Smartphone Viruses Propagation on Heterogeneous Composite Networks

Xuetao Wei[*] Nicholas C. Valler[‡] Michalis Faloutsos[×]
Iulian Neamtiu[*] B. Aditya Prakash[⋆] Christos Faloutsos[†]
[*]UC Riverside [†]Carnegie Mellon University
[⋆]Virginia Tech [‡]Crowdcompass, Inc. [×]U. of New Mexico
{xwei, neamtiu}@cs.ucr.edu christos@cs.cmu.edu
badityap@cs.vt.edu nvaller@crowdcompass.com michalis@cs.unm.edu

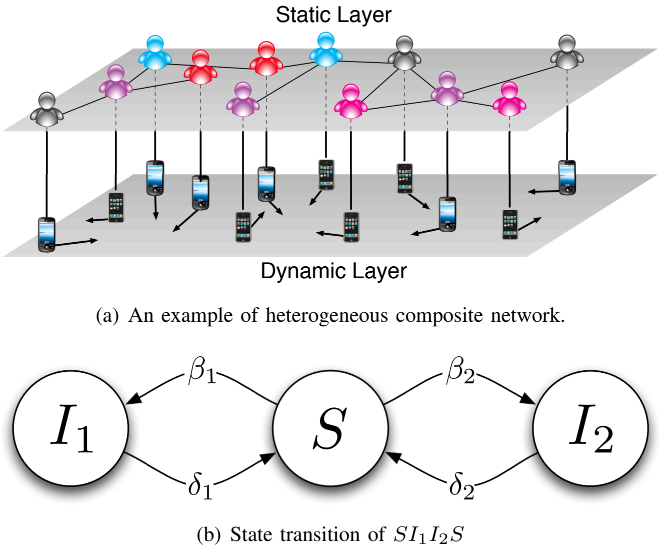
***Abstract*—Smartphones are now targets of malicious viruses. Furthermore, the increasing "connectedness" of smartphones has resulted in new delivery vectors for malicious viruses, including proximity-, social- and other technology-based methods. In fact, Cabir and CommWarrior are two viruses—observed in the wild—that spread, at least in part, using proximity-based techniques (line-of-sight bluetooth radio). In this paper, we propose and evaluate SI_1I_2S, a competition model that describes the spread of two mutually exclusive viruses across heterogeneous composite networks, one static (social connections) and one dynamic (mobility pattern). To approximate dynamic network behavior, we use classic mobility models from ad hoc networking, e.g., *Random Waypoint*, *Random Walk* and *Levy Flight*. We analyze our model using techniques from dynamic systems and find that the first eigenvalue of the system matrices $\lambda_{\mathbf{S1}}$,$\lambda_{\mathbf{S2}}$ of the two networks (static and dynamic networks) appropriately captures the competitive interplay between two viruses and effectively predicts the competition's "winner", which provides a feasible way to defend against smartphone viruses.**

Keywords*—*Epidemics, Competition, Social Networks, Mobile Networks

I. Introduction

Smartphones are a widely popular and growing space of computing technology, and are becoming an integral part of our daily life. In Q2 2012, over 153.9 million smartphones were sold worldwide [11]. Smartphones provide an ultra-portable interface to the Internet, and aid in their daily tasks with the wealth of device functionalities such as GPS, cameras, NFC(Near Field Communication) and accelerometers. As a direct result of their increasing popularity, functionality, and user reliance, smartphones are becoming common malware targets [4], [5], [12]. Smartphone "connectivity" forms a complex network of interaction, consisting of "layers" that represent various distinct social, technological and opportunistic contact networks. Each layer provides a unique, potential spreading vector for malware and other malicious viruses. Fortunately, emerging techniques from the field of *Network Science* provide a solid analytic framework to study such interconnections among varying, yet intertwined, network layers spanning the physical, information and social realms [6].

In this paper, we propose and evaluate the SI_1I_2S competition model to study the competition between two, mutually exclusive viruses actively spreading across heterogeneous composite networks. In our model, each layer corresponds to a different connectivity structure existing between a single set of nodes. Figure 1(a) illustrates a layered network. User social interaction forms a static network (top layer), whereas smartphones—carried by the users as they go about their daily activity—form a dynamic network (bottom layer). Notice that, previous work only studied the homogeneous composite network, namely, two layers are static [9]. In brief, our base propagation model SI_1I_2S—inspired by the popular compartmental models found in mathematical epidemiology—allows us to assert the mutual exclusivity condition between the competing viruses. In our model, each node transitions between states as indicated in Figure 1(b).



(a) An example of heterogeneous composite network.

(b) State transition of SI_1I_2S

Fig. 1. Example of Heterogeneous Composite Network and State Transition Diagram.

Using this framework, we seek to answer the following: *What are the parameters and conditions that determine the winner in heterogeneous composite networks?* Extending the work of [7]–[10], our analytic model shows the first eigenvalue of the system matrix corresponding to a layer ($\lambda_{\mathbf{S}}$) effectively determines the competition winner.

II. Model and Problem Definition

Our model consists of two primary components: *propagation* and *competition*. We detail them below, and for conve-

Symbol	Definition	Symbol	Definition
V_1, V_2	Virus #1, #2	$\mathbf{A_1}$, $\mathbf{A_2}$	Adjacency matrices
δ_1, δ_2	Viral Persistence of V_1, V_2	β_1, β_2	Viral Strength of V_1, V_2
S	Susceptible state	I_1, I_2	Infected state for V_1, V_2
$\mathbf{S_1}$, $\mathbf{S_2}$	System Matrix for $\mathbf{A_1}$, $\mathbf{A_2}$, where $\mathbf{S} = (1-\delta)\mathbf{I} + \beta\mathbf{A}$	$\lambda_{\mathbf{S_1}}$, $\lambda_{\mathbf{S_2}}$	Largest eigenvalue of $\mathbf{S_1}$, $\mathbf{S_2}$ in absolute value.

TABLE I. TERMINOLOGY

nience, provide terminology in Table I.

Propagation. We base our propagation model on the popular "flu-like" SIS compartmental model borrowed from mathematical epidemiology. Our model, SI_1I_2S, is composed of three distinct states: Susceptible (S), Infected with Virus #1 (I_1), and Infected with Virus #2 (I_2). A system agent (node) is in one of these states, with transitions as described in Fig. 1(b), where β, δ represent the viral strength and persistence, respectively.

Viral persistence describes the probability an node recovers from an infected state to the susceptible state, captured in an inverse manner, i.e., the higher δ, the less time a node remains infected. Viral strength is the probability that, when in contact with an infected node, an susceptible node is *attacked* by that infected neighbor. The attack only represents the potential for an infected node to propagate a virus to the susceptible node. Often, a susceptible node is attacked by multiple infected neighbors in a given timestep, possibly with different infections. To ultimately determine which infection is propagated to the susceptible node, we use the following algorithm. Let C_1 (C_2) be the number of neighbors that attack node i with virus #1 (#2) during a timestep Δt. Then, we have three possible scenarios for the susceptible node:

1) if $C_1 = 0$ and $C_2 = 0$, i remains in the susceptible state.
2) i transitions to I_1 with probability $\frac{C_1}{C_1+C_2}$.
3) i transitions to I_2 with probability $1 - \frac{C_1}{C_1+C_2}$.

Competition. Conceptually, we assume that the spreading viruses are mutually exclusive. That is, a node may never be infected by both viruses at the same time. This assumption is enforced in our model, as transitioning to an infected state only occurs when a node is in the susceptible state (see Figure 1(b)). The mutual exclusivity constraint is the basis of competition in our system.

We further assume that distinct viruses will have different attack vectors, corresponding to different layers of networked interaction across a set of nodes N. An attack vector corresponding to the static layer forms a connectivity matrix $\mathbf{A_1}$ that describes the possible edges a virus may traverse as it propagates through the population. Our dynamic layer is a set of connectivity matrices corresponding to the static connectivity at that moment in time. Each connectivity matrix is simply a layer in our connectivity model.

In our competition model, we assume that $\mathbf{A_1}$ is a static contact matrix over the duration of our evaluation. In contrast, we assume $\mathbf{A_2}$ is a dynamic contact matrix. To approximate dynamic graph behavior, we use the classic mobility models *Random Waypoint*, *Random Walk* and *Levy Flight*. To conserve space, we refer our reader to [2] and [3] for descriptions of these models.

Finally, we define the "winner" of the competition as the virus that successfully captures the largest fraction of nodes at steady state [7], [8].

Problem Definition. Given the propagation model described above, we now state the problem we address in this paper.

Given: (1) A static matrix $\mathbf{A_1}$; (2) A mobility model (random walk, levy flight or random waypoint mobility models); and (3) the propagation parameters of the competing viruses model (β_1, δ_1, β_2, δ_2).

Find: the parameters and conditions that determine the winner.

Ultimately, we believe this model is a reasonable starting point to analyze competition between viruses, and leave the analysis of other models as future work.

III. EIGENVALUE ANALYSIS

In this section, we briefly introduce the system matrix $\mathbf{S}$ and detail that why the first eigenvalue of the system matrix $\lambda_{\mathbf{S}}$ is the key parameter in determining the winner of our competition.

As in [6], we define the system matrix as $\mathbf{S} = (1-\delta)I + \beta\mathbf{A}$. One such matrix exists for each virus, but for convenience, here we drop the subscripts denoting individual virus and speak generally. As defined, the system matrix of a virus –and its primary eigenvalue– are a function of the topology *and* the virus propagation parameters. In more detail, given $\mathbf{S}$ formed from adjacency matrix $\mathbf{A}$ and the standard definition of an eigenvalue, we find:

$$\begin{aligned} \lambda_{\mathbf{S}}\vec{x} &= ((1-\delta)\mathbf{I} + \beta\mathbf{A})\vec{x} \\ \lambda_{\mathbf{S}} &= 1 - \delta + \beta\lambda_{\mathbf{A}} \end{aligned} \tag{1}$$

In conclusion, we could see that the system eigenvalue increases with the viral strength, β and the adjacency eigenvalue. Naturally, it decreases as the viral persistence, δ, increases.

IV. SIMULATION AND RESULTS

Simulation Methodology. We conduct a discrete-time simulation to evaluate the disease spread across our system. Each experimental run is composed of 100 trials, and the averaged results are reported below. For each run, we generate a power-law static graph across all the run's trials [1]. At the beginning of a trial, each virus initially infects a disjoint set of nodes, Ini_1 and Ini_2, populated randomly from the complete set of nodes, where $Ini_1 = 5$ and $Ini_2 = 1$. A trial completes upon reaching a relatively stable state, at which point, we determine the number of infected nodes for each virus.

Simulation Results. Figures 2,3, and 4 correspond to the case where the static graph has the higher connectivity as indicated by its eigenvalue λ_2. In these experiments, for each static graph, λ_2 is between 3.873–4.552, whereas the dynamic

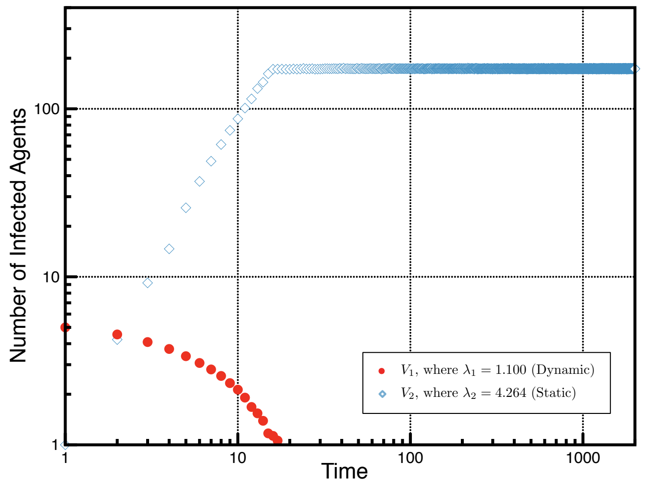


Fig. 2. Random Walk Mobility Model

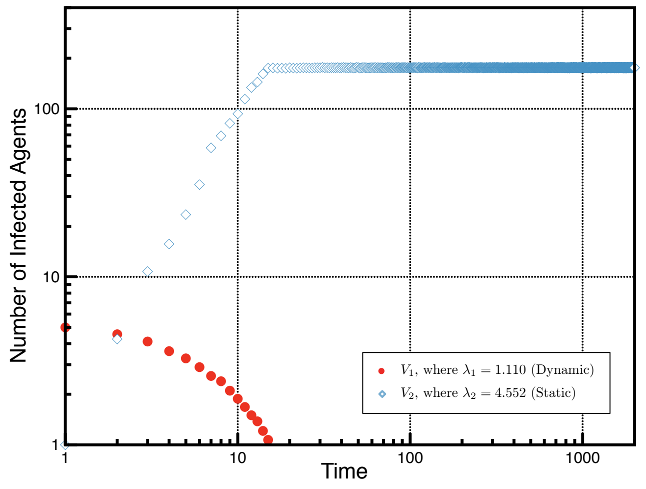


Fig. 4. Levy Flight Mobility Model

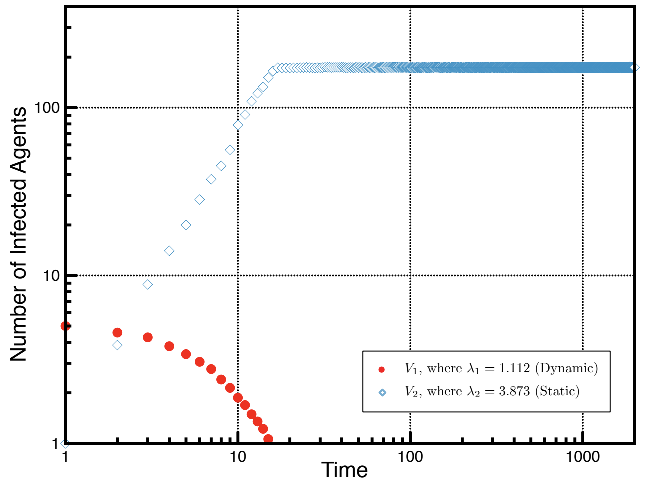


Fig. 3. Random Waypoint Mobility Model

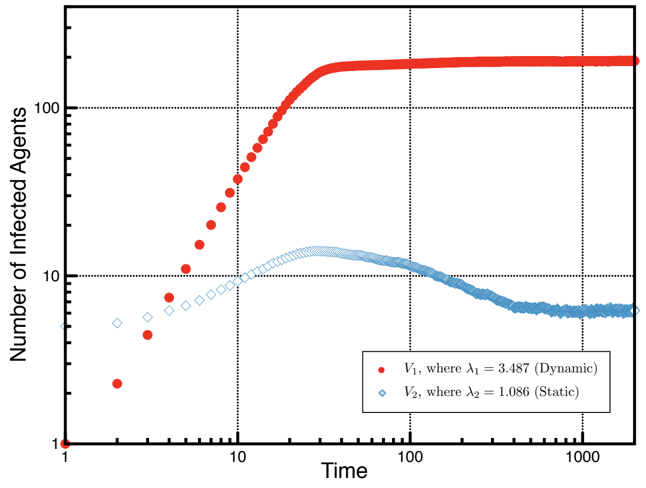


Fig. 5. Random Walk Mobility, note $\lambda_1 > \lambda_2$

graphs have a λ_1 between 1.100–1.112. As illustrated in Equation 1, assuming propagation parameters are equal, the disease operating on the graph with the highest λ corresponding to the system matrix is a stronger spreader, and will ultimately win the competition. Clearly, our experimental results support this assertion, as the static graph for each competitive scenario wins the competitions, regardless of mobility model.

Alternatively, Figure 5 illustrates the case where the eigenvalue of the dynamic layer is larger than the static layer (i.e., $\lambda_1 > \lambda_2$). Again, in this case we clearly observe the layer with the larger λ corresponding the its system matrix is the winner.

V. Conclusion

In this paper, we study the propagation of two mutually exclusive viruses competing for nodes across distinct static and dynamic network layers. Our problem is motivated by the growing concern over the spread of malicious viruses in smartphones. We extend on earlier work and show that λ_S is the essential parameter that effectively determines the winner of such a competition. Using simulation, we illustrate our result using various mobility model and find that the virus that has a larger λ_S tends to win the competition. This indicates that it is feasible to defend against smartphone virus propagation by disseminating anti-virus signature with a larger λ_S on heterogeneous composite networks. In the future, first, we will theoretically derive the epidemic threshold of competing viruses propagation on heterogeneous composite networks under the SIS model. Second, we will investigate other epidemic models(e.g., SIR) and explore how to resolve the parameters in our models into real smartphone virus propagation parameters. Finally, we would like to propose effective mitigation strategies to defend against smartphone malwares.

Acknowledgements

We would like to thank the anonymous reviewers for their feedback. This material is based upon work supported by the National Science Foundation under Grants No.IIS1017415, CNS-1064646 and NSF CISE NECO-0832069, and by U.S. Defense Advanced Research Projects Agency (DARPA) under the Social Media in Strategic Communication (SMISC) program, Agreement Number W911NF-12-C-0028 and Army Research Lab (ARL) under the Network Science CTA award W911NF-09-2-0053. This work is also partially supported by funds from the VT College of Engineering. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Laboratory, the National Science Foundation, or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation here on.

References

[1] C. Palmer and J. G. Steffan. Generating Network Topologies that Obey Power Laws. In GLOBECOMM, 2000.

[2] T.Camp, J. Boleng, and V. Davies. A Survey of Mobility Models for Ad Hoc Network Research. In Wirel. Commun. Mob. Comput, 2002.

[3] I. Rhee, M. Shin, S. Hong, K. Lee, S.J. Kim, and S. Chong. On the Levy-Walk Nature of Human Mobility. In IEEE/ACM Trans. Net., 2011.

[4] G. Zyba, G. Voelker, M. Liljenstam, A. Mehes, and P. Johansson. Defending Mobile Phones from Proximity Malware. In IEEE INFOCOM, 2009.

[5] F. Li, Y. Yang, and J. Wu. CPMC: An Efficient Proximity Malware Coping Scheme in Smartphone-based Mobile Networks. In IEEE INFOCOM, 2010.

[6] A.-L. Barabási. Linked: The New Science of Networks. Perseus Publishing, 2002.

[7] B. A. Prakash, H. Tong, N. Valler, M. Faloutsos, and C. Faloutsos. Virus propagation on time-varying networks: Theory and immunization algorithms. In ECML/PKDD, 2010.

[8] N.Valler, B. A. Prakash, H.Tong, M.Faloutsos, C. Faloutsos. Epidemic Spread in Mobile Ad Hoc Networks: Determining the Tipping Point. In IFIP NETWORKING 2011.

[9] X.Wei, N. Valler, B. A. Prakash, I. Neamtiu, M. Faloutsos, and C. Faloutsos. Competing Memes Propagation on Networks: A Case Study of Composite Networks. In ACM Sigcomm Computer Communication Review (CCR), October 2012.

[10] X.Wei, N. Valler, B. A. Prakash, I. Neamtiu, M. Faloutsos, and C. Faloutsos. Competing Memes Propagation on Networks: A Network Science Perspective. In IEEE Journal on Selected Areas in Communications(JSAC), 2013.

[11] Strong Demand for Smartphones in Second Quarter Continues to Drive the Worldwide Mobile Phone Market. www.idc.com

[12] A.P. Felt, M. Finifter, E. Chin, S. Hanna, and D. Wagner. A Survey of Mobile Malware in the Wild. In ACM SPSM 2011