On the Submodularity of Diffusion Models: Equivalent Conditions and Applications

Fangqi Li, Chong Di, Wenwen Xia
School of Cyber Science and Engineering,
Shanghai Jiao Tong University,
Shanghai, China.
solour_lfq@sjtu.edu.cn

***Abstract*—The diffusion model has been a crucial component in studies about social networks. Many studies, especially these about influence maximization concern the proof of the submodularity of particular diffusion models. Such proofs have been model-dependent and are somewhat ad hoc. In this paper, we prove a theorem that provides a necessary and sufficient condition for a diffusion model to be submodular. This theorem can be used to justify the submodularity of an arbitrary diffusion model. We also apply this theorem to build a projection operator that maps an arbitrary diffusion model into a submodular one. Moreover, we use the established theorem to propose a diffusion model of multiple heterogeneous pieces of information that partially features submodularity.**

Keywords: social network, diffusion model, submodularity.

I. Introduction

Diffusion model is an important part of studies concerning social networks and the propagation pattern of information within, it formalizes and answers this question: in what way does a piece of information propagate in a complex community? A community is a collection of participants that interacts with each other, and its properties are usually embedded in a weighted directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{W})$, given $\mathcal{G}$, a diffusion model computes what subset of participants can be influenced by some information released by another subset of participants. For example, in the problem of influence maximization, we are to select a subset $\mathcal{S}$ of $\mathcal{V}$ with no more than$K$ participants, such that an information propagation begins from it can affect as many participants as possible. The influenced subset in this scenario is denoted by $\sigma(\mathcal{S})$ and we are to maximized $|\sigma(\mathcal{S})|$. Here the propagation function $\sigma(\cdot)$ is given by a specific diffusion model.

Problems like influence maximization are innately combinatorial optimizations, whose space of possible solutions has the size $\mathcal{O}(|\mathcal{V}|^K)$ which grows exponentially with K, so efficient algorithms are hardly feasible. To reduce the cost of calculation in such problems, [1] proposed that: if we take advantage of the submodularity of diffusion model and the theorems in [2], then the greedy algorithm (during which K nodes that maximize the marginal propagation increment are selected sequentially) can result in a propagation range which is no smaller than $\left(1 - \frac{1}{e}\right)$ of the theoretical optimal propagation range. Studies hitherto [3][4][5] utilize other graph features or extra data structures to boost the calculation and have derived diversified and fruitful results. However, the proof of the submodularity of an arbitrary diffusion model remains challenging and inspiring, [1] proved the submodularity of two specific diffusion models, namely *independent cascade model* (IC) and *linear threshold model* (LT), but what about a general diffusion model? Moreover, does there exist a simple and feasible equivalent condition for a diffusion model to be submodular? For non-submodular diffusion models, is it possible to approximate them using submodular ones? What properties does this approximation method hold? At last, for information propagation in complex scenarios as competitive propagation or cooperative propagation, can we design diffusion models that partially feature submodularity? We try to provide answers to these questions in this paper. The contributions of this paper are three-folded:

1) We show that under some trivial assumptions, there exists a necessary and sufficient condition for a diffusion model to be submodular.
2) We propose a projection operator that transforms an arbitrary diffusion model into a submodular diffusion model. This operator is a legal mapping under a certain definition of equivalence between diffusion models.
3) The established results are utilized to coin a diffusion model of multiple pieces of infor-

mation within a network. This model partially features submodularity and can be applied in further studies concerning competitive influence maximization.

This paper proceeds as follows: Section II reviews the formulation of diffusion model, together with other elementary knowledge. Section III contains our main theorems about the equivalent condition of the submodularity of diffusion models. Section IV extends the dissusion in Section III and provides some practical corollaries. Section V concludes this paper.

II. Preliminaries

A. Formulation of the Network

The diffusion model formulates the information propagation within a network, the network here is a rather abstract and broad concept that can be instantiated into a social network, computer network, connected water body [4], etc. Therefore related researches have a broad range of possible applications. The network is usually formulated by a weighted directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{W})$, where the weight of an edge $w_e, e = (u, v) \in \mathcal{E} \subset \mathcal{V}^2$ reflects the intensity of the influence that node u exerts on node v (in this paper we use the terminologies *node*, *participant* and *vertex* interchangeably.) The physical interpretation of such intensity is model-dependent. It could be a function of probability, a relative measure of weight corresponding to some threshold metric, etc. In practice, it is impossible to provide a $\mathcal{G}$ without specifying a corresponding diffusion model which explains the meaning of its weights.

B. Formulation of the Diffusion Model

Given the network structure $\mathcal{G}$, a diffusion model M defines a conditional probability measure on the power set of vertices:

$$\mathcal{P}_M(\mathcal{T}|\mathcal{S}), \mathcal{S} \subset \mathcal{T} \subset \mathcal{V},$$

whose physical interpretation is: the probability that information propagation begins from $\mathcal{S} \subset \mathcal{V}$ terminates in $\mathcal{T} \subset \mathcal{V}$. To derive the theorems below, we exert two assumptions on diffusion models:

1) **Markovian**: We assume that for any vertex $v \in \mathcal{V}$, whether or not it is influenced/activated (in this paper, being *influenced* or *activated* means the same thing) is solely determined by the state of its parent nodes S_v, and is independent from the activation state of other vertices or the order of activation (by the *state* of its parent nodes S_v we mean the binary encoding of S_v, where 0 means a vertex is unactivated/does not receive the influence and 1 denotes otherwise). To put it in other words, the event of activation of v can be embedded by a conditional probability $\mathcal{P}_M(v'|S'_v)$, where S'_v is a specific activation state of S_v with altogether $2^{|S_v|}$ possible values (we also exert an order on all possible states of S_v, to do so, we firstly define an arbitrary order on S_v and encode every state S'_v using binary code so every S'_v can be mapped to a unique number between 0 and $(2^{|S_v|} - 1)$, hence a natural order is feasible), while v' is a binary variable that reflects whether vertex v is activated or not.
2) **Acylic**: We assume that in propagation, the graph $\mathcal{G}$ can be taken as an acyclic graph, hence there exists a topological order $\prec$ on $\mathcal{V}$. If we sort $\mathcal{V}$ using $\prec$, then for any vertex v, all its parent nodes $u \in S_v$ appear before v [6]. In other words, the unfolding process of the propagation can be conducted according to this topological order, for all the necessary information that determined whether vertex v can be activated or not are ready when v is visited according to $\prec$.

Under these two assumptions, the propagation induced by a diffusion model in $\mathcal{G}$ from $\mathcal{S} \subset \mathcal{V}$ to $\mathcal{T} \subset \mathcal{V}$ can be seen as proceeding as follows:

1) Exerting a topological order $\prec$ on $\mathcal{V}$.
2) Marking the vertices in $\mathcal{S}$ as activated ones.
3) Iterating over $\mathcal{V}$ according to $\prec$, for each $v \in \mathcal{V}$, determing whether or not it would be activated by sampling from $\mathcal{P}_M(v'|S'_v)$.
4) Finally, the probability $\mathcal{P}_M(\mathcal{T}|\mathcal{S})$ is given by the product of all conditional probabilities on vertices:

$$\mathcal{P}_M(\mathcal{T}|\mathcal{S}) = \prod_{v \in \mathcal{V}} \mathcal{P}_M(v' = 1|S_{v'})^{\mathbb{I}[v \in \mathcal{T}]} \cdot (1 - \mathcal{P}_M(v' = 1|S_{v'}))^{1 - \mathbb{I}[v \in \mathcal{T}]}.$$

where $\mathbb{I}[...]$ is an boolean indicator function.

The product in the equation above is taken according to $\prec$ so each term is ready to be computed according to the two assumptions given before.

C. Submodularity

A function that define on the power set of $\mathcal{V}$, $f : 2^{\mathcal{V}} \to \mathbb{R}$ is said to be **submodular** if $\forall \mathcal{S} \subset \mathcal{T} \subset \mathcal{V}, \forall v \in \mathcal{V}/\mathcal{T}$ the following inequality holds:

$$f(\mathcal{S} \cup \{v\}) - f(\mathcal{S}) \geq f(\mathcal{T} \cup \{v\}) - f(\mathcal{T}). \quad (1)$$

Submodularity is an elegant and desirable property that might be enjoyed by the propagation range of a diffusion model. Once the submodularity is established, the influence maximization problem with respect to the corresponding diffusion model can be

readily solved by greedy methods and a result lower bounded by $\left(1-\frac{1}{e}\right)$ of the theoretical optimum is secured. Therefore the proof of the submodularity of a diffusion model (i.e., the submodularity of the range of propagation according to this diffusion model) is of vital significance.

Most diffusion models include randomness in the propagation process (so the conditional probability $\mathcal{P}_M(\mathcal{T}|\mathcal{S})$ hardly degenerates), therefore the propagation range as a scalar function of the set of seeds $\mathcal{S}$ can be formulated as:

$$f(\mathcal{S}) = \mathbb{E}\left[|\sigma_M(\mathcal{S})|\right] = \sum_{\mathcal{T}\subset\mathcal{V}} \mathcal{P}_M(\mathcal{T}|\mathcal{S}) \cdot |\mathcal{T}|. \quad (2)$$

III. Main Theorems

A. Transforming to a New Measure

The submodularity of a function as (2) is difficult to analyze, so we resort to a transformation in measure, formally, consider the following form:

$$f(\mathcal{S}) = \sum_{G} \mathcal{P}_M(G) \cdot f(G, \mathcal{S}). \quad (3)$$

In which we try to transform the randomness in $P_M(\mathcal{T}|\mathcal{S})$ into another auxiliary variable G, and hope that G is independent from $\mathcal{S}$ but is capable of representing the distribution of propagation introduced by M. According to our discussion in Section 2.2., the propagation process is a cascade of application of activation rule for each v and its S_v *per se*, by activation rule we mean a reinterpretation of $\mathcal{P}_M(v'|S'_v)$, whose every instance is tantamount to a rule of activation that assigns each activation pattern/state of S_v (recall that there are altogether $2^{|S_v|}$ possibilities) a deterministic result about the state of v. So we can denote a specific response pattern of v by G_v, a binary vector of length $2^{|S_v|}$, whose the i-th component denotes whether v would be activated or not under the state of S_v encoded binarily as $(i-1)$. The collection of all possible G_v corrseponding to all v forms G. Section 2.2. has justified that $\mathcal{P}(G)$ preserves all information in $\mathcal{P}_M(\mathcal{T}|\mathcal{S})$, so we can study (3) instead of (2) without loss of generality [7]. Particularly, the following relationships hold, which might help to delve into this transformation between measures and variables:

$$\mathcal{P}_M(G_v) = \prod_{i=1}^{2^{|S_v|}} \mathcal{P}_M\left(v' = G_v(i) | S'_v \sim (i-1)\right),$$

$$\mathcal{P}_M(G) = \prod_{v\in\mathcal{V}} \mathcal{P}_M(G_v),$$

where $G_v(i)$ denotes the reaction of v according to G_v under the activation state S'_v and is a binary variable, 0 for unactivation and 1 for activation, $S'_v \sim (i-1)$ denotes that S'_v, an activation state of S_v, is binarily encoded as $(i-1)$. Since the $\mathcal{P}_M(v'|S'_v)$ for each v are independent from each other, there is no particular numeric constraints on $\mathcal{P}_M(G)$ except for normalization. Finally, the deterministic propagation function $f(G, \mathcal{S})$ conditioned on G can be evaluated as follows: iterating over the topological sequence on $\mathcal{V}$ and judging whether $v \in \mathcal{V}$ is activated or not according to $G_v \in G$, finally the number of activated vertices is returned. Note that since G_v has collected all possible states of S_v, so G is independent from $\mathcal{S}$. It is now obvious that given the set of seeds $\mathcal{S}$, (2) and (3) introduce the same distribution of activated nodes among $\mathcal{V}$.

B. The First Main Theorem

If for any possible G, $f(G,\cdot)$ appears to be a submodular function, then $f(\cdot)$, as a convex combination of submodular functions, turns out to be a submodular function. To check for this point, pluggin (1) into (3) and it is straightforward to see:

$$\begin{aligned} & f(\mathcal{S}\cup\{v\}) - f(\mathcal{S}) \\ =& \sum_{G} \mathcal{P}_M(G)\left[f(G, \mathcal{S}\cup\{v\}) - f(G,\mathcal{S})\right] \\ \geq& \sum_{G} \mathcal{P}_M(G)\left[f(G, \mathcal{T}\cup\{v\}) - f(G,\mathcal{T})\right] \\ =& f(\mathcal{T}\cup\{v\}) - f(\mathcal{T}). \end{aligned}$$

Unfortunately, there exists some assignment of G whose corresponding combination of activation rules refutes submodularity. Thus there is only a strict subset of all possible G on which submodularity holds. Therefore, if it turns out that a diffusion model M can be transform in a way such that all probability on this specific subset sums up to one, then $f(G,\cdot)$ would always be submodular with respect to this diffusion model and (3) would be submodular. Hence it is crucial to identify the class of G on which submodularity holds. Formally, we prove the following theorem:

Theorem 1. For a diffusion model M, $\forall v \in \mathcal{V}$, let v's expectations of being activated under altogether $2^{|S_v|}$ states of activation of S_v be collected in a vector $\boldsymbol{a}_M(v)$ of size $2^{|S_v|}$, where S'_v are sorted using binary code order. If $\forall v \in \mathcal{V}$, the following equation holds:

$$\exists \boldsymbol{b}, \boldsymbol{a}_M(v) = \boldsymbol{M}\boldsymbol{b}, \quad (4)$$

where $\boldsymbol{b}$ is a probability vector of length $2^{|S_v|}$, whose components are non-negative and sum up to unity. And the second-rank tensor $\boldsymbol{M}$ is defined by:

$$\boldsymbol{M}_{i,j} = \bigvee_{m=1}^{|S_v|} \left(i(m) \wedge j(m)\right), \quad (5)$$

*where $i(m), j(m)$ denotes the m-th index of the binary code for $(i-1), (j-1)$ respectively. **M** depends on $|S_v|$ while **b** depends on v. Then M is submodular in a sense that each G it supports introduces a submodular propagation range. The reverse theorem also holds.*

Proof: We first prove that this condition is sufficient for submodularity. Note that (5) collects all G_v of the following category: firstly, v is connected to S_v^*, a subset of S_v, secondly, if at least one vertex in S_v^* is activated then so is v, otherwise v is not going to be activated. The pattern of connection is encoded in the dimensionality of i, while the activation state of S_v is encoded in the dimensionality of j. Only if there exists at least one activated vertex in S_v and is connected to v will v be activated. Here *there exists at least one* and *activated and is connected to* are formally reflected in (5) by $\bigvee$ and $\wedge$. If the expectation to be activated at v under different states (i.e. $\mathbf{a}_M(v)$) satisfies (4), then it would be safe to transform the behavior of M at v by sampling from these G_v (the connection-activation pattern) by probability **b**. Since all the components of **b** sum up to one, the behavior of M is restricted to this subset of G. Under this condition, the proof in [1] can be revoked to show submodularity. A piece of information that begins from $\mathcal{S}$ affects v if and only if there is an activated path that connects one vertex in $\mathcal{S}$ to v, the activation of each path is controlled by **b**. In this way, M would turn out to be submodular.

Reversely, if the condition $\mathbf{a}_M(v) = \mathbf{Mb}$ fails to be met, then $\mathbf{a}_M(v)$ can be spanned on another set of basis $\mathbf{M}'$, which is different from **M** in (5), $\mathbf{a}_M(v) = \mathbf{M}'\mathbf{b}'$, then the propagation is tantamount to activating v according to a rule G_v different from (5)'s semantics. Explicitly, the activation rule at v can no longer be *always* written as a disjunctive clause of a subset of S_v. Therefore at least a conjunctive clause is included, i.e., v would be activated only if all vertices in a subset S_v^{**} of S_v (with strictly more than one node) are activated simultaneously. Otherwise v can not be activated (since **Mb** defined under (4) and (5) has exhausted the complimentary events). Now any conjunctive rule appears to be contradictive against submodularity (intuitively, in (1), conjunctive rules enable a larger $\mathcal{T}$ to cooperate with v so potentially more activation conditions are to be met, hence submodularity breaks down.) So once M has to be transformed so it might have to be spanned on this non-submodular basis, the submodularity of M is no longer preserved. Practically, it is no longer safe to sample from all space of G supported by M since a subset of it might introduce non-submodularity propagation. This completes the proof of the theorem.

To be illustrative, consider a case where $|S_v| = 3$, so there are altogether 8 different activation states: {000,001,010,011,100,101,110,111}. Under this symbolization, the matrix **M** takes the form:

$$\mathbf{M} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}.$$

To have a sight into the semantics of **M**, consider for example $\mathbf{M}_{3,5}$, the binary code of the row index and the column index is {011} and {101} respectively, using the same construction in the proof of the theorem, they are tantamount to the activation state of S_v as:

Figure 1. The third activation state of S_v, red marks activation, white marks unactivation.

While the connection looks like:

Figure 2. The fifth connection state of S_v, dashed line marks unactivation.

Combine figure.1 and figure.2 yields:

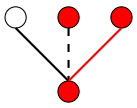

Figure 3. The evaluation of $\mathbf{M}_{3,5}$.

It can be observed that $\mathbf{M}_{3,5}$ takes the value 1 since there is an activated vertex that happens to be connected to v. By assuming $\mathbf{a}_M(v) = \mathbf{Mb}$, the diffusion model' s effect at v can be spanned onto only conjunctive rules as illustrated.

The proof for IC model and LT model in [1] can be included as special cases of Theorem 1, since IC model is corresponding to a specific assignment of **b**:

$$b_{S_v^*} = \prod_{i=1}^{|S_v|} p_i^{\mathbb{I}[v_i \in S_v^*]} (1-p_i)^{(1-\mathbb{I}[v_i \in S_v^*])},$$

where p_i are parameters in IC model. It is easy to verify that **b** is a probability vector, since it exhausts the space of pairwise independent connec-

tion. While LT model corresponds to the following **b**:

$$b_{v_i} = p_{v_i}, b_{\emptyset} = 1 - \sum_{i=1}^{|S_v|} w_{v_i},$$
$$\text{else } b_{\dots} = 0,$$

where w_{v_i} are parameters in LT model. It can be concluded that IC model and LT model cover two extreme cases in the class of all submodular diffusion models given by Theorem 1. IC model assumes the components of **b** are constrainted by independency, while LT model only normalized on a small subset of **b** (those corrsepond to only one activated connection edge). For general **b**, more degrees of freedom is left for the diffusion models that satisfy the two prerequsites in Section 2.2.

C. The Second Main Theorem

Theorem 1 takes the advantage of auxiliary variable G to algebraically express M, we can further demonstrate that the submodularity of a diffusion model in propagation range is equivalent to the submodularity of any individual vertex in activation expectation. Formally, we prove the following theorem:

Theorem 2: A diffusion model M is submodular if and only if for any $v \in \mathcal{V}$, the activation expectation E_M^v as a function of type: $2^{S_v} \to [0,1]$ (it maps a subset of S_v into the expectation/probability of v's being activated) is non-negative, monotonic and submodular.

Proof: We only need to show that the condition in Theorem 2 is equivalent to that in Theorem 1. For $v \in \mathcal{V}$, given the activation state of S_v be S_v^*, the expectation/probability of v being activated is $E_M^v(S_v^*)$ by definition. According to Theorem 1, this value is also the $\left(S_{v(2)}^* + 1\right)$-th component (decoding this binary expression) of $\mathbf{a}_M(v)$, i.e., the inner product of the $\left(S_{v(2)}^* + 1\right)$-th row of **M** and **b** ($S_{v(2)}^*$ denotes the value of $S_{v(2)}^*$ as a binary code.)

The non-negativity is conjugately implied by both definitions.

For monotonity, if $S_v^* \subset S_v^{**}$ then (5) declares that compared with the $\left(S_{v(2)}^* + 1\right)$-th line of **M**, the $\left(S_{v(2)}^{**} + 1\right)$-th line replaces some 0s with 1s, but not a single 1 is converted to 0. Therefore compared with $E_M^v(S_v^*)$, more components of **b** are added into $E_M^v(S_v^{**})$, therefore the monotonity holds.

For submodularity, consider the value of:

$$E_M^v(\mathcal{S} \cup \{u\}) - E_M^v(\mathcal{S}),$$

the value of which is the summation of the probability of this kind of G_v: such a G_v connects u and v, meanwhile, none vertex in $\mathcal{S} \subset S_v$ is connected to v. For$\mathcal{S} \subset \mathcal{T} \subset \mathcal{S}_v$, the value of:

$$E(\mathcal{T} \cup \{u\}) - E(\mathcal{T})$$

sums up the probability of that G_v connects u and v, while fails to connect any vertex in $\mathcal{T} \subset S_v$ to v. The second set of G_v is obvious a subset of the first one, since $\mathcal{S} \subset \mathcal{T}$, so a G_v that connects u and v while leaves $\mathcal{T}$ and v disconnects necessarilly disconnects $\mathcal{S}$ from v. Therefore the summation of the probability of the second set is no larger than that of the first. Hence the proof the Theorem 2 is finished.

Theorem 2 is symmetric and elegant in a sense that it shows: the submodularity of the propagation range is tantamount to the submodularity of the activation expectation of any vertex. This is a non-trivial observation since these two kinds of submodularity is different in nature. The submodularity of expectation for any vertex implies that the diffusion model can be intactly spanned on a set of submodular auxiliary variables, which further guarantees the submodularity of the diffusion model.

IV. Applications

In this section, we apply the two theorems in Section. III to yield some interesting results. First, note that different from the works in [1] where different ways of proving are applying to different diffusion models, it is possible to uniformly apply Theorem 1 and Theorem 2 to any diffusion model given its Markovian and acyclic properties.

A. Submodular Projection Operator

Considering this problem: given a non-submodular diffusion model, can we convert it to a submodular one while preserving its propagation property as well? We can solve this problem by applying Theorem 1, note that for fixed $|S_v|$, the set:

$$C = \{\mathbf{Mb} : \mathbf{b} \text{ is a probability vector, } \mathbf{M} \sim (5)\} \tag{6}$$

is convex, therefore if we minimize a convex error function as 2-norm, we ends up with a unique global optimum [8]:

$$\arg\min_{\mathbf{a}^* \in C} \left\{ |\mathbf{a}_M(v) - \mathbf{a}^*|_2^2 \right\}. \tag{7}$$

Now for an arbitrary diffusion model, we can conduct the projection in (7) to every $v \in \mathcal{V}$ (if at a particular v (4) holds, then the projection degenerates to identity) and we end up with a submodular diffusion model that is close to the original one.

This projection operator is a legal mapping that yields a unique image since (7) returns a unique result. Formally, an arbitrary diffusion model M can be projected into a submodular one M^* whose activation expectation at each v is:

$$\mathbf{a}^*(v) = \arg \min_{\mathbf{a}^*(v) \in C(v)} \left\{ |\mathbf{a}_M(v) - \mathbf{a}^*(v)|_2^2 \right\}, \quad (8)$$

where $\mathbf{a}_M(v)$ is given by the original M, while $C(v)$ as (6) depends solely on $|S_v|$. Here we consider two diffusion models M_1 and M_2 are equivalent if and only if:

$$\forall v \in \mathcal{V}, \mathbf{a}_{M_1}(v) = \mathbf{a}_{M_2}(v).$$

B. Diffusion Model for Multiple Information

When multiple pieces of information are transmitted simultaneously in the network. The whole scenario becomes much more complicated and hard to analyze. Despite some reported works [9][10][11], there is still no consensus on a diffusion model for multiple pieces of information.

The inherent problem within a diffusion model of multiple information is that the activation expectation vector/function takes value in a product space (we have to consider a vertex's response to different information at the same time.) So there is no direct generalization of order, not to mention submodularity.

So far the best we can do in carrying submodularity to diffusion models for multiple information is to consider submodularity for each kind of information. We propose a partial linear model in that different types of information enjoy submodularity in propagation range. Taking two types of information I_1 and I_2 for example. The idea behind is to adopt Theorem 2 and assume that for each $v \in \mathcal{V}$, its expectation to be activated by I_1 or I_2 is a submodular function of vertices activated by I_1 or I_2 in S_v. To ensure that such submodularity holds for both I_1, consider that:

$$E_1^v(\mathcal{S} \cup \{u\}) - E_1^v(\mathcal{S}) \geq E_1^v(\mathcal{T} \cup \{u\}) - E_1^v(\mathcal{T}), \quad (9)$$

where E_1^v is a function that computes the probability that v being activated by I_1, $\mathcal{S} \subset \mathcal{T}$ and $u \in S_v/\mathcal{T}$. Now assume that for each $S_v^* \subset S_v$ activated by I_1, let S_v/S_v^* be activated by I_2. Then it is intuitive to assume that

$$E_2^v(S_v/S_v^*) = 1 - E_1^v(S_v^*). \quad (10)$$

Combine (9) and (10):

$$E_2^v(\mathcal{T}' \cup \{u\}) - E_2^v(\mathcal{T}') \geq E_2^v(\mathcal{S}' \cup \{u\}) - E_2^v(\mathcal{S}'), \quad (11)$$

where $\mathcal{T}' = S_v/(\mathcal{S} \cup \{u\})$, $\mathcal{S}' = S_v/(\mathcal{T} \cup \{u\})$ so $\mathcal{S}' \subset \mathcal{T}$. However, by assumption the reverse of (11) has to held by the submodularity of I_2, so it turns out that:

$$E_2^v(\mathcal{T}' \cup \{u\}) - E_2^v(\mathcal{T}') = E_2^v(\mathcal{S}' \cup \{u\}) - E_2^v(\mathcal{S}'). \quad (12)$$

For I_1 and E_1^v, an analogous proposition holds. A naive realization of (12) is to assume that the increment of each vertex in S_v for either $E_{1,2}^v$ is independent of any other vertex. In this case the diffusion model is similar to LT model. To summarize, our **Parital Linear Model for Multiple Information**(PLMMI) operates as follows, assuming that there are N different types of information:

1) For S_v of $v \in \mathcal{V}$, each $u \in S_v$ is assigned N weights $\{w_{u,n}\}_{n=1}^N$ such that $\sum_{u \in S_v} \sum_{n=1}^N w_{u,n} = 1$.
2) Upon propagating, a vertex $v \in \mathcal{V}$ computes its expectation to be activated by any type of information I_n by summarizing over the weights of vertices activated by I_n, i.e.,

$$E_n^v = \sum_{u \in S_v} w_{u,n} \cdot \mathbb{I}[u \in S_v^n],$$

where S_v^n is the collection of vertices activated by information I_n in S_v.

This diffusion model partially enjoys submodularity. In a sense that the propagation range of each type of information as a function of seeds is submodular, which can be derived by applying Theorem 2 straightforwardly.

V. Conclusion

In this paper, we induce the equivalent condition for diffusion to be submodular. Compared with established paradigms that depend on specific models, our method is more general and concise. Moreover, we utilize the equivalent condition to coin a projection operator that maps an arbitrary diffusion model into a submodular one. At last, we propose a diffusion model for heterogeneous information that partially enjoys submodularity. In the future, we are going to research competitive influence maximization, during which we can utilize the PLMMI we have just presented.

VI. References

[1]. Kempe, David, Jon Kleinberg, and va Tardos. "Maximizing the spread of influence through a social network." Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2003.

[2]. Nemhauser, George L., Laurence A. Wolsey, and Marshall L. Fisher. "An analysis of approximations for maximizing submodular set functionsI." Mathematical programming 14.1 (1978): 265-294.

[3]. Han, Meng, and Yingshu Li. "Influence analysis: A survey of the state-of-the-art." Mathematical Foundations of Computing 1.3 (2018): 201-253.

[4]. LESKOVEC J, KRAUSE A, GUESTRIN C, et al. Cost-effective outbreak detection in networks[C], Proceedings of the 13th ACM SIGKDD International Conference on Knowledge discovery and data mining. San Jose, CA, USA: ACM, 2007: 420429.

[5]. Amit Goyal, Wei Lu, Laks V.S. Lakshmanan. CELF++: optimizing the greedy algorithm for influence maximization in social networks[C]. 2011.

[6]. Cormen T H, Leiserson C E, Rivest R L, et al. Introduction to Algorithms, Third Edition[M]. 2009.

[7]. Folland G B. Real analysis[M]. 1999.

[8]. Stephen Boyd, Lieven Vandenberghe. Convex Optimization[M], 2004.

[9]. Bharathi, Shishir, David Kempe, and Mahyar Salek. "Competitive influence maximization in social networks." International workshop on web and internet economics. Springer, Berlin, Heidelberg, 2007.

[10]. Carnes, Tim, et al. "Maximizing influence in a competitive social network: a follower's perspective." Proceedings of the ninth international conference on Electronic commerce. ACM, 2007.

[11]. Hui L, Bhowmick S S, Cui J, et al. GetReal: Towards Realistic Selection of Influence Maximization Strategies in Competitive Networks.[C] 2015.