On The Resilience of D2D-based Social Networking Service Against Random Failures

Sigit Aryo Pambudi Wenye Wang
Department of Electrical and Computer Engineering
North Carolina State University, Raleigh, NC 27606
Email: {sapambud,wwang}@ncsu.edu

Cliff Wang
Army Research Office
Research Triangle Park, NC 27709
Email: cliff.wang@us.army.mil

***Abstract*—Device-to-device (D2D)-based social networking service (SNS) is an emerging information system that enables users with social ties to exchange multimedia contents through multihop short-range wireless links. In the D2D-based SNS, a random initial node failure may lead to a cascade of failures, which is a series of events in which users become isolated from others over subsequent time instances. Different from previous studies that analyze whether network-wide connectivity can be preserved after a cascade of failures, *our study sheds light on the D2D-based SNS's resilience from the perspective of end-user connection experience*. In this paper, we first introduce a numerical method for calculating the mean fraction of nodes that are not affected by the cascading failures and the amount of time to reach the end of such sequence of failures. Then, we apply a probabilistic approach to derive the lower and upper bounds of a node resilience metric, which is the likelihood that an end-user will not be isolated during an ongoing social networking session. Our analysis and numerical simulations indicate that, compared to exponentially-distributed session times, user session times with Pareto (heavy-tailed) distribution results in poorer node resilience, which quickly deteriorates when the mean session time is high.**

I. Introduction

The massively-increasing popularity of social networking services (SNSs), such as Twitter and Facebook, and the recent advancements in device-to-device (D2D) communications for smartphones and tablets, made popular by the introduction of Apple iOS's Multipeer Connectivity Framework [1], have opened the new paradigm of providing SNS through D2D communications, referred here as *D2D-based SNS*. In October 2014 alone, about half million of smartphone users has been exchanging social messages through Bluetooth peer-to-peer and WiFi Direct using FireChat mobile app [2]. The sheer attention garnered by such a new paradigm is mainly due to its ability to alleviate traffic burden from the highly-congested cellular networks and to facilitate content exchange free from government censorship [2].

Despite the potentials, little has been known regarding the D2D-based SNS's *resilience* against numerous impairments. In the literature, a classical problem in the line of *network resilience* research is to study the network's connectivity under random failures [3], [4] and node misbehaviors [5], [6]. To this end, many existing works consider the D2D network only, neglecting the social interconnectivity between users, and investigate the occurrences of node isolations under random node and link failures. For example, the link resilience of ad-hoc networks with respect to random node failures has been analyzed as a fault tolerance measure [3]. The problem of node isolation due to misbehaving nodes, on the other hand, was studied in [5], from which the connectivity of the entire network is calculated using the probability of individual node's isolation [6]. The examined D2D-based SNS, however, is different than these conventional ad-hoc networks in that it consists of two coexisting SNS and D2D networks, as depicted in Fig. 1, which are interconnected through interdependence relationship (i.e., the vertical links). Although percolation theory has been employed to analyze the spatial and temporal properties of such interdependent networks [7], little has been known regarding the D2D-based MSN's ability to provide networking service from the perspective of end-user.

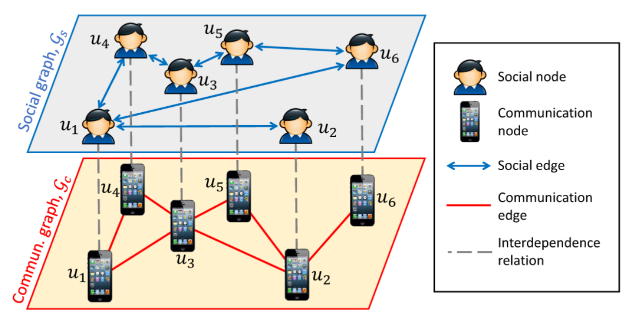


Fig. 1. D2D-based SNS architecture: social and communication graphs.

This work is motivated by considering the intrinsic properties of end-users (or nodes) in D2D-based SNSs. In these networks, failures that isolate a set of nodes from the rest occur initially at time $t = 0$ due to various reasons, such as severe co-channel interference or social malware attacks, and result in a *cascade of failures*, i.e., a sequence of additional nodes in isolation at time $t \geq 1$. Meanwhile, end-users may require a non-negative *session time*, i.e., the time until they respectively finish social content exchange through the D2D-based SNS. Thus, each end-user can be more interested in whether he/she will be isolated, either due to the initial node failure or the cascading failures, during his/her session time. In other words, after an end-user started its SNS session, as long as the user can enjoy a stable connection during his/her session, the networking service provided by the D2D-based SNS is said to be *resilient* to this node. Therefore, we are interested to investigate: "*What is the expected time before a node is isolated? What is the resilience of the D2D-based SNS in terms of end-users' satisfaction toward their social content exchange experience?*"

To answer the questions above, we start our analysis by treating the D2D-based SNS as an interdependent communication and social networks. Next, we introduce a mean fraction of functional nodes s_∞ that is proportional to the number of functional nodes at the end of the cascade of failures and a maximum isolation time TI_{max} that provides a temporal upper bound until a node becomes isolated, and then apply percolation theory to calculate their values. Furthermore, we define a node resilience metric Ψ_n that measures the likelihood that an end-user can finish his/her ongoing SNS session before being isolated, and propose a probabilistic approach for calculating the upper and lower bounds of Ψ_n. Numerical results show that the derived bounds are valid and the D2D-based SNS with exponentially-distributed session times is more resilient to random failures than that with pareto (heavy-tailed)-distributed session times.

The rest of this paper is organized as follows. Section II introduces the D2D-based SNS network model and the spatial-temporal metrics of cascading failures. Section III proposes an end user-based resilience metric and its bounds, which is validated through computer simulations in Section IV. Finally, Section V concludes this paper.

II. Network Model and Cascading Failures

In this section, we outline the D2D-based SNS model, the formation of the cascade of failures in such network, as well as the spatial and temporal impacts of such cascading failures.

A. D2D-Based Social Networking Service

We consider a D2D-based SNS containing users that are interconnected through short-range D2D wireless links such as Bluetooth peer-to-peer, WiFi Direct, or NFC. Let $\mathcal{V}$ be the set of users (or nodes), and $\mathcal{E}_c \subseteq \mathcal{V}\times\mathcal{V}$ be a set of communication edges with $e(u,v) \in \mathcal{E}_c \Leftrightarrow \forall u,v \in \mathcal{V}, ||X_u - X_v|| \leq r$, where $||X_u - X_v||$ is the Euclidean distance between nodes u and v, while r is the wireless transmission distance. Then, *communication graph* is defined as the tuple $\mathcal{G}_c := (\mathcal{V}, \mathcal{E}_c)$. In this paper, we assume $\mathcal{G}_c$ is connected[1]

Additionally, there exists an SNS that operates over the communication graph and enables people with common interest to meet virtually and exchange multimedia contents wirelessly. Let $\mathcal{E}_s \subseteq \mathcal{V}\times\mathcal{V}$ be a set of social edges between users. An edge exists between users $u \in \mathcal{V}$ and $v \in \mathcal{V}$, i.e., $e(u,v) \in \mathcal{E}_s$, if they both have social ties, which are similar to followee/follower relations in Twitter and friendships in Facebook. Then, *social graph* is defined as $\mathcal{G}_s := (\mathcal{V}, \mathcal{E}_s)$. We assume that $\mathcal{G}_s$ is connected[1] as well.

The structure of a D2D-based SNS is depicted in Fig. 1, in which users u_1 and u_2 are socially-connected (i.e., $e(u_1,u_2) \in \mathcal{E}_s$) and the former may want to send social contents to the latter. To facilitate this exchange, multi-hop communication [9] is enabled such that the contents are sent as packets that physically traverse through nodes (u_1,u_3,u_2) as the shortest path. As a result, social contents can be delivered from one node to another when they (i) have social relation, and (ii) are connected through D2D path(s).

[1]Graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is connected if there is a u,v-path for $\forall u,v \in \mathcal{V}$ [8].

B. Cascading Failures in D2D-based SNS

Users may experience many types of impairments that lead to *failure*, which is defined as a user's inability to exchange social contents through the D2D-based SNS. For example, users may be impaired by co-channel interference from nearby Bluetooth and WiFi devices, or they can be infected by malware that is disguised as legitimate social content. These severe impairments, which can prevent the users from exchanging contents, are modeled as follows. Let $N = |\mathcal{V}|$ be the number of nodes in the D2D-based SNS and let $t \in \mathbb{N}_0$ be a discrete time step, where $\mathbb{N}_0$ is the set of non-negative integers. Let $\mathcal{I}_f \subseteq \mathcal{V}$ be the set of nodes that experience failure at an initial time $t=0$ and let $p_f := \frac{|\mathcal{I}_f|}{N}$ be the initial fraction of failed nodes. We assume that a node $u \in \mathcal{V}$ experiences failure initially (i.e., $u \in \mathcal{I}_f$) according to a uniform distribution with probability $p_f \in (0,1)$. If u fails, then it becomes isolated, i.e., all communication and social edges incident[2] to u are removed from $\mathcal{G}_c$ and $\mathcal{G}_s$, respectively.

Unlike in conventional ad-hoc and mesh-based wireless networking [10], the impact of the initial failure in D2D-based SNSs will be exacerbated by dependencies between the communication and social graphs outlined as follows.

1) *Social-to-communication dependence:* Since social contents are delivered through D2D links, then a node will lose its social capability, i.e., have its social edges removed, when it cannot find any communication edge to its neighbor(s).
2) *Communication-to-social dependence:* The users participate in the D2D communications by relaying transitory packets in exchange for the ability to send/receive social contents to/from their social neighbors. When a user cannot find any of its social neighbors, e.g., does not have any social edge, then it will disconnect itself from the D2D communication network.

We present the following example to explain how these two-way dependencies, also known as an *interdependence* relation [7], [11], [12], and the initial node failure at $t=0$ may lead to a series of additional isolated nodes over subsequent time steps $t \geq 1$, which is referred as a *cascade of failures*.

In Fig. 2(a), we consider the D2D-based SNS structure in Fig. 1 in which node u_3 experiences failure initially and its incident communication and social edges are removed. Let $\mathcal{R}_c(t) \subseteq \mathcal{V}$ and $\mathcal{R}_s(t) \subseteq \mathcal{V}$ be the sets of *residual* nodes in $\mathcal{G}_c$ and $\mathcal{G}_s$, respectively, at time $t \geq 1$. Denote $\mathcal{F}_c(t) \subseteq \mathcal{R}_c(t)$ and $\mathcal{F}_s(t) \subseteq \mathcal{R}_s(t)$ as the sets of residual nodes that belongs to the largest connected component of $\mathcal{G}_c$ and $\mathcal{G}_s$, respectively. The following events occur.

[2]The communication and social edges incident to $u \in \mathcal{V}$ are defined as the sets $\{e(u,v) \in \mathcal{E}_c, \forall v \in \mathcal{V}\}$ and $\{e(u,v) \in \mathcal{E}_s, \forall v \in \mathcal{V}\}$, respectively.

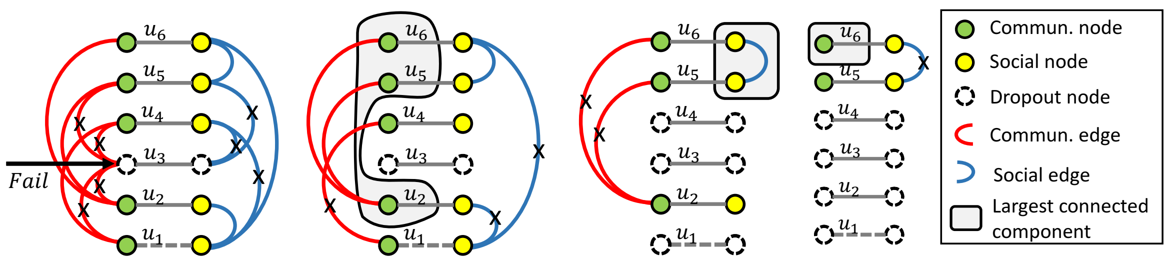


Fig. 2. Illustration of cascading failures in D2D-based SNS.

1) Firstly, the set of residual node is $\mathcal{R}_c(1) = \{u_1, u_2, u_4, u_5, u_6\}$ due to the initial failure of node u_3. Then, two connected components occur in $\mathcal{G}_c$ and nodes $\{u_1, u_4\}$, which do not belong to the largest connected component of $\mathcal{G}_c$ and cannot communicate with the remaining majority of the users, have their incident communication edges removed in Fig. 2(b). As a result, the set of functional residual nodes becomes $\mathcal{F}_c(1) = \{u_2, u_5, u_6\}$.
2) Secondly, due to the social-to-communication dependence, nodes $\{u_1, u_4\}$ also loses their social capability, i.e., have their incident social edges removed, such that $\mathcal{R}_s(1) = \mathcal{F}_c(1) = \{u_2, u_5, u_6\}$. As a result, node u_2 becomes isolated from the largest component of $\mathcal{G}_s$ in Fig. 2(c) and $\mathcal{F}_s(1) = \{u_5, u_6\}$. Note that the sequence of communication and social disconnections in Figs. 2(b)-2(c) are both considered as one step of the cascade of failure at time $t = 1$.
3) Thirdly, $\mathcal{R}_c(2) = \mathcal{F}_s(1) = \{u_5, u_6\}$ because node u_2 is also disconnected from $\mathcal{G}_c$, which results in $\mathcal{F}_c(2) = \{u_6\}$, i.e., a totally-disconnected $\mathcal{G}_c$ in Fig. 2(d).
4) Finally, we have $\mathcal{F}_s(2) = \{u_6\}$, i.e., $\mathcal{G}_s$ will also become totally disconnected after two steps of cascading failures.

Note how an initial failure on only node u_3 leads to a cascade of failures that totally disconnect the D2D-based SNS. Such a rampant cascading failures potentially hinders the quality-of-service rendered by the D2D-based SNS to its users.

C. Spatial Coverage of Cascading Failures

To better understand how the quality of service is impacted by an initial failure in D2D-based SNS, we study the coverage of the cascade of failures as follows. More specifically, we are interested in how many users that are not isolated at the end of a cascade of failures. Denote $s(t) := \mathbb{E}[\frac{|\mathcal{F}_s(t)|}{N}]$ and $s'(t) := \mathbb{E}[\frac{|\mathcal{R}_c(t)|}{N}]$ as the mean fractions of functional and residual nodes in $\mathcal{G}_s$ at time $t \in \mathbb{N}_0$. We assume that $\mathcal{F}_s(0) = \mathcal{I}_f$ such that $s(0) = (1 - p_f)$. Let $g_c(c'(t)) := \mathbb{E}[\frac{|\mathcal{F}_c(t)|}{|\mathcal{R}_c(t)|}]$ and let $g_s(s'(t)) := \mathbb{E}[\frac{|\mathcal{F}_s(t)|}{|\mathcal{R}_s(t)|}]$ be the mean fractions of the giant connected component in $\mathcal{R}_c(t)$ and $\mathcal{R}_s(t)$, respectively, at the subsequent time steps $t \geq 1$.

Proposition 1: The mean fraction of functional nodes at the end of the cascading failures is given as $s_\infty = \lim_{t\to\infty} s'(t) g_s(s'(t))$, where $s'(t)$ satisfies [7, Section C]

$$\begin{cases} s'(t) = (1 - p_f) g_c(c'(t)), \\ c'(t) = (1 - p_f) g_s(s'(t-1)). \end{cases} \tag{1}$$

To proceed, let $d_c(u) := |\{v : e(u,v) \in \mathcal{E}_c\}|$ and $d_s(u) := |\{v : e(u,v) \in \mathcal{E}_s\}|$ be the communication and social degrees of node $u \in \mathcal{V}$, respectively. Denote $p_c(k) := Pr\{d_c(u) = k\}$ and $p_s(k) := Pr\{d_s(u) = k\}$ as the communication and social degree distributions of a randomly-chosen node u, where $k \in \{0, 1, \ldots, N-1\}$. Then, we have the following proposition.

Proposition 2: Let $p \in [0, 1]$. The mean fractions of the giant connected component are given as [7]

$$g_c(p) = 1 - G_{c,0}(1 - p(1 - f_c)), \tag{2}$$
$$g_s(p) = 1 - G_{s,0}(1 - p(1 - f_s)), \tag{3}$$

where $G_{c,0}(x) := \sum_{k=0}^{\infty} p_c(k) x^k$, $G_{c,1}(x) := G'_{c,0}(x)/G'_{c,0}(1)$, $G'_{c,0}(k) := \frac{d}{dx} G_{c,0}(x)$, and $f_c \in [0, 1]$ satisfies $f_c = G_{c,1}(1 - (1 - \rho_0)(1 - f_c))$. Similar variables are also defined for $\mathcal{G}_s$ by replacing c with another symbol s.

From a different perspective, the mean fraction of functional nodes s_∞ is equal to the likelihood that a node will not be isolated at the end of a series of cascading failures. Given Propositions 1 and 2 as well as the degree distributions of $\mathcal{G}_c$ and $\mathcal{G}_s$, we can calculate s_∞ by setting $s'(t) = s'(t-1) = s'$, plugging (2)-(3) into (1), solving (1) numerically (e.g., using bisection method [13]), and employ $s_\infty = s' g_s(s')$.

D. Temporal Impact of Cascading Failures

From users' point-of-view, quality of service can also be determined by how long until the D2D-based SNS cannot provide social content exchange functionality. This is particularly important when the probability that a node will survive the cascading failures, s_∞, is fairly low so that the node will eventually be isolated with high probability. To quantify such a temporal quality of service, we define as follows.

Definition 1: Isolation time of node $u \in \mathcal{V}$ is the smallest time step in which node $u \in \mathcal{V}$ becomes isolated from the D2D-based SNS, which is defined as

$$TI_u := \inf\{t \geq 1 : u \in \mathcal{F}_s(t-1)\}. \tag{4}$$

We say that a node is *affected* if it has become isolated at the the end of the cascading failures. We assume that $TI_u = \infty$ if node u is not affected by the cascading failures. Note that the closed-form analysis of TI_u for all $u \in \mathcal{V}$ is hard. Thus, we focus on the affected nodes and examine the temporal upper bound of their disconnection time, instead.

Definition 2: Maximum isolation time is the largest time steps until an affected node becomes isolated by the cascading failures, defined as

$$TI_{max} := \sup \left\{ TI_u : u \in \lim_{t\to\infty} \mathcal{F}_s(t) \right\}. \quad (5)$$

In other words, the maximum isolation time TI_{max} quantifies the smallest time step until no more node (equivalently, less than one user) becomes isolated by the cascading failures, and can be re-defined as $TI_{max} := \inf\{t \geq 1 | s(t) - s(t-1) < \frac{1}{N}, t \in \mathbb{N}_0\}$. Based on this definition, TI_{max} can be computed by setting $c'(1) = s(0) = 1 - p_f$, calculating $s(t) = s'(t) g_s(s'(t))$ from (1), and finding the smallest t such that $s(t) - s(t-1) < \frac{1}{N}$ is satisfied.

Next, we discuss how to relate the discrete time steps to real time units (i.e., in seconds). Denote $T_c \in (0, \infty)$ as the time until additional nodes are isolated from $\mathcal{G}_c$, when the cascading failures have not yet ended. For example, T_c is equal to (`hello_interval*allowed_hello_loss`) milliseconds if AODV routing [9] is employed. We assume that users have immediate knowledge regarding their communication negihbors, such that node isolation due to social-to-communication dependence happens instantaneously. As a result, one time step in TI_u and TI_{max} is equal to T_c.

III. Resilience of D2D-based SNS

In this section, we define a node resilience metric in D2D-based SNS and study its lower and upper bounds.

A. Resilience From End-User Perspective

We examine how D2D-based SNS can support ongoing social networking session under initial node failure. Let $ST_u \in (0, \infty)$ be the *session time* of node $u \in \mathcal{V}$, i.e., the amount of time in which node u utilizes the SNS to exchange social contents. The session time ST_u is normalized to the length of one time step, T_c, and can either be exponentially- or heavy-tailed (such as Pareto)-distributed [14]. The metric ST_u is also known as the nodes' lifetime [6]. We assume that a node will not start an SNS session if it has detected any node failure in the network. Moreover, we consider the *worst-case* session time requirement by assuming that all the considered sessions start at $t = 0$, i.e., right when initial node failure occurs. Then, we are interested in: *"What is the likelihood that a D2D-based SNS sustains its users from potential isolation during their session times?"* The network-wide perspective of this probability has been referred as the network resilience, which is denoted by Ψ in [6]. Instead, we take an end-user view of resilience and define it as follows.

Definition 3: Node resilience is the likelihood that a randomly-chosen node $u \in \mathcal{V}$ does not become isolated during its lifetime, defined as

$$\Psi_n := Pr\{ST_u < TI_u\}. \quad (6)$$

The node resilience Ψ_n quantifies the D2D-based SNS's ability to support ongoing social networking session, and provides an end-user perspective of quality-of-service measure.

B. Upper and Lower Bounds of Resilience

Next, we discuss how the the mean fraction of functional nodes s_∞ and the maximum isolation time TI_{max}, obtained by following the procedures outlined in Sections II-C and II-D, can be utilized to calculate the lower and upper bounds of the node resilience Ψ_n. First, we examine the case of exponentially-distributed session times as follows.

Theorem 1: For exponentially-distributed node session times with the probability density function (PDF) of $Pr\{ST_u = x\} = \lambda e^{-\lambda x}$ and mean $\frac{1}{\lambda} > 0$, the node resilience is lower- and upper-bounded by

$$s_\infty + (1 - p_f - s_\infty)(1 - e^{-\lambda}) \leq \Psi_n \leq (p_f + s_\infty) e^{-\lambda TI_{max}} - p_f. \quad (7)$$

Proof: Let $\mathcal{C}_f := \lim_{t\to\infty} \mathcal{F}_s(t) \setminus \mathcal{I}_f$ be the set of nodes that becomes isolated at time $t \geq 1$. Then, we have

$$\begin{aligned} Pr\{ST_u > TI_u\} &= Pr\{ST_u > TI_u | u \in \mathcal{I}_f\} Pr\{u \in \mathcal{I}_f\} \\ &+ Pr\{ST_u > TI_u | u \in \mathcal{C}_f\} Pr\{u \in \mathcal{C}_f\} \\ &+ Pr\{ST_u > TI_u | u \notin \mathcal{C}_f \cup \mathcal{I}_f\} Pr\{u \notin \mathcal{C}_f \cup \mathcal{I}_f\} \quad (8) \end{aligned}$$

When node u is not affected by the cascade of failures (i.e., $u \notin \mathcal{I}_f \cup \mathcal{C}_f$), then the session time ST_u is always lower than isolation time TI_u, which equals ∞ by definition, such that the last summation term becomes zero. As a result, (8) can be re-stated as

$$\begin{aligned} &Pr\{ST_u > TI_u | u \in \mathcal{I}_f\} Pr\{u \in \mathcal{I}_f\} \\ &\quad + Pr\{ST_u > TI_u | u \in \mathcal{C}_f\} Pr\{u \in \mathcal{C}_f\} \quad (9) \\ &\geq p_f + (1 - p_f - s_\infty) Pr\{ST_u > TI_{max}\}, \quad (10) \end{aligned}$$

where (10) holds because $TI_u = 0$ when node u belongs to the set of initially-failed nodes (i.e., $u \in \mathcal{I}_f$) such that $Pr\{ST_u > TI_u | u \in \mathcal{I}_f\} = 1$, and $Pr\{ST_u > TI_u | u \in \mathcal{C}_f\} \leq Pr\{ST_u > TI_{max}\}$ because the node dropouts due to cascading failures always happen before the maximum isolation time TI_{max}. By using $\Psi_n = 1 - Pr(ST_u > TI_u)$ and plugging in the complementary cumulative distribution function (CCDF) of exponentially-distributed ST_u, we obtain the upper bound.

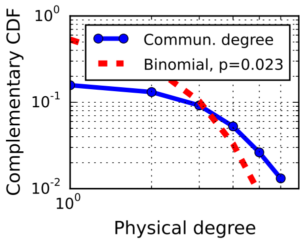
To derive the lower bound, we re-state the node resilience as follows.

$$\begin{aligned} \Psi_n &= Pr\{ST_u < TI_u\} \\ &= Pr\{ST_u < TI_u | u \in \mathcal{I}_f\} Pr\{u \in \mathcal{I}_f\} \\ &\quad + Pr\{ST_u < TI_u | u \in \mathcal{C}_f\} Pr\{u \in \mathcal{C}_f\} \quad (11) \\ &\quad + Pr\{ST_u < TI_u | u \notin \mathcal{C}_f \cup \mathcal{I}_f\} Pr\{u \notin \mathcal{C}_f \cup \mathcal{I}_f\}. \end{aligned}$$

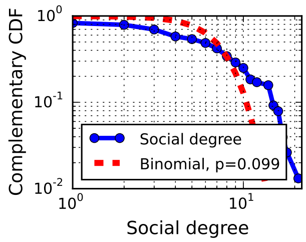
When $u \in \mathcal{I}_f$, the isolation time is always equal to $TI_{max} = 0$ by definition, such that the first summation term disappears and the node resilience Ψ_n can be re-stated as

$$\begin{aligned} &Pr\{ST_u < TI_u | u \in \mathcal{C}_f\} Pr\{u \in \mathcal{C}_f\} \\ &\quad + Pr\{ST_u < TI_u | u \notin \mathcal{C}_f \cup \mathcal{I}_f\} Pr\{u \notin \mathcal{C}_f \cup \mathcal{I}_f\} \quad (12) \\ &\geq s_\infty + (1 - s_\infty - p_f) Pr\{ST_u < 1\}, \quad (13) \end{aligned}$$

where (13) applies because $Pr\{ST_u < TI_u | u \in \mathcal{C}_f\} \leq Pr\{ST_u < 1\}$, i.e., the node dropouts due to cascading failures always happen at $t \geq 1$. ■



(a) Communication degree.



(b) Social degree.

Fig. 3. Degree distribution of sigcomm2009 trace [15] at $t = 1,500$ min.

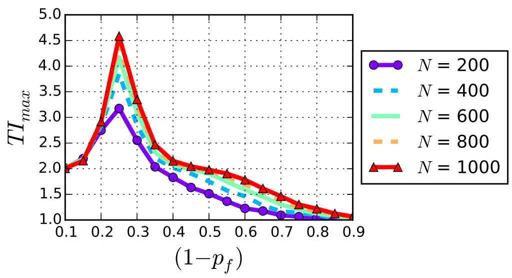


Fig. 4. Maximum isolation time TI_{max} for various number of nodes N.

Next, by following the same steps and employing the ccdf of the heavy-tailed Pareto distribution [14], we can obtain the following corollary.

Corollary 1: For Pareto-distributed ST_u with PDF $Pr\{ST_u = x\} = \frac{\alpha x_m^\alpha}{x^{\alpha+1}}$, shape parameter $\alpha > 0$, and scale $x_m > 0$, the node resilience is bounded by

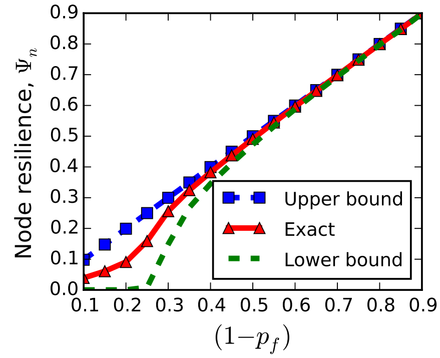
$$s_\infty + (1 - p_f + s_\infty)(1 - x_m^\alpha) \leq \Psi_n \leq (p_f + s_\infty)(x_m/TI_{max})^\alpha - p_f. \quad (14)$$

Remark 1: The results in Theorem 1 and Corollary 1 can be easily extended to the more general cases by substituting $1 - Pr\{ST_u < 1\}$ and $Pr\{ST_u > TI_{max}\}$ with the CCDFs of the session time ST_u under examination.

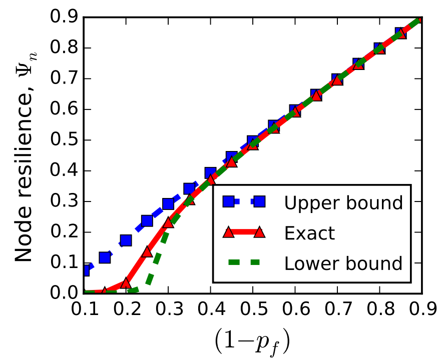
IV. Numerical Results

In this section, the analyses in Theorem 1 and Corollary 1 are validated through numerical simulations in `python`. First of all, we determine the structure of $\mathcal{G}_c$ and $\mathcal{G}_s$ by employing the Bluetooth sighting traces and Facebook friendship between 76 mobile phone users during the SIGCOMM 2009 conference [15]. We assume that a communication edge occurs between two nodes if they come into contact with each other at least once during the previous hour. From the resulting degree distributions and the curve fittings in Figs. 3(a) and 3(b), we observe that (i) the degrees are Binomially-distributed, and (ii) the communication degree $\langle k \rangle_c$ is smaller than the social degree $\langle k \rangle_s$.

Based on the trace analysis above, two Erdös-Rényi graphs, which has Binomially-distributed degrees, are generated for $\mathcal{G}_c$ and $\mathcal{G}_s$, respectively. To ensure that $\langle k \rangle_c < \langle k \rangle_s$ and the graphs $\mathcal{G}_c$ and $\mathcal{G}_s$ are respectively connected, we set $\langle k \rangle_c = 8$, $\langle k \rangle_s = 10$, and the number of nodes to $N = 10^3$. For every plot, we generate 100 unique combinations of $\mathcal{G}_c$ and $\mathcal{G}_s$, in which 10 different initial set of failed nodes $\mathcal{I}_f$ are drawn for every network realization. The maximum isolation time TI_{max} versus the fraction of initial non-isolated nodes $(1 - p_f)$ is depicted in Fig. 4. We observe the followings. Firstly, TI_{max} increases with respect to $(1 - p_f)$ up to a critical fraction [7] of $p_c = (1 - p_f^c) \approx 0.25$ and decreases exponentially above this point. It is known that the mean fraction of functional nodes is $s_\infty \approx 0$ below this critical point [7], such that for a cascade of failures to happen near p_c, all the residual nodes $\mathcal{R}_c(0)$ needs to be removed, which requires a large number of steps. Secondly, TI_{max} increases with respect to the number of nodes N, since the required gap $s(t) - s(t-1)$ for calculating TI_{max} is proportional to $\frac{1}{N}$ (see Section II-D for more details).



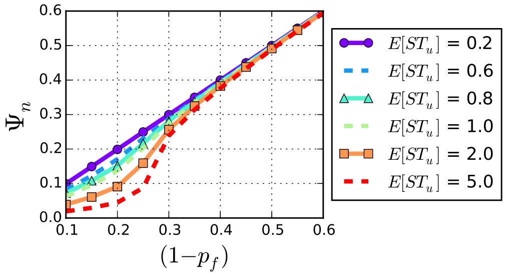
(a) Exponential distribution.



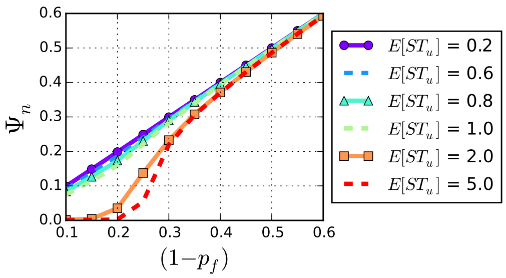
(b) Pareto (heavy-tailed) distribution.

Fig. 5. Node resilience Ψ_n for different session time distributions.

Next, we validate the bounds of the node resilience Ψ_n in Theorem 1 and Corollary 1. We set the average session time to $\mathbb{E}[ST_u] = 2$ time steps and re-use the same parameters as in the simulation of Fig. 4 to obtain the exact Ψ_n. For exponentially-distributed session times, we set $\lambda = 1/\mathbb{E}[ST_u]$. On the other hand, we fix $\alpha = 2$ to ensure that the mean is finite and assign $x_m = \frac{\alpha-1}{\alpha}\mathbb{E}[ST_u] = 1$ to get the desired mean of $\mathbb{E}[ST_u] = 2$ for Pareto-distributed session times. Due to the Binomial-distributed communication and social degree distributions, we employ the generating functions of



(a) Exponentially-distributed ST_u.



(b) Pareto-distributed ST_u.

Fig. 6. Node resilience for various mean session times $\mathbb{E}[ST_u]$'s.

$G_{c,0}(x) = e^{\langle k \rangle_c (x-1)}$ and $G_{s,0}(x) = e^{\langle k \rangle_s (x-1)}$ for calculating (2) and (3). Then, the mean fraction of functional nodes s_∞ and the maximum isolation time TI_{max} are calculated using the methods outlined in Sections II-C and II-D, which are then substituted into Theorem 1 and Corollary 1 to get the lower and upper bounds of Ψ_n. From the plots in Figs. 5(a) and 5(b), we observe that the bounds in (7) and (14) are valid and becomes tighter as the fraction of initial non-isolated nodes $(1 - p_f)$ increases. Although the TI_{max} in Fig. 4 exhibits a *double-scaling*, i.e., TI_{max} increases when $(1-p_f)$ is small and decreases against $(1-p_f)$, otherwise, we observe that the bounds of the node resilience Ψ_n in Theorem 1 and Corollary 1 are dominated by the mean fraction of functional nodes at the end of the cascading failures s_∞, which is monotonically-increasing with respect to $(1 - p_f)$ [7, Fig. 3(a)]. Thus, Ψ_n becomes an increasing function of $(1-p_f)$. Moreover, some users have very long session times when ST_u follows Pareto (heavy-tailed) distribution such that we observe the Ψ_n in Fig. 5(b) is lower than that when ST_u is exponentially-distributed in Fig. 5(a).

Finally, we examine the impact of various mean session time $\mathbb{E}[ST_u]$ to the resilience of the D2D-based SNS. We employ exponentially-distributed session times in Fig. 6(a). Since $(1 - p_f) > s_\infty > 0$, the lower and upper bounds in (7) is proportional to $\Psi_n \sim 1 - e^{-\lambda}$. Thus, $\lambda = 1/\mathbb{E}[ST_u]$ becomes smaller and the node resilience Ψ_n will also become smaller when $\mathbb{E}[ST_u]$ increases. The case of Pareto-distributed session times, on the other hand, is examined in Fig. 5(b). In this case, the bounds in Corollary 1 is proportional to $1 - x_m^\alpha$. Since we fix α to 2 throughout the simulations and set $x_m = \frac{\alpha-1}{\alpha}\mathbb{E}[ST_u]$, then, the session time Ψ_n is inversely-proportional to the mean session time $\mathbb{E}[ST_u]$. Note that, with respect to the mean session time, $1 - x_m^\alpha|_{x_m = \frac{\alpha-1}{\alpha}\mathbb{E}[ST_u]}$ decreases slower than $1 - e^{-\lambda}|_{\lambda = 1/\mathbb{E}[ST_u]}$, such that the resilience of D2D-based SNS with Pareto-distributed ST_u in Fig. 6(b) is lower than that with exponential ST_u in Fig. 6(a) for large mean session time $\mathbb{E}[ST_u]$.

V. Conclusion

In this paper, we analyzed the resilience of D2D-based SNS against cascade of failures induced by random initial node failures, from an end-user perspective. We outlined methods for calculating the mean fraction of surviving nodes and the maximum isolation time, which quantify the number of nodes not affected by the cascade of failures and the time until the end of such sequence of failures, respectively. By using these metrics, we derive the lower and upper bounds of resilience for exponentially- and Pareto (heavy-tailed)-distributed social networking session times. The numerical and analytical results indicate that D2D-based SNS with Pareto-distributed session times has lower resilience and thus provide poor quality-of-service, compared to that with exponential session times.

References

[1] Apple, "Multipeer Connectivity Framework Reference." https://developer.apple.com/library/ios/documentation/MultipeerConnectivity/Reference/MultipeerConnectivityFramework/. Accessed: 2016-03-28.

[2] CNN Tech, "FireChat in Hong Kong: How an app tapped its way into the protests." http://www.cnn.com/2014/10/16/tech/mobile/tomorrow-transformed-firechat/. Accessed: 2016-03-28.

[3] T. Dimitar, F. Sonja, M. Jani, and G. Aksenti, "Connection resilience to nodes failures in ad hoc networks," in *Proc. 12th IEEE MELECON*, vol. 2, pp. 579–582, IEEE, 2004.

[4] C. Bettstetter, "On the minimum node degree and connectivity of a wireless multihop network," in *Proc. 3rd ACM MobiHoc*, pp. 80–91, ACM, 2002.

[5] F. Xing and W. Wang, "Modeling and analysis of connectivity in mobile ad hoc networks with misbehaving nodes," in *Proc. IEEE ICC'06*, vol. 4, pp. 1879–1884, IEEE, 2006.

[6] F. Xing and W. Wang, "On the expected connection lifetime and stochastic resilience of wireless multi-hop networks," in *Proc. IEEE GLOBECOM'07*, pp. 1263–1267, IEEE, 2007.

[7] S. V. Buldyrev, R. Parshani, G. Paul, H. E. Stanley, and S. Havlin, "Catastrophic cascade of failures in interdependent networks," *Nature*, vol. 464, no. 7291, pp. 1025–1028, 2010.

[8] D. B. West *et al.*, *Introduction to graph theory*, vol. 2. Prentice hall Upper Saddle River, 2001.

[9] C. Perkins, E. Belding-Royer, and S. Das, "Ad hoc on-demand distance vector (AODV) routing," tech. rep., 2003.

[10] Z. Lu, Y. Wen, and G. Cao, "Information diffusion in mobile social networks: The speed perspective," in *Proc. 2014 IEEE INFOCOM*, pp. 1932–1940, IEEE, 2014.

[11] X. Huang, J. Gao, S. V. Buldyrev, S. Havlin, and H. E. Stanley, "Robustness of interdependent networks under targeted attack," *Physical Review E*, vol. 83, no. 6, p. 065101, 2011.

[12] R. Parshani, S. V. Buldyrev, and S. Havlin, "Interdependent networks: Reducing the coupling strength leads to a change from a first to second order percolation transition," *Physical review letters*, vol. 105, no. 4, p. 048701, 2010.

[13] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge UP, 2004.

[14] G. Casella and R. L. Berger, *Statistical inference*, vol. 2. Duxbury Pacific Grove, CA, 2002.

[15] A.-K. Pietilainen, "CRAWDAD data set thlab/sigcomm2009 (v. 2012-07-15)." Downloaded from http://crawdad.org/, July 2012.