Heterogeneous Projection of Disruptive Malware Prevalence in Mobile Social Networks

Aldiyar Dabarov, Madiyar Sharipov, Aresh Dadlani, *Member, IEEE*, Muthukrishnan Senthil Kumar, *Member, IEEE*, Walid Saad, *Fellow, IEEE*, and Choong Seon Hong, *Senior Member, IEEE*

***Abstract*— Segregating the latency phase from the actual disruptive phase of certain mobile malware grades offers more opportunities to effectively mitigate the viral spread in its early stages. Inspired by epidemiology, in this letter, a stochastic propagation model that accounts for infection latency of disruptive malware in both personal and spatial social links between constituent mobile network user pairs is proposed. To elucidate the true impact of unique user attributes on the virulence of the proposed spreading process, heterogeneity in transition rates is also considered in an approximated mean-field epidemic network model. Furthermore, derivations for the system equilibrium and stability analysis are provided. Simulation results showcase the viability of our model in contrasting between latent and disruptive infection stages with respect to a homogeneous population-level benchmark model.**

***Index Terms*— Mobile social networks, heterogeneous epidemic model, disruptive virus, mean-field theory, equilibrium analysis.**

I. Introduction

PREDOMINANTLY driven by the prevalent use of personal mobile devices, the evolution of assortative interactions between mobile social network (MSN) users has initiated an array of research topics [1]. The heterogeneity of users and of the multi-faceted relations among them however, further complicate the characterization of information flow intertwined with the underlying network structure. While information diffusion can take many forms, the precision of modeling frameworks in foreseeing malware outbreaks in MSNs remains a key challenge.

With the expanding smartphone market, the vectors exploited by different malware strains to infect susceptible smart devices have also grown in variety [2], [3]. Some commonly adopted vectors for cyber attacks include Bluetooth (BT), email attachments, and SMS/MMS messaging services (MS). Digital viruses can exploit both *personal* and *spatial* social links to propagate in MSNs [4]–[6]. Personal links are established based on the contact lists and anonymized call records in each handset, while standard short-range communication protocols such as wireless BT define the spatial ties between neighboring mobile users within a given transmission range.

Epidemic models have been instrumental in quantitatively predicting malware outbreaks in generic social networks. In regard with population-level models, the authors in [5] characterize hybrid mobile viruses that exploit BT and MS protocols to target the susceptible user base. In [6], the mixed behaviors of long-range infection spreading pattern through MS and ripple-based infection via BT using ordinary differential equations is investigated. Epidemic-based information dissemination in MSNs using opportunistic peer-to-peer links has been studied in [7]. More recently, mean-field approximations of individual-based spreading processes have been compelling in exposing the relationship between epidemic thresholds and the spectral radius of contact networks. However, relatively fewer agent-based epidemic models exist that describe the time evolution of the state occupancy probabilities in terms of the number of Markovian users over complex networks [8]–[11]. The authors in [8] proved that the infection-free equilibrium in aggregated-Markovian random graph processes is almost surely exponential. To account for user tendency to switch between alternative social preferences, the authors in [9] generalized the seminal work in [10] to derive the steady-state phase transition thresholds between extinction, co-existence, and absolute dominance of memes. In [11], the authors proposed a continuous-time bi-layer network model with generic infection rates to analyze the dynamics of competitive spreading in multi-virus scenarios.

Unlike conventional virus models that impair the functionality of mobile gadgets immediately after being transmitted, *disruptive* viruses such as Commwarrior, Mellisa, CIH, and BlueBorne have a two-phase life cycle: the *latent phase* succeeded by the *disruptive phase* [12]. In the former phase, the malware infects other connected susceptible nodes by replicating itself while residing in the victimized host, whereas the functionality of the infected host is hindered only in the latter phase. All the above efforts ([5]–[9], [11]) fail to discern between the two phases therefore, resulting in over-estimated predictions on the fraction of mobile users infected by disruptive malware. To our best knowledge, there exists no prior analytical work quantifying the infection latency for hybrid disruptive malware spread in MSNs, where microscopic user-level dynamics are incorporated and the state transition rates are subjective to each user.

To fill this gap, the main contribution of this letter is a novel mean-field approximated epidemic model to characterize the

Manuscript received February 8, 2020; revised March 23, 2020; accepted May 1, 2020. Date of publication May 6, 2020; date of current version August 12, 2020. This research was supported by the Faculty Development Competitive Research Grant (No. 240919FD3918), Nazarbayev University. The associate editor coordinating the review of this letter and approving it for publication was T. Han. *(Corresponding author: Aresh Dadlani.)*

Aldiyar Dabarov, Madiyar Sharipov, and Aresh Dadlani are with the Department of ECE, Nazarbayev University, Nur-Sultan 010000, Kazakhstan (e-mail: aldiyar.dabarov@nu.edu.kz; madiyar.sharipov@nu.edu.kz; aresh.dadlani@nu.edu.kz).

Muthukrishnan Senthil Kumar is with the Department of Applied Mathematics and Computational Sciences, PSG College of Technology, Coimbatore 641004, India (e-mail: msk@amc.psgtech.ac.in).

Walid Saad is with the Bradley Department of Electrical and Computer Engineering, Virginia Tech, Blacksburg, VA 24061 USA, and also with the Department of Computer Engineering, Kyung Hee University, South Korea (e-mail: walids@vt.edu).

Choong Seon Hong is with the Department of Computer Engineering, Kyung Hee University, Yongin 449-701, South Korea.

Digital Object Identifier 10.1109/LCOMM.2020.2992562

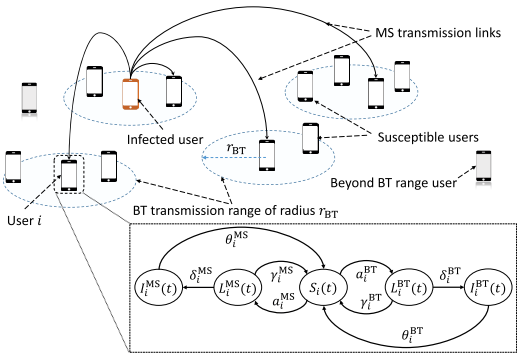


Fig. 1. Propagation mechanism and the proposed epidemic model of a disruptive malware in MSNs.

spreading pattern of disruptive malware promoted through BT and MS protocols in generic MSNs by subsuming the heterogeneity in user-level interactions. To ensure the validity of the resulting prediction in steady-state, we prove the existence of a unique viral equilibrium and investigate the asymptotic stability of the viral-free equilibrium for our model. We then demonstrate the precision of the proposed network model by benchmarking it against a homogeneously-mixing population-level epidemic model via simulations. Results obtained show the efficacy of the model in differentiating between users infected via the two transmission modes. In absence of large empirical data samples, such a fine-tuned projection model would help devise effective control strategies in significantly shorter time and minimize the investment costs incurred.

II. System Model Description

Consider a typical MSN of size N, where each mobile user, labeled from 1 to N, interact with each other using smart devices. To distinguish between social links created by spatial BT and personal MS links in the network, we define respectively, graphs $G_1(\mathcal{V},\mathcal{E}_1)$ and $G_2(\mathcal{V},\mathcal{E}_2)$, where $\mathcal{V}=\{1,2,\ldots,N\}$. Link $(i,j)\in\mathcal{E}_1$ if users i and j are within the BT transmission range of radius r_{BT}. Similarly, link $(i,j)\in\mathcal{E}_2$ if user i is connected to user j in the personal social network. Let $\boldsymbol{A}\triangleq[a_{ij}]_{N\times N}$ and $\boldsymbol{B}\triangleq[b_{ij}]_{N\times N}$ be the irreducible adjacency matrices corresponding to G_1 and G_2, respectively. We assume that the network is undirected and thus, matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ are symmetric.

Consistent with definitions in the literature, each user is either in the *susceptible* (S), *latent* (L), or *disruptive* (I) state at any given time as shown in Fig. 1. User i is said to be susceptible if it is healthy and not yet infected by the malware. Upon receiving the malware, the user becomes latent if the infection is in the latent phase and then turns disruptive in the successive phase [12]. To represent the state of the network at time t, we define the stochastic process $\{X_i(t);t\geq 0\}$, where $\forall i\in\mathcal{V}$, $X_i(t)$ is:

$$X_i(t)=\begin{cases}0; & \text{if user } i \text{ is susceptible at } t,\\ 1; & \text{if user } i \text{ is latent by BT at } t,\\ 2; & \text{if user } i \text{ is disruptive by BT at } t,\\ 3; & \text{if user } i \text{ is latent by MS at } t,\\ 4; & \text{if user } i \text{ is disruptive by MS at } t.\end{cases} \tag{1}$$

Using (1), we now denote the probability of user i being in any one of the five possible states as $S_i(t)=\Pr\{X_i(t)=0\}$, $L_i^{\text{BT}}(t)=\Pr\{X_i(t)=1\}$, $I_i^{\text{BT}}(t)=\Pr\{X_i(t)=2\}$, $L_i^{\text{MS}}(t)=\Pr\{X_i(t)=3\}$, and $I_i^{\text{MS}}(t)=\Pr\{X_i(t)=4\}$, where $1\leq i\leq N$ and $S_i(t)+L_i^{\text{BT}}(t)+I_i^{\text{BT}}(t)+L_i^{\text{MS}}(t)+I_i^{\text{MS}}(t)=1$.

A susceptible user i is infected by user j via BT asynchronous connectionless link (ACL) in the *discoverable mode* at the constant rate of $\beta_j^{L_{\text{BT}}}>0$ if user j is in the latent infection state and with rate $\beta_j^{I_{\text{BT}}}>0$ if user j is in the disruptive state. Likewise, in MS-mediated propagation, user i is infected by latent (disruptive) user j at the constant rate of $\beta_j^{L_{\text{MS}}}(\beta_j^{I_{\text{MS}}})>0$. The latency time for latent user i affected through BT (MS) is assumed to be exponentially distributed with the latency rate of $\delta_i^{\text{BT}}(\delta_i^{\text{MS}})>0$ [5], [6]. Due to frequent updates of mobile operating systems and anti-viruses, a latent (disruptive) user i recovers back to susceptibility at rates $\gamma_i^{\text{BT}}(\theta_i^{\text{BT}})>0$ and $\gamma_i^{\text{MS}}(\theta_i^{\text{MS}})>0$ in the BT and MS settings, respectively.

The corresponding continuous-time Markov process of the proposed model becomes analytically intractable as the state space size grows exponentially with increase in $|\mathcal{V}|$. Approximation techniques are normally applied to resolve the state space size problem at the expense of accuracy. Deductions from mean-field approximated network models are shown to be asymptotically almost exact to sufficiently large real-world networks as they provide an upper bound for the exact probability of user infection [13]. Accordingly, we adopt a first-order mean-field approximation to reduce the dimensionality of the exact Markovian network model down to $4N$ space.

Given a sufficiently small time interval of $(t+\Delta t)$, for $\Delta t>0$, the following state transition probabilities for mobile user i hold valid, where the remaining invalid state transition conditional probabilities are denoted by the asymptotic notation $o(\Delta t)$:

$$\begin{cases}\Pr\{X_i(t+\Delta t)=1|X_i(t)=0\}=a_i^{\text{BT}}\Delta t+o(\Delta t),\\ \Pr\{X_i(t+\Delta t)=3|X_i(t)=0\}=a_i^{\text{MS}}\Delta t+o(\Delta t),\\ \Pr\{X_i(t+\Delta t)=0|X_i(t)=1\}=\gamma_i^{\text{BT}}\Delta t+o(\Delta t),\\ \Pr\{X_i(t+\Delta t)=2|X_i(t)=1\}=\delta_i^{\text{BT}}\Delta t+o(\Delta t),\\ \Pr\{X_i(t+\Delta t)=0|X_i(t)=2\}=\theta_i^{\text{BT}}\Delta t+o(\Delta t),\\ \Pr\{X_i(t+\Delta t)=0|X_i(t)=3\}=\gamma_i^{\text{MS}}\Delta t+o(\Delta t),\\ \Pr\{X_i(t+\Delta t)=4|X_i(t)=3\}=\delta_i^{\text{MS}}\Delta t+o(\Delta t),\\ \Pr\{X_i(t+\Delta t)=0|X_i(t)=4\}=\theta_i^{\text{MS}}\Delta t+o(\Delta t),\end{cases} \tag{2}$$

The linear infection rates via BT and MS links, denoted by a_i^{BT} and a_i^{MS}, respectively, are defined as follows:

$$a_i^{\text{BT}}\triangleq\sum_{j=1}^{N}a_{i,j}\left[\beta_j^{L_{\text{BT}}}L_j^{\text{BT}}(t)+\beta_j^{I_{\text{BT}}}I_j^{\text{BT}}(t)\right] \tag{3}$$

$$\text{and } a_i^{\text{MS}}\triangleq\sum_{j=1}^{N}b_{i,j}\left[\beta_j^{L_{\text{MS}}}L_j^{\text{MS}}(t)+\beta_j^{I_{\text{MS}}}I_j^{\text{MS}}(t)\right]. \tag{4}$$

Undertaking the approach in [11], we use (2) to derive $L_i^{\text{BT}}(t+\Delta t)$, $I_i^{\text{BT}}(t+\Delta t)$, $L_i^{\text{MS}}(t+\Delta t)$, and $I_i^{\text{MS}}(t+\Delta t)$ based on the law of total probability. We then linearize the resultant system in what follows to facilitate our analysis in closed form.

A. BT-Mediated Spreading Dynamics

Transition of user i to the latent state in a BT network at time $(t+\Delta t)$ can occur only if (i) user i was susceptible or (ii) latent at time t. Mathematically, this is expressed as:

$$\begin{aligned}L_i^{\text{BT}}(t+\Delta t)=S_i(t)\cdot\Pr\{X_i(t+\Delta t)=1|X_i(t)=0\}\\ +L_i^{\text{BT}}(t)\cdot\Pr\{X_i(t+\Delta t)=1|X_i(t)=1\}.\end{aligned} \tag{5}$$

Similarly, user i enters the disruptive state at $(t+\Delta t)$ via BT communication only if (i) latent or (ii) disruptive at time t, i.e.,

$$I_i^{\text{BT}}(t+\Delta t) = L_i^{\text{BT}}(t)\cdot \Pr\{X_i(t+\Delta t)=2|X_i(t)=1\} + I_i^{\text{BT}}(t)\cdot \Pr\{X_i(t+\Delta t)=2|X_i(t)=2\}. \quad (6)$$

By substituting (2) in (5) and (6), dividing both sides by Δt, and letting $\Delta t \to 0$, we arrive at the following system:

$$\begin{cases} \dfrac{dL_i^{\text{BT}}(t)}{dt} = \left[1 - L_i^{\text{BT}}(t) - I_i^{\text{BT}}(t) - L_i^{\text{MS}}(t) - I_i^{\text{MS}}(t)\right] a_i^{\text{BT}} \\ \qquad\qquad - (\gamma_i^{\text{BT}} + \delta_i^{\text{BT}}) L_i^{\text{BT}}(t), \quad \forall i = 1,2,\ldots,N, \\ \dfrac{dI_i^{\text{BT}}(t)}{dt} = \delta_i^{\text{BT}} L_i^{\text{BT}}(t) - \theta_i^{\text{BT}} I_i^{\text{BT}}(t), \quad \forall i = 1,2,\ldots,N. \end{cases} \quad (7)$$

B. MS-Mediated Spreading Dynamics

The MS spreading model is derived in a similar manner, with the resulting system having the same structure as (7), except for the matrix $\boldsymbol{B}$ corresponding to the personal social network, where a_i^{BT} is replaced by a_i^{MS}:

$$\begin{cases} \dfrac{dL_i^{\text{MS}}(t)}{dt} = \left[1 - L_i^{\text{BT}}(t) - I_i^{\text{BT}}(t) - L_i^{\text{MS}}(t) - I_i^{\text{MS}}(t)\right] a_i^{\text{MS}} \\ \qquad\qquad - (\gamma_i^{\text{MS}} + \delta_i^{\text{MS}}) L_i^{\text{MS}}(t), \quad \forall i = 1,2,\ldots,N, \\ \dfrac{dI_i^{\text{MS}}(t)}{dt} = \delta_i^{\text{MS}} L_i^{\text{MS}}(t) - \theta_i^{\text{MS}} I_i^{\text{MS}}(t), \quad \forall i = 1,2,\ldots,N. \end{cases} \quad (8)$$

Hence, the approximated network model is a system of $4N$ differential equations represented by (7) and (8) collectively.

III. Equilibrium and Stability Analysis

We now postulate a theorem related to the global stability of the trivial infection-free equilibrium, given by $\boldsymbol{E}_0$, and then derive the unique non-trivial virulent equilibrium, $\boldsymbol{E}^*$. In steady-state, such analysis ensures that our non-linear model reaches an equilibrium point irrespective of the initial number of infected users. This is thus, necessary to justify the prediction fidelity of our model by showing that it stabilizes in a positively invariant state space [11]. To this end, we define vector $\boldsymbol{D}(t)$ as:

$$\begin{aligned} \boldsymbol{D}(t) &\triangleq \big(L_1^{\text{BT}}(t),\ldots,L_N^{\text{BT}}(t), I_1^{\text{BT}}(t),\ldots,I_N^{\text{BT}}(t), \\ &\qquad L_1^{\text{MS}}(t),\ldots,L_N^{\text{MS}}(t), I_1^{\text{MS}}(t),\ldots,I_N^{\text{MS}}(t)\big)^T \\ &= \big(L_{1\ldots N}^{\text{BT}}(t), I_{1\ldots N}^{\text{BT}}(t), L_{1\ldots N}^{\text{MS}}(t), I_{1\ldots N}^{\text{MS}}(t)\big)^T. \end{aligned} \quad (9)$$

Also, let the reduced state space, Ω, be given as:

$$\Omega = \Big\{\big(L_{1\ldots N}^{\text{BT}}(t), I_{1\ldots N}^{\text{BT}}(t), L_{1\ldots N}^{\text{MS}}(t), I_{1\ldots N}^{\text{MS}}(t)\big)^T \in \mathbb{R}_+^{4N} \,\Big|\, L_i^{\text{BT}}(t) + I_i^{\text{BT}}(t) + L_i^{\text{MS}}(t) + I_i^{\text{MS}}(t) \le 1, i = 1,\ldots,N\Big\}. \quad (10)$$

Since $L_i^{\text{BT}}(t)$, $I_i^{\text{BT}}(t)$, $L_i^{\text{MS}}(t)$, and $I_i^{\text{MS}}(t)$ are probabilistic values in $[0,1]$ that sum up to one for all $t \ge 0$, Ω is positively invariant for the model in (7) and (8) [12]. In other words, $\boldsymbol{D}(0) \in \Omega$ implies that $\boldsymbol{D}(t) \in \Omega$ for all t values. Our proposed system has a trivial steady-state equilibrium $\boldsymbol{E}_0 = (0,0,\ldots,0)^T$ which is always infection-free. An equilibrium is said to be globally stable if it is both, *asymptotically stable* and *globally attracting*. For matrices $\boldsymbol{Y_1} \triangleq \boldsymbol{A}\cdot \text{diag}(\beta_i^{L_{\text{BT}}})$ and $\boldsymbol{Z_1} \triangleq \boldsymbol{B}\cdot \text{diag}(\beta_i^{L_{\text{MS}}})$, the following theorem examines the global stability condition for $\boldsymbol{E}_0$, where $c \triangleq \min_{1\le i\le N}\{\gamma_i^{\text{BT}}, \gamma_i^{\text{MS}}\}$, $\boldsymbol{I}$ is the identity matrix of order N, and $\lambda_1(\cdot)$ is the spectral radius of a square matrix.

Theorem 1: Equilibrium $\boldsymbol{E}_0$ is globally asymptotically stable with respect to Ω if $\lambda_1(\boldsymbol{Y_1} + \boldsymbol{Z_1} - c\boldsymbol{I}) < 0$.

Proof: Let $C_i(t)$ be the sum $L_i^{\text{BT}}(t) + I_i^{\text{BT}}(t) + L_i^{\text{MS}}(t) + I_i^{\text{MS}}(t)$. For all $i \in \mathcal{V}$, taking the derivative of $C_i(t)$ yields:

$$\begin{aligned} \frac{dC_i(t)}{dt} &= \left[1 - C_i(t)\right]\left(a_i^{\text{BT}} + a_i^{\text{MS}}\right) - \gamma_i^{\text{BT}} L_i^{\text{BT}}(t) - \theta_i^{\text{BT}} I_i^{\text{BT}}(t) \\ &\quad - \gamma_i^{\text{MS}} L_i^{\text{MS}}(t) - \theta_i^{\text{MS}} I_i^{\text{MS}}(t) \\ &\le \sum_{j=1}^{N} a_{i,j}\beta_j^{L_{\text{BT}}} C_j(t) + \sum_{j=1}^{N} b_{i,j}\beta_j^{L_{\text{MS}}} C_j(t) - c\ C_i(t), \end{aligned}$$

For $\boldsymbol{w}(t) \triangleq \big(w_1(t), w_2(t), \ldots, w_N(t)\big)^T$ and $w_i(0) = C_i(0)$, $\forall i \in \mathcal{V}$, the comparison system can be expressed as:

$$\frac{dw_i(t)}{dt} = \sum_{j=1}^{N} a_{i,j}\beta_j^{L_{\text{BT}}} w_j(t) + \sum_{j=1}^{N} b_{i,j}\beta_j^{L_{\text{MS}}} w_j(t) - c\ w_i(t),$$

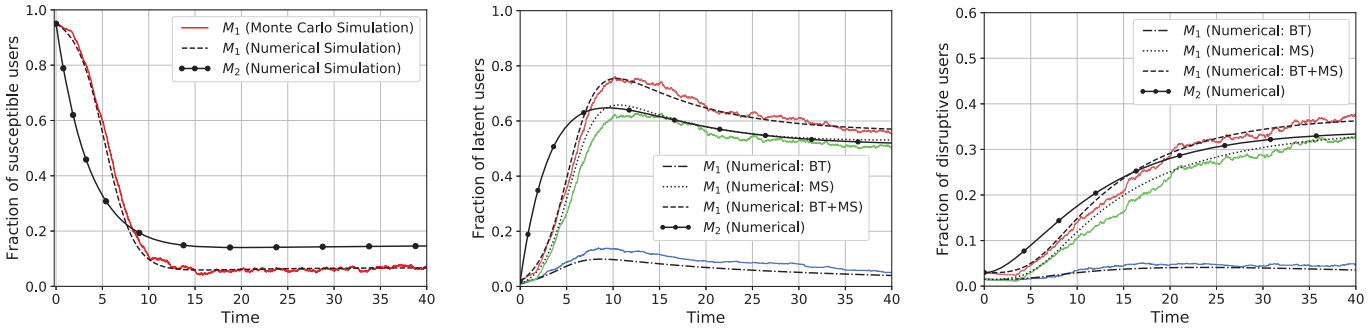
and re-written in matrix form as $\boldsymbol{w}'(t) = (\boldsymbol{Y_1} + \boldsymbol{Z_1} - c\boldsymbol{I})\boldsymbol{w}(t)$. Since $\lambda_1(\boldsymbol{Y_1} + \boldsymbol{Z_1} - c\boldsymbol{I}) < 0$, it follows from the fundamental theory on linear differential systems that $\boldsymbol{w}(t) \to \boldsymbol{0}$. Consequently, according to Chaplygin lemma on differential inequalities, we have $\boldsymbol{D}(t) \le \boldsymbol{w}(t)$ for all $t > 0$ values. Thus, as t approaches infinity, $\boldsymbol{D}(t) \to \boldsymbol{0}$, which completes the proof. □

In epidemiology, the existence of the non-trivial viral equilibrium $\boldsymbol{E}^*$ is determined by the outbreak threshold, commonly known as the *basic reproduction ratio* ($\mathcal{R}_0$). In particular, the infection eventually dies out in the network (i.e., reaches $\boldsymbol{E}_0$) if $\mathcal{R}_0 < 1$ and persists (i.e., converges to $\boldsymbol{E}^*$) if $\mathcal{R}_0 > 1$. Such an equilibrium can now be obtained by considering (7) and (8) together in steady-state. Thus, for all $i = 1,2,\ldots,N$, setting the left-side derivatives of the equations to zero yields the following:

$$\begin{cases} I_i^{\text{BT}} = \dfrac{\delta_i^{\text{BT}}}{\theta_i^{\text{BT}}} L_i^{\text{BT}}, \\ I_i^{\text{MS}} = \dfrac{\delta_i^{\text{MS}}}{\theta_i^{\text{MS}}} L_i^{\text{MS}}, \\ \epsilon_i^{\text{BT}} L_i^{\text{BT}} = \left(1 - L_i^{\text{BT}} - \dfrac{\delta_i^{\text{BT}}}{\theta_i^{\text{BT}}} L_i^{\text{BT}} - L_i^{\text{MS}} - \dfrac{\delta_i^{\text{MS}}}{\theta_i^{\text{MS}}} L_i^{\text{MS}}\right) a_i^{\text{BT}}, \\ \epsilon_i^{\text{MS}} L_i^{\text{MS}} = \left(1 - L_i^{\text{BT}} - \dfrac{\delta_i^{\text{BT}}}{\theta_i^{\text{BT}}} L_i^{\text{BT}} - L_i^{\text{MS}} - \dfrac{\delta_i^{\text{MS}}}{\theta_i^{\text{MS}}} L_i^{\text{MS}}\right) a_i^{\text{MS}}, \end{cases} \quad (11)$$

where ϵ_i^{BT} and ϵ_i^{MS} denote $(\delta_i^{\text{BT}} + \gamma_i^{\text{BT}})$ and $(\delta_i^{\text{MS}} + \gamma_i^{\text{MS}})$, respectively. By solving (11) for L_i^{BT}, I_i^{BT}, L_i^{MS}, and I_i^{MS}, it can be easily deduced that $\boldsymbol{D}(t)$ is a non-trivial equilibrium of the proposed model if and only if $\forall i \in \mathcal{V}$:

$$\begin{cases} I_i^{\text{BT}} = \dfrac{a_i^{\text{BT}}\delta_i^{\text{BT}}\epsilon_i^{\text{MS}}\theta_i^{\text{MS}}}{\epsilon_i^{\text{MS}}\theta_i^{\text{MS}}(\epsilon_i^{\text{BT}}\theta_i^{\text{BT}} + a_i^{\text{BT}}\nu_i^{\text{BT}}) + a_i^{\text{MS}}\epsilon_i^{\text{BT}}\nu_i^{\text{MS}}\theta_i^{\text{BT}}}, \\ I_i^{\text{MS}} = \dfrac{a_i^{\text{MS}}\delta_i^{\text{MS}}\epsilon_i^{\text{BT}}\theta_i^{\text{BT}}}{\epsilon_i^{\text{MS}}\theta_i^{\text{MS}}(\epsilon_i^{\text{BT}}\theta_i^{\text{BT}} + a_i^{\text{BT}}\nu_i^{\text{BT}}) + a_i^{\text{MS}}\epsilon_i^{\text{BT}}\nu_i^{\text{MS}}\theta_i^{\text{BT}}}, \\ L_i^{\text{BT}} = \dfrac{a_i^{\text{BT}}\theta_i^{\text{BT}}\epsilon_i^{\text{MS}}\theta_i^{\text{MS}}}{\epsilon_i^{\text{MS}}\theta_i^{\text{MS}}(\epsilon_i^{\text{BT}}\theta_i^{\text{BT}} + a_i^{\text{BT}}\nu_i^{\text{BT}}) + a_i^{\text{MS}}\epsilon_i^{\text{BT}}\nu_i^{\text{MS}}\theta_i^{\text{BT}}}, \\ L_i^{\text{MS}} = \dfrac{a_i^{\text{MS}}\theta_i^{\text{BT}}\epsilon_i^{\text{BT}}\theta_i^{\text{MS}}}{\epsilon_i^{\text{MS}}\theta_i^{\text{MS}}(\epsilon_i^{\text{BT}}\theta_i^{\text{BT}} + a_i^{\text{BT}}\nu_i^{\text{BT}}) + a_i^{\text{MS}}\epsilon_i^{\text{BT}}\nu_i^{\text{MS}}\theta_i^{\text{BT}}}, \end{cases} \quad (12)$$



(a) Time evolution of susceptible population. (b) Time evolution of latent population. (c) Time evolution of disruptive population.

Fig. 2. Transient and steady-state comparison between the proposed (M_1) and benchmark (M_2) models, where $r_{\text{BT}} = 10$ meters and the initial number of infected users are $L^{\text{BT}}(0) = L^{\text{MS}}(0) = 10$ and $I^{\text{BT}}(0) = I^{\text{MS}}(0) = 15$ users as in [6] and [12].

where ν_i^{BT} and ν_i^{MS} represent $(\delta_i^{\text{BT}} + \theta_i^{\text{BT}})$ and $(\delta_i^{\text{MS}} + \theta_i^{\text{MS}})$, respectively. Proof of the sufficient conditions for (12) to exist has been excluded due to the page limitation. Nonetheless, we refer the reader to [12] for details on a similar derivation.

In summary, Theorem 1 showed that the state of the network model derived in (7) and (8) will always belong to Ω. If the network approaches the trivial equilibrium $\boldsymbol{E}_0$ on the long-run, then the malware spread in the user population would eventually die out leaving all devices susceptible. There however, exists another unique equilibrium $\boldsymbol{E}^*$ at which some constant fraction of the population will always remain infected. Hence, if the network reaches $\boldsymbol{E}^*$, our model is capable of not only distinguishing latent users from disruptive users, but also identifying the transmission protocol promoting the malware spread using (12). In turn, such information allows for early and effective implementation of cost-aware control measures.

IV. Simulation Results and Discussions

In this section, Monte Carlo and numerical simulations are conducted to validate the accuracy of our model (M_1) derived in (7) and (8). An arbitrary MSN of $N = 1000$ mobile users is implemented using the GEMFsim tool [14]. For comparison, we consider N homogeneously mixing users distributed randomly in a 100×100 geographical area with density σ similar to [6] as the benchmark. Unlike M_1, where the user interactions are governed by contact matrices $\boldsymbol{A}$ and $\boldsymbol{B}$, all users in the benchmark model (M_2) have equal probability to receive the malware via MS while each infected user can contact $\sigma\pi r_{\text{BT}}^2$ neighboring nodes in discoverable BT mode. It is noteworthy to mention that M_2 is a limiting case of M_1 and the two models would converge in the case of a fully connected network. The M_1 approximation is thus, more reliable for statistical empirical data analysis as it spans over various network structures.

Without loss of generality, the transition rates for all $i \in \mathcal{V}$ are taken to be fixed by dropping the subscripts. That is to say, $\beta_i^{L_{\text{BT}}} = \beta^{L_{\text{BT}}}$, $\beta_i^{L_{\text{MS}}} = \beta^{L_{\text{MS}}}$, $\beta_i^{I_{\text{BT}}} = \beta^{I_{\text{BT}}}$, $\beta_i^{I_{\text{MS}}} = \beta^{I_{\text{MS}}}$, $\gamma_i^{\text{BT}} = \gamma^{\text{BT}}$, $\gamma_i^{\text{MS}} = \gamma^{\text{MS}}$, $\delta_i^{\text{BT}} = \delta^{\text{BT}}$, $\delta_i^{\text{MS}} = \delta^{\text{MS}}$, $\theta_i^{\text{BT}} = \theta^{\text{BT}}$, and $\theta_i^{\text{MS}} = \theta^{\text{MS}}$. To mimic the disruptive behavior of malware, we also set $\beta^{L_{\text{BT}}} > \beta^{I_{\text{BT}}}$, $\beta^{L_{\text{MS}}} > \beta^{I_{\text{MS}}}$, $\theta^{\text{BT}} > \gamma^{\text{BT}}$, and $\theta^{\text{MS}} > \gamma^{\text{MS}}$ as in [12]. Unless explicitly specified, the simulation parameters and initial network conditions are given in Table I.

Fig. 2 shows the population size distribution for each epidemic class with respect to time. In contrast to the exponential decay exhibited by the benchmark model in Fig. 2(a), the fraction of susceptible mobile users decreases to a relatively lower value of approximately 18% in steady-state. Such behavior can be explained by the increase in users with latent infection through both, BT and MS services shown in Fig. 2(b). As evident in this figure, the malware infects nearly 75% of the total population at $t = 10$ before stabilizing to a steady value of around 58%. More specifically, for a BT transmission range of $r_{\text{BT}} = 10$ meters, 10% of the users experience infection latency via BT and about 65% through MS service at $t = 10$. In agreement with the findings in [5], MS is therefore, more effective in spreading the malware as the underlying contact graph is not limited to any spatial constraints. While the population of latent users in Fig. 2(b) increases rapidly to its maximum in the transient period before slowly descending towards the infection-chronic equilibrium point, Fig. 2(c) shows a gradual growth in the number of disruptive users. This is a clear indication of the impact of infection latency on delaying the disruptive phase of the malware in affected user handsets. From these figures, we observe that the benchmark either underestimates or overestimates the behavioral dynamics of our malware model which is corroborated with results from stochastic simulations (colored lines) averaged over 10 runs.

TABLE I
Network Simulation Parameters and Settings

Transition rates	Value
Latent infection rate, $\beta^{L_{\text{BT}}}$ $(=\beta^{L_{\text{MS}}})$	0.015
Disruptive infection rate, $\beta^{I_{\text{BT}}}$ $(=\beta^{I_{\text{MS}}})$	0.01
Latency rate, δ^{BT} $(=\delta^{\text{MS}})$	0.04
Latent user recovery rate, γ^{BT} $(=\gamma^{\text{MS}})$	0.03
Infected user recovery rate, θ^{BT} $(=\theta^{\text{MS}})$	0.06
Initial latent users, $L^{\text{BT}}(0)$ $(=L^{\text{MS}}(0))$	10
Initial infected users, $I^{\text{BT}}(0)$ $(=I^{\text{MS}}(0))$	15

To highlight the contribution of wireless BT as a short-range malware spreading vector, Fig. 3 compares the fraction of BT-mediated infected mobile users (latent as well as disruptive) for different r_{BT} values with respect to time. For the same initial conditions given in Fig. 2, the number of users affected by the malware through BT enabled connections rises with increase in the transmission range. In most portable devices equipped with broadcast communication technology, the distance at which the information can be exchanged reaches up to 50 meters, if the devices are in direct line of sight of each other, and between 10 to 20 meters in buildings. Taking into account these extreme cases, we observe that under

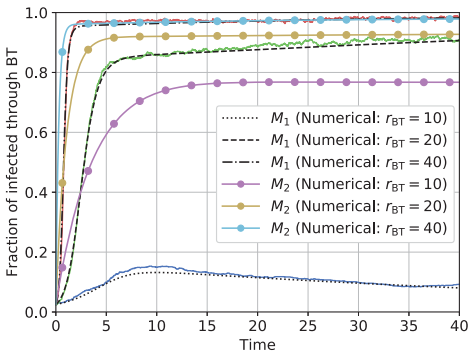


Fig. 3. Time evolution of BT-mediated infection $(L^{BT} + I^{BT})$ for different values of r_{BT}.

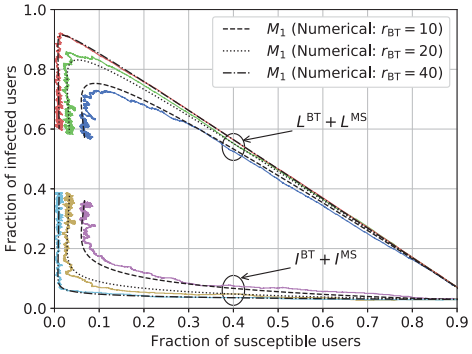


Fig. 4. Susceptible versus infected user populations w.r.t. r_{BT}.

ideal environmental conditions, the malware proliferates over the network in lesser time when r_{BT} is large. For instance, nearly 95% of the user devices in the network host the malware before $t = 5$ when $r_{BT} = 40$ meters, whereas a shorter range of $r_{BT} = 10$ meters would result in less than 10% of the network being infected within the same time period. Such behavior is due to the fact that increasing r_{BT} would cover a wider area thus, more likely allowing the malware to compromise a larger set of the susceptible users in the defined proximity. As a result, the average connectivity of each user increases which explains why our model coincides with the benchmark for larger r_{BT} values.

The stationary relationship between the susceptible and infected user groups is illustrated under different settings of r_{BT} in Fig. 4. As time progresses, the number of susceptible users decreases with increase in infected users. By separating latent users from disruptive users, the figure reveals that the latent population increases at a faster rate in comparison to the disruptive population. This is because disruptive malware codes are more active in infecting neighboring users while in their latent period and mainly distort the user data stored in devices during the disruptive phase. The dynamics of the latent and disruptive users about the viral equilibrium is also worth noting. Unlike the latent population that reaches its maximum before falling towards equilibrium $\boldsymbol{E}^*$, the fraction of disruptive users steeply rises. Moreover, increasing r_{BT} further raises the peak at which latent infection outbreak occurs. For example, extending the BT range from 10 to 40 meters increases the maximum population of latent users by nearly 21% which in turn, suppresses the disruptive population growth as $\beta^{L_{BT}} > \beta^{I_{BT}}$ and $\beta^{L_{MS}} > \beta^{I_{MS}}$ are specific only to disruptive malware. Such distinctions are obscure in existing malware models that undermine the impact of infection propagation latency.

V. Conclusion

In this letter, we introduced a modeling framework for effective projection of disruptive malware epidemics in MSNs. Unlike most existing virus epidemic models, we incorporated infection delay specific to disruptive malware programs to differentiate between the steady-state fraction of infected users in the latent and disruptive stages. Specifically, we proposed a tractable mean-field approximation network model for the underlying Markovian process to capture the impact of user-level interaction dynamics on the spreading pattern of the malware through personal and spatial social connections. By considering heterogeneity in the state transition rates, global stability and existence of the system equilibrium points were investigated to justify the steady-state behavior, based on which more effective containment measures can be devised. With respect to the benchmark model built on uniform user interactions, we demonstrated that infection latency can profoundly impact the accuracy of the proposed model in not only assessing the spreading risks of the hybrid malware via spatial and personal communication links in short time, but also in optimizing investments needed to control the spread by targeting devices that are in the latent infection stage.

References

[1] X. Hu, T. H. S. Chu, V. C. M. Leung, E. C.-H. Ngai, P. Kruchten, and H. C. B. Chan, "A survey on mobile social networks: Applications, platforms, system architectures, and future research directions," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 3, pp. 1557–1581, 3rd Quart., 2015.

[2] S. Peng, S. Yu, and A. Yang, "Smartphone malware and its propagation modeling: A survey," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 2, pp. 925–941, 2nd Quart., 2014.

[3] S. Sen, E. Aydogan, and A. I. Aysan, "Coevolution of mobile malware and anti-malware," *IEEE Trans. Inf. Forensics Security*, vol. 13, no. 10, pp. 2563–2574, Oct. 2018.

[4] N. Abuzainab and W. Saad, "A multiclass mean-field game for thwarting misinformation spread in the Internet of battlefield things," *IEEE Trans. Commun.*, vol. 66, no. 12, pp. 6643–6658, Dec. 2018.

[5] P. Wang, M. C. Gonzalez, C. A. Hidalgo, and A.-L. Barabasi, "Understanding the spreading patterns of mobile phone viruses," *Science*, vol. 324, no. 5930, pp. 1071–1076, May 2009.

[6] S.-M. Cheng, W. C. Ao, P.-Y. Chen, and K.-C. Chen, "On modeling malware propagation in generalized social networks," *IEEE Commun. Lett.*, vol. 15, no. 1, pp. 25–27, Jan. 2011.

[7] Q. Xu, Z. Su, K. Zhang, P. Ren, and X. S. Shen, "Epidemic information dissemination in mobile social networks with opportunistic links," *IEEE Trans. Emerg. Topics Comput.*, vol. 3, no. 3, pp. 399–409, Sep. 2015.

[8] M. Ogura and V. M. Preciado, "Stability of spreading processes over time-varying large-scale networks," *IEEE Trans. Netw. Sci. Eng.*, vol. 3, no. 1, pp. 44–57, Jan. 2016.

[9] A. Dadlani, M. S. Kumar, M. G. Maddi, and K. Kim, "Mean-field dynamics of inter-switching memes competing over multiplex social networks," *IEEE Commun. Lett.*, vol. 21, no. 5, pp. 967–970, May 2017.

[10] F. Darabi Sahneh, C. Scoglio, and P. Van Mieghem, "Generalized epidemic mean-field model for spreading processes over multilayer complex networks," *IEEE/ACM Trans. Netw.*, vol. 21, no. 5, pp. 1609–1620, Oct. 2013.

[11] L.-X. Yang, X. Yang, and Y. Y. Tang, "A bi-virus competing spreading model with generic infection rates," *IEEE Trans. Netw. Sci. Eng.*, vol. 5, no. 1, pp. 2–13, Jan. 2018.

[12] Y. Wu, P. Li, L.-X. Yang, X. Yang, and Y. Y. Tang, "A theoretical method for assessing disruptive computer viruses," *Phys. A, Stat. Mech. Appl.*, vol. 482, pp. 325–336, Sep. 2017.

[13] P. V. Mieghem, *Graph Spectra for Complex Networks*. Cambridge, U.K.: Cambridge Univ. Press, 2011.

[14] F. D. Sahneh, A. Vajdi, H. Shakeri, F. Fan, and C. Scoglio, "GEMFsim: A stochastic simulator for the generalized epidemic modeling framework," *J. Comput. Sci.*, vol. 22, pp. 36–44, Sep. 2017.