tag{9.5}$$

In the above equations, the parameter $\beta$ denotes the infection rate of these *susceptible* users by the *infected* users in a unit time, and $\gamma$ represents the recovery rate. Generally, all the users in the social network will belong to these three categories, and the total number of users in these three categories will sum to $|\mathcal{V}|$ at any time in the diffusion process. Therefore, we can also get the derivatives of the summation term (i.e., $S(t) + I(t) + R(t)$) with regarding to the time parameter $t$ will be 0. At a unit time, the number of users transit from the *susceptible* status to the *infection* status depends on the available susceptible and infected users at the same time. For each infected user, the number of users he/she can infect is proportional to the available susceptible users, which can be denoted ad $\beta S(t)$. For all the infected users, the total number of users can get infected will be $\beta S(t) I(t)$. For the number of users who are recovered in a unit time, it depends on the number of total infected users $I(t)$ as well as the recovery rate $\gamma$, which can be represented as $\gamma I(t)$. Meanwhile, as to the number of infected user changes in a unit time, it is determined by both the number of susceptible users who get infected and the infected users who get recovered.

We have parameters $\beta, \gamma \geq 0$, and the numbers $S(t), I(t), R(t) \geq 0$ to be positive at any time. Therefore, we can know that (1) $\frac{\mathrm{d}S(t)}{\mathrm{d}t} \leq 0$, and users in the *susceptible* group is non-increasing; (2) $\frac{\mathrm{d}R(t)}{\mathrm{d}t} \geq 0$, and users in the *recovered* group is non-decreasing; while (3) the sign of term $\frac{\mathrm{d}I(t)}{\mathrm{d}t}$ can be either positive, zero or negative depending on the parameters $\beta, \gamma$ and the users in the *susceptible* and *infected* groups:

- *positive*: if $\beta S(t) > \gamma$;
- *zero*: if $\beta S(t) = \gamma$ or $I(t) = 0$;
- *negative*: if $\beta S(t) < \gamma$.

#### 9.2.3.2 Susceptible-Infected-Susceptible (SIS) Diffusion Model

In some cases, the users cannot get immune to the information and don't exist the *recovery* status actually. For the users, who get infected, they can go to the *susceptible* status and can get *infected* again in the future. To model such a phenomenon, another diffusion model very similar to the SIR model has been proposed, which is called the susceptible-infected-susceptible (SIS) model [12].

In the SIS model, the individual status flow is provided as follows:

$$S \to I \to S. \tag{9.6}$$

Such a status flow will continue, and individuals will switch their status between *susceptible* and *infected* in the information diffusion process. Therefore, the absolute number changes of individuals in these two categories will be the same in unit time.

$$\begin{cases} (1) \quad S(t) + I(t) + R(t) = |\mathcal{V}|, \\ (2) \quad \frac{\mathrm{d}S(t)}{\mathrm{d}t} + \frac{\mathrm{d}I(t)}{\mathrm{d}t} + \frac{\mathrm{d}R(t)}{\mathrm{d}t} = 0, \\ (3) \quad \frac{\mathrm{d}S(t)}{\mathrm{d}t} = -\beta S(t) I(t) + \gamma I(t), \\ (4) \quad \frac{\mathrm{d}I(t)}{\mathrm{d}t} = \beta S(t) I(t) - \gamma I(t). \end{cases} \tag{9.7}$$

#### 9.2.3.3 Susceptible-Infected-Recovered-Susceptible (SIRS) Diffusion Model

The susceptible-infected-recovered-susceptible (SIRS) diffusion model to be introduced in this part is another type of epidemic model, where the individuals in the *recovery* category can lose their immunity and transit to the *susceptible* category. These individuals have the potential to get infected again. Therefore, the individual status flow will be as follows:

$$S \to I \to R \to S. \tag{9.8}$$

We can denote the rate of individuals who lose the immunity as $f$, and the total number of individuals who may lose the immunity will be $f \cdot R(t)$. Therefore, we can have the derivative of the individual numbers belonging to different categories as follows:

$$\frac{\mathrm{d}S(t)}{\mathrm{d}t} = -\beta S(t)I(t) + fR(t), \tag{9.9}$$

$$\frac{\mathrm{d}I(t)}{\mathrm{d}t} = \beta S(t)I(t) - \gamma I(t), \tag{9.10}$$

$$\frac{\mathrm{d}R(t)}{\mathrm{d}t} = \gamma I(t) - fR(t). \tag{9.11}$$

Besides these epidemic diffusion models introduced in this subsection, there also exist many different versions of the epidemic diffusion models, which consider many other factors in the diffusion process, like the birth/death of individuals. It is also very common in the real-world online social networks, since new users will join in the social network, and existing users will also delete their account and get removed from the social network. Involving such factors will make the diffusion model more complex, and we will not introduce them here due to the limited space. More information about these different epidemic diffusion models is available in [26].

### 9.2.4 Heat Diffusion Models

Heat diffusion is a well-observed physical phenomenon. Generally, in a medium, heat will always diffuse from regions with a high temperature to the region with a lower temperature. In this subsection, we will talk about the *heat diffusion model* [21] and introduce how to adapt it to model the information diffusion process in online social networks.

#### 9.2.4.1 General Heat Diffusion

Throughout a geometric manifold, let function $f(x, t)$ denote the temperature at region $x$ at time $t$, and we can represent the initial temperature at different regions as $f_0(x)$. The heat flows with initial conditions can be described by the following differential equation:

$$\begin{cases} \frac{\partial f(x,t)}{\partial t} - \Delta f(x, t) = 0 \\ f(x, 0) = f_0(x), \end{cases} \tag{9.12}$$

where $\Delta f(x, t)$ is a *Laplace-Beltrami operator* on function $f(x, t)$.

Many existing works on the heat diffusion studies are mainly focused on the heat kernel matrix. Formally, let $\mathbf{K}_t$ denote the heat kernel matrix at timestamp $t$, which describes the heat diffusion among different regions in the medium. In the matrix, entry $K_t(x, y)$ denotes the heat diffused from the original region $y$ to region $x$ at time $t$. However, it is very difficult to represent the medium as a regular geometry with known dimensions. In the next part, we will introduce how to apply the heat diffusion observations to model the information diffusion in the online social networks.

#### 9.2.4.2 Heat Diffusion Model

Given a homogeneous network $G = (\mathcal{V}, \mathcal{E})$, for each node $u \in \mathcal{V}$ in the network, we can represent the information accumulated at $u$ in timestamp $t$ as $f(u, t)$. The initial information available at each of the node can be denoted as $f(u, 0)$. The information can propagate among the nodes in the network if there exists a pipe (i.e., a link) between them. For instance, with a link $(u, v) \in \mathcal{E}$ in the network, information can propagate between $u$ and $v$.

Generally, in the diffusion process, the amount of information propagated between different nodes in the network depends on (1) the difference of information available at these two nodes, and (2) the thermal conductivity, i.e., the heat diffusion coefficient $\alpha$. For instance, at timestamp $t$, we can represent the amount of information reaching nodes $u, v \in \mathcal{V}$ as $f(u, t)$ and $f(v, t)$. If $f(u, t) > f(v, t)$, information tends to propagate from $u$ to $v$ in the network, and the propagated information amount is denoted as $\alpha \cdot (f(u, t) - f(v, t))$. The propagation direction will be reversed if $f(u, t) < f(v, t)$. The information amount changes at node $u$ at timestamps $t$ and $t + \Delta t$ can be represented as

$$\frac{f(u, t + \Delta t) - f(u, t)}{\Delta t} = - \sum_{v \in \Gamma(u)} \alpha \cdot (f(u, t) - f(v, t)) . \tag{9.13}$$

Let's use a vector $\mathbf{f}(t)$ to represent the amount of information available at all the nodes in the network at timestamp $t$. The above information amount changes can be rewritten with the following equation:

$$\frac{\mathbf{f}(t + \Delta t) - \mathbf{f}(t)}{\Delta t} = \alpha \mathbf{H} \mathbf{f}(t), \tag{9.14}$$

where in the matrix $\mathbf{H} \in \mathbb{R}^{|\mathcal{V}| \times |\mathcal{V}|}$, entry $H(u, v)$ has value

$$H(u, v) = \begin{cases} 1, & \text{if } (u, v) \in \mathcal{E} \vee (v, u) \in \mathcal{E}, \\ -D(u), & \text{if } u = v, \\ 0, & \text{otherwise}, \end{cases} \tag{9.15}$$

where $D(u)$ denotes the degree of node $u$ in the network.

In the limit case $\Delta t \to 0$, we can rewrite the equation as

$$\frac{\mathrm{d}\mathbf{f}(t)}{\mathrm{d}t} = \alpha \mathbf{H} \mathbf{f}(t). \tag{9.16}$$

By solving the above function, we can represent the amount of information at each node in the network as

$$\begin{aligned} \mathbf{f}(t) &= \exp^{t \alpha \mathbf{H}} \mathbf{f}(0) \\ &= \left( \mathbf{I} + \alpha t \mathbf{H} + \frac{\alpha^2 t^2}{2!} \mathbf{H}^2 + \frac{\alpha^3 t^3}{3!} \mathbf{H}^3 + \cdots \right) \mathbf{f}(0), \end{aligned} \tag{9.17}$$

where term $\exp^{t\alpha \mathbf{H}}$ is also called the diffusion kernel matrix, and it can be expanded as indicated in the above equation according to the Taylor's theorem.

## 9.3 Intertwined Diffusion Models

For the information diffusion models introduced in the previous section, they are all proposed for modeling the propagation of information in social network with one single diffusion channel and one type of information topic only. However, in the real world, multiple types of information can be propagated within the network simultaneously, relationships among which can be quite intertwined, including *competitive*, *complimentary*, and *independent*. Furthermore, within the networks, even the network structure is homogeneous but the social links among users may be associated with polarities indicating the relationship among the users. For instance, for some of the social links, they denote friendship, while for the other links, they indicate the user pairs are enemies. Formally, the social network structure with polarities associated with the social links are called *signed networks* [18], where the link polarities can affect the information diffusion in them greatly.

In this section, we will introduce two different *intertwined diffusion models* proposed to describe the information propagation process about both (1) the information entities with intertwined relationships [46], and (2) for network structures with links attaching different polarities [50], respectively.

### 9.3.1 Intertwined Diffusion Models for Multiple Topics

Traditional information diffusion studies mainly focus on one single online social network and have extensive concrete applications in the real world, e.g., product promotion [5, 24] and opinion spread [4]. In the traditional viral marketing setting [7,13], only one product/idea is to be promoted. However, in the real scenarios, the promotions of multiple products can co-exist in the social networks at the same time, which is referred to as the *intertwined information diffusion problem* [46].

*Example 9.1* The relationships among the products to be promoted in the network can be very complicated. For example, in Fig. 9.2, we show 4 different products to be promoted in an online social network and HP printer is our target product. At the product level, the relationships among these products can be:

- *independent*: promotion activities of some products (e.g., HP printer and Pepsi) can be *independent* of each other.
- *competing*: products having common functions will *compete* for the market share [2, 3] (e.g., HP printer and Canon printer). Users who have bought a HP printer are less likely to buy a Canon printer again.
- *complementary*: product cross-sell is also very common in marketing [24]. Users who have bought a certain product (e.g., PC) will be more likely to buy another product (e.g., HP printer) and the promotion of PC is said to be *complementary* to that of HP printer.

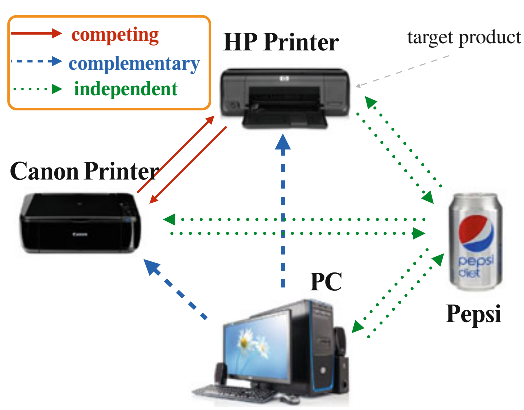


**Fig. 9.2** Intertwined relationships among products

In this section, we will study the information diffusion problem in online social networks, where multiple products are being promoted simultaneously. The relationships among these product can be obtained in advance via effective market research, which can be *independent*, *competitive*, or *complementary*. A novel information diffusion model interTwined Linear Threshold (TLT) proposed in [46] will be introduced in this section. TLT quantifies the *impacts* among products with the *intertwined threshold updating strategy* and can handle the intertwined diffusions of these products at the same time.

Before talking about the detailed diffusion models, we first introduce the definitions of several important terminologies as follows, which will be used in this section.

**Definition 9.1 (Social Network)** An online *social network* can be represented as $G = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ is the set of users and $\mathcal{E}$ contains the interactions among users in $\mathcal{V}$. The set of $n$ different products to be promoted in network $G$ can be represented as $\mathcal{P} = \{p^1, p^2, \ldots, p^n\}$.

**Definition 9.2 (User Status Vector)** For a given product $p^j \in \mathcal{P}$, users who are influenced to buy $p^j$ are defined to be "*active*" to $p^j$, while the remaining users who have not bought $p^j$ are defined to be "*inactive*" to $p^j$. User $u_i$'s status towards all the products in $\mathcal{P}$ can be represented as "*user status vector*" $\mathbf{s}_i = (s_i^1, s_i^2, \ldots, s_i^n)$, where $s_i^j$ is $u_i$'s status to product $p^j$. Users can be activated by multiple products at the same time (even competing products), i.e., multiple entries in *status vector* $\mathbf{s}_i$ can be "*active*" concurrently.

**Definition 9.3 (Independent, Competing, and Complementary Products)** Let $P(s_i^j = 1)$ (or $P(s_i^j)$ for simplicity) denote the probability that $u_i$ is activated by product $p^j$ and $P(s_i^j|s_i^k)$ be the conditional probability given that $u_i$ has been activated by $p^k$ already. For products $p^j, p^k \in \mathcal{P}$, the promotion of $p^k$ is defined to be (1) *independent* to that of $p^j$ if $\forall u_i \in \mathcal{V}$, $P(s_i^j|s_i^k) = P(s_i^j)$, (2) *competing* to that of $p^j$ if $\forall u_i \in \mathcal{V}$, $P(s_i^j|s_i^k) < P(s_i^j)$, and (3) *complementary* to that of $p^j$ if $\forall u_i \in \mathcal{V}$, $P(s_i^j|s_i^k) > P(s_i^j)$.

#### 9.3.1.1 TLT Diffusion Model

To depict the intertwined diffusions of multiple independent, competing or complementary products, we will introduce a new information diffusion model TLT. In the existence of multiple products $\mathcal{P}$, user $u_i$'s influence to his neighbor $u_k$ in promoting product $p^j$ can be represented as $w_{i,k}^j \geq 0$. Similar to the traditional LT model, in TLT, the influence of different products can propagate within the network step by step. User $u_i$'s threshold for product $p^j$ can be represented as $\theta_i^j$ and $u_i$ will be activated by his neighbors to buy product $p^j$ if

$$\sum_{u_l \in \Gamma_{out}(u_i)} w_{l,i}^j \geq \theta_i^j. \tag{9.18}$$

Different from traditional LT model, in TLT, users in online social networks can be activated by multiple products at the same time, which can be either *independent*, *competing* or *complementary*. As shown in Fig. 9.2, we observe that users' chance to buy the HP printer will be (1) unchanged given that they have bought Pepsi (i.e., the *independent* product of HP printer), (2) increased if they own PCs (i.e., the *complementary* product of HP printer), and (3) decreased if they already have the Canon printer (i.e., the *competing* product of HP printer).

To model such a phenomenon in TLT, we introduce the following *intertwined threshold updating strategy*, where users' *thresholds* to different products will change *dynamically* as the influence of other products propagates in the network.

**Definition 9.4 (Intertwined Threshold Updating Strategy)** Assuming that user $u_i$ has been activated by $m$ products $p^{\tau_1}, p^{\tau_2}, \ldots, p^{\tau_m} \in \mathcal{P} \setminus \{p^j\}$ in a sequence, then $u_i$'s *threshold* towards product $p^j$ will be updated as follows:

$$\left(\theta_i^j\right)^{\tau_1} = \theta_i^j \frac{P\left(s_i^j\right)}{P(s_i^j|s_i^{\tau_1})}, \left(\theta_i^j\right)^{\tau_2} = \left(\theta_i^j\right)^{\tau_1} \frac{P\left(s_i^j|s_i^{\tau_1}\right)}{P\left(s_i^j|s_i^{\tau_1}, s_i^{\tau_2}\right)}, \cdots \tag{9.19}$$

$$\left(\theta_i^j\right)^{\tau_m} = \left(\theta_i^j\right)^{\tau_{m-1}} \frac{P\left(s_i^j|s_i^{\tau_1}, \ldots, s_i^{\tau_{m-1}}\right)}{P\left(s_i^j|s_i^{\tau_1}, \ldots, s_i^{\tau_{m-1}}, s_i^{\tau_m}\right)}, \tag{9.20}$$

where $(\theta_i^j)^{\tau_k}$ denotes $u_i$'s threshold to $p^j$ after he has been activated by $p^{\tau_1}$, $p^{\tau_2}$, …, $p^{\tau_k}$, $k \in \{1, 2, \ldots, m\}$.

In this section, we do not focus on the order of products that activate users [4] and to simplify the calculation of the *threshold updating strategy*, we assume only the most recent activation has an effect on updating current thresholds, i.e.,

$$\frac{P\left(s_i^j|s_i^{\tau_1}, \ldots, s_i^{\tau_{m-1}}\right)}{P\left(s_i^j|s_i^{\tau_1}, \ldots, s_i^{\tau_{m-1}}, s_i^{\tau_m}\right)} \approx \frac{P\left(s_i^j\right)}{P\left(s_i^j|s_i^{\tau_m}\right)} = \phi_i^{\tau_m \to j}. \tag{9.21}$$

**Definition 9.5 (Threshold Updating Coefficient)** Term $\phi_i^{l\to j} = \frac{P(s_i^j)}{P(s_i^j|s_i^l)}$ is formally defined as the "*threshold updating coefficient*" of product $p^l$ to product $p^j$ for user $u_i$, where

$$\phi_i^{l\to j} \begin{cases} < 1, & \text{if } p^l \text{ is } \textit{complementary} \text{ to } p^j, \\ = 1, & \text{if } p^l \text{ is } \textit{independent} \text{ to } p^j, \\ > 1, & \text{if } p^l \text{ is } \textit{competing} \text{ to } p^j. \end{cases} \tag{9.22}$$

The *intertwined threshold updating strategy* can be rewritten based on the *threshold updating coefficients* as follows:

$$(\theta_i^j)^{\tau_m} \approx \theta_i^j \cdot \phi_i^{\tau_1\to j} \cdot \phi_i^{\tau_2\to j} \cdots \phi_i^{\tau_m\to j}. \tag{9.23}$$

Based on the above *threshold updating coefficients* equation together with the detailed information diffusion process described in the LT model, the TLT model introduced in this part can depict the information diffusion process about multiple products with *intertwined relationships*.

### 9.3.2 Diffusion Models for Signed Networks

In this part, we will introduce another type of *intertwined diffusion model* proposed for the *signed networks* [31, 48] involving polarized links among the nodes specifically. In recent years, signed networks have gained increasing attention because of their ability to represent diverse and contrasting social relationships. Some examples of such contrasting relationships include friends vs enemies [35], trust vs distrust [36], positive attitudes vs negative attitudes [37], and so on. These contrasting relationships can be represented as links of different polarities, which result in signed networks. Signed social networks can provide a meaningful perspective on a wide range of social network studies, like *user sentiment analysis* [34], *social interaction pattern extraction* [18], *trustworthy friend recommendation* [17], and so on.

Information dissemination is common in social networks [29]. Due to the extensive social links among users, information on certain topics, e.g., politics, celebrities and product promotions, can propagate leading to a large number of nodes reporting the same (incorrect) observations rapidly in online social networks. In particular, the links in signed networks are of different polarities and can denote trust and distrust relationships among users [20], which will inevitably have an impact on information propagation inside the networks.

*Example 9.2* In Fig. 9.3, an example is provided to help illustrate the information diffusion problem in signed networks more clearly. In the example, users are connected to one another with signed links, depending on their trust and distrust relations. It is noteworthy that the conventions used for the direction of information diffusion in this network are slightly different from traditional influence analysis, because they represent signed links. For instance, if Alice trusts (or follows) Bob, a directed edge exists from Alice to Bob, but the information diffusion direction will be from Bob to Alice. Via the signed links, inactive users in the network can get infected by certain information propagated from their neighbors with either a positive or negative opinion about the information (i.e., the green or red states in the figure). Considering the fact that it is often difficult to directly identify all the user infection states in real settings, we allow for the possibility of some user states in the network to be unknown. Activated users can propagate the information to other users. In general, if a user is activated with a positive or negative opinion about the information, she might activate one or more of her incoming neighbors to trust or distrust the information, depending on the sign of the incoming link.

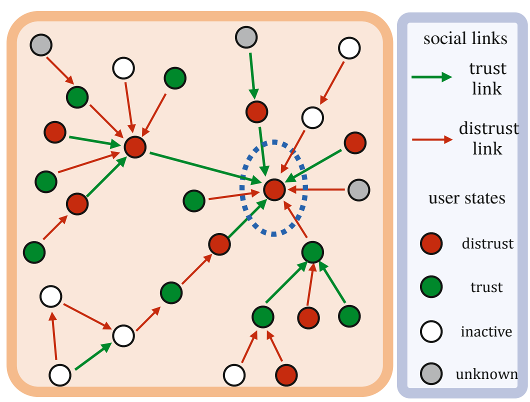


**Fig. 9.3** Example of the information diffusion problem in signed networks

The edges in the network are directed and signed, and they represent trust or distrust relationships. For example, when node $i$ trusts or distrusts node $j$, we will have a corresponding positive or negative link from node $i$ to node $j$. In this setting, nodes are associated with states corresponding to a prevailing opinion about the truth of a fact. These states can be drawn from $\{-1, +1, 0, ?\}$, where $+1$ indicates their agreement with a specific fact, $-1$ indicates their disagreement, 0 indicates the fact that they have no opinion of the fact at hand, and ? indicates their opinion is unknown. The last of these states is necessary to model the fact that the states of many nodes in large-scale networks are often unknown. Note that the use of multiple states of nodes in the network is different from traditional influence analysis. Users are influenced with varying opinions of the fact in question, based on their observation of their neighbors (i.e., states of neighborhood nodes), and their trust or distrust of their neighbor's opinions (i.e., signs of links with them). This model is essentially a signed version of influence propagation models, because the sign of the link plays a critical role in how a specific bit of information is transmitted.

Most existing information diffusion models are designed for unsigned networks. In signed networks, information diffusion is also related to actor-centric trust and distrust, in which notions of node states and the signs on links play an important role. To depict how information propagates in the signed networks, we will introduce the *asyMmetric Flipping Cascade* (MFC) diffusion model proposed in [50] for signed networks specifically.

#### 9.3.2.1 Terminology Definition

Traditional social networks are unsigned in the sense that the links are assumed, by default, to be positive links. Signed social networks are a generalization of this basic concept.

**Definition 9.6 (Weighted Signed Social Network)** A *weighted signed social network* can be represented as a graph $G = (\mathcal{V}, \mathcal{E}, s, w)$, where $\mathcal{V}$ and $\mathcal{E}$ represent the nodes (users) and directed edges (social links), respectively. In signed networks, each social link has its own polarity (i.e., the sign) and is associated with a weight indicating the intimacy among users, which can be represented with the mappings $s : \mathcal{E} \to \{-1, +1\}$ and $w : \mathcal{E} \to [0, 1]$, respectively.

As discussed before, we interpret the signs from a trust-centric point of view. Information propagated among users is highly associated with the intimacy scores [43] among them: information tends to propagate among close users. To represent the information diffusion process in trust-centric networks, we define the concept of *weighted signed diffusion network* as follows:

**Definition 9.7 (Weighted Signed Diffusion Network)** Given a *signed social network* $G$, its corresponding *weighted signed diffusion network* can be represented as $G_D = (\mathcal{V}_D, \mathcal{E}_D, s_D, w_D)$, where $\mathcal{V}_D = \mathcal{V}$ and $\mathcal{E}_D = \{(v, u)\}_{(u,v)\in\mathcal{E}}$. Diffusion links in $\mathcal{E}_D$ share the same sign and weight mappings as those in $\mathcal{E}$, which can be obtained via mappings $s_D : \mathcal{E}_D \to \{-1, +1\}$, $s_D(v, u) = s(u, v), \forall (v, u) \in \mathcal{E}_D$ and $w_D : \mathcal{E}_D \to [0, 1]$, $w_D(v, u) = w(u, v), \forall (v, u) \in \mathcal{E}_D$. For any directed diffusion link $(u, v) \in \mathcal{E}_D$, we can represent its sign and weight to be $s_D(u, v)$ and $w_D(u, v)$, respectively.

Note that we have reversed the direction of the links because of the trust-centric interpretation, in which information diffuses from A to B, when B trusts A. However, in networks with other semantic interpretations, this reversal does not need to be performed. The overall algorithm is agnostic to the specific preprocessing performed in order to fit a particular semantic interpretation of the signed network.

#### 9.3.2.2 MFC Diffusion Model

The signs associated with diffusion links denote the "positive" and "negative" relationships, e.g., trust and distrust, among users. In everyday life, people tend to believe information from people they trust and not believe the information from those they distrust. For example, if someone we trust says that "Hillary Clinton will be the new president," we believe it to be true. However, if someone we distrust says the same thing, we might not believe it. In addition, when receiving contradictory messages, information obtained from the trusted people is usually given higher weights. In other words, the effects of trust and distrust diffusion links are asymmetric in activating users. For instance, when various actors assert that "Hillary Clinton will be the new president," we may tend to follow those we trust, even though the distrusted ones also say it. In addition, if someone we distrust says that "Hillary Clinton will be the new president," we may think it to be false and will not believe it. However, after being activated to distrust it, if we are exposed to contradictory information from a trusted party, we might be willing to change our minds. To model such cases, which are unique to signed and state-centric networks, a number of basic principles are introduced in the MFC model, (1) the effects of positive links in activating users is boosted to give them higher weights in activating users, and (2) users who are activated already will stay active in the subsequential rounds but their activation states can be flipped to follow the people they trust.

In MFC, users have 3 unique known states in the information diffusion process: $\{+1, -1, 0\}$ (i.e., trust, distrust, and inactive, respectively). Users with unknown states are automatically taken into account during the model construction process by assuming states as necessary. For simplicity, we use $s(\cdot)$ to represent both the sign of links and the states of users. If user $u$ trusts the information, then user $u$ is said to have a positive state $s(u) = +1$ towards the information. The initial states of all users in MFC are assigned a value of 0 (i.e., inactive to the information). A set of information seed users $\mathcal{I} \subseteq \mathcal{V}$ activated by the information at the very beginning will have their own attitudes towards the information based on their judgments, which can be represented with $\mathcal{S} = \{+1, -1\}^{|\mathcal{I}|}$. Information seed users in $\mathcal{I}$ spread the information to other users in signed networks step by step. At step $\tau$, user $u$ (activated at $\tau - 1$) is given only one chance to activate (1) inactive neighbor $v$, as well as (2) active neighbor $v$ but $v$ has different state from $u$ and $v$ trusts $u$, with the boosted success probability $\overline{w_D}(u, v)$, where $\overline{w_D}(v, u) \in [0, 1]$ can be represented as

$$\overline{w_D}(v, u) = \begin{cases} \min\{\alpha \cdot w_D(v, u), 1\} & \text{if } s_D(v, u) = +1, \\ w_D(v, u), & \text{otherwise.} \end{cases} \tag{9.24}$$

In the above equation, parameter $\alpha > 1$ denotes the boosting of information from $u$ to $v$ and is called the *asymmetric boosting coefficient*.

If $u$ succeeds, $v$ will become active in step $\tau + 1$, whose states can be represented as $s(v) = s(u) \cdot s(u, v)$. For example, if user $u$ thinks the information to be real (i.e., $s(u) = +1$) and $v$ trusts $u$ (i.e., $s(u, v) = +1$), once $v$ get activated by $u$ successfully, the state of $v$ will be $s(v) = +1$ (i.e., believe the information to be true). Otherwise, $v$ will keep its original state (either inactive or activated) and $u$ cannot make any further attempts to activate $v$ in subsequent rounds. All activated users will stay active in the following rounds and the process continues until no more activations are possible.

*Example 9.3* MFC can model the information diffusion process in signed social networks much better than traditional diffusion models, such as IC. To illustrate the advantages of MFC, we also give an example in Fig. 9.4, where two different cases: "simultaneous activation" (i.e., the left two plots) and "sequential activation" (i.e., the right two plots) are shown. In the "simultaneous activation" case,

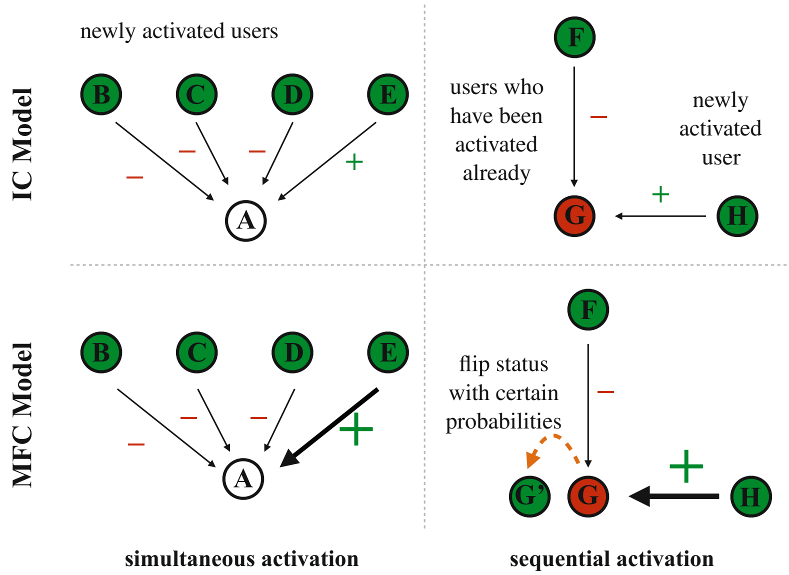


**Fig. 9.4** Example of the binary tree transformation

multiple users ($B$, $C$, $D$, and $E$) are all just activated at step $\tau$, who all think information to be true and at step $\tau + 1$, $B$-$E$ will activate their inactive neighbor $A$. Among these users, $A$ trusts $E$ and distrusts the remaining users. In traditional IC models, signs on links are ignored and $B$-$E$ are given equal chance to activate $A$ in random order with activation probabilities $w_D(\cdot, A), \cdot \in \{B, C, D, E\}$. However, in the MFC model, signs of links are utilized and the activation probability of positive diffusion $(E, A)$ will be boosted and can be represented as $\min\{\alpha \cdot w_D(E, A), 1\}$. As a result, user $A$ is more likely to be activated by $E$ in MFC. Meanwhile, in the sequential activation case, once a user (e.g., $F$) succeeds in activating $G$, $G$ will remain active and other users (e.g., H) cannot reactivate $A$ any longer in traditional IC model. However, in the MFC model, we allow users to flip their activation state by people they trust. For example, if $G$ has been activated by $F$ with state $s(G) = -1$ already, the trusted user $H$ can still have the chance to flip $G$'s state with probability $\min\{\alpha \cdot w_D(H, G), 1\}$.

The pseudo-code of the MFC diffusion model is provided in Algorithm 1.

## 9.4 Inter-Network Information Diffusion

The information diffusion models introduced in the previous sections are mostly based on one single network, assuming that information will propagate within the network only. However, in the real-world, users are involved in multiple social sites simultaneously, and cross-platform information diffusion is happening all the time. Users, as the bridges, can receive information from one social sites, and share with their friends in another network intentionally. Meanwhile, due to some social network settings, sometimes the activities happening in one social site (e.g., Foursquare) can be reposted to other social sites (e.g., Twitter) automatically, if the users login Foursquare with their Twitter account.

Cross-network information sharing and reposting renders the inter-network information diffusion ubiquitous and very common in the real-world online social networks. By involving in multiple online social networks simultaneously, users can also be exposed to more information from multiple social sites at the same time. In this section, we will study the information diffusion across multiple social platforms, and introduce two inter-network diffusion models across multiple online social

**Algorithm 1** MFC Information Diffusion Model

**Require:** input rumor initiators $\mathcal{I}$ with states $\mathcal{S}$
diffusion network $G_D = (\mathcal{V}_D, \mathcal{E}_D, s_D, w_D)$
**Ensure:** infected diffusion network $G_I$
initialize infected user set $\mathcal{U} = \mathcal{I}$, state set $\mathcal{S}_{\mathcal{U}} = \mathcal{S}$
let recently infected user set $\mathcal{R} = \mathcal{I}$
**while** $\mathcal{R} \neq \emptyset$ **do**
  new recently infected user set $\mathcal{N} = \emptyset$
  **for** $u \in \mathcal{R}$ **do**
    let the set of users that $u$ can activate to be $\Gamma(u)$
    **for** $v \in \Gamma(u)$ **do**
      **if** $s(v) = 0$ or $\left(s_D(u, v) = +1 \text{ and } s(u) \neq s(v)\right)$ **then**
        **if** $s_D(u, v) = +1$ **then**
          $p = \min\{1.0, \alpha \cdot w_D(u, v)\}$
        **else**
          $p = w_D(u, v)$
        **end if**
        **if** u activates $v$ with probability $p$ **then**
          $\mathcal{U} = \mathcal{U} \cup \{v\}$, $\mathcal{S}_{\mathcal{U}} = \mathcal{S}_{\mathcal{U}} \cup \{s(v) = s(u) \cdot s_D(u, v)\}$
          $\mathcal{N} = \mathcal{N} \cup \{v\}$
        **end if**
      **end if**
    **end for**
  **end for**
  $\mathcal{R} = \mathcal{N}$
**end while**
extract infected diffusion network $G_I$ consisting of infected users $\mathcal{U}$

networks [41, 49]. Generally, different social network platforms will create different information diffusion sources, and the interactions available among users in each of the sources can all propagate information among users.

### 9.4.1 Network Coupling Based Cross-Network Information Diffusion

In the online social world, once a user has been activated in one of the social sites, the user account owner will receive the information and diffuse it to other users in the other networks. The network coupling model to be introduced in this part proposes to combine multiple social networks together, and treat the information diffusion in each of the networks independently. If one of a user's account has been activated in a network, the user will be treated as activated. To simplify the activation checking criterion, the *network coupling based information diffusion model* [49] also introduces a relaxed criterion.

#### 9.4.1.1 Single Network Diffusion Model

Formally, let $G^{(1)}, G^{(2)}, \ldots, G^{(k)}$ denote the $k$ online social networks that we are focusing on in the information diffusion, whose network structures are all homogeneous involving users and friendship links only. For each of the network, e.g., $G^{(i)}$, we can represent its structure as $G^{(i)} = (\mathcal{V}^{(i)}, \mathcal{E}^{(i)})$, where $\mathcal{V}^{(i)}$ denotes the set of users in the network. Information diffusion process in network $G^{(i)}$ can be modeled with some existing models. In this part, we will use the LT model as the base diffusion model for each of the networks.

Based on network $G^{(i)}$, each user $u$ in the network is associated with a threshold $\theta_u^{(i)}$ indicating the minimal amount of required information to activate the users. Meanwhile, the amount of information sent between the users (e.g., $u$ and $v$) can be denoted as weight $w_{u,v}^{(i)}$, whose value can be determined in the same way as the LT model introduced before. For an inactive user $u$, he/she can be activated iff the amount of information propagated from their friends to him/her is greater than $u$'s threshold, i.e.,

$$\sum_{v \in \Gamma(u; G^{(i)})} \mathbb{I}(v, t) \cdot w_{v,u}^{(i)} \geq \theta_u^{(i)}, \tag{9.25}$$

where $\Gamma(u; G^{(i)})$ represents the neighbors of user $u$ and $\mathbb{I}(v, t)$ indicates whether $v$ has been activated or not at time $t$.

#### 9.4.1.2 Network Coupling Based Information Diffusion Model

Meanwhile, among these $k$ different online social sites $G^{(1)}, G^{(2)}, \ldots, G^{(k)}$, if there exists one network $G^{(i)}$, in which the above equation holds, user $u$ will become activated by the information. In other words, to determine whether user $u$ has been activated or not, we need to check his/her status in all these $k$ networks one by one as follows:

$$\begin{aligned}
\text{Network } G^{(1)} &: \sum_{v \in \Gamma(u; G^{(1)})} \mathbb{I}(v, t) \cdot w_{v,u}^{(1)} \geq \theta_u^{(1)}, \\
\text{Network } G^{(2)} &: \sum_{v \in \Gamma(u; G^{(2)})} \mathbb{I}(v, t) \cdot w_{v,u}^{(2)} \geq \theta_u^{(2)}, \\
&\cdots, \\
\text{Network } G^{(k)} &: \sum_{v \in \Gamma(u; G^{(k)})} \mathbb{I}(v, t) \cdot w_{v,u}^{(k)} \geq \theta_u^{(k)}.
\end{aligned} \tag{9.26}$$

Such a status checking process can be very time-consuming. To reduce the activation checking works, in the lossy network coupling scheme, the users' activation checking criterion is relaxed to

$$\sum_{i=1}^{k} \alpha^{(i)} \cdot \sum_{v \in \Gamma(u)} \mathbb{I}(v, t) \cdot w_{v,u}^{(i)} \geq \sum_{i=1}^{k} \alpha^{(i)} \cdot \theta_u^{(i)}, \tag{9.27}$$

where $\alpha^{(1)}, \alpha^{(2)}, \ldots, \alpha^{(k)} > 0$ denote the weight parameters representing the importance of different networks in user activation.

**Theorem 9.1** *Given the k networks, $G^{(1)}, G^{(2)}, \ldots, G^{(k)}$, if equation*

$$\sum_{i=1}^{k} \alpha^{(i)} \cdot \sum_{v \in \Gamma(u)} \mathbb{I}(v, t) \cdot w_{v,u}^{(i)} \geq \sum_{i=1}^{k} \alpha^{(i)} \cdot \theta_u^{(i)}, \tag{9.28}$$

*holds, user u will be activated.*

*Proof* The theorem can be proven with by contradiction. Let's assume the equation holds but $u$ has not been activated in networks $G^{(1)}, G^{(2)}, \dots, G^{(k)}$, then we have

$$\sum_{v \in \Gamma(u)} \mathbb{I}(v, t) \cdot w_{v,u}^{(i)} < \theta_u^{(i)}, \tag{9.29}$$

hold for all these networks.

By multiplying both sides of the inequality with a positive weight $\alpha^{(i)}$, and sum the equations across all these $k$ networks, we have

$$\sum_{i=1}^{k} \alpha^{(i)} \cdot \sum_{v \in \Gamma(u)} \mathbb{I}(v, t) \cdot w_{v,u}^{(i)} < \sum_{i=1}^{k} \alpha^{(i)} \cdot \theta_u^{(i)}, \tag{9.30}$$

which contradicts the equation in the theorem.

Therefore, if the new activation criterion holds, user $u$ will be activated (in at least one of these online social networks).

The relaxed activation criterion is actually a sufficient but not necessary condition when determining whether $u$ is activated or not. In some cases, $u$ has already been activated in some of the networks, but the criterion cannot meet, which will lead to some latency in status checking. One way to solve the problem is to assign an appropriate weight $\alpha^{(i)}$ by increasing the value proportion of these networks in the summation equation $\sum_{v \in \Gamma(u)} \mathbb{I}(v, t) \cdot w_{v,u}^{(i)}$. In the special case that user $u$ can be activated in network $G^{(i)}$ already, we can assign the weight $\alpha^{(i)}$ with a very large value, where $\alpha^{(i)} \gg \alpha^{(j)}, j \in \{1, 2, \dots, k\}, j \neq i$ is way larger compared with the remaining networks. So far, there don't exist any methods to adjust the parameters automatically in the *network coupling based information diffusion model*, and heuristics are applied in most of the cases.

### 9.4.2 Random Walk Based Cross-Network Information Diffusion

Different online social networks usually have their own characteristics, and users tend to have different status regarding the same information in different platforms. For instance, information about personal entertainments (like movies, pop stars) can be widely spread among users in Facebook, and users who are interested in them will be activated very easily and also share the information to their friends. However, such a kind of information is relatively rare in the professional social network LinkedIn, where people seldom share personal entertainment to their colleagues, even though they may have been activated already in Facebook. What's more, the structures of these online social networks are usually heterogeneous, containing various types of connections. Besides the direct follow relationships among the users, these diverse connections available among the users may create different types of communication channels for information diffusion. To model such an observation in information diffusion across multiple heterogeneous online social sites, in this part, we will introduce a new information diffusion model, IPATH [41], based on random walk. Since there exist multiple networks here and users will also have different status in different networks, we can denote the *objective target network* that we study as $G^t$ and the external aligned social platform as $G^s$. Between them, there exists a set of anchor links connecting the shared anchor users, which can be denoted as $\mathcal{A}^{(t,s)}$. For the scenarios with more than one external source network, a simple extension to the following model will be applicable to depict the diffusion process across these networks.

#### 9.4.2.1 Intra-Network Propagation

The traditional research works on homogeneous networks assume that information can only be spread by the social links among users. If user $v$ follows user $u$, i.e., $(v, u) \in \mathcal{E}$, the message can spread from $u$ to $v$, i.e. $u \rightarrow v$. However in a heterogeneous social network, the multi-typed and interconnected entities can create various information propagation channels among the users. For instance, if user $u$ recommends a good restaurant to his friend $v$ by checking in at this place, information will flow from $u$ to $v$ through the location entity $l$, which can be expressed by $u \xrightarrow[l]{check-in} v$. Similarly, we can represent the information diffusion routes among users via other information entities, which can be formally represented as the diffusion route set $\mathcal{R} = \{r_1, r_2, \ldots, r_m\}$, where $m$ is the total route number.

According to each diffusion route, we can represent the connections among users as an adjacency matrix actually. We can take the source network $G^s = (\mathcal{V}^s, \mathcal{E}^s)$ as an example. For any relation $r_i \in \mathcal{R}^s$, the adjacency matrix defined based on $r_i$ among the set of users (i.e., $\mathcal{U}^s$) can be represented as $\mathbf{A}_i^s \in \mathbb{R}^{|\mathcal{U}^s| \times |\mathcal{U}^s|}$, where $A_i^s(u, v)$ is a binary-value variable and $A_i^s(u, v) = 1$ iff $u$ and $v$ are connected with each other via relation $r_i$. The weighted diffusion matrix can be represented as the normalized matrix of $\mathbf{A}_i^s$, i.e., $\mathbf{W}_i^s = \mathbf{A}_i^s \mathbf{D}^{-1}$, where $\mathbf{D}$ is a diagonal matrix with $D(u, u) = \sum_{v \in \mathcal{U}^s} A_i^s(v, u)$denoting the in-degree of $u$ on its diagonal. The entry $W_i^s(u, v)$ denotes the probability of going from $v$ to $u$ in one step. In a similar way, we can represent the weighted diffusion matrices for other relations, which altogether can be represented as $\{\mathbf{W}_1^s, \mathbf{W}_2^s, \ldots, \mathbf{W}_m^s\}$. To fuse the information diffused from different relations, iPath uses linearly combination to integrate these weighted matrices as follows:

$$\mathbf{W}^s = \lambda_1 \times \mathbf{W}_1^s + \lambda_2 \times \mathbf{W}_2^s + \cdots + \lambda_m \times \mathbf{W}_m^s, \tag{9.31}$$

where $\lambda_i$ denotes the aggregation weight of matrix corresponding to relation $r_i$. In real scenarios, different relations play different roles in the information propagation for different users. However, to simplify the settings, in this part, we treat all these relations to be equally important, and the aggregated matrix $\mathbf{W}^s$ takes the average of all these weighted diffusion matrices. In a similar way, we can define the weight matrix $\mathbf{W}^t$ for the target network $G^t$.

#### 9.4.2.2 Inter-Network Propagation

Across the aligned networks, information can propagate not only within networks but also across networks. Based on the known anchor links between networks $G^t$ and $G^s$, we can define the binary adjacency matrix $\mathbf{A}^{s \rightarrow t} \in \mathbb{R}^{|\mathcal{U}^s| \times |\mathcal{U}^t|}$, where $A^{s \rightarrow t}(u, v) = 1$ if $(u^s, v^t) \in \mathcal{A}^{(t,s)}$. In iPath, we assume that each anchor user in $G^s$ only has one corresponding account in $G^t$. Therefore $A^{s \rightarrow t}$ is already normalized (into binary) and the weight matrix $\mathbf{W}^{s \rightarrow t} = \mathbf{A}^{s \rightarrow t}$, denoting the chance of information propagating from $G^s$ to $G^t$. Furthermore, we can represent the weighted diffusion matrix from networks $G^s$ to $G^t$ as $\mathbf{W}^{t \rightarrow s} = (\mathbf{W}^{s \rightarrow t})^\top$, considering that the anchor links are undirected.

Both the intra-network propagation relations, represented by weight matrices $\mathbf{W}^s$ and $\mathbf{W}^t$ in networks $G^s$ and $G^t$, respectively, and the inter-network propagation relations, represented by weight matrix $\mathbf{W}^{s \rightarrow t}$ and $\mathbf{W}^{t \rightarrow s}$, have been constructed already in the previous subsection. As shown in Fig. 9.5, to model the cross-network information diffusion process involving both the intra- and inter-network relations simultaneously, iPath proposes to combine these weighted diffusion matrices to build an integrated matrix $\mathbf{W} \in \mathbb{R}^{(|\mathcal{U}^s|+|\mathcal{U}^t|)^2}$. In the integrated matrix $\mathbf{W}$, the parameter $\alpha \in [0, 1]$ denotes the probability that the message stay in the original network, thus $1 - \alpha$ represents the chance of being transmitted across networks (i.e., the probability of activated anchor user passing the influence to the target network). In real scenarios, the probabilities for different users to repost

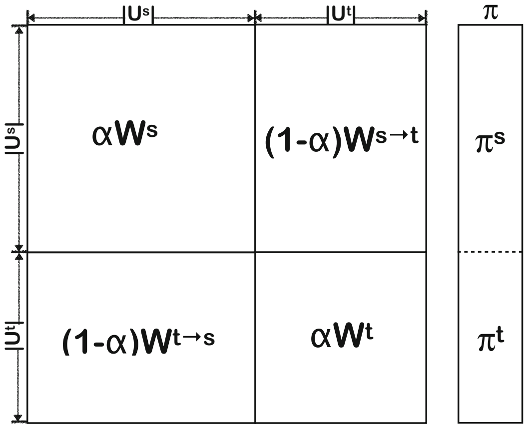


**Fig. 9.5** The weight matrix and the information distribution vector

information across aligned networks can be quite diverse. However, to simplify the problem setting, in IPATH, we will unify these probabilities with parameter $\alpha$.

Let vector $\pi^{(k)} \in \mathbb{R}^{(|\mathcal{U}^s|+|\mathcal{U}^t|)}$ represent the information that users in $G^s$ and $G^t$ can receive after $k$ steps. As shown in Fig. 9.5, vector $\pi^{(k)}$ consists of two parts $\pi^{(k)} = [\pi^{s,(k)}, \pi^{t,(k)}]$, where $\pi^{s,(k)} \in \mathbb{R}^{|\mathcal{U}^s|}$ and $\pi^{t,(k)} \in \mathbb{R}^{|\mathcal{U}^t|}$. The initial state of the vector can be denoted as $\pi^{(0)}$, which is defined based on the seed user set $\mathcal{Z}$ with function $g(\cdot)$ as follows:

$$\pi^{(0)} = g(\mathcal{Z}), \text{ where } \pi^{(0)}(u) \quad = \begin{cases} 1 & \text{if } u \in \mathcal{Z}, \\ 0 & \text{otherwise.} \end{cases} \tag{9.32}$$

Seed set $\mathcal{Z}$ can also be represented as $\mathcal{Z} = g^{-1}(\pi^{(0)})$. Users from $G^s$ and $G^t$ both have the chance of being selected as seeds, but when the structure information of $G^t$ is hard to obtain, the seed users will be only chosen from $G^s$. In IPATH, the information diffusion process is modeled by *random walk*, because it is widely used in which the total probability of the diffusing through different relations remains constant 1 [8,32]. Therefore in the information propagation process, vector $\pi$ will be updated stepwise with the following equation:

$$\pi^{(k+1)} = (1-a) \times \mathbf{W}\pi^{(k)} + a \times \pi^{(0)}, \tag{9.33}$$

where constant $a$ denotes the probability of returning to the initial state. By keeping updating $\pi$ according to (9.33) until convergence, we can present the stationary state of vector $\pi$ to be $\pi^*$,

$$\pi^* = a[\mathbf{I} - (1-a)\mathbf{W}]^{-1}\pi^{(0)}, \tag{9.34}$$

where matrix $\mathbf{I} \in \{0, 1\}^{(|\mathcal{U}^s|+|\mathcal{U}^t|)\times(|\mathcal{U}^s|+|\mathcal{U}^t|)}$ is an identity matrix. The value of entry $\pi^*[u]$ denotes the activation probability of $u$, and user $u$ will be activated if $\pi^*[u] \geq \theta$, where $\theta$ denotes the threshold of accepting the message. In IPATH, parameter $\theta$ is randomly sampled from range $[0, \theta_{bound}]$. The threshold bound $\theta_{bound}$ is a small constant value, as the amount of information each user can get at the stationary state in IPATH can be very small. In addition, we can further represent the activation status of user $u$ as vector $\pi'$, where

$$\pi'[u] = \begin{cases} 1 & \text{if } \pi^*[u] \geq \theta, \\ 0 & \text{otherwise.} \end{cases} \tag{9.35}$$

In Eq. (9.35), $\pi'[u] = 1$ denotes that user $u$ is activated. In practice, the value of $\pi^*[u]$ is usually in $[0, 1]$ when the networks are sparse and the size of the seed set is small, and it can be represented approximately as following:

$$\pi'[u] \approx \lfloor \pi^*[u] - \theta + 1 \rfloor. \tag{9.36}$$

Based on this, we define the mapping function $h$ between two vectors, where the floor function is applied to each element in the vector.

$$\pi' = h(\pi^*) = \lfloor \pi^* + \mathbf{c} \rfloor. \tag{9.37}$$

Here, $\mathbf{c}$ is a constant vector where each entry equals to $1 - \theta$.

## 9.5 Information Diffusion Across Online and Offline World

Besides the online world, information can also propagate within the offline word as well as between the online and offline world. In this section, we will use the workplace as one example to illustrate the information diffusion across both the online and offline world simultaneously.

### 9.5.1 Background Knowledge

On average, people nowadays need to spend more than 30% of their time at work everyday. According to the statistical data in [16], the total amount of time people spent at workplace in their life is tremendously large. For instance, a young man who is 20 years old now will spend 19.1% of his future time working [16]. Therefore, workplace is actually an easily neglected yet important social occasion for effective communication and information exchange among people in our social life.

Besides the traditional offline contacts, like face-to-face communication, telephone calls and messaging, to facilitate the cooperation and communications among employees, a new type of online social networks named enterprise social networks (ESNs) has been launched inside the firewalls of many companies [44, 45]. A representative example is Yammer, which is used by over 500,000 leading businesses around the world, including 85% of the Fortune 500.[1] Yammer provides various online communication services for employees at workplace, which include instant online messaging, write/reply/like posts, file upload/download/share, etc. In summary, the communication means existing among employees at workplaces are so diverse, which can generally be divided into two categories [33]: (1) offline communication means, and (2) online virtual communication means.

In this section, we will study how information diffuses via both online and offline communication means among employees at workplace, which is formally defined as the "Information Diffusion in Enterprise" (IDE) problem [47].

*Example 9.4* To help illustrate the IDE problem more clearly, we also give an example in Fig. 9.6. The left plot of Fig. 9.6 is about an online ESN, employees in which can perform various social activities. For instances, employees can follow each other, can write/reply/like posts online, and posts written by them can also @certain employees to send notifications, which create various online information diffusion channels (i.e., the green lines) among employees. Meanwhile, the relative management

[1] https://about.yammer.com/why-yammer/.

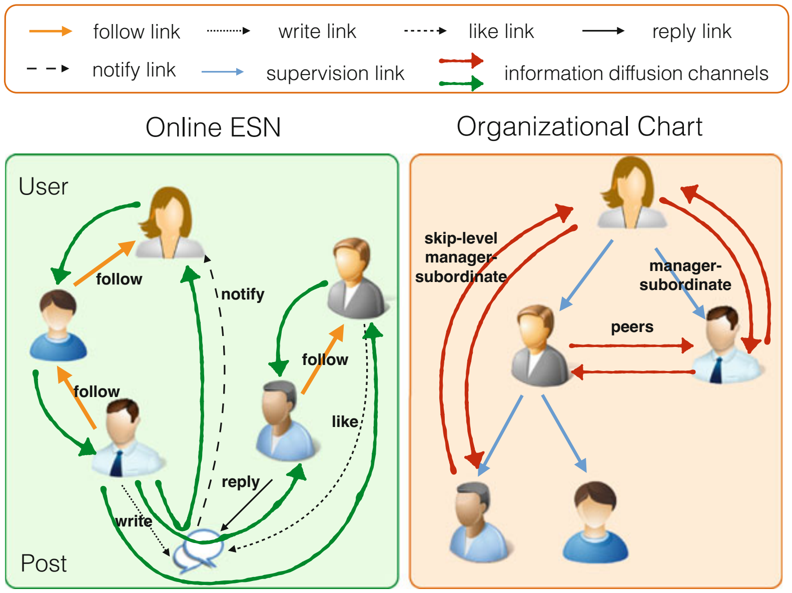


**Fig. 9.6** An example of information diffusion at workplace

relationships among the employees in the company can be represented with the organizational chart (i.e., the right plot), which is a tree-structure diagram connecting employees via supervision links (from managers to subordinates). Colleagues who are physically close in the organizational chart (e.g., peers, manager-subordinates) may have more chance to meet in the offline workplace. For example, subordinates need to report to their managers regularly, peers may co-operate to finish projects together, which can form various offline information diffusion channels (i.e., the red lines) among employees at workplace.

**Definition 9.8 (Enterprise Social Networks (ESNs))** Online *enterprise social networks* are a new type of online social networks used in enterprises to facilitate employees' communications and daily work, which can be represented as *heterogeneous information networks* $G = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \bigcup_i \mathcal{V}_i$ is the set of different kinds of nodes and $\mathcal{E} = \bigcup_j \mathcal{E}_j$ is the union of complex links in the network.

In this section, we will use Yammer as an example of online ESNs. Yammer can be represented as $G = (\mathcal{V}, \mathcal{E})$, where node set $\mathcal{V} = \mathcal{U} \cup \mathcal{O} \cup \mathcal{P}$ and $\mathcal{U}$, $\mathcal{O}$, and $\mathcal{P}$ are the sets of users, groups, and posts, respectively; link set $\mathcal{E} = \mathcal{E}_s \cup \mathcal{E}_j \cup \mathcal{E}_w \cup \mathcal{E}_r \cup \mathcal{E}_l$ denoting the union of social, group membership, write, reply, and like links in Yammer, respectively. At the workplace, information of various topics can propagate among the employees simultaneously.

**Definition 9.9 (Organizational Chart)** *Organizational chart* is a diagram outlining the structure of an organization as well as the relative ranks of employees' positions and jobs, which can be represented as a rooted tree $C = (\mathcal{N}, \mathcal{L}, root)$, where $\mathcal{N}$ denotes the set of employees and $\mathcal{L}$ is the set of directed *supervision links* from managers to subordinates in the company, *root* usually represents the CEO by default.

Each employee in the company can create exactly one account in Yammer with valid employment ID, i.e., there is *one-to-one* correspondence between the users in Yammer and employees in the organization chart. For simplicity, in this section, we assume the user set in online ESN to be identical to the employee set in the organizational chart (i.e., $\mathcal{U} = \mathcal{N}$) and we will use "Employee" to denote individuals in both online ESN and offline organizational chart by default.

To address all the above challenges, a novel information diffusion model **MUSE** (Multi-source Multi-channel Multi-topic diffUsion SElection) proposed in [47] will be introduced in this section. **MUSE** extracts and infers sets of online, offline, and hybrid (of online and offline) diffusion channels among employees across online ESN and offline organizational structure. Information propagated via different channels can be aggregated effectively in **MUSE**. Different diffusion channels will be weighted according to their importance learned from the social activity log data with optimization techniques and top-K effective diffusion channels will be selected in **MUSE** finally.

### 9.5.2 Preliminary

Before we talk about the detailed components involved in **MUSE**, we will introduce the general framework of the **MUSE** model in this part. Formally, we denote the set of topics diffusing in the workplace as set $\mathcal{T}$. Three different diffusion sources will be our main focus in this section: online source, offline source, and the hybrid source (across online and offline sources). The diffusion channel set of all these three sources can be represented as $\mathcal{C}^{(on)}$, $\mathcal{C}^{(off)}$ and $\mathcal{C}^{(hyb)}$, respectively, whose sizes are $\left|\mathcal{C}^{(on)}\right| = k^{(on)}$, $\left|\mathcal{C}^{(off)}\right| = k^{(off)}$, $\left|\mathcal{C}^{(hyb)}\right| = k^{(hyb)}$.

In **MUSE**, a set of users are activated initially, whose information will propagate in discrete steps within the network to other users. Let $v$ be an employee at workplace who has been activated by topic $t \in \mathcal{T}$. For instance, at step $\tau$, $v$ will send an amount of $w^{(on),i}(v, u, t)$ information on topic $t$ to $u$ via the $i_{th}$ channel in the online source (i.e., channel $c^{(on),i} \in \mathcal{C}^{(on)}$), where $u$ is an employee following $v$ in channel $c^{(on),i}$. The amount of information that $u$ receives from $v$ via all the channels in the online source at step $\tau$ can be represented as vector $\mathbf{w}^{(on)}(v, u, t) = [w^{(on),1}(v, u, t), w^{(on),2}(v, u, t), \ldots, w^{(on),k^{(on)}}(v, u, t)]$. Similarly, we can also represent the vectors of information $u$ receives from $v$ through channels in offline source and hybrid source as vectors $\mathbf{w}^{(off)}(v, u, t)$ and $\mathbf{w}^{(hyb)}(v, u, t)$, respectively.

Meanwhile, users in **MUSE** are associated thresholds to different topics, which are selected at random from the uniform distribution in range [0, 1]. Employee $u$ can get activated by topic $t$ if the information received from his active neighbors via diffusion channels of all these three sources can exceed his *activation threshold* $\theta(u, t)$ to topic $t$:

$$f\left(\mathbf{w}^{(on)}(\cdot, u, t), \mathbf{w}^{(off)}(\cdot, u, t), \mathbf{w}^{(hyb)}(\cdot, u, t)\right) \geq \theta(u, t), \tag{9.38}$$

where aggregation function $f(\cdot)$ maps the information $u$ receives from all the channels to $u$'s *activation probability* in range [0, 1]. Here, the vector $\mathbf{w}^{(on)}(\cdot, u, t) = [w^{(on),1}(\cdot, u, t), w^{(on),2}(\cdot, u, t), \ldots, w^{(on),k^{(on)}}(\cdot, u, t)]$, where $w^{(on),i}(\cdot, u, t)$ denotes the information received from all the employees $u$ follows in channel $c^{(on),i}$, i.e.,

$$w^{(on),i}(\cdot, u, t) = \sum_{v \in \Gamma_{out}^{(on),i}(u)} w^{(on),i}(v, u, t). \tag{9.39}$$

Vectors $\mathbf{w}^{(off)}(\cdot, u, t)$ and $\mathbf{w}^{(hyb)}(\cdot, u, t)$ can be represented in a similar way. Once being activated, a user will stay active in the remaining rounds and each user can be activated at most once. Such a process will end if no new activations are possible.

Considering that individuals' *activation thresholds* $\theta(u, t)$ to topic $t$ is pre-determined by the uniform distribution, next we will focus on studying the information received via channels of the *online*, *offline*, and *hybrid* sources and the *aggregation function* $f(\cdot)$ in detail.

### 9.5.3 Online Diffusion Channel

Online ESNs provide various communication tools for employees to contact each other, where individuals who have no social connections can still pass information via many other connections. Each connection among employees can form an information diffusion channel in online ESN. In this section, we will introduce the various diffusion channels among employees extracted based on a set of *online social meta paths* [30] from the heterogeneous information in the online ESN. Before that, we first introduce the schema of enterprise social network as follows.

Based on an online ESN $G = (\mathcal{V}, \mathcal{E})$, we can represent its network schema as $S_G = (\mathcal{T}_G, \mathcal{R}_G)$, where $\mathcal{T}_G$ and $\mathcal{R}_G$ represent the sets of node types and link types in network $G$, respectively. For example, for the Yammer network, we can define its schema as $S_G$, where the node type set $\mathcal{T}_G = \{Employee, Post\}$, link type set $\mathcal{R}_G = \{Social^{1/-1}, Write^{1/-1}, Reply^{1/-1}, Like^{1/-1}, Notify^{1/-1}\}$, and the superscript $-1$ denotes the reverse of the corresponding link type in online ESN.

In enterprise social networks, individuals can (1) get information from employees they follow (i.e., their followees) and (2) people that their "followees" follow (i.e., 2nd level followees), and obtain information from employees by (3) viewing and replying their posts, (4) viewing and liking their posts, as well as (5) getting notified by their posts (i.e., explicitly @ certain users in posts). MUSE extracts 5 different *online social meta paths* from the online ESN, whose physical meanings, representations, and abbreviated notations are listed as follows:

- Followee: $Employee \xleftarrow{Social^{-1}} Employee$, whose notation is $\Phi_1$.
- Followee-Followee: $Employee \xleftarrow{Social^{-1}} Employee \xleftarrow{Social^{-1}} Employee$, whose notation is $\Phi_2$.
- Reply Post: $Employee \xleftarrow{Reply^{-1}} Post \xleftarrow{Write} Employee$, whose notation is $\Phi_3$.
- Like Post: $Employee \xleftarrow{Like^{-1}} Post \xleftarrow{Write} Employee$, whose notation is $\Phi_4$.
- Post Notification: $Employee \xleftarrow{Notify} Post \xleftarrow{Write} Employee$, whose notation is $\Phi_5$.

The direction of the links denotes the information diffusion direction and end of the diffusion links (i.e., the first employee of the above paths) represents the target employee to receive the information. For example, $\Phi_1$ denotes the target user receives information from his followees, while $\Phi_5$ means that the target employee receives information from employees who have ever written posts @ the target employee.

Each of the above *online social meta path* defines an information diffusion channel among individuals in online ESN. As a result, in this section, $\mathcal{C}^{(on)} = \{\Phi_1, \Phi_2, \Phi_3, \Phi_4, \Phi_5\}$ and $k^{(on)} = 5$ and $\Phi_i$ is identical to $c^{(on),i}$ mentioned before (denoting the $i_{th}$ online diffusion channel). Based on each of these online social meta paths, we can extract the corresponding path instances connecting employees $u$ and $v$ (i.e., the concrete information diffusion traces from $v$ to $u$), which can be represented as set $\mathcal{P}^{(on)}_{\Phi_i}(v \rightsquigarrow u)$, for $\forall \Phi_i \in \mathcal{C}^{(on)}$. Furthermore, let $\mathcal{P}^{(on)}_{\Phi_i}(v \rightsquigarrow \cdot)$ and $\mathcal{P}^{(on)}_{\Phi_i}(\cdot \rightsquigarrow u)$ be the sets of path instances of $\Phi_i$ going out from $v$ and going into $u$, respectively, with which we can define the amount of information propagating from $v$ to $u$ via diffusion channel $c^{(on),i} = \Phi_i$ to be

$$w^{(on),i}(v,u,t) = \frac{2\left|\mathcal{P}^{(on)}_{\Phi_i}(v \rightsquigarrow u)\right| \cdot I(v,t)}{\left|\mathcal{P}^{(on)}_{\Phi_i}(v \rightsquigarrow \cdot)\right| + \left|\mathcal{P}^{(on)}_{\Phi_i}(\cdot \rightsquigarrow u)\right|}, \tag{9.40}$$

where binary function $I(v,t) = 1$ if $v$ has been activated by topic $t$ and 0 otherwise.

In the above definition, the proportion of information propagated from $v$ to $u$ via the communication channels (i.e., $w^{(on),i}(v,u,t)$) can denote how close these two users are, which depends on (1) the number of concrete diffusion path instances between them (i.e., $\mathcal{P}^{(on)}_{\Phi_i}(v \rightsquigarrow u)$); (2) the out-degree in the channel from $v$ (i.e., $\mathcal{P}^{(on)}_{\Phi_i}(v \rightsquigarrow \cdot)$); and (3) the in-degree in the channel to $u$ (i.e., $\mathcal{P}^{(on)}_{\Phi_i}(\cdot \rightsquigarrow u)$).

### 9.5.4 Offline Diffusion Channel

Employees' offline interactions are actually confidential to both companies and the public, which is very hard to know exactly. To infer the potential offline information diffusion channels at workplace, a set of potential information diffusion channels among individuals are extracted based on the organizational chart of the company. Similar to online enterprise social networks, we can define the schema of the organization chart as $S_C = (\mathcal{T}_C, \mathcal{R}_C)$, where $\mathcal{T}_C = \{Employee\}$ and $\mathcal{R}_C = \{Supervision^{1/-1}\}$. In offline workplace, the most common social interaction should happen between close colleagues, e.g., peers, manager-subordinate, and skip-level manager-subordinates, etc. The physical meaning and notations of offline social meta paths extracted in this section are listed as follows:

- Manager: $Employee \xleftarrow{Supervision} Employee$, whose notation is $\Omega_1$.
- Subordinate: $Employee \xleftarrow{Supervision^{-1}} Employee$, whose notation is $\Omega_2$.
- Peer: $Employee \xleftarrow{Supervision} Employee \xleftarrow{Supervision^{-1}} Employee$, whose notation is $\Omega_3$.
- 2nd-Level Manager: $Employee \xleftarrow{Supervision} Employee \xleftarrow{Supervision} Employee$, whose notation is $\Omega_4$.
- 2nd-Level Subordinate: $Employee \xleftarrow{Supervision^{-1}} Employee \xleftarrow{Supervision^{-1}} Employee$, whose notation is $\Omega_5$.

Similarly, the direction of links represents the information flow direction and the ending employees of the paths denotes the target employee, who receives information. For instance, meta path $\Omega_1$ means that the target employee receives information from his manager, while $\Omega_3$ denotes that the target employee receives information from his peers.

Each employee at the workplace can be influenced by both his manager and his subordinates (if exist) and to clarify the difference between these two different diffusion channels, we define both $\Omega_1$ and $\Omega_2$ (as well as $\Omega_4$ and $\Omega_5$). Based on the above introduced offline social meta paths, the offline diffusion channel can be represented as set $\mathcal{G}^{(off)} = \{\Omega_1, \Omega_2, \Omega_3, \Omega_4, \Omega_5\}$ and $k^{(off)} = 5$, where $\Omega_i$ denotes the $i_{th}$ offline diffusion channel among employees.

Based on offline social meta path, e.g., $\Omega_i$, the amount of information on topic $t$ propagating from employee $v$ to $u$ can be represented as

$$w^{(off),i}(v,u,t) = \frac{2\left|\mathcal{P}^{(off)}_{\Omega_i}(v \rightsquigarrow u)\right| \cdot I(v,t)}{\left|\mathcal{P}^{(off)}_{\Omega_i}(v \rightsquigarrow \cdot)\right| + \left|\mathcal{P}^{(off)}_{\Omega_i}(\cdot \rightsquigarrow u)\right|}, \tag{9.41}$$

where $\mathcal{P}^{(off)}_{\Omega_i}(v \rightsquigarrow u)$ denotes the offline social meta path instance set of $\Omega_i$ connecting $v$ to $u$ in the chart.

### 9.5.5 Hybrid Diffusion Channel

Besides the pure online/offline diffusion channels, information can also propagate across both online and offline worlds simultaneously. Consider, for example, two employees $v$ and $u$ who are not connected by any diffusion channels in online ESN or offline workplace, $v$ can still influence $u$ by activating $u$'s manager via online contacts and the manager will further propagate the influence to $v$ via offline interactions. To represent such a kind of diffusion channels, a set of *hybrid social meta paths* are also extracted in MUSE. As proposed in [16], every pair of people in the worlds can get connected via 6 hops (i.e., six degrees of separation theory). To avoid connecting all the employees by hybrid diffusion channels, we limit its length (i.e., the number of relations in the meta path) to 3 only. The set of hybrid social meta path used in this section, together with their physical meanings, notations are listed as follows:

- Followee-Manager: $Employee \xleftarrow{Social^{-1}} Employee \xleftarrow{Supervision} Employee$, whose notation is $\Psi_1$,
- Followee-Subordinate: $Employee \xleftarrow{Social^{-1}} Employee \xleftarrow{Supervision^{-1}} Employee$, whose notation is $\Psi_2$,
- Manager-Followee: $Employee \xleftarrow{Supervision} Employee \xleftarrow{Social^{-1}} Employee$, whose notation is $\Psi_3$,
- Subordinate-Followee: $Employee \xleftarrow{Supervision^{-1}} Employee \xleftarrow{Social^{-1}} Employee$, whose notation is $\Psi_4$,
- Followee-Peer: $Employee \xleftarrow{Social^{-1}} Employee \xleftarrow{Supervision} Employee \xleftarrow{Supervision^{-1}} Employee$, whose notation is $\Psi_5$,
- Peer-Followee: $Employee \xleftarrow{Supervision} Employee \xleftarrow{Supervision^{-1}} Employee \xleftarrow{Social^{-1}} Employee$, whose notation is $\Psi_6$.

where meta path, e.g., $\Psi_1$, denotes that the target employee receives information from his followee in online ESN, who gets information from his manager in the offline workplace. We can get $\mathcal{C}^{(hyb)} = \{\Psi_1, \Psi_2, \Psi_3, \Psi_4, \Psi_5, \Psi_6\}$ and $k^{(hyb)} = 6$. Based on each hybrid diffusion channel, e.g., $\Psi_i$, the amount of information on topic $t$ that $v$ sends to $u$ can be represented as

$$w^{(hyb),i}(v,u,t) = \frac{2\left|\mathcal{P}^{(hyb)}_{\Psi_i}(v \rightsquigarrow u)\right| \cdot I(v,t)}{\left|\mathcal{P}^{(hyb)}_{\Psi_i}(v \rightsquigarrow \cdot)\right| + \left|\mathcal{P}^{(hyb)}_{\Psi_i}(\cdot \rightsquigarrow u)\right|}. \tag{9.42}$$

### 9.5.6 Channel Aggregation

Different diffusion channels deliver various amounts of information among employees via the online communications in ESN and offline contacts. In this subsection, we will focus on aggregating information propagated via different channels with the information aggregation function $f(\cdot) : \mathbb{R}^{n\times 1} \to [0, 1]$, which can map the amount of information received by employees to their activation probabilities. Generally, any function that can map real number to probabilities in range [0, 1] can be applied and without loss of generality, we will use the logistic function $f(x) = \frac{e^x}{1+e^x}$ [6] in this section.

Based on the information on topic $t$ received by $u$ via the online, offline, and hybrid diffusion channels, we can represent $u$'s activation probability to be:

$$f\left(\mathbf{w}^{(on)}(\cdot, u, t), \mathbf{w}^{(off)}(\cdot, u, t), \mathbf{w}^{(hyb)}(\cdot, u, t)\right)$$
$$= \frac{e^{\left(g(\mathbf{w}^{(on)}(\cdot,u,t))+g(\mathbf{w}^{(off)}(\cdot,u,t))+g(\mathbf{w}^{(hyb)}(\cdot,u,t))+\theta_0\right)}}{1+e^{\left(g(\mathbf{w}^{(on)}(\cdot,u,t))+g(\mathbf{w}^{(off)}(\cdot,u,t))+g(\mathbf{w}^{(hyb)}(\cdot,u,t))+\theta_0\right)}}, \tag{9.43}$$

where function $g(\cdot)$ linearly combines the information in different channels belonging to certain sources and $\theta_0$ denotes the weight of the constant factor. Terms $g(\mathbf{w}^{(on)}(\cdot, u, t))$, $g(\mathbf{w}^{(off)}(\cdot, u, t))$, and $g(\mathbf{w}^{(hyb)}(\cdot, u, t))$ can be represented as follows:

$$g(\mathbf{w}^{(on)}(\cdot, u, t)) = \sum_{i=1}^{k^{(on)}} \alpha_i \cdot \sum_{v \in \Gamma_{out}^{(on),i}(u)} w^{(on),i}(v, u, t), \tag{9.44}$$

$$g(\mathbf{w}^{(off)}(\cdot, u, t)) = \sum_{i=1}^{k^{(off)}} \beta_i \cdot \sum_{v \in \Gamma_{out}^{(off),i}(u)} w^{(off),i}(v, u, t), \tag{9.45}$$

$$g(\mathbf{w}^{(hyb)}(\cdot, u, t)) = \sum_{i=1}^{k^{(hyb)}} \gamma_i \cdot \sum_{v \in \Gamma_{out}^{(hyb),i}(u)} w^{(hyb),i}(v, u, t), \tag{9.46}$$

where $\alpha_i, \beta_i, \gamma_i$ are the weights of different *online*, *offline*, and *hybrid* diffusion channels, respectively and $\sum_{i=1}^{k^{(on)}} \alpha_i + \sum_{i=1}^{k^{(off)}} \beta_i + \sum_{i=1}^{k^{(hyb)}} \gamma_i + \theta_0 = 1$. Depending on the roles of different diffusion channels, the weights can be

- $> 0$, if positive information in the channel will increase employees' activation probability;
- $= 0$, if positive information in the channel will not change employees' activation probability;
- $< 0$, if positive information in the channel will decrease employees' activation probability.

In **MUSE**, weights of certain diffusion channels can be negative. As a result, the likelihood for a node to become active will no longer grow monotonically in the **MUSE** diffusion model. The optimal weights of different diffusion channels can be learned from the users' infection record data. Different diffusion channels will be ranked according to their importance and top-$k$ diffusion channels which can increase individuals' activation probabilities will be selected in the next subsection.

### 9.5.7 Channel Weighting and Selection

The historical users' infection records data can be represented as a set of tuples $\{(u, t)\}_{u,t}$, where tuple $(u, t)$ represents that user $u$ gets activated by topic $t$. Such a tuple set can be split into three parts according to ratio 3:1:1 in the order of the timestamps, where threefolds are used as the training set, onefold is used as the validation set and onefold as the test set. We will use the training set data to calculate the activation probabilities of individuals getting activated by topics in both the validation set and test set, while validation set is used to learn the weights of different diffusion channels and test set is used to evaluate the learned model.

Let $\mathcal{V} = \{(u, t)\}_{u,t}$ be the validation set. Based on the amount of information propagating among employees in the workplace calculated with the training set, we can infer the probability of user $u$'s (who has not been activated yet) get activated by topic $t$, for $\forall(u, t) \in \mathcal{V}$, which can be represented with matrix $\mathbf{F} \in \mathbb{R}^{|\mathcal{U}|\times|\mathcal{T}|}$, where $F(i, j)$ denotes the inferred activation probability of tuple $(u_i, t_j)$ in the validation set. Meanwhile, based on the validation set itself, we can get the ground-truth of users' infection records, which can be represented as a binary matrix $\mathbf{H} \in \{0, 1\}^{|\mathcal{U}|\times|\mathcal{T}|}$. In matrix $\mathbf{H}$, only entries corresponding tuples in the validation set are filled with value 1 and the remaining entries are all filled with 0. The optimal weights of information delivered in different diffusion channels (i.e., $\boldsymbol{\alpha}^*$, $\boldsymbol{\beta}^*$, $\boldsymbol{\gamma}^*$, $\theta_0^*$) can be obtained by solving the following objective function:

$$\boldsymbol{\alpha}^*, \boldsymbol{\beta}^*, \boldsymbol{\gamma}^*, \theta_0^* = \arg \min_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\theta}_0} \|\mathbf{F} - \mathbf{H}\|_F^2$$

$$s.t. \sum_{i=1}^{k^{(on)}} \alpha_i + \sum_{i=1}^{k^{(off)}} \beta_i + \sum_{i=1}^{k^{(hyb)}} \gamma_i + \theta_0 = 1. \tag{9.47}$$

The final objective function is not convex and can have multiple local optima, as the aggregation function (i.e., the logistic function) is not convex actually. **M**USE solves the objective function and handle the non-convex issue by using a two-stage process to ensure the robust of the learning process as much as possible.

(1) Firstly, the above objective function can be solved by using the method of Lagrange multipliers [1], where the corresponding Lagrangian function of the objective function can be represented as

$$\begin{aligned}
&\mathcal{L}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \theta_0, \eta) \\
&= \|\mathbf{F} - \mathbf{H}\|_F^2 + \eta \left( \sum_{i=1}^{k^{(on)}} \alpha_i + \sum_{i=1}^{k^{(off)}} \beta_i + \sum_{i=1}^{k^{(hyb)}} \gamma_i + \theta_0 - 1 \right), \\
&= \mathrm{Tr}(\mathbf{F}\mathbf{F}^\top - \mathbf{F}\mathbf{H}^\top - \mathbf{H}\mathbf{F}^\top + \mathbf{H}\mathbf{H}^\top) + \eta \left( \sum_{i=1}^{k^{(on)}} \alpha_i + \sum_{i=1}^{k^{(off)}} \beta_i + \sum_{i=1}^{k^{(hyb)}} \gamma_i + \theta_0 - 1 \right).
\end{aligned} \tag{9.48}$$

By taking the partial derivatives of the Lagrange function with regard to variable $\alpha_i, i \in \{1, 2, \ldots, k^{(on)}\}$, we can get

$$\begin{aligned}
\frac{\partial \mathcal{L}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \theta_0, \eta)}{\partial \alpha_i} &= \frac{\partial \mathrm{Tr}(\mathbf{F}\mathbf{F}^\top)}{\partial \alpha_i} - \frac{\partial \mathrm{Tr}(\mathbf{F}\mathbf{H}^\top)}{\partial \alpha_i} - \frac{\partial \mathrm{Tr}(\mathbf{H}\mathbf{F}^\top)}{\partial \alpha_i} + \frac{\partial \mathrm{Tr}(\mathbf{H}\mathbf{H}^\top)}{\partial \alpha_i} \\
&\quad + \frac{\partial \eta \left( \sum_{i=1}^{k^{(on)}} \alpha_i + \sum_{i=1}^{k^{(off)}} \beta_i + \sum_{i=1}^{k^{(hyb)}} \gamma_i + \theta_0 - 1 \right)}{\partial \alpha_i}.
\end{aligned} \tag{9.49}$$

Term

$$\frac{\partial \eta \left( \sum_{i=1}^{k^{(on)}} \alpha_i + \sum_{i=1}^{k^{(off)}} \beta_i + \sum_{i=1}^{k^{(hyb)}} \gamma_i + \theta_0 - 1 \right)}{\partial \alpha_i} = \eta \tag{9.50}$$

$$\begin{aligned} \frac{\partial \mathrm{Tr}(\mathbf{F}\mathbf{F}^\top)}{\partial \alpha_i} &= \sum_{j=1}^{|\mathcal{U}|} \sum_{l=1}^{|\mathcal{T}|} \frac{\partial \mathbf{F}^2(j,l)}{\partial \alpha_i} \\ &= \sum_{j=1}^{|\mathcal{U}|} \sum_{l=1}^{|\mathcal{T}|} 2f(\mathbf{w}^{(on)}(\cdot, u_j, t_l), \mathbf{w}^{(off)}(\cdot, u_j, t_l), \mathbf{w}^{(hyb)}(\cdot, u_j, t_l)) \cdot \frac{e^y}{(1+e^y)^2} \cdot \frac{\partial y}{\partial \alpha_i}, \end{aligned} \tag{9.51}$$

where the introduced term $y$ denotes

$$y = g(\mathbf{w}^{(on)}(\cdot, u_j, t_l)) + g(\mathbf{w}^{(off)}(\cdot, u_j, t_l)) + g(\mathbf{w}^{(hyb)}(\cdot, u_j, t_l)) + \theta_0 \tag{9.52}$$

and its derivative can be represented as

$$\frac{\partial y}{\partial \alpha_i} = \frac{\partial g(\mathbf{w}^{(on)}(\cdot, u_j, t_l))}{\partial \alpha_i} = \sum_{v \in \Gamma_{out}^{(on),i}(u)} w^{(on),i}(v, u_j, t_k). \tag{9.53}$$

Similarly, we can obtain terms $\frac{\partial \mathrm{Tr}(\mathbf{FH}^\top)}{\partial \alpha_i}$, $\frac{\partial \mathrm{Tr}(\mathbf{HF}^\top)}{\partial \alpha_i}$, and $\frac{\partial \mathrm{Tr}(\mathbf{HH}^\top)}{\partial \alpha_i}$. By making $\frac{\partial \mathcal{L}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \theta_0, \eta)}{\partial \alpha_i} = 0$, we can obtain an equation involving variables $\alpha_i$, $\beta_i$, $\gamma_i$, $\theta_0$ and $\eta$. Furthermore, we can calculate the partial derivatives of the Lagrange function with regards to variable $\beta_i$, $\gamma_i$, $\theta_0$, and $\eta$, respectively, and make the equation equal to 0, which will lead to an equation group about variables $\alpha_i$, $\beta_i$, $\gamma_i$, $\theta_0$, and $\eta$. The equation group can be solved with open source toolkits, e.g., SciPy Nonlinear Solver,[2] effectively. By giving the variables with different initial values, multiple solutions (i.e., multiple local optimal points) can be obtained by resolving the objective function.

(2) Secondly, the local optimal points obtained are further applied to the objective function and the one achieving the lowest objective function value is selected as the final results (i.e., the weights of different channels).

According to the learned weights, different diffusion channels can be ranked according to their importance in delivering information to activate employees in the workplace. Considering that, some diffusion channels may not perform very well in information propagation (e.g., those with negative or zero learned weights), top-$k$ channels that can increase employees' activation probabilities are selected as the effective channels used in MUSE model finally. In other words, $k$ equals to the number of diffusion channels with positive weights learnt from the above objective function. Such a process is formally called diffusion channel weighting and selection in this section.

## 9.6 Summary

In this section, we introduced various information diffusion models, which can depict the information propagation process in online social networks. Via the interactions among users, information of

[2] http://docs.scipy.org/doc/scipy-0.14.0/reference/optimize.nonlin.html.

various topics can diffuse among users in online social networks. In the information diffusion process, the information topics, the propagation channels, the target users, and the whole network properties can all affect the information diffusion and lead to different observations.

For the traditional single-homogeneous networks, we introduced several classic information diffusion models, including the threshold based diffusion model, cascade based diffusion model, epidemic diffusion model, and heat diffusion model. For these diffusion models, they assume there exist one single type of diffusion channels among users inside the network, and they can work well for the single-network scenarios with one single topic.

To model the information diffusion of multiple topics with intertwined relationships, we introduced an intertwined information diffusion model TLT by extending the classic LT model. The TLT diffusion model can update users' thresholds towards different topics with considerations about the relationships among the topics. To describe the information diffusion process in signed networks, we introduced the MFC diffusion model, which incorporates the link polarities into the model.

To model the information diffusion across multiple social networks, we introduced two different information diffusion models based on network coupling and cross-network random walk, respectively. In the network coupling based diffusion model, we introduced an efficient method to check the status of users in the diffusion process. In the random walk based diffusion model, users may propagate information either within networks or across networks with a certain chance.

Finally, at the last section, we introduced the information diffusion model **MUSE** at the workplace across the online and offline world. The introduced **MUSE** model is built based on the meta path concept. By aggregating the information propagated from the diverse meta paths, **MUSE** is able to compute the optimal weights assigned for these diverse channels and sources, respectively.

## 9.7 Bibliography Notes

Information diffusion has been a hot research topic in the last decade and dozens of papers have been published on this topic so far. Domingos and Richardson [7, 28] are the first to propose to study the influence propagation based on knowledge-sharing sites. Kempe et al. [13] are the first to study the influence propagation problem through social networks and propose two famous diffusion models: Independent cascade (IC) model and linear threshold (LT) model, which have been the basis of many diffusion models proposed later.

In recent years, signed networks have gained increasing attention. Li et al. [19] studied the influence diffusion dynamics in social networks with friend and foe relationships. Polarity related influence maximization problem in signed social networks is studied in [20], where a new diffusion model, corresponding to the polarity independent cascade (P-IC) model, is proposed. If the readers are interested in these models proposed for the signed networks, you may take a look at these articles to get more information.

Meanwhile, some works have also been done on studying information diffusion problems by considering multiple networks/sources. Zhan et al. [38] propose to study the information diffusion problem across two partially aligned social networks based on the extracted multi-relations among users. Zhan et al. [41] also introduce a novel information diffusion model based on random walk, which can be used to depict the cross-network information diffusion process. Nguyen et al. [25] propose a coupling-based diffusion models to study the information diffusion problem in multiplex social networks. Myers et al. [23] and Lin et al. [22] present two different information diffusion models incorporating both external influence sources and the internal influence among users in online social networks.

In addition, a comprehensive survey about the existing information diffusion models and research works has been provided in [11, 42] and the readers may also refer to these articles when reading this chapter.

## 9.8 Exercises

1. (Easy) In both LT and IC diffusion models, let $\mathcal{V}^a(t)$ denote the set of activated users at step $t$ in the diffusion process. Please prove that the following statement holds for both LT and IC models:

$$\mathcal{V}^a(t-1) \subseteq \mathcal{V}^a(t), \forall t \in \{1, 2, \ldots\} \tag{9.54}$$

2. (Easy) Please think about the *heat diffusion model*, and explain what is the potential problem of this model when applied to depict the information diffusion process in online social networks (Hint: What will be the equilibrium status when the diffusion ends? Can the heat diffusion equilibrium status model the information diffusion ending status?)
3. (Easy) Please explain how the MFC handles the asymmetrical effects of positive and negative links in the information diffusion process.
4. (Easy) Please explain why the relaxed activation criterion adopted in the *network coupling based information diffusion* model will create some latency in users' status checking.
5. (Easy) Please briefly explain why the meta path selection is necessary in the **M**USE model, especially when there exist a large number of meta paths defined in the model.
6. (Medium) Please implement the IC and LT diffusion models, and output the expected number of infected users based on a synthetic social network dataset.
7. (Medium) Please try to implement the TLT diffusion model with a preferred programming language, and test the algorithm with a synthetic social network dataset.
8. (Medium) Please implement the *random walk* based information diffusion model IPATH with a preferred programming language.
9. (Hard) Please revise the SIRS model by incorporating the population birth and death into considerations, and try to derive the partial derivative of the population at different stages regarding the time.
10. (Hard) Please try to implement the **M**USE model together with its meta path weighting and selection algorithm.

## References


1. D. Bertsekas, *Constrained Optimization and Lagrange Multiplier Methods (Optimization and Neural Computation Series)* (Athena Scientific, Nashua, 1996)
2. S. Bharathi, D. Kempe, M. Salek, Competitive influence maximization in social networks, in *Internet and Network Economics. WINE 2007* (Springer, Berlin, 2007)
3. T. Carnes, R. Nagarajan, S. Wild, A. Zuylen, Maximizing influence in a competitive social network: a follower's perspective, in *Proceedings of the Ninth International Conference on Electronic Commerce (ICEC '07)* (ACM, New York, 2007)
4. W. Chen, A. Collins, R. Cummings, T. Ke, Z. Liu, D. Rincon, X. Sun, Y. Wang, W. Wei, Y. Yuan, Influence maximization in social networks when negative opinions may emerge and propagate – Microsoft research, in *Proceedings of the Eleventh SIAM International Conference on Data Mining (SDM 2011)*, 2011.
5. S. Datta, A. Majumder, N. Shrivastava, Viral marketing for multiple products, in *2010 IEEE International Conference on Data Mining* (IEEE, Piscataway, 2010)
6. J.S. deCani, R.A. Stine, A note on deriving the information matrix for a logistic distribution. Am. Stat. **40**(3), 220–222 (1986)

7. P. Domingos, M. Richardson, Mining the network value of customers, in *Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '01)* (ACM, New York, 2001)
8. R. Ghosh, K. Lerman, Non-conservative diffusion and its application to social network analysis. arXiv preprint arXiv:1102.4639 (2011)
9. J. Goldenberg, B. Libai, E. Muller, Talk of the network: a complex systems look at the underlying process of word-of-mouth. Mark. Lett. **12**(3), 211–223 (2001)
10. M. Granovetter, Threshold models of collective behavior. Am. J. Sociol. **83**(6), 1420–1443 (1978)
11. A. Guille, H. Hacid, C. Favre, D. Zighed, Information diffusion in online social networks: a survey. SIGMOD Rec. **42**(2), 17–28 (2013)
12. H. Hethcote, *Three Basic Epidemiological Models* (Springer, Berlin, 1989)
13. D. Kempe, J. Kleinberg, É. 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 (KDD '03)* (ACM, New York, 2003)
14. D. Kempe, J. Kleinberg, É. Tardos, Influential nodes in a diffusion model for social networks, in *Proceedings of the 32nd International Conference on Automata, Languages and Programming (ICALP'05)* (Springer, Berlin, 2005)
15. W. Kermack, A. McKendrick, A contribution to the mathematical theory of epidemics. Proc. Roy. Soc. Lond. A Math. Phys. Eng. Sci. **115**(772), 700–721 (1927)
16. J. Kleinberg, The small-world phenomenon: an algorithmic perspective, in *Proceedings of the Thirty-Second Annual ACM Symposium on Theory of Computing (STOC '00)* (ACM, New York, 2000)
17. J. Leskovec, D. Huttenlocher, J. Kleinberg, Predicting positive and negative links in online social networks, in *Proceedings of the 19th International Conference on World Wide Web (WWW '10)* (ACM, New York, 2010)
18. J. Leskovec, D. Huttenlocher, J. Kleinberg, Signed networks in social media, in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)* (ACM, New York, 2010)
19. Y. Li, W. Chen, Y. Wang, Z. Zhang, Influence diffusion dynamics and influence maximization in social networks with friend and foe relationships, in *Proceedings of the Sixth ACM International Conference on Web Search and Data Mining (WSDM '13)* (ACM, New York, 2013)
20. D. Li, Z. Xu, N. Chakraborty, A. Gupta, K. Sycara, S. Li, Polarity related influence maximization in signed social networks. PLoS One **9**(7), e102199 (2014)
21. J. Lienhard, *A Heat Transfer Textbook: Fourth Edition* (Dover Civil and Mechanical Engineering. Dover Publications, New York, 2013)
22. S. Lin, F. Wang, Q. Hu, P. Yu, Extracting social events for learning better information diffusion models, in *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '13)* (ACM, New York, 2013)
23. S. Myers, C. Zhu, J. Leskovec, Information diffusion and external influence in networks, in *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '12)* (ACM, New York, 2012)
24. R. Narayanam, A. Nanavati, Viral marketing for product cross-sell through social networks, in *Machine Learning and Knowledge Discovery in Databases (ECML PKDD 2012)* (Springer, Berlin, 2012)
25. D. Nguyen, H. Zhang, S. Das, M. Thai, T. Dinh, Least cost influence in multiplex social networks: model representation and analysis, in *2013 IEEE 13th International Conference on Data Mining* (IEEE, Piscataway, 2013)
26. R. Pastor-Satorras, C. Castellano, P. Mieghem, A. Vespignani, Epidemic processes in complex networks. CoRR, abs/1408.2701 (2014)
27. D. Peleg, Local majority voting, small coalitions and controlling monopolies in graphs: a review, in *Proceedings 3rd Colloquium on Structural Information and Communication Complexity* (Carleton University Press, Siena, 1996)
28. M. Richardson, P. Domingos, Mining knowledge-sharing sites for viral marketing, in *Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '02)* (ACM, New York, 2002)
29. D. Shah, T. Zaman, Rumors in a network: who's the culprit? IEEE Trans. Inf. Theory **57**(8), 5163–5181 (2011)
30. Y. Sun, J. Han, X. Yan, P. Yu, T. Wu, PathSim: meta path-based top-k similarity search in heterogeneous information networks, in *Proceedings of the VLDB Endowment* (2011)
31. J. Tang, Y. Chang, C. Aggarwal, H. Liu, A survey of signed network mining in social media. *ACM Computing Surveys, to appear*, CoRR abs/1511.07569 (2015)
32. H. Tong, C. Faloutsos, J.-Y. Pan, Fast random walk with restart and its applications (2006)
33. T. Turner, P. Qvarfordt, J. Biehl, G. Golovchinsky, M. Back, Exploring the workplace communication ecology, in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)* (ACM, New York, 2010)
34. R. West, H. Paskov, J. Leskovec, C. Potts, Exploiting social network structure for person-to-person sentiment analysis. Trans. Assoc. Comput. Linguist. **2**, 297–310 (2014)
35. K. Wilcox, A. T. Stephen, Are close friends the enemy? Online social networks, self-esteem, and self-control. J. Consum. Res. **40**(1), 90–103 (2012)

36. Y. Yao, H. Tong, X. Yan, F. Xu, J. Lu, Matri: a multi-aspect and transitive trust inference model. in *Proceedings of the 22nd International Conference on World Wide Web* (ACM, New York, 2013)
37. J. Ye, H. Cheng, Z. Zhu, M. Chen, Predicting positive and negative links in signed social networks by transfer learning, in *Proceedings of the 22nd International Conference on World Wide Web (WWW '13)* (ACM, New York, 2013)
38. Q. Zhan, J. Zhang, S. Wang, P. Yu, J. Xie, Influence maximization across partially aligned heterogeneous social networks, in *Pacific-Asia Conference on Knowledge Discovery and Data Mining* (Springer, Cham, 2015)
39. Q. Zhan, J. Zhang, X. Pan, P. Yu, Discover tipping users for cross network influencing, in *2016 IEEE 17th International Conference on Information Reuse and Integration (IRI)* (IEEE, Piscataway, 2016)
40. Q. Zhan, J. Zhang, P. Yu, S. Emery, J. Xie, Inferring social influence of anti-tobacco mass media campaigns, in *2016 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)* (IEEE, Piscataway, 2016)
41. Q. Zhan, J. Zhang, P. Yu, J. Xie, Viral marketing through aligned networks (2018)
42. J. Zhang, Social network fusion and mining: a survey (2018)
43. J. Zhang, P. Yu, Community detection for emerging networks, in *2015 SIAM International Conference on Data Mining* (2015)
44. J. Zhang, Y. Lv, P. Yu, Enterprise social link recommendation, in *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management (CIKM '15)* (ACM, New York, 2015)
45. J. Zhang, P. Yu, Y. Lv, Organizational chart inference, in *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '15)* (ACM, New York, 2015)
46. J. Zhang, S. Wang, Q. Zhan, P. Yu, Intertwined viral marketing in social networks, in *2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)* (IEEE, Piscataway, 2016)
47. J. Zhang, P. Yu, Y. Lv, Q. Zhan, Information diffusion at workplace, in *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management (CIKM '16)* (ACM, New York, 2016)
48. J. Zhang, Q. Zhan, L. He, C. Aggarwal, P. Yu, Trust hole identification in signed networks, in *Machine Learning and Knowledge Discovery in Databases (ECML PKDD 2016)* (Springer, Cham, 2016)
49. H. Zhang, D. Nguyen, H. Zhang, M. Thai, Least cost influence maximization across multiple social networks. IEEE ACM Trans. Netw. **24**(2), 929–939 (2016)
50. J. Zhang, C. Aggarwal, P. Yu, Rumor initiator detection in infected signed networks, in *2017 IEEE 37th International Conference on Distributed Computing Systems (ICDCS)* (IEEE, Piscataway, 2017)