Evolution Matters: Content Transmission in Evolving Wireless Social Networks

Chen Feng, Luoyi Fu, Xudong Wu, Xiaoying Gan, Xinbing Wang, Guihai Chen, and Jun Xu

***Abstract*—With the popularization of smart devices and social platforms (Facebook, Instagram and etc.), content transmission is becoming an increasingly prevalent form of human interaction with transmission time being a critical issue in miscellaneous applications. Extensive works have been devoted to this problem with the typical assumption that the network is non-evolving. In contrast, real-world networks which the transmission runs through are widely observed to be evolving over time and thus display distinctive properties.**

In this paper, we make the first study of content transmission in evolving networks. Particularly, we focus on the specific transmission time of the content, which is an important performance metric. For a comprehensive analysis, we consider both static and mobile cases. The network evolves in the following sense: each new user randomly chooses a geographic location while establishing social relations with existing users according to the evolving model, called Affiliation Networks. The transmission scheme running on the network exploits both social relationships and geographic information. While revealing the microscopic property of evolving social networks as the basis, in both static and mobile cases we manage to bound the transmission time in evolving networks which departs from results in non-evolving networks. In both cases, we find that the transmission time in evolving networks is smaller than the non-evolving counterpart under our scheme, and the gap is constantly increasing over time. Especially, the mobile case obtains such gain with less geographic knowledge. The theoretical findings are confirmed by experiments on both synthetic and real networks under different settings. We find that transmission time in evolving networks is smaller than non-evolving counterparts under the tested settings, and our proposed algorithms outperform the baselines.

***Index Terms*—Evolving wireless social networks, transmission time, mobility.**

I. Introduction

CONTENT transmission is becoming an increasingly prevalent form of human interaction and plays an important role in sharing information, diffusing social influence and affecting human behavior. Nowadays, with the proliferation of mobile devices and social applications therein, content transmission is constantly accomplished in wireless social networks. For example, a large event called Hajj is supported by the pilgrimage platform built in [1], where users know about interesting content (like documents, videos and pictures) by their social relations and share it via wireless transmission. In many applications, transmission time acts as one of the most concerned metrics, as users' experience on many mobile services depends largely on the length of transmission time.

Thus, a large body of works have been dedicated to studying the transmission time and many efficient schemes are devised for various applications, where the network topology is assumed to be invariant without the advent of new users or formation of new links [2] [3] [4]. However, it is widely observed that real-world networks are evolving over time [5] [6] [7]. For instance, Facebook was reported to have nearly 2.2 billion monthly active users in the first quarter of 2018, a 50% increase compared with that of 2015 [8]. The evolution of networks does not simply imply the increase of the number of users but also structural transition. Massive experiments in [9] reveal that evolving social networks exhibit distinctive properties:

(1) The social relations densify over time, i.e., the number of edges grows super-linearly with the number of users.
(2) The diameter [1] is shrinking or stabilizing over time.

Such structural changes will affect the transmission efficiency since the awareness of content spreads via social relationships. With denser social connections, it is easier for the content to be known by users. Accordingly, more users will be active in obtaining the content at a time and each user may have more possibilities to find a holder to get the content. Meanwhile, since transmissions need to be accomplished in wireless networks, with more users eager to obtain the content, the physical transmission may suffer from higher interference, which directly affects the transmission efficiency. In view of the above factors, this question arises naturally: *What is the transmission time in evolving wireless social networks?* Moreover, what is the impact of structural parameters on transmission time? In order to answer the two questions, we are thus motivated to make the first attempt to analyze the transmission time in evolving networks.

In this paper, we focus on the specific transmission time, defined as the time when the content will be transmitted throughout the whole network. For a comprehensive analysis, we consider both static and mobile cases. To approach this problem, we consider the wireless social network coupled with the social layer and the physical layer. The coupled network evolves in the following sense: Every once in a while, a new user joins in, then it randomly chooses a geographic location and connects to some existing users according to the evolving model called Affiliation Networks [11], which was proven to capture the two aforementioned properties of real evolving

Chen Feng, Luoyi Fu, Xudong Wu, Xiaoying Gan, Xinbing Wang and Guihai Chen are with the School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai, 20040, China. Jun Xu is with the School of Computer Science, Georgia Institute of Technology.
Email: {fengchen, yiluofu, xudongwu, ganxiaoying, xwang8}@sjtu.edu.cn, gchen@cs.sjtu.edu.cn, jx@cc.gatech.edu.

[1]The diameter of a network is the greatest distance between any two nodes [10].

networks. The transmission scheme running on the network follows the idea that each interested user will search its social neighbors to find a holder within geographic range to request the content.

When approaching this problem, we are faced with several difficulties. First of all, most macroscopic properties of evolving social networks have been analyzed in [11], such as degree distribution, diameter. However, the microscopic property (the size of neighborhood within certain social hops) remains under-explored, which is proven to be critical in the derivation of transmission time. Second, in both static and mobile cases, even though corresponding transmission schemes are designed, it is even more challenging to determine the geographic range properly which efficiently satisfies all interested users. Third, in the mobile case, the acquisition of content becomes probabilistic, which not only demands exploration of the probability for users to obtain the content, but also raises hurdles in analyzing the basic case – transmission time for users of the same social hops to the source.

To address such difficulties, we begin with the analysis of basic properties of evolving social networks (e.g., user degree) and further reveal the desired microscopic property, which provides theoretical basis for parameter setting of schemes and derivation of transmission time. The geographic range in the scheme is set in a manner that all the interested users could efficiently obtain the content with theoretical guarantee. In both static and mobile cases, through a series of analysis, we managed to derive both the upper bound and lower bound of the transmission time over evolving networks. For comparison, we further look into the non-evolving mobile network, whose transmission time has not been unraveled. In both cases, we find that, with the same ability to search the neighborhood, the transmission time in evolving networks is smaller than non-evolving networks under our schemes, and the gap constantly increases over time, while the mobile case obtains such gain with smaller geographic range.

The contributions of this paper are summarized as follows:

- We make the first attempt to identify and comprehensively study the specific transmission time in evolving wireless social networks. To quantify the number of social neighbors exploited in the transmission, we theoretically bound the number of users within certain social distance, which turns out to be larger than non-evolving networks in order sense. And by this result we mathematically delineated the densification property of evolving social networks.
- In the static case, an efficient transmission scheme is designed with guarantee for each user to find a geographically near holder for request. Both upper bound and lower bound of transmission time are derived with the impact of structural parameters revealed. Thus, we managed to answer the questions proposed in the third paragraph. And we find that the transmission time of non-evolving networks is substantially larger than that of evolving networks where the ratio is a power of the network size.
- We initially study transmission time of evolving networks in the mobile case. The proposed scheme strikes a balance between the interference and the probability to meet a content holder. On this basis, we obtain closed-form transmission time in both evolving and non-evolving networks, by which we find that evolving networks still have smaller transmission time and the geographic range is smaller than the static case.
- Extensive experiments are carried out to validate our theoretical results on both synthetic and real networks. Three extra algorithms are included for comparison. And all algorithms are evaluated under various settings for a comprehensive evaluation. The results show that the transmission time in evolving networks is smaller than non-evolving counterparts, under all the tested datasets and network settings. Moreover, our proposed algorithms achieve better performance than the baselines.

The reminder of this paper is organized as follows. We review the works related to content transmission in Section II. In Section III, we present the model and formulate the problem. Section IV reveals the microscopic property of evolving social networks. The detailed analysis of transmission time in static networks and mobile networks is given in Section V and Section VI respectively. We discuss the experiment results in Section VII. And discussions are presented in Section VIII. Finally, we conclude the paper in Section IX. For readability, some proofs are deferred to the supplementary material [2].

II. Related Work

The research of content transmission could be traced back to the work of Frieze and Grimmet [3], which puts forward the problem that how many rounds are needed to transmit the content throughout the whole network. The problem is then progressively studied in different networks [4] [12], such as hypercubes, complete graphs and ER graphs. Till now, the past decades have witnessed a flourish of study on content transmission [2] [13] [14] [15].

Faced with the growing mobile traffic [16], wireless ad hoc networks, which are extensively studied [17] [18], have been widely applied to tackle this issue. Gupta *et al.* pioneered the study of network capacity to provide the bound of transmission rate [19]. Subsequently, [20] considered the mobile case. Besides the previous unicast scheme, multicast and motioncast are studied in [21] [22] and [23]. Routing schemes are also studied to support the transmission [24].

Since mobile devices are carried and operated by human beings, users are coupled not only in the physical domain, but also in the social domain due to the social relationship, forming wireless social networks. Various applications are realized in wireless social networks to cope with different scenarios, such as the pilgrimage platform [1], the applications BASA [25] and AdSocial [26]. [27] and [28] provide theoretical results for transmissions in wireless social networks. Ma *et al.* recognize the social factor in D2D caching and further propose an efficient caching strategy [29]. File sharing mechanisms are designed in [14] [13] to improve users' experience. The authors in [2] attempt to speed up the content transmission by exploiting the social relationship.

[2]The file could be found in https://tinyurl.com/y62popzc, https://github.com/Planet-B612/TW-Jun-19-0770 and https://jbox.sjtu.edu.cn/l/uoaCFb.

The underlying networks are shown to be evolving over time with new users joining and new edges coming into being [6] [7]. Empirical studies in [9] reveal that evolving networks exhibit distinctive properties: shrinking diameter and densifying edges. These properties are well-captured by the model proposed in [11], based on which several studies [30] [31] are conducted to investigate evolving networks.

In this work, we focus on the specific transmission time in evolving networks. Different from previous work on wireless social networks, such as [27] [28] and [2], we notice the evolving nature of real-world networks and characterize it by the state-of-the-art model – Affiliation Network [11]. Although the evolution model is also considered in [30] [31], our work is hardly related to them. Specifically, [30] did not explore the properties of Affiliation Network and aimed at the capacity of evolving networks instead of transmission time. While [31] just ignores the physical transmission and is not interested in the transmission performance. Contrarily, we consider both the physical and social layers and aim for the transmission time by studying transmission schemes and the microscopic properties of evolving social networks.

III. Model and Problem Formulation

We consider a network coupled with the physical layer and the social layer. Denote the users at time t as V_t with $|V_t| = n_t$. In the physical layer, the users are independently and uniformly distributed on a square of width $\sqrt{n_t}$. As the network evolves, the square will extend to accommodate new users with the physical density remaining constant 1, i.e., we adopt the extended network setting [2, 30, 32], which finds its application in various scenarios. For example, when a new user joins in a conference, he/she may be interested in different tracks and thus goes to a different venue. As a result, the network area is expanded. In the social layer, the social relationships are encoded by the undirected graph $G_t(V_t, E_t)$, where E_t records the social relationships. The neighbors of users $S \subseteq V_t$ are denoted as $N(S)$. A piece of content, e.g., a short video, is generated from the source with length F.

A. Communication Model

The content transmission takes place in the physical layer and is described by the well-known protocol model introduced in [19]. Let r_i denote the transmission range of user i and x_i denote the location of i. A transmission from user i could be successfully received by user j, if the following two constraints are met: (1) The geographic distance between user i and j is no more than r_i, namely, $|x_i - x_j| \leq r_i$. (2) For each user k which transmits simultaneously except i, it satisfies $|x_k - x_j| \geq (1 + \Delta) r_i$, where Δ ensures a guard zone to protect i from interference. And we adopt the TDMA (Time Division Multiple Access) scheme in the transmission, as in much previous work [19, 27, 32].

In each transmission, the transmission time is composed of two parts actually: propagation delay and the receiving time. The former is the time for the first bit to reach the receiver since it leaves the transmitter. The latter is the time needed to receive the last bit since the first bit arrives. Seeing that the electromagnetic wave travels at the speed of light, the former is negligible and we mainly consider the receiving time henceforth.

Algorithm 1: Evolution of $B_t(V_t, I_t)$

Input: Positive integers c_v, c_i; $\beta \in (0, 1)$; initial bipartite graph $B_0(V_0, I_0)$, with at least $c_v c_i$ edges, where each node in V_0 has at least c_v edges and each node in I_0 has at least c_i edges.

Output: A temporal graph $B_t(V_t, I_t)$

while $k < t$ **do**
- (Evolution of V_k)
- With probability β, a new node v is added to V_k;
- v chooses a prototype v' from V_k with probability proportional to the degree of v';
- v chooses c_v neighbors of v' uniformly at random without replacement and connects with them;
- (Evolution of I_k)
- With probability $1 - \beta$, a new node i is added to I_k;
- i connects to c_i nodes in I_k according to a symmetrical process;
- $k = k + 1$;

Return $B_t(V_t, I_t)$;

B. Model of Evolving Social Networks

The social layer is independent of the physical layer, although they share the same set of users. To depict the evolving social network, we adopt the Affiliation Networks model proposed in [11], which not only finds a good agreement with real-world evolving networks by capturing the edge densification phenomenon, shrinking/stabilizing diameter [9] and the well-received power-law degree distribution, but also provides mathematical tractability for further analysis.

Specifically, the affiliation network consists of two types of nodes: users (V_t) and groups (I_t). The former is affiliated to the latter. The affiliation relationship is viewed as the bipartite graph $B_t(V_t, I_t)$ naturally. On this basis, the social network $G_t(V_t, E_t)$ regarding users is generated from $B_t(V_t, I_t)$, since the social relationship among humans is closely related to the collectivity they belong.

At any time, the social network $G_t(V_t, E_t)$ is a one-to-one correspondence to $B_t(V_t, I_t)$ and shares the same node set. Specifically, for users u and v in V_t, there is an edge between u, v in $G_t(V_t, E_t)$ if they have a common neighbor in I_t, since acquaintanceships among people often stem from common affiliations. For example, u and v belong to the same club, then it is likely that they are socially connected through memberships. The intuition is that co-affiliation provides the conditions for the development of social ties, such as frequent contact [33] and physical proximity [34]. In view of such correspondence, it is sufficient to describe the evolving process of $B_t(V_t, I_t)$ in particular and the social network $G_t(V_t, E_t)$ could be accordingly derived.

In the evolving process of $B_t(V_t, I_t)$, at each time slot, a new user v joins V_t with probability β. It preferentially chooses a user v' from existing users as a prototype with probability proportional to the degree of v'. The intuition behind is that a user tends to follow the behavior of an influential user with high degree. Then, it uniformly and randomly selects c_v groups, which contain v', to join. With probability $1 - \beta$,

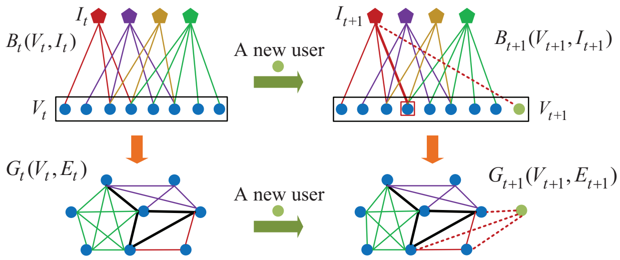


Fig. 1. An illustration of the evolving model. Different colors of edges in $B_t(V_t, I_t)$ indicate that users in V_t connect to different nodes in I_t. The black edge in $G_t(V_t, E_t)$ and $G_{t+1}(V_{t+1}, E_{t+1})$ indicates that it is a compound of several edges.

a new group u comes into being [3]. It randomly chooses a group u' from existing groups as a prototype with probability proportional to the degree of u'. Then, it uniformly and randomly selects c_i users, which join u', to recruit members. We assume that the initial graph $B_0(V_0, I_0)$ is connected. More details could be found in [11]. Based on the above description, the evolving process is summarized in Algorithm 1.

We summarize mathematical properties of the model as follows [11], which are experimentally verified in Section VII.

- After time ϕt ($\phi \in (0, 1)$ is a constant, indicating that the time is at the same scale of t), when $t \to \infty$, the diameter of the social network D_t is non-increasing, i.e., $D_t \geq D_{t+1}$.
- For the node $v \in V_t$ in $G_t(V_t, E_t)$ with degree larger than $(\log n_t)^{2+\epsilon}$ ($\epsilon > 0$ is a constant) and smaller than n_t^{γ}, when $t \to \infty$, its degree distribution follows $P(\text{degree of } v = k) = \frac{C}{k^{1+\tau}}$, where C is the normalization factor, constant $\tau = \frac{c_i(1-\beta)}{c_v\beta} < 1$, γ should be smaller than $\frac{1}{4+\tau}$ (specifically, we assume that γ is smaller than $\min\{\frac{1}{4+\tau}, \frac{1}{D_t}\}$ in this work).
- For the node $i \in I_t$ (resp. $u \in V_t$) in $B_t(V_t, I_t)$ with degree smaller than n_t^{γ} (resp. $n_t^{\frac{\tau}{1+4\tau}}$), its degree distribution follows $P(\text{degree of } i = k) = \frac{C}{k^{2+\tau}}$ (resp. $P(\text{degree of } u = k) = \frac{C}{k^{2+\frac{1}{\tau}}}$).

Remark. There are two time scales in our network model: the time slot for TDMA transmission and the time interval for the network evolution. We assume that the evolution interval is in a much larger scale than the transmission time slot. In other words, the social network structure remains the same when the content is being transmitted through the network. This assumption makes sense since, with the advancement of wireless communication technology, the transmission over users in an area only takes several hours, while it often takes several days before an evident change of the social network structure. And thus, for simplicity, we omit the subscript t in the following sections.

C. Problem Formulation

As the content is being transmitted through the network, a user will go through three states. (1) Initially, the user is *inactive*, meaning unaware of the content. (2) When one of its social neighbors has the content, the user becomes *eager* to hold it[4]. (3) After the user receives the content, it turns *active*. Especially, the creator of the content is called the *source*, denoted as s. The process is progressive, i.e., users can not change back to previous states.

A content of size F is generated by the source s, and the demand for it spreads through the social network with diameter D. Each user is permitted to explore its social neighbors within i hops (called social depth) and L distance (called geographic range). Let $S_k(s)$ denote the set of users whose social distance to the source s is within k hops. The time when all the eager users in $S_k(s)$ receive the content is denoted as T_k. Our goal is to figure out the specific time when all the eager users receive the content (i.e., T_D). If the social network is connected, it is equivalent to asking when all the users receive the content, since the demand will spread throughout the whole network.

Subsequently, we first quantify the neighborhood size of evolving networks, since social relations are exploited in the transmission. Then, we derive the transmission time in both static and mobile cases, where our emphasis lies.

IV. Analysis of Evolving Social Networks

To obtain the size of i-hop neighborhood, we begin with basic properties of evolving social networks. The neighborhood size turns out to be substantially larger than non-evolving networks. By the term "non-evolving networks", we mean networks without the advent of new users or formation of new edges. Specifically, when we are comparing results with "non-evolving networks", theoretical bounds in [2] are frequently used as the baseline. For the interest of space, the proofs of Lemmata 1, 3, 4 in this section are deferred to Sections I.A, I.B, I.C of the supplementary material respectively.

For ease of exposition, we introduce some notations here. The total degree of a set of users S in $G(V, E)$ (resp. $B(V, I)$) is denoted as $vol^G(S) = \sum_{u \in S} d^G(u)$ (resp. $vol^B(S) = \sum_{u \in S} d^B(u)$). Similarly, we define the sum of the k-th power of users' degree in $G(V, E)$ (resp. $B(V, I)$) as $vol_k^G(S) = \sum_{u \in S} \left[d^G(u)\right]^k$ (resp. $vol_k^B(S) = \sum_{u \in S} \left[d^B(u)\right]^k$).

Definition 1. *(Connected Network). A network G is called connected if it is non-empty and there exists at least one path between any two different nodes in G.*

Since the interest propagates along social networks, we first examine the connectivity of $G(V, E)$ to see whether all the users will become eager for the content. By Lemma 1, since $B_0(V_0, I_0)$ is assumed to be connected, we know that $G(V, E)$ is connected. Thus the transmission time is equivalent to the time when all users receive the content.

Lemma 1. *The social network $G(V, E)$ is connected if and only if the initial bipartite graph $B_0(V_0, I_0)$ is connected.*

[3]The affiliation model assumes non-decreasing users, which is reasonable in many applications, e.g., the active users of Facebook in constantly growing over the years [8].

[4]Considering practical scenarios, it is assumed that the change of state is instant. For example, in a public event, when a user happens to know something interesting via their social APPs, e.g., the agenda of the conference. The user would become eager to obtain such content immediately, otherwise he/she may miss important information.

To obtain the size of neighborhood, we need to explore the degree relationship between $G(V,E)$ and $B(V,I)$, and the connection probability of users, which are shown in Lemma 2 and Lemma 3 respectively.

Lemma 2. *For a user $u \in V$, the degree of u in $G(V,E)$ and $B(V,I)$ follow the equation*[5]

$$E\left[d^G(u)\right] = E\left[d^B(u)\right]\Theta(n^{\gamma(1-\tau)}). \quad (1)$$

Proof. Randomly select an edge from $B(V,I)$, let random variable Y_i denote the degree of the node i in I pointed by the edge. According to Lemma 4 in [11], we have $P(Y_i \geq k) = Ck^{-\tau}$, where C is a constant. Thus, the expectation of Y_i is

$$E[Y_i] = \sum_{k=1}^{n^\gamma} P(Y_i \geq k) = \sum_{k=1}^{n^\gamma} Ck^{-\tau} = \Theta(n^{\gamma(1-\tau)}). \quad (2)$$

Recall the generation process of $G(V,E)$ from $B(V,I)$. If an edge of user $u \in V$ in $B(V,I)$ points to a node $i \in I$, then the edge brings $Y_i - 1$ neighbors for u. Thus, we have $d^G(u) = \sum_{i=1}^{d^B(u)}(Y_i - 1)$. According to Wald's equation, the expectation of $d^G(u)$ is

$$\begin{aligned} E\left[d^G(u)\right] &= E\left[d^B(u)\right]E[Y_i - 1] \\ &= E\left[d^B(u)\right](n^{\gamma(1-\tau)} - 1) \\ &= E\left[d^B(u)\right]\Theta(n^{\gamma(1-\tau)}). \end{aligned} \quad (3)$$

Thus, we complete the proof. □

It is easy to obtain Corollary 1 below by multiplying the number of users $|V|$ in both sides of Equation 1, where the total degree is evaluated in an average level.

Corollary 1. *There is a gap of $\Theta(n^{\gamma(1-\tau)})$ between the total degree of V in $G(V,E)$ and that in $B(V,I)$, i.e., $E[vol^G(V)] = E[vol^B(V)]\Theta(n^{\gamma(1-\tau)})$.*

We next derive the probability that two users in $G(V,E)$ are connected by Lemma 3.

Lemma 3. *Consider two users $v, w \in V$ with degrees $d^B(v)$ and $d^B(w)$ in $B(V,I)$ respectively, they are connected in $G(V,E)$ with probability*

$$P(v,w) = \frac{d^B(v)d^B(w)vol_2^B(I)}{n^2\bar{d}^2}, \quad (4)$$

where $\bar{d} = E[d^B(u)]$, $u \in V$.

Before deriving the neighborhood size, we present Lemma 4, which states the increase ratio of the sum of degrees between $S \subseteq V$ and $N(S)$ in $G(V,E)$.

[5]In this equation, we adopt the notation Θ, which is one of the Knuth notations. We give the definition of three Knuth notations frequently used in this paper. Let f and g be two non-negative real valued functions defined on natural numbers $\mathbb{N}$, $n \in \mathbb{N}$, then

$$f(n) = O(g(n)) \Leftrightarrow \limsup_{n\to\infty} \frac{f(n)}{g(n)} < \infty, f(n) = \Omega(g(n)) \Leftrightarrow \liminf_{n\to\infty} \frac{f(n)}{g(n)} < \infty,$$
$$f(n) = \Theta(g(n)) \Leftrightarrow f(n) = O(g(n)) \text{ and } g(n) = O(f(n)).$$

Lemma 4. *For the social network $G(V,E)$, given two subsets $S \subseteq V, T \subseteq V$ and $S \cap T = \emptyset$, if*

$$\frac{2c\,vol_3^G(T)}{\varepsilon^2(vol_2^G(T))^2}\frac{n^{2+2\gamma(1-\tau)}\bar{d}^2}{vol_2^B(I)} \leq vol^G(S), \quad (5)$$

$$vol^G(S) \leq \frac{\varepsilon\, vol_2^G(T)}{vol_3^G(T)}\frac{n^{2+2\gamma(1-\tau)}\bar{d}^2}{vol_2^B(I)}, \quad (6)$$

then, with probability $1 - e^{-c}$ we have

$$vol^G(N(S) \cap T) \geq (1-2\varepsilon)\frac{vol_2^G(T)vol^G(S)vol_2^B(I)}{n^{2+2\gamma(1-\tau)}\bar{d}^2}. \quad (7)$$

Based on the existing results, we are ready to study the number of users in $S_i(u)$. Let $N_i(u)$ denote the i-hop neighbors of u, it is easy to see that $S_i(u) = \bigcup_{k=0}^{i} N_k(u)$.

Theorem 1. *In the social network $G(V,E)$, for an arbitrary user u, $|S_i(u)|$ is bounded by*
(1) $|S_i(u)| \leq n^{\frac{i}{D}}$;
(2) $|S_i(u)| \geq \mu n^{\gamma(i-\tau)}$ with probability $1 - o\left(\frac{1}{n}\right)$;
where $i \leq D-1$ and μ is a constant related to τ and i.

Proof. (1) We prove the first case by induction.

(i) When $i = 0$, $S_i(u) = S_0(u)$ is u itself, so $|S_i(u)| = 1$. The right side is $n^0 = 1$. Thus, the inequality holds. When $i = 1$, $|S_i(u)|$ is the degree of u plus 1. Since the largest degree in the network is $n^\gamma < n^{\frac{1}{D}}$, the inequality holds.

(ii) Assume $|S_i(u)| \leq n^{\frac{i}{D}}$, the goal is to verify $|S_{i+1}(u)| \leq n^{\frac{i+1}{D}}$, where $|S_{i+1}(u)| = |S_i(u)| + |N(S_i(u))|$. Consider the best situation where each user in $S_i(u)$ has n^γ neighbors in $V \setminus S_i(u)$ and arbitrary two users have no common neighbor. Note that $\gamma < \frac{1}{D}$, then $|S_{i+1}(u)| \leq |S_i(u)| + n^\gamma n^{\frac{i}{D}} \leq n^{\frac{i}{D}} + n^{\frac{i}{D}+\gamma} < n^{\frac{i+1}{D}}$. Thus, $|S_{i+1}(u)| \leq n^{\frac{i+1}{D}}$.

Summarizing (i) and (ii), we can draw the conclusion that $|S_i(u)| \leq n^{\frac{i}{D}}$.

(2) To prove the lower bound, we first try to obtain the increase ratio of neighborhood between adjacent layers by proving $vol(N_{i+1}(u)) \geq (1-2\varepsilon)n^\gamma vol(N_i(u))$ with Lemma 4. Then, we could accordingly derive the size of $S_i(u)$.

To apply Lemma 4, we have to determine the parameters and verify the two inequalities. Specifically, we set $c = 2\log n, S = N_i(u), T = T_i(u) = V \setminus S_i(u)$ and $\varepsilon = 1/4$. It is easy to see that $N(S) \cap T = N_{i+1}(u)$. We proceed to show that $T_i(u) \approx V$. Since $|S_i(u)| \leq n^{\frac{i}{D}}$, we have $|T_i(u)| = n - |S_i(u)| \geq n - n^{\frac{i}{D}} = n\left(1 - n^{\frac{i}{D}-1}\right)$. Due to $i \leq D-1$, we have $|T_i(u)| = n(1-o(1))$. Thus, $T_i(u) \approx V$. Then, we validate the two conditions in Lemma 4.

For Inequality 5, we have

$$\begin{aligned} &\frac{2c\,vol_3^G(T_i(u))}{\varepsilon^2(vol_2^G(T_i(u)))^2}\frac{n^{2+2\gamma(1-\tau)}\bar{d}^2}{vol_2^B(I)} = \\ &\frac{2c\bar{d}^2 n^{2+2\gamma(1-\tau)}\int_{\log^{2+\epsilon}n}^{n^\gamma} n\frac{C}{x^{1+\tau}}x^3dx}{\varepsilon^2\left[\int_{\log^{2+\epsilon}n}^{n^\gamma} n\frac{C}{x^{1+\tau}}x^2dx\right]^2\int_1^{n^\gamma} n\frac{C}{x^{2+\tau}}x^2dx} \\ &= \frac{2c\bar{d}^2 n^{2+2\gamma(1-\tau)}nn^{\gamma(3-\tau)}}{\varepsilon^2 n^2 n^{2\gamma(2-\tau)}nn^{\gamma(1-\tau)}} = \frac{2c\bar{d}^2}{\varepsilon^2} = \Theta(\log n). \end{aligned} \quad (8)$$

Since all users in $G(V,E)$ have a degree greater than $\log^{2+\epsilon} n$, we have $vol^G(N_i(u)) > \Theta(\log n)$. Thus, the first condition is validated.

For Inequality 6, it follows that

$$\frac{\varepsilon vol_2^G(T_i(u))}{vol_3^G(T_i(u))}\frac{n^{2+2\gamma(1-\tau)}\bar{d}^2}{vol_2^B(I)} = \frac{\varepsilon \bar{d}^2 n^{2+2\gamma(1-\tau)}\int_{\log^{2+\epsilon}n}^{n^\gamma} n\frac{C}{x^{1+\tau}}x^2dx}{\int_{\log^{2+\epsilon}n}^{n^\gamma} n\frac{C}{x^{1+\tau}}x^3dx\int_1^{n^\gamma} n\frac{C}{x^{2+\tau}}x^2dx}$$
$$= \frac{\varepsilon \bar{d}^2 n^{2+2\gamma(1-\tau)} n n^{\gamma(2-\tau)}}{n n^{\gamma(3-\tau)} n n^{\gamma(1-\tau)}} = \varepsilon \bar{d}^2 n^{1-\gamma\tau}. \tag{9}$$

Since the number of users in $N_i(u)$ is upper bounded by $n^{\frac{i}{D}}$ and the expected degree in $G(V,E)$ is $n^{\gamma(1-\tau)}$. By the definition of $vol(\cdot)$, we have that the value of $vol^G(N_i(u))$ is upper bounded by $n^{\frac{i}{D}}n^{\gamma(1-\tau)}$. Since $i \le D-1$ and $\gamma < \frac{1}{D}$, we have $vol^G(N_i(u))$ is upper bounded by $\varepsilon \bar{d}^2 n^{1-\gamma\tau}$. Thus, the second condition is verified.

Since Inequalities 5 and 6 are satisfied, we have

$$vol^G(N_{i+1}(u)) \ge (1-2\varepsilon)\frac{vol_2^G(T_i(u))vol^G(N_i(u))vol_2^B(I)}{n^{2+2\gamma(1-\tau)}\bar{d}^2}$$
$$= \frac{(1-2\varepsilon)vol^G(N_i(u))\int_1^{n^\gamma} n\frac{C}{x^{2+\tau}}x^2dx}{n^{2+2\gamma(1-\tau)}\bar{d}^2}\int_{\log^{2+\epsilon}n}^{n^\gamma} n\frac{C}{x^{1+\tau}}x^2dx$$
$$= (1-2\varepsilon)vol^G(N_i(u))n^\gamma/\bar{d}^2. \tag{10}$$

Recall that $\varepsilon = 1/4$ and $c = 2\log n$, the sum of degrees of users i hops away from u is at least $n^\gamma/(2\bar{d}^2)$ times that of users $i-1$ hops away, with probability $1-o\left(\frac{1}{n}\right)$. The expected degree of $N_0(u)$ (i.e., u), is $\sum_{k=1}^{n^\gamma} k\frac{C}{k^{1+\tau}} = \Theta(n^{\gamma(1-\tau)})$. Then, $vol^G(N_{i-1}(u)) \ge \Theta(n^{\gamma(i-\tau)})$. Following similar arguments in proof of Lemma 12 in [2], we derive that there are at least $\mu n^{\gamma(i-\tau)}$ users in $S_i(u)$, where μ is a constant related to τ and i. This completes the proof. □

With the same social depth i, based on the results in [2], we can derive the value of $|S_i(u)|$ in non-evolving networks, which is about $n^{\frac{i}{2\log n}}$, a value substantially smaller than evolving networks. The explanation is that, due to the shrinking diameter and edge densification, connections in evolving social networks become much closer and expectedly result in a larger neighborhood.

After deriving the structural properties as the basis, we next proceed to the main part – transmission analysis in both static and mobile cases.

V. Content Transmission in Static Networks

In this section, we focus on static networks, where users' location will not change once they are located (i.e., static). We first present the transmission scheme and then analyze the time needed to disseminate the content throughout the whole network.

A. The Transmission Scheme

The scheme exploits both the social relationship and geographic information, which consists of users' behavior at each time slot and the transmission strategy.

Inactive users: Inactive users will become eager once one of its social neighbors has the content, e.g., a user would be stimulated when learning about the content via social applications.

Eager users: Eager users (e.g., v) will try to request the content from legitimate active users, denoted as $H(v;i,L)$, which satisfy two constraints. (1) Social constraint: any user in $H(v;i,L)$ should be within i hops of v in $G(V,E)$. (2) Physical constraint: for any user in $H(v;i,L)$, its Euclidean distance to v should be within geographic range L in the network area. If $H(v;i,L)$ is not empty, then v randomly selects one from $H(v;i,L)$ to send a request (called $\mathcal{N}$-request). If $H(v;i,L)$ is empty and v is in $S_{\lceil \frac{i}{2} \rceil}(s)$, v randomly selects a one-hop active neighbor and sends a request (called $\mathcal{R}$-request). Otherwise, v will wait for the next time slot.

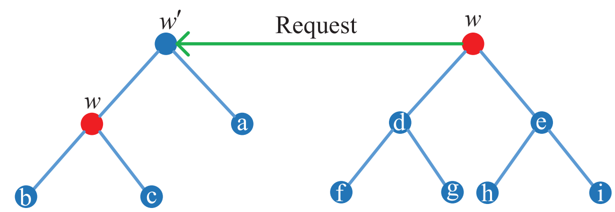


Fig. 2. An illustration of the request tree.

Active users: Consider an active user w, since the degree distribution is power law, it may be crowded with requests. Severe congestion and delay will occur if w has to respond all the requests. Thus, load balancing is needed. User w will sequentially add the received requests into a balanced binary tree (called the request tree) with itself being the root, as illustrated in Fig. 2. Then, user w will serve its children d and e, i.e., the parent will serve its two children. Note that, user w is initially inactive and receives the content by requesting another active user w'. Hence, user w is also in the request tree of w'. And it has to serve another two children b and c. In this way, user w only has to serve four users at most (b, c, d and e). For each user becoming active, its behavior will follow the same mechanism. We adopt a TDMA transmission mechanism. The time slot is divided into four mini-slots to serve the four users respectively.

A transmission is called an $\mathcal{N}$-transmission if both the transmitter and the receiver get the content by sending $\mathcal{N}$-requests. Otherwise, it is called an $\mathcal{R}$-transmission.

Transmission strategy: To transmit the content from an active user to the designated eager user, we adopt the highway system proposed in [32] which enjoys favorable properties, such as achieving tight transmission bound, compatible to both extended and dense networks. And we would like to introduce it from three aspects: formation, mechanism and properties. For more details, the readers may refer to [32].

Formation: The area is divided into non-overlapping squares of constant length c, as shown in Fig. 3. Each square is said to be open with probability $p = 1 - e^{-c^2}$ (i.e., at least one node inside it). We draw a blue horizontal line across half of the squares and a vertical line across the others. Each edge is also open with probability p. For c large enough, there are numerous horizontal and vertical paths composed of open edges crossing the whole area (e.g., green lines in Fig. 3). For each open edge, we select one node in the square as a relay. Replacing each edge with the relay node, we obtain chains of nodes, which form the highway system.

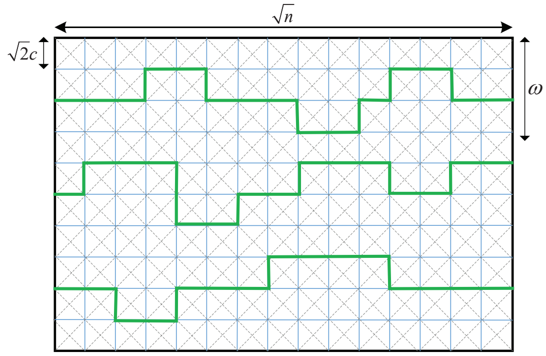


Fig. 3. The highway system.

Mechanism: For each transmission, the source first drains the content into the highway. Then, the content is transmitted along horizontal and vertical highways. Finally, the content is delivered from the highway to the destination.

Properties: We mainly apply three properties of the highway system, i.e., (1) Assuming a large constant c, the highways are almost straight lines; (2) The transmission rate between two relays is constant R; (3) Each relay only serves nodes within a slab of constant width ω.

Algorithm 2: The Transmission Scheme in Static networks

```
Input: Social depth i, geographic range L, social network
       G(V,E), the source s
while True do
    for each inactive user u do
        if one of users in N(u) is active then
            u becomes eager;
    for each eager user v do
        if H(v;i,L) ≠ ∅ then
            Request a random user in H(v;i,L);
        else if v ∈ S_⌈i/2⌉(s) then
            Request a random one-hop active neighbor;
        else
            Wait for the next time slot;
    for each active user w do
        Add all requests received into the binary tree;
        Route the content to its children with the highway
          sysytem;
    if no packet is transmitting then
        break;
```

Based on the above description, the pseudo-code is given in Algorithm 2 and an illustrative example could be found in Fig. 4. Social depth i is an integer with $i \in [1, D-1]$. According to the value of i, the geographic range L is set to be $8\sqrt{n^{1-\gamma(\lfloor i/2 \rfloor - \tau)}\log n/\pi\mu}$, such that each user could find at least one neighbor within L for request. Intuitively, with a larger neighborhood as i increases, it is easier to find a geographically near neighbor, and thus L is smaller. The rigorous derivation of L is given in Lemma 5.

Note that L and r_u are two different parameters, where L is the geographic searching range of users, while r_u is the transmission range of a user. For example, an eager user v attempts to find a holder for request. The holder should be within L. Suppose user w responds to the request and a user in w's tree is arranged to serve v via the highway system. Then, the relays (e.g., u) in the highway communicate with each other according to the protocol model, with r_u being the transmission range.

B. Transmission Time in Static Networks

Before proceeding to the transmission time, we first explain the rationality behind the value of L. According to the social distance to the source, we divide the users into two groups, i.e., $S_{\lceil \frac{i}{2} \rceil}(s)$ and $V \backslash S_{\lceil \frac{i}{2} \rceil}(s)$. For users in $S_{\lceil \frac{i}{2} \rceil}(s)$, according to the scheme, they will first try to find a holder within i and L to send an $\mathcal{N}$-request. If failed, they will send an $\mathcal{R}$-request to get the content. In either way, they will finally obtain the content. For users in $V \backslash S_{\lceil \frac{i}{2} \rceil}(s)$, we prove that they will obtain the content with high probability by Lemma 5.

Lemma 5. *Under the transmission scheme in Algorithm 2, let* $L = 8\sqrt{n^{1-\gamma(\lfloor i/2 \rfloor - \tau)}\log n/\pi\mu}$, *the content will be held by all the users in* $V \backslash S_{\lceil \frac{i}{2} \rceil}(s)$ *with probability* $1 - o\left(\frac{1}{n}\right)$.

Proof. According to the scheme, each request will be satisfied via the highway system after its parent becomes active. Thus, we only need to prove that users can find at least one user within i and L, which is closer to the source, to request the content.

To this end, we first consider how to set a distance l such that when a user has m neighbors available, at least one neighbor could be found within l. Then, we substitute m with the number of neighbors which are closer to the source and accordingly obtain L.

Consider a user w with m neighbors, let X_k be the indicator variable that the distance between w and the k-th user is within l. Considering the edge effect where the user is near the periphery of the square and that a range greater than the width is unnecessary, we see that the probability of $X_k = 1$ is always greater than $\frac{\pi l^2}{4n}$. Let $X = \sum_{k=1}^{m} X_k$ denote the number of users within l to w, we have $E[X] = \sum_{k=1}^{m} E[X_k] \geq \frac{\pi m l^2}{4n}$. If we set $l = 8\sqrt{n \log n/\pi m}$, by Lemma 2.1 in [35], there are at least $8 \log n$ users within l among the m users with probability $1 - o\left(\frac{1}{n}\right)$.

For a user u in $V \backslash S_{\lceil \frac{i}{2} \rceil}(s)$, we next quantify the number of its neighbors closer to the source, i.e., m. As illustrated in Fig. 5, we assume that u is h ($h > \lceil \frac{i}{2} \rceil$) hops away from the source s, then along the h-hop path there must exist a user w that is $\lceil \frac{i}{2} \rceil$ hops away from u and $h - \lceil \frac{i}{2} \rceil$ hops away from the source. According to Lemma 1, there are at least $\mu n^{\gamma(\lfloor i/2 \rfloor - \tau)}$ users in $S_{\lfloor i/2 \rfloor}(w)$, who are also in $S_i(u)$ exactly. If we replace m by the lower bound of $S_i(u)$, i.e., $l = L = 8\sqrt{n^{1-\gamma(\lfloor i/2 \rfloor - \tau)}\log n/\pi\mu}$, there is at least one user in $S_{\lfloor i/2 \rfloor}(w)$ that is within L of u. Moreover, the social distance of users in $S_{\lfloor i/2 \rfloor}(w)$ to the source is at most $h - \lceil \frac{i}{2} \rceil + \lfloor \frac{i}{2} \rfloor$, which is smaller than h. Hence, each user $u \in V \backslash S_{\lceil \frac{i}{2} \rceil}(s)$ could find at least one user within i and L, which is closer to the source, to send $\mathcal{N}$-request and get the content.

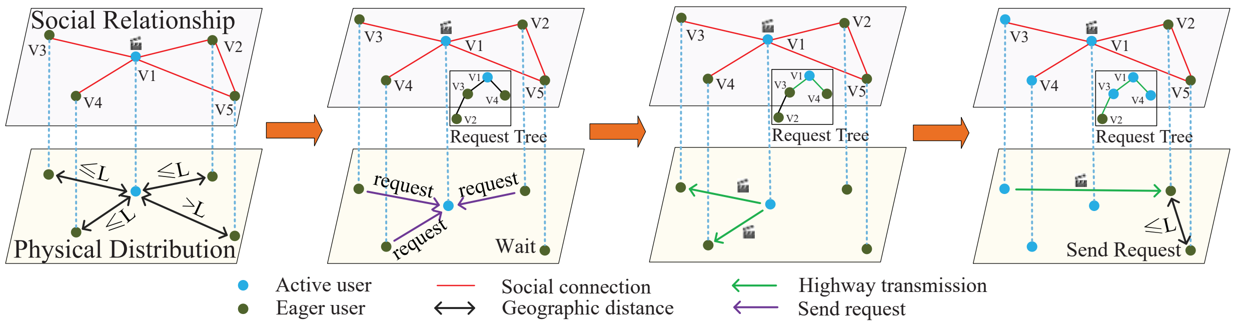


Fig. 4. An illustration of the transmission scheme with $i = 1$ and V1 being the source. Initially, V2-5 will become eager because of having an active neighbor V1. V2-4 then send requests to V1 forming a request tree, since $H(V2, V3, V4; i, L) = V1$. However, V5 can not send the request to V1 since $H(V5; i, L) = \emptyset$. Then, V1 will serve its direct children V3, V4, and V3 has to serve its child V2. After V2 obtains the content, V5 will find that $H(V5; i, L) = V2$. Thus, V5 will send a request to V2.

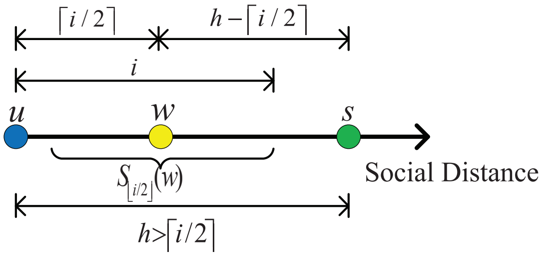


Fig. 5. An illustration of social distances.

Combining the two parts, we prove that the content will finally be held by all the users. □

According to the transmission scheme, $\mathcal{R}$-requests are only sent by users in $S_{\lceil \frac{i}{2} \rceil}(s)$, while all other users send $\mathcal{N}$-requests. Thus, the amount of $\mathcal{R}$-requests is upper bounded by $|S_{\lceil \frac{i}{2} \rceil}(s)|$, which is smaller than $n^{\lceil i/2 \rceil / D}$. Since each user who sends $\mathcal{R}$-request will serve at most four users, the number of $\mathcal{R}$-transmissions is bounded by $O(n^{\lceil i/2 \rceil / D})$ as well, which is negligible compared with that of $\mathcal{N}$-transmissions. Thus, we will mainly consider the $\mathcal{N}$-transmission.

We next derive the transmission rate from the amount of transmissions. According to the transmission scheme, the transmitter and the receiver of a transmission are in the same request tree. And their geographic distances to the root node are no bigger than L. Then, their distance is no bigger than $2L$. Furthermore, since the highway is almost straight, the transmission passes through a relay on the horizontal (vertical) highway only if the horizontal (vertical) distance between the relay and transmitter (receiver) is no bigger than $2L$. Thus, transmissions passing through a relay must fall in a rectangle of area $2L \times \omega$. That is to say, relays will only serve the transmissions in a rectangle of area $O(L)$. To see how many users a relay has to serve, we derive the number of users in such an area by Lemma 6.

Lemma 6. *Suppose n users are independently and uniformly distributed on a square of width $\sqrt{n}$, consider a region of area $A = \omega(\log n)$ in the square, the number of users located in this region is bounded by $\Theta(A)$ with high probability.*

Proof. Please refer to Section I.D of the supplementary material. □

Moreover, each user establishes at most four transmissions. Thus, the relay only has to serve at most $O(L)$ transmissions. Then, the bit rate of each transmission is $\Omega(R/L)$. Thus, each transmission takes at most cLF/R time slots to finish, where c is a constant. Recall that $|S_i(u)| = m$ is considerably large in evolving social networks, the geographic range L is a relatively small value. Thus, the transmission rate $\Omega(R/L)$ is particularly high in evolving networks.

Finally, we derive the upper and lower bound of transmission time by Theorems 2 and 3.

Theorem 2. *Suppose the social depth is i, apply the transmission scheme in the evolving social network with n users, the transmission time is upper bounded by*

$$O\left(DF\sqrt{n^{1-\gamma(\lfloor i/2 \rfloor - \tau)}}\log^{\frac{3}{2}} n/R\right)$$

with high probability.

Proof. Let T_k denote the time when all users in $S_k(s)$ get the content. We try to prove $T_k \leq kc\log nLF/R$ by induction, then the transmission time T_D could be derived by letting $k = D$.

(i) When $k = 0$, $S_k(s) = s$, i.e., the source itself, thus $T_k = 0$. The inequality holds.

When $k = 1$, T_1 is the time when all the neighbors of the source s get the content. According to the scheme, the neighbors will all be in the request tree of s. Since the degree of s is at most n^γ, the last neighbor will get the content after at most $\log(n^\gamma)$ transmissions. Recall that each transmission takes at most cLF/R time slots, we have $T_1 \leq c\log(n^\gamma)LF/R < c\log nLF/R$.

(ii) We assume that $T_{k-1} \leq (k-1)c\log nLF/R$ holds. At time T_{k-1}, all the users in $S_{k-1}(s)$ have the content, and all the users in $N_k(s)$ become eager at the same time. According to the proof of Lemma 5, users are able to find a holder closer to the source to send requests. Thus, they all join the binary tree of users in $S_k(s)$. When all the users in $N_k(s)$ obtain the content, we reach time T_k, i.e., T_k is determined by the user who gets the content latest. The height of request trees is at most $\log n$, so the latest user has to wait at most $\log n$ transmissions to get the content, while each transmission takes cLF/R time slots. Thus, $T_k \leq T_{k-1} + c\log nLF/R \leq kc\log nLF/R$.

Combining (i) and (ii), we can draw the conclusion that $T_k \leq kc\log nLF/R$. Let $k = D$ and replace L, we derive time $T_D = O\left(DF\sqrt{n^{1-\gamma(\lfloor i/2 \rfloor - \tau)}}\log^{\frac{3}{2}} n/R\right)$, the time when the whole network gets the content. □

Theorem 3. *The lower bound of the transmission time is*

$$\Omega\left(F\sqrt{n^{1-\frac{2i}{D}-\eta}}/R\right),$$

with high probability, where $\eta > 0$ is an arbitrary constant.

Proof. We obtain the lower bound of transmission time by analyzing the bit-meter product. For each transmission, the bit-meter product is defined as the product of the distance it travels and the bits it conveys. The total bit-meter product is the sum of the bit-meter product of all transmission pairs. To derive a lower bound of transmission time, we could divide the total bit-meter product by the maximal bit-meter rate.

To this end, we first derive the bit-meter rate, i.e., the bit-meter product transmitted within one time slot. According to the formation of the highway system, the distance between two adjacent relays is bounded by a constant $\sqrt{5}c$ and their transmission rate is constant R. Assuming all the users are transmitting or receiving packets simultaneously, we obtain the maximal bit-meter product the network could transmit in one time slot, which is $O(Rn)$.

We then proceed to derive the lower bound of the total bit-meter product. We further optimistically assume that users obtain content from the nearest neighbor. According to the request tree, a user u will receive the content from a holder at most $2i$ hops away in the social network. Let $X_u = 1$ denote that the geographic distance from u to $S_{2i}(u)$ is smaller than $L' = \sqrt{n^{1-\frac{2i}{D}}/\pi\log n}$. Then, we have

$$\begin{aligned} P(X_u = 1) &= 1 - P(X_u = 0) \\ &= 1 - \left(1 - \frac{\pi(L')^2}{n}\right)^{|S_{2i}(u)|} \\ &= |S_{2i}(u)|/\left(n^{\frac{2i}{D}}\log n\right) < \frac{1}{\log n}. \end{aligned}$$

The expectation of the number of users having a neighbor within L' is $E\left[\sum_{u\in V} X_u\right] \le n/\log n$. By Hoeffding's inequality, we further have $P\left(\left|\sum_{u\in V} X_u - E\left[\sum_{u\in V} X_u\right]\right| > \sqrt{n\log n}\right) < o\left(\frac{1}{n}\right)$.

We only consider the bit-meter product of users whose transmitter is more than $\sqrt{n^{1-\frac{2i}{D}}/\pi\log n}$ away from them, to guarantee a lower bound. Since more than $n - n/\log n$ users are taken into account, the total bit-meter product is $\Omega\left(F\sqrt{n^{3-\frac{2i}{D}-\eta}}\right)$, where η is an arbitrary positive constant. Recall that the bit-meter product the network transmits in one time slot is $O(Rn)$, thus the transmission time is $\Omega\left(F\sqrt{n^{1-\frac{2i}{D}-\eta}}/R\right)$. □

Remark. We further make a comparison with that the non-evolving network, whose transmission time is proven to be $O(\sqrt{n^{1-\epsilon}}\log^{2.5} nF)$ under social depth $i = 2\epsilon\log n + 1$ [2]. Then, given the same social depth i, the ratio of transmission time between non-evolving and evolving networks is $n^{\frac{\gamma}{2}(\lfloor i/2\rfloor-\tau)-\frac{i-1}{4\log n}}D/\log n$, a positive power of n increasing as the network evolves [6]. Similar conclusions also hold in the lower bound. *Therefore, we obtain a smaller transmission time in evolving networks than non-evolving counterparts.*

Intuitively, in our scheme, due to the larger neighborhood in evolving networks, it is more probable for users to find a legitimated user to request, which facilitates the transmission. To explain, with the same probability to find a neighbor within L, the geographic constraint L in evolving networks could be smaller due to the large neighborhood. Recall that the bit rate for each transmission pair is inversely proportional to L (i.e., $1/L$). Thus, the transmission rate in evolving networks is much higher, implying a smaller transmission time.

VI. Content Transmission in Mobile Networks

Since in some scenarios users tend to be mobile during the transmission (e.g., business district), for a comprehensive analysis, in this section we study the case where users are mobile. Just like the static case, we consider both non-evolving networks and evolving networks to reveal the impact of evolution.

A. Mobile Settings

Except the mobility and transmission strategy, the basic settings are the same as the last section, including the communication model and social networks. The users move according to the two-dimensional i.i.d. mobility model [7]:

The time is slotted with each time slot enough to transmit F bits. At each time slot, the users are uniformly and randomly distributed in the square. Their positions are independent of each other and for any user, its positions in different time slots are independent as well. That is to say, users are totally shuffled at each time slot. Notice that the time slot of mobility is the same as the communication model, which ensures a successful transmission before the positions of transceivers change.

The transmission scheme in mobile networks inherits the framework of the static one. Algorithm 3 defines the way users obtain the content. Specifically, at each time slot, users move according to the i.i.d. mobility model. The inactive users will become eager if one of its neighbors has the content. For each eager user, say v, it will search its active social neighbors within i hops and geographically within L, which form the set $H(v; i, L)$. If there is no transmitter within $(1 + \Delta)L$ and $H(v; i, L)$ is not empty, v randomly selects one user from $H(v; i, L)$ and sends a request. For active users, they will establish a transmission with the eager user directly upon receiving the request. Otherwise, user v has to wait for the next time slot. In the wireless communication, the transmission range r_u of user u is set to be the geographic range of the user it serves. Note that, here we do not need to care about the case where an eager user can not find an active user to request at

[6]In the derivation of L, when $i = 1$, we replace the number of users in $S_{\lfloor i/2\rfloor}(w)$ with $\mu n^{-\gamma\tau}$. In fact, $S_{\lfloor i/2\rfloor}(w)$ should be at least 1. Thus, the ratio of $i = 1$ should be $\log^2 n$. In the mobile case, such argument makes sense as well.

[7]For mathematical tractability, we adopt the i.i.d. model, which is widely applied in mobile networks, such as [20] [23] [36].

each time slot. Because it is always possible for an eager user to obtain the content at next time slot.

Algorithm 3: The Transmission Scheme in Mobile Networks

Input: Social depth i, geographic range L, social network $G(V,E)$, the source s

while *True* **do**
  for *each inactive user u* **do**
    if *one of users in $N(u)$ is active* **then**
      u becomes eager;
  for *each eager user v* **do**
    if *$H(v;i,L)\neq\emptyset$ and no transmitter is within $(1+\Delta)L$* **then**
      Request a random user in $H(v;i,L)$;
    else
      Wait for the next time slot;
  for *each active user w* **do**
    Transmit content to the eager user which requests w;
  if *no eager user exists* **then**
    break;

In the mobile case, the social depth i takes value in $[1,D-1]$ as well. As for the geographic range L, its value is a little more delicate, related to the social distance to the source. For users in $S_{\lceil\frac{i}{2}\rceil}(s)$, where s is the source, they adopt the same geographic range $L_c=\sqrt{n\Big/\pi(1+\Delta)\sqrt{|S_{\lceil\frac{i}{2}\rceil}(s)|\log n}}$. For users out of $S_{\lceil\frac{i}{2}\rceil}(s)$, e.g., user u is $h>\lceil\frac{i}{2}\rceil$ hops away from the source, it is allowed to search users at most $L_r=\sqrt{n\Big/\pi(1+\Delta)\sqrt{|S_{\lfloor\frac{i}{2}\rfloor}(w)||N_h(s)|\log n}}$ away. The geographic range is set such that the probability of interference is minimized, i.e., $o(1)$. The rigorous derivation could be found in the proof of Theorem 4.

B. Transmission Time in Evolving Networks

We first derive the upper bound of transmission time by calculating and summing the time of each layer of users. To this end, we first study a primary case in Lemma 7, where there are K users to be served. And the proof of Lemma 7 as well as the subsequent theorem on transmission time could be found in Section I.E and I.F of the supplemental file respectively.

Lemma 7. *Suppose the probability for each user to obtain the content is p at each time slot, where p is a function of n and tends to 0 as $n\to\infty$, the number of time slots needed for K eager users to become active is upper bounded by*

$$\frac{\log K\log n}{p}$$

with probability $1-o(1)$ when $n\to\infty$.

With the above results, we continue to bound the transmission time in the whole network.

Theorem 4. *In the evolving mobile network, under the transmission scheme in Algorithm 3, the transmission time over the whole network is upper bounded by*

$$O\left(4(1+\Delta)\log^{2.5}n\sqrt{n^{1-\gamma(\lfloor i/2\rfloor-\tau)}D/\mu}\right)$$

with probability $1-o(1)$.

Here we would like to explain the reason behind the setting of L_r and L_c. By intuition, with a larger L_r (resp. L_c), more users will be within reach, while the interference between users will be severer. Formally speaking, setting L_r (resp. L_c) large increases the probability of meeting an active neighbor (denoted as p_1), while decreases the probability of not being interfered (denoted as p_2). Thus the setting of L_r (resp. L_c) is a tradeoff between the two contradictory probabilities, which raises the probability to meet an active neighbor while prevents interference (p_2 is $1-o(1)$).

Based on the idea that transmission consumes area, we derive the lower bound of transmission time in Theorem 5.

Theorem 5. *The lower bound of transmission time in evolving mobile networks under Alg. 3 is*

$$\Omega\left(\frac{\Delta^2}{16(1+\Delta)}\sqrt{n^{1-\lfloor\frac{i}{2}\rfloor/D}/\log n}\right).$$

Proof. Consider the protocol model illustrated in Fig. 6. For any two transmissions, say i to j and k to l, according to the constraints of the communication model, we have $|x_i-x_j|\le L_j$, $|x_k-x_l|\le L_l$, $|x_k-x_j|\ge(1+\Delta)L_j$ and $|x_i-x_l|\ge(1+\Delta)L_l$, where L_j and L_l denote the geographic range of j and l respectively. Since users i,j,l form a triangle, according to the triangle inequality, we have

$$|x_j-x_l|\ge|x_i-x_l|-|x_i-x_j|\ge(1+\Delta)L_l-L_j. \tag{11}$$

Similarly, users i,j,l form a triangle, thus

$$|x_j-x_l|\ge|x_k-x_j|-|x_l-x_k|\ge(1+\Delta)L_j-L_l. \tag{12}$$

Summing Inequalities 11 and 12, we have

$$|x_j-x_l|\ge\frac{1}{2}\left(\Delta L_j+\Delta L_l\right). \tag{13}$$

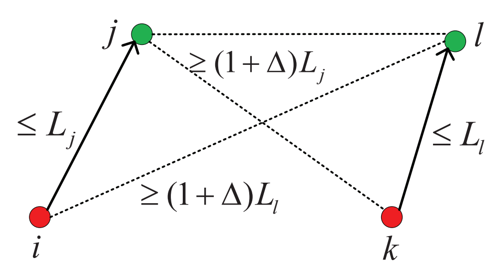


Fig. 6. An illustration of the protocol model.

We can infer from Inequality 13 that disks of radius $\Delta/2$ times the geographic range centered at the receivers in the same time slot are substantially disjoint. Thus, we can see that transmissions actually consume area. Considering the edge effect that a user is near the boundary of the square, at least a quarter of the disk is inside the square. Hence, each transmission takes up at least $\pi(\Delta L)^2/16$.

Suppose after T_D slots all the users obtain the content. Since the area available (nT_D in total) must be greater than the need of the transmissions (at least $\frac{\pi n(\Delta L)^2}{16}$), we have $nT_D\ge\frac{\pi n(\Delta L)^2}{16}$.

Since we are deriving the lower bound of transmission time, we could ignore the difference between L_r, L_c and replace them with $L = \sqrt{n \Big/ \pi(1+\Delta)\sqrt{|S_{\lfloor \frac{i}{2} \rfloor}(w)| \cdot n \cdot \log n}}$, which is in a lower order than L_c and L_r. Thus, we have

$$\begin{aligned} T_D &\geq \frac{\pi(\Delta L)^2}{16} \\ &= \frac{\Delta^2}{16(1+\Delta)}\sqrt{\frac{n}{|S_{\lfloor \frac{i}{2} \rfloor}(s)| \log n}} \\ &\geq \frac{\Delta^2}{16(1+\Delta)}\sqrt{n^{1-\lfloor \frac{i}{2} \rfloor / D} / \log n}. \end{aligned}$$

This completes the proof. □

C. Transmission Time in Non-evolving Networks

For comparison with the results in evolving networks, we further study the transmission time in non-evolving mobile networks. The mobility model and the transmission scheme are the same with evolving networks. Differently, the geographic ranges L_c and L_r are renewed by replacing $|S_{\lfloor i/2 \rfloor}(s)|$, $|S_{\lceil i/2 \rceil}(s)|$ and $|N_h(s)|$ with the value in non-evolving social networks. Since we did not specify the value of $|S_{\lfloor i/2 \rfloor}(w)|$, $|S_{\lceil i/2 \rceil}(s)|$ and $|N_h(s)|$ when deriving the bounds in the evolving networks, the arguments of deriving transmission time are still valid in the non-evolving case. Thus, following similar method, we obtain the bounds of transmission time in non-evolving networks in the two theorems below.

Theorem 6. *In the non-evolving mobile networks, under the transmission scheme in Algorithm 3, the transmission time is upper bounded by*

$$O\left(4(1+\Delta)\log^3 n\sqrt{n^{1-\lfloor i/2 \rfloor / D_s}/\sigma}\right),$$

with probability $1 - o(n^{-1})$*, where* $D_s = \Theta(\log n)$ *is the diameter of the non-evolving network and* σ *is a constant.*

Proof. The outline of the proof is the same as Theorem 4. We only focus on some critical points. In the derivation of T_a, we replace $|S_{\lfloor i/2 \rfloor}(w)|$ with the lower bound $\sigma n^{\lfloor i/2 \rfloor / D_s}$ [2] in the non-evolving network and notice that $D_s = \Theta(\log(n)$. We derive the time for users in $V \backslash S_{\lceil i/2 \rceil}(s)$ to obtain the content $T_a < 4(1+\Delta)\log^3 n\sqrt{n^{1-\lfloor i/2 \rfloor / D_s}/\sigma}$.

In terms of the time for users in $S_{\lceil i/2 \rceil}(s)$ to obtain the content T_b, we replace $|S_{\lceil i/2 \rceil}(s)|$ with the upper bound $2Wn^{\lceil i/2 \rceil / D_s}/\log n$ [2] and have $T_b < 4(1+\Delta)\log^2 n\sqrt{2Wn^{\lceil i/2 \rceil / D_s}}$, where W is a term of $\log n$ with constant order at least 2. Since $i < D_s$, we derive that T_b is inferior to T_a in order, i.e., $T_a = \omega(T_b)$. Thus, the upper bound of the transmission time T_D is

$$O\left(4(1+\Delta)\log^3 n\sqrt{n^{1-\lfloor i/2 \rfloor / D_s}/\sigma}\right).$$

□

Similarly, since transmission consumes area, we obtain the lower bound in Theorem 7.

Theorem 7. *In the non-evolving mobile networks, under the transmission scheme in Algorithm 3, the lower bound of the transmission time is*

$$\Omega\left(\frac{\Delta^2}{16(1+\Delta)}\sqrt{n^{1-\lfloor i/2 \rfloor / D_s}/2W}\right),$$

where W *is a term of* $\log n$ *with constant order at least* 2*.*

The proof is similar to Theorem 5. We only need to replace $|S_{\lfloor i/2 \rfloor}(s)|$ with its upper bound $2Wn^{\lfloor i/2 \rfloor / D_s}/\log n$ in non-evolving networks [2]. For space limitation, we omit the proof here

Remark. In comparison with non-evolving networks, with the same social depth i, regarding the upper bound of transmission time, the gap between evolving and non-evolving networks is roughly $\sqrt{n^{\gamma(\lfloor i/2 \rfloor - \tau) - \frac{i}{2\log n}} D / \log n}$, which increases as the network evolves. Similar results can be found in the lower bound. *Thus, we see that the transmission time of evolving networks is smaller than that of non-evolving networks.*

The intuition is similar to the static case. Namely, the large neighborhood of evolving networks makes it easier for users to find a neighbor for request. Then, fixing the social depth i and probability to find a holder, the physical constraint L of evolving networks could be smaller, resulting in less interference. And thus, the transmission time in evolving networks is improved.

D. Mobile Case versus Static Case

Recall the upper bound in evolving networks, we found that the transmission time of the mobile case is greater than that of static case by a factor of $\log n$. In terms of the lower bound, the gap becomes a small power of n, i.e., nearly $n^{\frac{3i}{4D}}$. In non-evolving networks, the transmission time follows the similar discipline. As for the upper bound, the factor is $\sqrt{\log n}$. And the gap of lower bound is roughly $n^{\frac{3i}{4D_s}}$. In summary, the transmission time in the mobile case is slightly greater than the static case. This is possibly due to the simple physical transmission pattern of the mobile scheme where direct communication is considered. A better physical transmission strategy could be designed to reduce the transmission time (e.g., the permission of relays), which however is not the emphasis of this work and could be studied in the future. The geographic range is restricted in both static (L) and mobile cases (L_c or L_r), where L_c and L_r are in fact smaller that L.

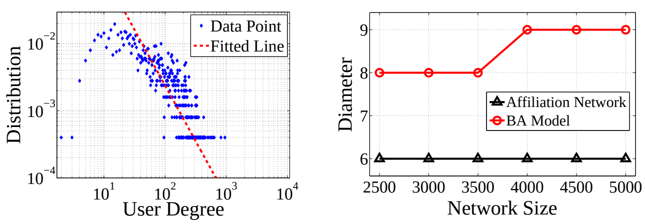
VII. EXPERIMENTS

In this section, we validate our theoretical results on both synthetic networks and real-world datasets. Meanwhile, we include three baseline algorithms and two more mobility models to demonstrate the performance of our proposed algorithms. We begin with the description of the datasets and experimental settings. Then, we present the performance of different algorithms in static and mobile cases respectively.

A. Dataset Description

Synthetic Networks. The evolving synthetic networks are generated according to the affiliation network model with the probability $\beta = 0.5$, parameters $c_i = 2$ and $c_v = 3$. Thus, degrees of users in the generated social network $G(V, E)$ follow a power law distribution with exponent $\frac{5}{3}$. For comparison, we form non-evolving networks by the Barab*á*si-Albert model (BA model), which guarantees the connectivity of the generated network and preserves the power-law degree distribution that the work of [2] builds on at the same time.

We verify the degree distribution and diameter of the evolving social network in Fig. 7, where we observe favorable agreement with theoretical results in both properties. Particularly, the degree distribution of the affiliation network with 2500 nodes is shown in Fig. 7(a). The nodal degree is represented by blue dots, which is well fitted by the red dashed line of slope 1.67. It can be seen from Fig. 7(b) that the diameter of the evolving social network is stabilizing over the network size, while the diameter of BA model is slowly increasing.



(a) Degree Distribution (b) Diameter

Fig. 7. Validation of the evolving model.

Real Networks. Since both social and geographic information are embodied in our work, we choose Gowalla and Brightkite [37] as real networks, which are location-based social networks.

- **Gowalla.** This dataset is collected from the homonymic location-based social networking website where users share their locations by checking-in. The network is undirected and consists of 196,591 nodes and 950,327 edges.
- **Brightkite.** This dataset is collected from the homonymic location-based social networking service provider, where users share their locations by checking-in. The network is undirected as well with 58,228 nodes and 214,078 edges.

Note that the original real datasets are non-evolving networks. Thus, it is necessary to transform them into evolving networks. Based on the observation that the earlier a node arrives, the greater the degree is (theoretical results could be found in [31]), we rank all the nodes by the degree in a non-increasing order and consecutively choose nodes whose subgraph is connected and take the subgraph as the evolving network. For comparison, we select the corresponding number of nodes by random walk starting from a random node and take the subgraph as the non-evolving network (the connectivity is certainly guaranteed).

B. Experimental Settings

For each dataset, we carry out transmission on networks with $\{2500, 3000, 3500, 4000, 4500, 5000\}$ nodes to emulate the evolving process. We report the results under social depth $i \in \{1, 2, 3\}$. The source is randomly selected by choosing the 500-th node of each network. In synthetic networks, the location of users is generated according to the i.i.d. model. As for real networks, we rescale the locations of users to put them into a square of length $\sqrt{|V|}$.

We next elaborate the setting of geographic range L. In synthetic networks, the geographic range is set according to the theoretical value. Regarding real-world datasets, we analyze the network to estimate the parameters in L. Specifically, we first collect the degree of each user to obtain the degree distribution and then the power law parameter τ could be derived. With τ and i, the constant μ is accordingly specified. Since the largest degree is n^γ according to the social network model, γ could be further obtained by identifying the maximum degree. The number of users in each layer is also recorded.

In the static case, to derive the transmission time, inspired by the proof of Theorem 2, we calculate the number of rounds it takes to propagate the content throughout the network and multiply the round by the geographic range L to obtain the transmission time. In the mobile case, the actual transmission time could be obtained by multiplying the number of slots by the duration of time slots F/R.

We next introduce the baselines and two more mobility models tested in the experiment.

- **Algorithm Description**

Flooding. This is a classic transmission scheme in wireless networks. Upon receiving the content, each user transmits it to all of its social neighbors until all the users have the content.

Converge Multicast [21]. The network is divided into clusters each with $k = 25$ users. The transmitter first splits the content into k pieces equally and distributes to the other $k - 1$ users. Then, the content pieces are transmitted between neighboring clusters until all destination clusters receive the content. Finally, the receiver collects the pieces to obtain the whole content.

Relay-based Algorithm [20]. This is a typical scheme in mobile wireless networks. Each transmission is divided into two phases: (1) the transmitter delivers the content to random relays, and (2) the relay transmits the content to the receiver. The number of relay is set to be 10.

- **Mobility Model**

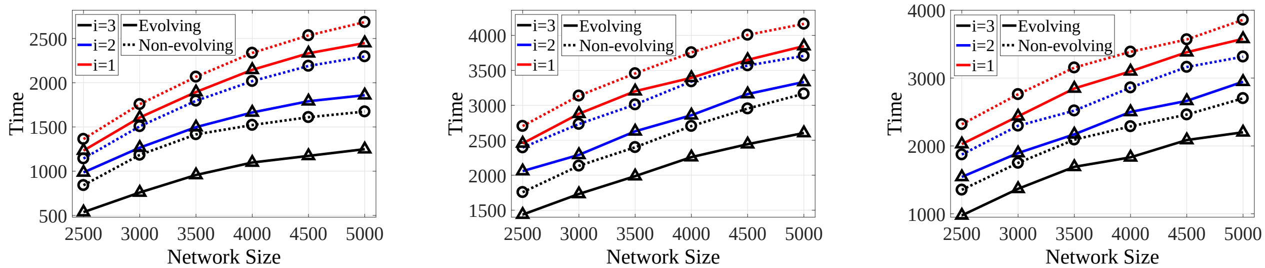
Random Walk Model. The network area is equally partitioned into lattices with $\lceil 10\sqrt{n} \rceil$ cells in each edge. In each time slot, users randomly select a neighboring vertex to arrive.

Random Waypoint Model. At the beginning of a time slot, each user moves according to a two-dimensional vector (X, Y), where the value of X and Y is uniformly and randomly selected from $[1, 5]$. Then the user moves a distance of X horizontally and Y vertically.

C. Experimental Results

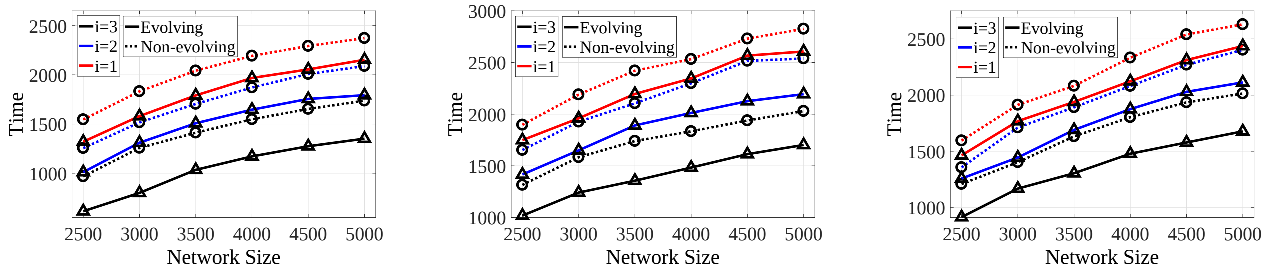
We first validate our theoretical findings by Fig. 8 and Fig. 9. Then, we compare our algorithms with baselines, where the experimental figures are deferred to Sec. III of the supplemental file.

We verify our analytical results in the static case by providing the transmission time of the proposed Alg. 2 in Fig. 8. Our



(a) Synthetic Networks (b) Gowalla (c) Brightkite

Fig. 8. Transmission Time of Alg. 2 in the Static Case.



(a) Synthetic Networks (b) Gowalla (c) Brightkite

Fig. 9. Transmission Time of Alg. 3 in the Mobile Case.

first observation is that in both synthetic and real networks, the transmission time is decreasing as the social depth i becomes larger, which verifies the time bounds in Theorems 2 and 3 where the parameter i is subtracted in the exponent of n. Secondly, under the same social depth, there is evident gap (roughly 25% when $i = 3$) between evolving and non-evolving networks, which could be explained by comparing their theoretical bounds on transmission time. The underlying factors of the above two observations could be derived from Lemma 5. Specifically, with a larger social depth i or in an evolving social network, a user would have more neighbors to quest and thus the physical constraint of requesting neighbors could be relaxed, i.e., the geographic range L could be smaller. Accordingly, the transmission rate of each transceiver $\Omega(R/L)$ would become larger and thus the resultant transmission time is reduced.

The performance of Alg. 3 is shown in Fig. 9 to validate the analytical results in the mobile case. In accordance with our analytical results in Theorems 4-7, the transmission time of both evolving and non-evolving networks decreases as the social depth i grows. However, there is an evident gap between evolving networks and non-evolving networks, which could be anticipated by comparing analytical results in evolving (Theorems 4, 5) and non-evolving (Theorems 6, 7) networks. Moreover, as shown in the figure, the transmission time grows sublinearly with the network size, which is consistent with the results of Theorems 4-7.

From the above illustration, we can see that the theoretical findings are well validated by the experiments. And we would like to remark that the transmission time reported in the static case and mobile case is not comparable, since constant factors c and F in the static case and the length of a time slot F/R in the mobile case all need to be measured in practice to obtain the exact transmission time.

The proposed schemes are further tested against three existing algorithms. The social depth of our algorithm is set to be $i = 2$. Results of the static case are shown in Fig. S1 of the supplementary file, where the transmission time in evolving networks is still smaller than the non-evolving counterparts due to similar reason in Fig. 8. Moreover, our algorithm outperforms the baselines, since it is easier for users to find a holder when social neighbors are explored. And the multicast converge algorithm shows evident advantages over flooding. The reason is that the transmission in flooding spans arbitrary distance, i.e., relays in the highway system have to serve more users resulting in lower transmission rate.

In the mobile case, algorithms are carried out in both random walk model and random waypoint model. The results are presented in Fig. S2 and Fig. S3 of the supplementary file respectively. Similar conclusions about evolving networks and our algorithm can be drawn. Our algorithm obtains a smaller transmission time than the baselines since users have higher probability to meet a holder in the large neighborhood. The relay based algorithm is superior to flooding, since the adoption of relay increases the probability for the receiver to obtain the content.

VIII. Discussion

In this work, we are mainly concerned with primary settings, while the analysis also applies to many other practical settings. For examples, (1) in the case of fixed network area, we only need to rescale the terms about the network size in the transmission scheme and proofs. And we could find that the results are the same as before. (2) The generation process of $G(V,E)$ could be probabilistic (i.e., users of the same affiliation i are connected with some probability), as well as the state transition of users. In this case, we ask the probability to be reciprocal to $[d^B(i)]^\alpha$ ($\alpha \in (0,1)$ is a constant), which also results in densifying edges, shrinking diameter and heavy-tailed distribution according to Section 8 of [11]. We further update the theoretical results about the social network and find that the neighborhood size is the same as before. Then, previous derivations of transmission time could be applied seamlessly. (3) Regarding the case of non-trivial user reaction time (i.e., the time interval between state transitions), we only need to append an extra term to indicate the time for users to react, which is in a smaller scale than the transmission time. (4) Comparing our results with much previous work of similar settings (e.g., [19, 20, 32]), our schemes achieve smaller transmission time and even comparable results with schemes that have additional aids (e.g., relays in the mobile case). For detailed derivations, readers may refer to Section II of the supplemental file.

IX. Conclusion and Future Work

In this paper, we made the first attempt to study the specific transmission time in evolving wireless social networks. The microscopic property of evolving social networks was initially explored in this work, which provided the basis for scheme design and theoretical analysis. For a comprehensive analysis, both static and mobile cases were considered, where the transmission schemes are deliberately designed respectively. Subsequently, we managed to bound the transmission time in both cases through a series of analysis. The transmission in non-evolving mobile networks was further studied for comparison. The results indicate that, under our schemes, the transmission in evolving networks takes less time than non-evolving networks, and the gap increases as network evolves. Moreover, by mathematically delineating transmission time, the impact of structural parameters is also revealed. Finally, numerous experiments on both synthetic and real datasets were conducted to validate the theoretical results.

Some desirable future directions still remain to be explored. First, the adopted Affiliation Network assumes non-decreasing users, while it is possible that users will leave. Thus, it is necessary to study the transmission time under a more practical model when available. Second, due to space limitation, only the transmission time is studied in detail, while many other practical metrics are worthy of exploration, such as transmission error, re-transmission, network traffic. Third, although probabilistic connections between users are discussed, for more general cases, we need to make further investigations.

Acknowledgment

This work was supported by National Key R&D Program of China 2018YFB2100302, NSF China under Grant (No. 61822206, 61960206002, 61832013, 62041205, 61532012), Tencent AI Lab Rhino-Bird Focused Research Program JR202034. And we would like to express our special thanks to Chengyang Wu and Jiaqi Liu for their helpful discussions and assistance.

References

[1] M. A. Rahman and M. S. Hossain, "A location-based mobile crowdsensing framework supporting a massive ad hoc social network environment," vol. 55, no. 3, pp. 76–85, Mar. 2017.

[2] Y. Chen, C. Caramanis, and S. Shakkottai, "On sharing viral video over an Ad Hoc wireless network," *arXiv preprint arXiv:1101.5088*, 2011.

[3] A. M. Frieze and G. R. Grimmett, "The shortest-path problem for graphs with random arc-lengths," *Discrete Appl. Math.*, vol. 10, no. 1, pp. 57–77, Jan. 1985.

[4] B. Pittel, "On spreading a rumor," *SIAM J. Appl. Math.*, vol. 47, no. 1, pp. 213–223, Feb. 1987.

[5] S. N. Dorogovtsev and J. F. Mendes, "Evolution of networks," *Advances in Phys.*, vol. 51, no. 4, pp. 1079–1187, 2002.

[6] M. Atzmueller, A. Ernst, F. Krebs, C. Scholz, and G. Stumme, "On the evolution of social groups during coffee breaks," in *Proc. ACM Int. Conf. World Wide Web*, Apr. 2014, pp. 631–636.

[7] N. Z. Gong, W. Xu, L. Huang, and *et al.*, "Evolution of social-attribute networks: measurements, modeling, and implications using google+," in *Proc. ACM SIGCOMM Internet Meas. Conf., IMC*, Nov. 2012, pp. 131–144.

[8] "Facebook: Number of active users," https://www.statista.com/, (accessed May 4, 2020).

[9] J. Leskovec, J. Kleinberg, and C. Faloutsos, "Graphs over time: densification laws, shrinking diameters and possible explanations," in *Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.*, Aug. 2005, pp. 177–187.

[10] R. Diestel, *Graph theory*. Germany: Springer, 2018.

[11] S. Lattanzi and D. Sivakumar, "Affiliation networks," in *Proc. Annu. ACM Symp. Theory Comput., STOC*, May 2009, pp. 427–434.

[12] U. Feige, D. Peleg, P. Raghavan, and *et al.*, "Randomized broadcast in networks," *Random Struct. & Algorithms*, vol. 1, no. 4, pp. 447–460, Aug. 1990.

[13] Z. Zheng, T. Wang, L. Song, Z. Han, and J. Wu, "Social-aware multi-file dissemination in device-to-device overlay networks," in *IEEE Conf. Comput. Commun. Workshops, INFOCOM WKSHPS*, Apr. 2014, pp. 219–220.

[14] L. Wang, L. Gao, A. Zhang, and M. Chen, "Social-aware file-sharing mechanism for device-to-device communications," in *IEEE Int. Conf. Wirel. Commun. Signal Process., WCSP*, Nov. 2015, pp. 1–5.

[15] L. Shangguan, Z. Yang, A. X. Liu, Z. Zhou, and Y. Liu, "STPP: Spatial-temporal phase profiling-based method for relative RFID tag localization," *IEEE/ACM Trans. Netw.*, vol. 25, no. 1, pp. 596–609, Feb. 2017.

[16] Cisco, "Global mobile data traffic forecast update, 2018–2023 white paper," https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.html, (accessed May 4, 2020).

[17] Y. Song, L. Liu, H. Ma, and A. V. Vasilakos, "A biology-based algorithm to minimal exposure problem of wireless sensor networks," *IEEE Trans. Netw. Service Manag.*, vol. 11, no. 3, pp. 417–430, Sep. 2014.

[18] L. Liu and H. Ma, "On coverage of wireless sensor networks for rolling terrains," *IEEE Trans. Parallel Distrib. Syst.*, vol. 23, no. 1, pp. 118–125, Jan. 2012.

[19] P. Gupta and P. R. Kumar, "The capacity of wireless networks," *IEEE Trans. Inf. Theory*, vol. 46, no. 2, pp. 388–404, Mar. 2000.

[20] M. Grossglauser and D. Tse, "Mobility increases the capacity of ad-hoc wireless networks," in *IEEE Conf. Comput. Commun., INFOCOM*, Apr. 2001, pp. 1360–1369.

[21] X. Wang, L. Fu, and C. Hu, "Multicast performance with hierarchical cooperation," *IEEE/ACM Trans. Netw.*, vol. 20, no. 3, pp. 917–930, Jun. 2012.

[22] L. Fu and X. Wang, "Multicast scaling law in multichannel multiradio wireless networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 24, no. 12, pp. 2418–2428, Dec. 2013.

[23] X. Wang, W. Huang, S. Wang, J. Zhang, and C. Hu, "Delay and capacity tradeoff analysis for motioncast," *IEEE/ACM Trans. Netw.*, vol. 19, no. 5, pp. 1354–1367, Oct. 2011.

[24] H. Luo, J. Luo, Y. Liu, and S. K. Das, "Adaptive data fusion for energy efficient routing in wireless sensor networks," *IEEE Trans. Comput.*, vol. 55, no. 10, pp. 1286–1299, Oct. 2006.

[25] D. Zhang, D. Zhang, H. Xiong, C.-H. Hsu, and A. V. Vasilakos, "BASA: Building mobile Ad-Hoc social networks on top of android," *IEEE Netw.*, vol. 28, no. 1, pp. 4–9, Feb. 2014.

[26] E. Sarigöl, O. Riva, P. Stuedi, and G. Alonso, "Enabling social networking in ad hoc networks of mobile phones," *Proc. VLDB Endow.*, vol. 2, no. 2, pp. 1634–1637, Aug. 2009.

[27] L. Fu, W. Huang, X. Gan, F. Yang, and X. Wang, "Capacity of wireless networks with social characteristics," *IEEE Trans. Wireless Commun.*, vol. 15, no. 2, pp. 1505–1516, Feb. 2016.

[28] J. Liu, L. Fu, J. Zhang, and *et al.*, "Modeling multicast group in wireless social networks: A combination of geographic and non-geographic perspective," *IEEE Trans. Wireless Commun.*, vol. 16, no. 6, Apr. 2017.

[29] C. Ma, M. Ding, H. Chen, Z. Lin, G. Mao, Y. C. Liang, and B. Vucetic,

"Socially aware caching strategy in device to device communication networks," *IEEE Trans. Veh. Technol.*, vol. 67, no. 5, pp. 4615–4629, May 2018.

[30] L. Fu, J. Zhang, and X. Wang, "Evolution-cast: Temporal evolution in wireless social networks and its impact on capacity," *IEEE Trans. Parallel Distrib. Syst.*, vol. 25, no. 10, pp. 2583–2594, Oct. 2014.

[31] J. Liu, L. Fu, Z. Liu, X.-Y. Liu, and X. Wang, "Interest-aware information diffusion in evolving social networks," *IEEE Trans. Wireless Commun.*, vol. 17, no. 7, pp. 4593–4606, Jul. 2018.

[32] M. Franceschetti, O. Dousse, N. David, and P. Thiran, "Closing the gap in the capacity of wireless networks via percolation theory," *IEEE Trans. Inf. Theory*, vol. 53, no. 3, pp. 1009–1018, Mar. 2007.

[33] S. L. Feld, "The focused organization of social ties," *Amer. J. Sociol.*, vol. 86, no. 5, pp. 1015–1035, Mar. 1981.

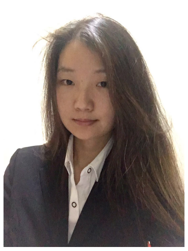
[34] T. J. Allen *et al.*, *Managing the flow of technology: Technology transfer and the dissemination of technological information within the R&D organization*. Cambridge, MA, USA: MIT Press, 1984.

[35] F. Chung and L. Lu, "Connected components in random graphs with given expected degree sequences," *Ann. Combinatorics*, vol. 6, no. 2, pp. 125–145, Nov. 2002.

[36] L. Ying, S. Yang, and R. Srikant, "Optimal delay–throughput tradeoffs in mobile ad hoc networks," *IEEE Trans. Inf. Theory*, vol. 54, no. 9, pp. 4119–4143, Sep. 2008.

[37] J. Leskovec and A. Krevl, "SNAP Datasets: Stanford large network dataset collection," http://snap.stanford.edu/data, Jun. 2014.

Chen Feng received his B.E. degree in Communication Engineering from Tianjin University, China, in 2016. He is currently pursuing the Ph.D. degree in Electronic Engineering at Shanghai Jiao Tong University, Shanghai, China. His current research interests are in the area of social networks and information diffusion.



Luoyi Fu received her B.E. degree in Electronic Engineering from Shanghai Jiao Tong University, China, in 2009 and Ph.D. degree in Computer Science and Engineering in the same university in 2015. She is currently an Assistant Professor in Department of Computer Science and Engineering in Shanghai Jiao Tong University. Her research of interests are in the area of social networking and big data, scaling laws analysis in wireless networks, connectivity analysis and random graphs.

Xudong Wu received his B. E. degree in Information and Communication Engineering from Nanjing Institute of Technology, China, in 2015. He is currently pursuing the Ph.D. degree in Computer Science and Engineering in Shanghai Jiao Tong University. His research interests include mobile computing systems and social networks.

Xiaoying Gan received the Ph.D. degree in Electronic Engineering from Shanghai Jiao Tong University, Shanghai, China, in 2006. From 2009 to 2010, she was a Visiting Researcher with the California Institute for Telecommunications and Information, University of California at San Diego, San Diego, CA, USA. She is currently an Associate Professor with the Department of Electronic Engineering, Shanghai Jiao Tong University. Her current research interests include network economics, social aware networks, heterogeneous cellular networks, multiuser multi-channel access, and dynamic resource management.

Xinbing Wang received the B.S. degree (with hons.) in automation from Shanghai Jiao Tong University, China, in 1998, the M.S. degree in computer science and technology from Tsinghua University, China, in 2001, and the Ph.D. degree with a major in electrical and computer engineering and minor in mathematics from North Carolina State University, in 2006. Currently, he is a Professor in the Department of Electronic Engineering, and Department of Computer Science, Shanghai Jiao Tong University, China. Dr. Wang has been an Associate Editor for IEEE/ACM TRANSACTIONS ON NETWORKING, IEEE TRANSACTIONS ON MOBILE COMPUTING, and ACM Transactions on Sensor Networks. He has also been the Technical Program Committees of several conferences including ACM MobiCom 2012,2014, ACM MobiHoc 2012-2017, IEEE INFOCOM 2009-2017.

Guihai Chen received the B.S. degree from Nanjing University, the M.E. degree from Southeast University, and the Ph.D. degree from The University of Hong Kong. He visited the Kyushu Institute of Technology, in 1998, as a Research Fellow, and the University of Queensland, in 2000, as a Visiting Professor. From 2001 to 2003, he was a Visiting Professor with Wayne State University. He is currently a Distinguished Professor and a Deputy Chair with the Department of Computer Science, Shanghai Jiao Tong University. His research interests include wireless sensor networks, peer-to-peer computing, and performance evaluation. He has served on technical program committees of numerous international conferences.

Jun Xu received the Ph.D. degree in computer and information science from The Ohio State University, in 2000. He is currently a Professor with the College of Computing, Georgia Institute of Technology. His current research interests include data streaming algorithms for the measurement and monitoring of computer networks and hardware algorithms, and data structures for high-speed routers. He received the US National Science Foundation (NSF) CAREER Award, in 2003, the ACM Sigmetrics Best Student Paper Award, in 2004, and the IBM Faculty Awards, in 2006 and 2008, respectively. He was named an ACM Distinguished Scientist, in 2010.