

A Dynamic Game Solution to Malware Attack

M.H. R. Khouzani
Electrical and Systems Engineering
University of Pennsylvania
Email: khouzani@seas.upenn.edu

Saswati Sarkar
Electrical and Systems Engineering
University of Pennsylvania
Email: swati@seas.upenn.edu

Eitan Altman
INRIA, Sophia Antipolis, France
Email: altman@sophia.inria.fr

Abstract—Given the flexibility that software-based operation provides, it is unreasonable to expect that new malware will demonstrate a fixed behavior over time. Instead, malware can dynamically change the parameters of their infective hosts in response to the dynamics of the network, in order to maximize their overall damage. However, in return, the network can also dynamically change its counter-measure parameters in order to attain a robust defense against the spread of malware while minimally affecting the normal performance of the network. The infinite dimension of freedom introduced by variation over time and antagonistic and strategic optimization of malware and network against each other demand new attempts for modeling and analysis. We develop a zero-sum dynamic game model and investigate the structural properties of the saddle-point strategies. We specifically show that saddle-point strategies are simple threshold-based policies and hence, a robust dynamic defense is practicable.

I. INTRODUCTION

a) Motivation and Overture: New wireless technologies with increasing communication and computation capabilities transcend our mere person-to-person mobile communication needs. Sensitive and critical applications are rapidly developed and popularized, thanks to the software-based operation of wireless devices. The added flexibility, however, comes at a price: malware writers are expected to launch malicious applications which threaten to compromise critical security, privacy and in case of e-health, vitality of the users.

Worms spread during data or control message transmission between nodes that are infected (*infectives*) and those that are vulnerable, but not yet infected (*susceptibles*). Worms can disrupt the normal functionalities of the hosts, steal their private information, and use them to eavesdrop on other nodes. The worm can also render the host dysfunctional by deliberately draining its battery, or by executing a pernicious code that incurs irretrievable critical hardware or software damage, e.g., by re-fleshing the BIOS corrupting the bootstrap program required to initialize the OS [21]. We call these inoperative nodes *dead*. Upon an outbreak of a new malware, anomaly detection techniques can be used to identify the presence of malicious activities and generate security patches [23] that can then be distributed among the nodes on a transmission-upon-contact basis. Such patches either *immunize* susceptible nodes against future attacks, by rectifying their underlying vulnerability, or *heal* the infectives of the infection and render them robust against future attacks. Nodes that have been immunized or healed are denoted as *recovered*. In the meanwhile, reducing the communication rates in the network can quarantine the

worm by slowing down its spread. Specifically, the hosts can simply drop packets sent to them before processing them, or even refuse some connection requests.

Given the flexibility that software-based operation provides, it is unreasonable to expect that new malware will demonstrate a fixed behavior over time. Instead, malware can dynamically change its modus operandi in response to the dynamics of the network, in order to maximize the overall damage it inflicts. However, in return, the network can also dynamically change its counter-measure policy to more effectively oppose the spread of the infection. The infinite dimension of freedom introduced by variation over time and antagonistic optimization of malware and network against each other demand new attempts for modeling and analysis of their strategic confrontation. This paper investigates such confrontations and identifies maximum damage dynamic strategies of attack and devises robust dynamic defense before such threats emerge.

b) Defense and Attack Decision Problems: Since the media in the wireless network is common and the channels are unreliable, the bandwidth consumed for distribution of the security patches can itself disrupt the normal functionality of the network. Excessive quarantining through reception rate reduction also deteriorates the quality of service (QoS) for the data traffic. Such quarantining can not usually discriminate based on the identity of the transmitters, since the hosts applying the reception rate control in general do not know which other nodes are infected; the reception rate itself may however be judiciously selected. The network's challenge now is to achieve a guaranteed performance by selecting the instantaneous (a) rate of patching, and (b) reception rate that jointly minimize the overall damage due to (i) the subversive activities of the malware that is capable of annihilating infectives, and (ii) the additional resource consumption and deterioration of QoS owing to the application of the countermeasures. The design must adapt over time remaining cognizant of the malware's ability to dynamically optimize its spread in response to the network's dynamic strategy.

The malware also faces an interesting tradeoff: should it kill its host as soon as feasible after infecting it? While a quick annihilation of a host inflicts a high cost on the network right away by rendering it completely dysfunctional, it also rules out the use of that node in infecting the remaining susceptibles. Thus, early mutilation of infective nodes may thwart the spread of malware. Moreover, killing a node deprives the malware of the other malicious activities the node can be used for, such

as eavesdropping, stealing private information, etc. Deferral of killing, on the other hand, is at the risk of losing that node through installation of security patches and recovery of that node by the network. The annihilation strategy should therefore depend on relative benefits for the malware and the damages for the network incurred by each of the above factors. For instance, if the malware is primarily interested in stealing a node's private information or eavesdropping on others, it ought to defer killing for some time, however, not too long lest the node recovers. If on the other hand its primary goal is to degrade network functions by disabling as many nodes as possible and as soon as possible, it ought to start the slaughter as soon as it has infected a sizeable population of hosts.

A *robust* counter-measure is one that seeks to minimize the damage inflicted by the malware assuming that the malware chooses its strategy so as to maximize this damage with full knowledge of the counter-measure. Due to the above trade-offs and since an optimal strategy of the malware depends on the strategy of the network and *vice versa*, determination of the robust strategies of either is non-trivial. This paper proposes a method to answer these questions.

c) Contributions: First, we construct a mathematical framework which cogently models the strategic confrontations between the malware and the network as a zero-sum (minimax) dynamic game (§III-A) drawing from (i) existing epidemic models that have earlier been proposed and validated for worm propagation in wireless networks (§II-A), and (ii) damage functions that we introduce to investigate the trade-offs resulting from different decisions of the entities concerned (§II-B). To the best of our knowledge, this is the first paper that combines epidemiological models with dynamic game formulations for propagation of worms in wireless networks. We are also able to prove the existence of the *robust*, or *saddle-point* strategies of the network and the malware (§II-A), and compute them (§III-B). Existence of such strategies and also their computations are not clear a priori, since the strategy set of each player is uncountably infinite and consists of functions of time.

We prove that the robust defense strategy has a simple two-phased structure (§III-C): (i) patch at the maximum possible rate until a threshold time, and then stop patching (ii) choose the minimum possible reception rate (i.e., the maximum packet drop rate at the receivers) until a threshold time and subsequently revert to the normal reception rate. The initial aggressive defense limits the spread of infection and thereby the pool of nodes that can potentially be exploited or killed; this guarantees an upper bound on the damage inflicted irrespective of the malware's choice of annihilation rates. Given its simple structure, the defense control can readily be implemented in resource constrained wireless devices. From a game-theoretical point of view, the structural results are somewhat surprising given the non-linear dynamics of state evolutions and the non-monotonicity of the state functions, and their proofs rely on non-standard techniques.

The game formulations, and in particular the epidemic models, rely on some abstractions, that have been made for

analytical tractability. Using simulations, we validate the formulations when these assumptions are relaxed (§V). Our numerical computations reveal that our robust dynamic defense strategy attains substantially lower value of the maximum damage inflicted by the malware as compared to that for heuristic static choice of defense parameters.

d) Related Works: Malware outbreaks in wireless networks constitute an emerging research topic (e.g., [24]) Epidemic modeling based on the classic Kermack-Mckendrick model [5] has extensively been used to analyze the spread of malware in wired and cellular networks [3], [24], [25] *etc.*, and more recently in wireless networks [19]. These works show, through simulations and matching with actual data, that when the number of nodes in a network is large, the deterministic epidemic models can successfully represent the dynamics of the spread of the malware.

Dynamic control of parameters of the network or the worm have been investigated in several papers [10]–[14], [22]. These papers, however, allow only one of the network or malware to dynamically change their parameters, and assume that the other's choice of parameters is not only static but also known to the opponent. In contrast, we consider a dynamic game where the network chooses its patching and communication rates dynamically so as to minimize the overall damage when the malware also intelligently varies its parameters, specifically, killing the infective nodes, over time so as to maximize this damage; also each player remains cognizant of the other player's ability to optimally respond to the opponent's choices. Specifically, each player (say player A) selects its strategy without any knowledge of the others' strategy but being prepared for the eventuality in which the other (i.e., B) selects its strategy after learning A's strategy.

Game theory has been used in the context of security in networks as it is apt to model the interactions of attackers and defenders, e.g. in [1], [7], [9], [17]. [17] presents models for the inference of the intents, objectives and strategies of a new attacker and apply it to a DDoS attack. In their work, however, the sets of actions of both the attack and the defense are finite, and structural property of Nash Equilibriums or saddle-points have not been obtained; the work focuses on the modeling and numerical evaluations. Algorithmic implementations of (variations of) models in [17] are pursued in [9], [7], *etc.* We apply dynamic zero-sum games to model the strategic confrontations of a malware and the defense in a wireless network, and delve into the structural properties of the saddle-point strategies, when the attack and defense can intelligently choose the annihilation, patching and reception rates respectively. Thus, unlike most of the existing work, the defense operates also at the MAC and physical layers, as opposed to only at the routing or application layers. Indeed, we analyze not only the security risks (fraction of infectives, dead nodes), but also the QoS degradations (packet drops) and the lower layer bandwidth consumptions (in transmission of patches) associated with the tradeoffs. Also, the strategy sets of each player is uncountably infinite since the strategies are functions of continuous time with continuous ranges. The

differences in the contexts and the nature of choices require a substantially different analytical approach. Our contributions complement [1], which focuses on detecting the intrusion of a worm that dynamically controls the intensity of its activity, but does not investigate subsequent defense.

II. SYSTEM MODEL

A. Dynamics of State Evolution

A **susceptible** node is a mobile wireless device¹ which is not contaminated by the worm, but is prone to infection. A node is **infective** if it is contaminated by the worm. An infective spreads the worm to a susceptible while transmitting data or control messages to it. The worm can *kill* an infective host, i.e., render it completely dysfunctional - such nodes are denoted **dead**. A functional node that is immune to the worm is referred to as **recovered**.

Nodes are roaming in a vast 2-D region of area A with an average velocity v . No node is aware of the state of other nodes. Specifically, if a susceptible node knew a priori which nodes are infective, then it would have just blacklisted them. It is also difficult for the malware to constantly measure network states given that a large number of nodes are roaming over a large area, and given that the set of neighbors of the infectives are constantly changing owing to node mobility.

Let the total number of nodes in the network be N . Let the number of susceptible, infective, recovered and dead nodes at time t be respectively denoted by $n_S(t), n_I(t), n_R(t)$ and $n_D(t)$, and the corresponding fractions be $S(t) = n_S(t)/N$, $I(t) = n_I(t)/N$, $R(t) = n_R(t)/N$, and $D(t) = n_D(t)/N$. Then, $S(t) + I(t) + R(t) + D(t) = 1$. At the time of the outbreak of the infection, that is at time zero, some nodes are infected: $0 < I(0) = I_0 \leq 1$. For simplicity, we assume $R(0) = D(0) = 0$. Thus, $S(0) = 1 - I_0$.

We now model the dynamics of the propagation of the infection as an epidemic model that has been validated for mobile wireless networks through experiments as well as network simulations (see e.g. [4], [20]). A susceptible is infected whenever it receives a message from an infective. The epidemic models consider homogenous mixing (which we later relax using simulations) where an infective is equally likely to initiate communication with each node, and hence each susceptible, say at rate $\hat{\beta}$. This represents worm propagation in 3G and 4G cellular networks where infective mobiles try to infect randomly and uniformly generated addresses. Note that in any such mobile to mobile communication, irrespective of the locations of the mobiles, there are two wireless communications between access points and mobiles and the rest of the communications are through the backbone network where the delays and congestions are relatively limited. The homogenous mixing can also be justified in delay tolerant networks (DTN) where the infectives initiate communication

only with nodes that are within their transmission ranges². In fact, under mobility models such as the random waypoint or the random direction model [2], Groenevelt *et al.* [6] have mathematically proven the homogenous mixing assumption³. In both scenarios, $\tilde{\beta}$ depends only on the rates at which the infectives scan for the susceptibles, node velocities, transmission ranges, node densities, and uplink and downlink communication rates (the last two for cellular networks).

A susceptible accepts a communication request with a probability $u^{N_r}(t)$ ⁴. At any given t , there are $n_S(t)n_I(t)$ infective-susceptible pairs. Susceptibles are therefore transformed to infectives at rate $\hat{\beta}u^{N_r}(t)n_S(t)n_I(t)$. Infection propagation, therefore, can be contained through appropriate regulation of $u^{N_r}(t)$ ⁵ subject to:

$$0 < u_{\min}^{N_r} \leq u^{N_r}(t) \leq u_{\text{norm}}^{N_r} \text{ at each } t.$$

The lower bound $u_{\min}^{N_r}$ arises due to the minimum quality of service (QoS) requirements for data traffic (since the acceptance probability is the same irrespective of whether the request arrives from another infective, susceptible, or recovered node). The upper bound $u_{\text{norm}}^{N_r}$ (which can be normalized to 1) provides the reception rate that nodes use for providing the desired QoS in absence of security considerations.

We now consider the dissemination of security patches in the network. A pre-determined set of nodes, referred to as dispatchers (e.g., BS for cellular and exit-points for delay-tolerant networks) are pre-loaded with the patches. We assume that the dispatchers can not be infected, and that there are NR_0 dispatchers where the network parameter R_0 , is between 0 and 1. Each node communicates with the dispatchers, and thereby fetches security patches, at an overall rate $\tilde{\beta}NR_0u^{N_i}(t)$ at any time t . The parameter $\tilde{\beta}$ depends on node density, mobility parameters, allowable transmission rates etc., whereas $u^{N_i}(t)$ ⁶ is a control function which can be used to regulate the bandwidth consumed in propagation of patches - the higher the value of $u^{N_i}(t)$, the higher is the recovery rate but so is the resource consumption in patch transmission. Clearly,

$$0 \leq u^{N_i}(t) \leq 1 \text{ at each } t.$$

If the node that receives the patch is a susceptible node, it installs the patch and its state changes to recovered. If an infective receives the patch, the patch may fail to heal it, or, the worm may prevent its installation. We capture the above possibility, by introducing a coefficient $0 \leq \pi \leq 1$: $\pi = 0$ occurs when the patch is completely unable to remove the

²Infectives do not initiate connection with other nodes as in DTNs mobile nodes roam a vast area which is much larger than their communication ranges, and there is no backbone network and more often than not end-to-end connectivity does not exist.

³The result has been proven when the communication range of the nodes is small compared to the total area of the region and node velocity v is sufficiently high. Numerical computations in [6] show that the result mostly extend even when these assumptions are relaxed.

⁴The subscript r represents reception.

⁵Superscript N designates control functions of the network, and M designates control functions of the malware.

⁶The subscript i denotes immunization.

¹Similar state dynamics can be motivated for a p2p network (c.f. e.g. [18]).

worm from infectives and only immunizes the susceptibles, whereas $\pi = 1$ represents the other extreme scenario where a patch can equally well immunize and heal susceptibles and infective nodes⁷ Now, if the patch heals an infective, its state changes to recovered, else it continues to remain an infective.

The worm at an infective host kills the host with rate proportional to $u^M(t)$ at a given time t ; this is accomplished by executing specific codes with a probability of choice. The worm regulates the death process by appropriately choosing $u^M(t)$ at each t , subject to:

$$0 \leq u^M(t) \leq u_{\max}^M \text{ at each } t.$$

The upper bound arises due to processor constraints and the resulting limitations on the maximum rate of execution of such codes.

$$\text{Let, } \beta_0 := N\hat{\beta}, \quad \beta_1 := N\tilde{\beta}R_0,$$

Our discussions lead to⁸ the following system of differential equations representing the dynamics of the system:⁹

$$\dot{S}(t) = -\beta_0 u^{Nr} I(t) S(t) - \beta_1 u^{Ni}(t) R_0 S(t) \quad (1a)$$

$$\dot{I}(t) = \beta_0 u^{Nr} I(t) S(t) - \pi \beta_1 u^{Ni}(t) R_0 I(t) - u^M(t) I(t) \quad (1b)$$

$$\dot{D}(t) = u^M(t) I(t) \quad (1c)$$

with initial constraints:

$$I(0) = \lim_{N \rightarrow \infty} n_I(0)/N = I_0, \quad S(0) = 1 - I_0, \quad D(0) = 0 \quad (2)$$

and also satisfy the following constraints at all t :

$$0 \leq S(t), I(t), D(t) \quad (3a)$$

$$S(t) + I(t) + D(t) \leq 1. \quad (3b)$$

Thus, $(S(\cdot), I(\cdot), D(\cdot))$ constitute the system state functions, $u^N(\cdot) = (u^{Nr}(\cdot), u^{Ni}(\cdot))$ constitutes the network control func-

⁷In order to avoid immediate detection and blacklisting, the infectives may choose not to refuse all connection requests from the dispatchers.

⁸The introduction of the set of differential equations system as the dynamics of the system can be made rigorous if further technical assumptions are made. Specifically, if $(n_S(t), n_I(t), n_D(t))$ constitutes a Continuous-Time Markov Chain (CTMC), then according to the results of [16], as N grows, $S(t)$, $I(t)$ and $D(t)$ converge to the solution of the the system of differential equation in the following sense:

$$\forall \epsilon > 0 \forall t > 0, \quad \lim_{N \rightarrow \infty} \Pr\{\sup_{\tau \leq t} |\frac{n_S(\tau)}{N} - S(\tau)| > \epsilon\} = 0$$

Likewise for $I(t)$ and $D(t)$. Note that the CTMC property entails assuming that the inter-contact times are exponentially distributed. For DTN networks, this property is shown for by Groenevelt *et al.* [6] under a number of mobility models such as random waypoint or random direction model [2]. Also, while considering the limits, β_0, β_1 are limits of the respective R.H.S. According to the results of [6], $\hat{\beta}, \tilde{\beta}$ are inversely proportional to the area of the roaming region (A). Thus, the limits $\lim_{N \rightarrow \infty} N\hat{\beta}$, $\lim_{N \rightarrow \infty} N\tilde{\beta}$ exist as long as the node density $\lim_{N \rightarrow \infty} N/A$ exists for large N , and are also positive since the node density $\lim_{N \rightarrow \infty} N/A, \hat{\beta}, \tilde{\beta}$ are all positive.

⁹Throughout the paper, variables with dot marks (e.g., $\dot{S}(t)$) will represent their time derivatives (e.g., time derivative of $S(t)$).

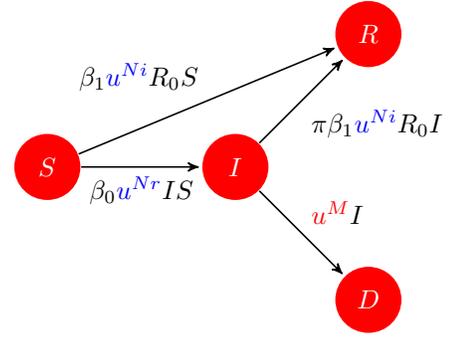


Fig. 1: State transitions. $u^{Ni}(t)$ and $u^{Nr}(t)$ are the control parameters of the network while $u^M(t)$ is the control parameter of the malware.

tions and $u^M(\cdot)$ constitutes the malware's control function¹⁰. Note that nodes use *identical* reception, patching and killing rate functions irrespective of the states in their neighborhoods since they do not know these states. Nevertheless, since these rates are allowed to vary with time, they can be chosen in accordance with how the network states are *expected* to evolve.

Henceforth, wherever not ambiguous, we drop the dependence on t and make it implicit. Fig. 1 illustrates the transitions between different states of nodes and the notations used.

B. Defense and Attack objectives

We first quantify the total damage inflicted by the malware during the network operation interval $[0, T]$. This damage is due to the presence of infectives, the death of nodes, the resources consumed for spreading the security patches, and the QoS deterioration due to the reduction of reception rate. Infectives can perform harmful activities over time, e.g., they can (i) eavesdrop and analyze traffic that is generated or relayed by the infected hosts, or the traffic that traverses in the hosts' vicinity, and (ii) alter or destroy the traffic that is generated or relayed by the infected hosts. Dead nodes are inoperative and thus inflict a time-accumulative cost on the network. The bandwidth overhead at time t due to the media scanning and transmission of the security packets by the dispatchers is $R_0 u^{Ni}(t)$. Due to the reception rate control, the susceptibles lose a $u_{\text{norm}}^{Nr} - u^{Nr}(t)$ fraction of packets transmitted by all nodes which degrades the overall QoS. We therefore consider the aggregate network damage at time t as a combination of $I(t)$, $D(t)$, $u^{Ni}(t)$, $u^{Nr}(t)$.¹¹

Note that the damage function can be scaled so that one of the coefficients may be chosen as unity: we choose the one associated with the instantaneous bandwidth overhead. Thus,

¹⁰Note that in the real system $S(t), I(t), D(t) \in \{0, 1/N, 2/N, \dots, 1\}$, i.e., are discrete, but the solutions of the above differential equations are in the continuum. The resulting error however reduces fast as N increases, and vanishes in the limit $N \rightarrow \infty$. Formally, from mean field approximations [16], $\forall \epsilon > 0 \forall t > 0, \quad \lim_{N \rightarrow \infty} \Pr\{\sup_{\tau \geq t} |\frac{n_S(\tau)}{N} - S(\tau)| > \epsilon\} = 0$, and likewise for $I(t)$ and $D(t)$. Also, while considering the limits, β_0, β_1 are limits of the respective R.H.S. It can be shown that these limits exist as long as the node density $\lim_{N \rightarrow \infty} N/A$ exists for large N , and are also positive since the node density $\lim_{N \rightarrow \infty} N/A, \hat{\beta}, \tilde{\beta}$ are all positive.

¹¹We adopt a linear structure for analytical tractability, and also because non-linear functions may be approximated by (piece-wise) linear versions.

the damage over the time horizon $[0, T]$ is¹²:

$$J(u^N, u^M) = \int_0^T [\kappa_I I(t) + \kappa_D D(t) + R_0 u^{N_i}(t) - \kappa_r u^{N_r}(t)] dt + K_D D(T), \quad (4)$$

$K_D D(T)$ relates to the final tally of the dead nodes. The coefficients are all non-negative and represent the relative importance of each corresponding term in the overall damage, e.g., if the worm gains the most by killing, and thereby completely disabling nodes, $\kappa_D \gg \kappa_I$. Let $\kappa_I > 0, \kappa_r > 0$.

The network seeks to choose its control vector $u^N(\cdot)$ so as to minimize the above while the malware seeks to choose its control $u^M(\cdot)$ so as to maximize the above, subject to satisfying the state constraints (3) and ensuring that

$$u_{\min}^{N_r} \leq u^{N_r}(t) \leq u_{\text{norm}}^{N_r}, \quad 0 \leq u^{N_i}(t) \leq 1, \quad (5a)$$

$$0 \leq u^M(t) \leq u_{\max}^M. \quad (5b)$$

In §III, we model their interactions resulting from opposing objectives as a dynamic game. The formulation relies on the following result (which we prove in Appendix-A) that allows us to ignore the state constraints without any loss of generality.

Lemma 1: Any pair of strategies $(u^N(\cdot), u^M(\cdot))$ that satisfy the control constraints (5a), (5b), satisfy the state constraints (3) and ensure that $I(t) > 0, S(t) > 0$ for all $t \in [0, T]$.

III. NETWORK-MALWARE DYNAMIC GAME

A. Formulation

Consider a system with two players N (network) and M (malware), specified by a system of n differential equations [15, P.83]:

$$\dot{x}(t) = f(t, x(t), u^N(t), u^M(t)) \quad t \in [t_0, T], \quad (6a)$$

$$u^N \in U^N \subset \mathcal{R}^m, \quad u^M \in U^M \subset \mathcal{R}^s, \quad (6b)$$

and initial condition $x(t_0) = x_0$, and a damage function

$$J[u^P, u^E] = g(x(T)) + \int_{t_0}^T h(x, u^P, u^E, t) dt. \quad (6c)$$

where $x(t)$ is the n -dimensional state vector. *Player N* seeks to minimize J by controlling the m dimensional *control function* $u^N(\cdot)$, and *player M* seeks to maximize J by controlling the s -dimensional control function $u^M(\cdot)$.¹³ The game is therefore referred to as a dynamic two-player minimax game. The players' payoffs, and the set of strategies available to them are called *rules of the game*. Both players know the rules of the game and each player knows that its opponent knows the rule and ad infinitum¹⁴.

In our context, (1) provides the $f(\cdot)$ functions, the initial conditions are provided by (2), (4) provides the $g(\cdot), h(\cdot)$

¹²Note that $(u_{\text{norm}}^{N_r} - u^{N_r})$ inside the integral is replaced with $-u^{N_r}$ as $\kappa_r u_{\text{norm}}^{N_r} T$ does not depend on the evolution of the states or the controls.

¹³Equivalently, if M attains a reward of J , and N a reward of $-J$, and M, N both seek to maximize their individual rewards - the game is referred to as zero-sum since their rewards always sum to 0.

¹⁴each player knows that each player knows that the opponent knows *etc.*

functions, (5a), (5b) provide U^N, U^M . Also, we have, $n = 3, m = 2, s = 1$. Note that the $f(\cdot), h(\cdot)$ functions in our context depends on time t only implicitly, that is through the state and control functions. Also, the formulation does not capture any other constraints on the state functions, and in our context it does not need to either, owing to Lemma 1.

We now consider the values of the game. The *lower value* denoted by V_* , is the overall damage when the minimizing player (N) is given the upperhand, i.e., selects its strategy after learning its opponent's strategy. Mathematically:

$$V_* = \max_{u^M} \min_{u^N} J[u^N, u^M]$$

Conversely, the *upper value* of the game V^* is defined as

$$V^* = \min_{u^N} \max_{u^M} J[u^N, u^M]$$

Thus, V_* (V^* , resp.) is the maximum (minimum, resp.) damage that the malware (network, resp.) can inflict (incur, resp.) if the other player has the upper-hand. Also, $V_* \leq V^*$. A pair of strategies (u^{N*}, u^{M*}) is called a *saddle-point* if

$$J(u^{N*}, u^M) \leq J(u^{N*}, u^{M*}) = V \leq J(u^N, u^{M*})$$

for any strategy u^N of the network and u^M of the malware, and then V is the value of the game, and $V = V_* = V^*$.

Thus, if the network selects its saddle-point strategy u^{N*} , irrespective of the strategy of the malware, the damage it incurs is at most V , which is also the minimum damage that the malware can inflict if it has the upper-hand. Thus, the network's saddle-point strategy is also its *robust* strategy, in the sense, that it minimizes the maximum possible damage it can incur. Conversely, the malware's saddle point strategy is also its robust strategy, since it maximizes the minimum possible damage it can inflict. Also, the network's and the malware's saddle point strategies are their respective best responses to the other's robust strategy.

Theorem 1: The minimax game defined above has a saddle-point pair of strategies.

We prove this theorem in Appendix-C.

B. A framework for computation of the saddle-point strategies

Since the set of deterministic strategies of each player is uncountably-infinite, the saddle-point strategies and the value of the game can not be computed using convex or linear programming. We now present a framework for numerical computation of the saddle-point strategies.

Define the *Hamiltonian* for a given policy pair (u^N, u^M) in an arbitrary two-person minimax dynamic game as follows:

$$\mathcal{H}(u^N, u^M) = \langle \lambda, f(x, u^N, u^M, t) \rangle + h(x, u^N, u^M, t)$$

where the state functions $x(\cdot)$ are those that correspond to the strategy pair (u^N, u^M) , and λ , the *co-state* (or *adjoint*) functions, are continuous and piecewise differentiable functions of time that satisfy the following system of differential equations wherever the controls (u^N, u^M) are continuous:

$$\dot{\lambda} = -\frac{\partial}{\partial x} \mathcal{H}(x, \lambda, t)$$

and the final value (transversality) condition

$$\lambda(T) = \frac{\partial(g(x))}{\partial(x)} \Big|_{x=x(T)}$$

In our context,

$$\begin{aligned} \mathcal{H}(u^M, u^N) = & \kappa_I I + \kappa_D D + u^{N_i} R_0 - \kappa_r u^{N_r} + (\lambda_I - \lambda_S) \beta_0 u^{N_r} I S \\ & - \lambda_S \beta_1 R_0 u^{N_i} S - \lambda_I \beta_2 R_0 u^{N_i} I + (\lambda_D - \lambda_I) u^M I \end{aligned}$$

where again the state functions ($S(\cdot), I(\cdot), D(\cdot)$) are obtained from (1) with $(u^N(\cdot), u^M(\cdot))$ as the control functions, and the co-state functions ($\lambda_S(\cdot), \lambda_I(\cdot), \lambda_D(\cdot)$) are obtained from the following system of differential equations (with $u^N(\cdot), u^M(\cdot)$ as the control functions)

$$\dot{\lambda}_S = -\frac{\partial \mathcal{H}}{\partial S} = -(\lambda_I - \lambda_S) \beta_0 u^{N_r} I + \lambda_S \beta_1 R_0 u^{N_i} \quad (7a)$$

$$\begin{aligned} \dot{\lambda}_I = -\frac{\partial \mathcal{H}}{\partial I} = & -\kappa_I - (\lambda_I - \lambda_S) \beta_0 u^{N_r} S + \lambda_I \beta_2 u^{N_i} R_0 \\ & - (\lambda_D - \lambda_I) u^M \end{aligned} \quad (7b)$$

$$\dot{\lambda}_D = -\frac{\partial \mathcal{H}}{\partial D} = -\kappa_D \quad (7c)$$

with the final conditions

$$\lambda_S(T) = 0, \quad \lambda_I(T) = 0, \quad \lambda_D(T) = K_D. \quad (8)$$

Then, following [15, P.31], a necessary condition for the pair (u^N, u^M) to be a saddle-point strategy pair is that for all $t \in [0, T]$:

$$(u^N, u^M) \in \arg \min_{\tilde{u}^N} \max_{\tilde{u}^M} \mathcal{H}(\tilde{u}^N, \tilde{u}^M) \quad \text{and} \quad (9a)$$

$$(u^N, u^M) \in \arg \max_{\tilde{u}^M} \min_{\tilde{u}^N} \mathcal{H}(\tilde{u}^N, \tilde{u}^M). \quad (9b)$$

Henceforth, we denote the saddle point strategy pair as $(u^N(\cdot), u^M(\cdot))$, and $(S(\cdot), I(\cdot), D(\cdot)), (\lambda_S(\cdot), \lambda_I(\cdot), \lambda_D(\cdot))$ as the corresponding state and co-state functions and \mathcal{H} as the corresponding Hamiltonian. We now express $(u^N(\cdot), u^M(\cdot))$ in terms of $(S(\cdot), I(\cdot), D(\cdot)), (\lambda_S(\cdot), \lambda_I(\cdot), \lambda_D(\cdot))$ using the necessary conditions (9). Let

$$\begin{cases} \psi^{N_r} := (\lambda_I - \lambda_S) \beta_0 I S - \kappa_r \\ \psi^{N_i} := R_0 - \lambda_S \beta_1 R_0 S - \lambda_I \beta_2 R_0 I \\ \psi^M := (\lambda_D - \lambda_I) I \end{cases}$$

Now, the Hamiltonian can be rewritten as:

$$\mathcal{H} = \kappa_I I + \kappa_D D + \psi^{N_r} u^{N_r} + \psi^{N_i} u^{N_i} + \psi^M u^M. \quad (10)$$

Thus, the Hamiltonian is a separable function of different components of the defense controls $(u^{N_r}(\cdot), u^{N_i}(\cdot))$ and the attack control $u^M(\cdot)$, that is, each of these appear in different terms in the R.H.S of the above characterization. Now, from the necessary conditions in (9) subject to the control

constraints in (5), the saddle-point strategies are derived as:

$$u^{N_r} = \begin{cases} u_{\min}^{N_r} & \text{if } \psi^{N_r} > 0, \\ u_{\text{norm}}^{N_r} & \text{if } \psi^{N_r} < 0 \end{cases} \quad (11)$$

$$u^{N_i} = \begin{cases} 0 & \text{if } \psi^{N_i} > 0, \\ 1 & \text{if } \psi^{N_i} < 0 \end{cases} \quad (12)$$

$$u^M = \begin{cases} u_{\max}^M & \text{if } \psi^M > 0. \\ 0 & \text{if } \psi^M < 0, \end{cases} \quad (13)$$

Since $\psi^{N_r}, \psi^{N_i}, \psi^M$ are uniquely specified once the state and the co-state functions are known, the above relations express the saddle-point strategies in terms of the state and co-state functions. The strategies $u^{N_r}(\cdot), u^{N_i}(\cdot), u^M(\cdot)$ can be substituted by the above characterizations in (1) and (7), resulting in a system of 6 differential equations involving only the state and the co-state functions. Using standard numerical methods for solving differential equations, this system can be solved (very fast) using the initial and final conditions (2), (8). The state and co-state functions obtained as solutions will now provide the $\psi^{N_r}, \psi^{N_i}, \psi^M$ functions, and thereby the saddle-point strategies via (11), (12), (13). The resulting set of differential equations is non-linear and a close-form solution is unknown. However, as we will show in the next section, using novel techniques, even without access to the closed-form solution, we can establish the type of behavior that the saddle-point strategies exhibit.

C. Structural Properties of Saddle-Point Defense Strategy

We establish that the saddle-point defense strategy has a simple threshold-based structure that ought to facilitate its implementation in a localized manner in resource constrained wireless devices. Specifically, we prove that:

Theorem 2: For the saddle-point defense strategy $u^N(\cdot) = (u^{N_r}(\cdot), u^{N_i}(\cdot))$, there exists times $t_1, t_2, 0 \leq t_1 < T, 0 \leq t_2 < T$ such that:

- $u^{N_r}(t) = u_{\min}^{N_r}$ for $0 < t < t_1$, and $u^{N_r}(t) = u_{\text{norm}}^{N_r}$ for $t_1 < t < T$.
- $u^{N_i}(t) = 1$ for $0 < t < t_2$, and $u^{N_i}(t) = 0$ for $t_2 < t < T$.

The overall strategy therefore has the following three phases. In the initial *aggressive defense* phase, i.e., during $(0, \min(t_1, t_2))$, the susceptibles select the minimum possible reception rate, and the dispatchers transmit the patches whenever they are in contact with any other node. Thus, the quarantining is the most stringent, and the recovery most rapid during this phase. Then, in the interim *watchful* phase, i.e., during $(\min(t_1, t_2), \max(t_1, t_2))$, one of the defense controls subside while the other continues as before. If $t_1 < t_2$, then the reception rate control subsides (i.e., the susceptibles select their normal reception rate $u_{\text{norm}}^{N_r}$). If, however $t_1 > t_2$, then instead the dispatchers stop transmitting the patches. If $t_1 = t_2$, there is no watchful period. Finally, in the terminal *relaxed* phase, i.e., in $(\max(t_1, t_2), T)$, both defense controls subside, that is, the susceptibles select their normal reception rate and the dispatchers do not transmit the patches. Thus, the

QoS in data traffic is back to its normal value and the resource consumption overhead due to patch transmission ends.

The durations of the phases (i.e., the values of the threshold times t_1, t_2) and which defense subsides in the interim watchful period, depend on the damage coefficients $\kappa_I, \kappa_D, K_D, \kappa_r, \kappa_i$. For example, if the last two are very high (relative to the first three), t_1 and/or t_2 may turn out to be zero. If $t_1 = 0$, then the susceptibles always select their normal reception rate, and the system never quarantines the infection. Similarly, if $t_2 = 0$, the dispatchers never transmit patches and hence there is no immunization, nor healing. Finally, note that the defense strategy always chooses either the maximum or the minimum values of the parameters except possibly in a set of measure zero (i.e., except possibly at t_1, t_2). Such strategies are referred to as *bang-bang* in the control literature. We conclude this sub-section by proving Theorem 2.

Proof: The continuity of $\psi^{N_r}(\cdot), \psi^{N_i}(\cdot)$ follows from those of the co-state functions. From the final conditions on the co-state functions, i.e., (8), $\psi^{N_r}(T) = -\kappa_r < 0$, $\psi^{N_i}(T) = R_0 > 0$. We show that $\psi^{N_r}(\cdot)$ ($\psi^{N_i}(\cdot)$, resp.) are strictly decreasing (increasing, resp.) functions of time. Thus, each has at most one zero-crossing point in $(0, T)$; denote these as t_1, t_2 . If ψ^{N_r} (ψ^{N_i} , resp.) has no zero crossing point in $(0, T)$, $t_1 = 0$ ($t_2 = 0$, resp.). Thus, from the continuity of the $\psi(\cdot)$ functions, and from their terminal values, (i) $\psi^{N_r}(\cdot)$ is negative in (t_1, T) and positive in $(0, t_1)$, and (ii) $\psi^{N_i}(\cdot)$ is positive in (t_2, T) and negative in $(0, t_2)$. The theorem follows from (11) and (12).

We prove the strict monotonicity of $\psi^{N_r}(\cdot), \psi^{N_i}(\cdot)$, using:

Lemma 2: $\lambda_S > 0$ and $\lambda_I > \lambda_S, \lambda_D \geq 0 \forall t, 0 < t < T$.

The lemma is intuitive since the shadow prices (i.e., co-state variables) associated with the susceptibles and dead nodes ought to be positive, and also the shadow price associated with the infectives ought to be at least as high as that associated with susceptibles. The proof (provided in Appendix-B) requires detailed analysis of the state and co-state differential equations (1), (7) respectively, are less direct.

1) *Strict monotonicity of $\psi^{N_r}(\cdot)$:* We show that $\dot{\psi}^{N_r}(t)$ is strictly negative at all $t \in (0, T)$ ¹⁵

$$\dot{\psi}^{N_r} = \frac{\partial}{\partial t} \psi^{N_r} = (\dot{\lambda}_I - \dot{\lambda}_S)\beta_0 I S + (\lambda_I - \lambda_S)\beta_0 \dot{I} S + (\lambda_I - \lambda_S)\beta_0 I \dot{S}$$

which after replacement and simplification yields

$$\begin{aligned} \Rightarrow \frac{\dot{\psi}^{N_r}}{\beta_0 I S} &= -\kappa_I - u^M(\lambda_D - \lambda_S) - \beta_1 \lambda_I u^{N_i} + \beta_2 \lambda_S u^{N_i} \\ &\quad - \kappa_I - (\lambda_D - \lambda_I)u^M - (\lambda_I - \lambda_S)u^M - (\beta_1 - \beta_2)\lambda_I u^{N_i} \\ &\quad - (\lambda_I - \lambda_S)\beta_2 u^{N_i} \end{aligned}$$

From (13), lemma 2 and since $\kappa_I > 0, \beta_1 \geq \beta_2, u^M(t) \geq 0, u^{N_i}(t) \geq 0$ at all t , the right hand side is negative. The result follows since $\beta_0 > 0$ and $S(t) > 0, I(t) > 0$ at all t (lemma 1).

¹⁵partial derivative w.r.t time, only because of the dependence also on the initial values for the states. Otherwise, t is the only independent variable.

2) *Strict monotonicity of $\psi^{N_i}(\cdot)$:*

$$\begin{aligned} \dot{\psi}^{N_i} &= \frac{\partial}{\partial t} \psi^{N_i} = \\ &= (\dot{\lambda}_I - \dot{\lambda}_S)\beta_0 I S + (\lambda_I - \lambda_S)\beta_0 \dot{I} S + (\lambda_I - \lambda_S)\beta_0 I \dot{S} \\ &\Rightarrow \frac{\dot{\psi}^{N_i}}{\beta_0 I} = \kappa_I \beta_2 + \beta_2 u^M \lambda_D + \beta_0 \beta_1 S u^{N_r} (\lambda_I - \lambda_S) \end{aligned}$$

The R.H.S is positive from lemma 2 and since $\kappa_I > 0, \beta_0 > 0$. Thus, $\dot{\psi}^{N_i} > 0$ since $\beta_0 > 0$ and $I(t) > 0$ at all t (lemma 1). ■

D. Structure results for the saddle point attack strategy

The saddle-point attack has a simple *first-amass, then slaughter* structure in the special case that the worm benefits from killing only through the final tally of the dead (i.e., $\kappa_D = 0$), and the patches can only immunize the susceptibles, but can not heal the infectives (i.e., $\beta_2 = 0$). Specifically:

Theorem 3: For the saddle-point attack strategy $u^M(\cdot)$, there exists a time $t_3, 0 \leq t_3 < T$ such that $u^M(t) = 0$ for $0 < t < t_3$, and $u^M(t) = u_{\max}^M$ for $t_3 < t < T$.

Thus, the worm does not kill any infective during the initial amass period of $(0, t_1)$ when it uses them to spread the infection; it slaughters them at the maximum rate subsequently. The intuition behind this structure is as follows. Once the worm infects a host, it never loses it to the recovery process, and thus, since it benefits from killing a host only because this enhances the final tally of the dead, it ought to kill hosts towards the end and utilize them before. The proof follows.

Proof: Note that $\psi^M(T) = K_I I(T) > 0$ (because of lemma 1). Thus, as in the proof of Theorem 2, the result follows if we can show that $\psi^M(t)$ crosses zero at most once. We establish this slightly differently: we show that $\dot{\psi}^M$ is strictly positive at its zero-crossing point (as opposed to showing it for all t). But this is also sufficient to conclude ψ^M has at most one zero-crossing point.

$$\begin{aligned} \dot{\psi}^M &= I(\dot{\lambda}_D - \dot{\lambda}_I) + \dot{I}\psi^M = \kappa_I - \kappa_D + u^M(\lambda_D - \lambda_I) \\ &\quad - \beta_2 \lambda_I u^{N_i} + S\beta_0 u^{N_r}(\lambda_I - \lambda_S) + \dot{I}\psi^M \end{aligned}$$

At a zero-crossing point of ψ^M , the last term vanishes. Now, $\kappa_D = \beta_2 = 0$, and the remaining terms are all non-negative because of (13) and lemma 2. The result follows since $\kappa_I > 0$. ■

The saddle-point attack strategy may however be more involved when either $\beta_2 > 0$ or $\kappa_D > 0$. For example, fig. 2 depicts the saddle point strategies and the state evolution in an example scenario where $\kappa_D = 20, \beta_2 = 0.109$. The initial infection is relatively high ($I_0 = 0.3$) and the dispatchers are relatively few ($R_0 = 0.1$). The malware starts killing the nodes from the beginning, but around the time that the defense strategy relaxes the reception rate of the nodes to normal, the malware stops the killing and infects the newly accessible susceptible nodes, boosting the fraction of the infective and shortly, starts to kill them all again.

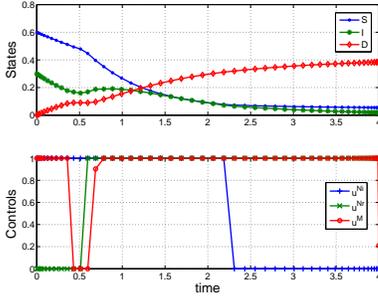


Fig. 2: State evolution and saddle-point strategies. The parameters of the game are as follows: $\kappa_I = 10$, $\kappa_D = 20$, $\kappa_u = 10$, $\kappa_r = 15$, $\beta_2 = \beta_1 = \beta_0 = 0.109$, $K_I = K_D = 0$, and initial fractions $I_0 = 0.3$, $R_0 = 0.1$, $D_0 = 0$, and $T = 4$ hours.

IV. ISSUES RELATED TO IMPLEMENTATION

The simple structure of the saddle-point defense strategies, as established in Theorem 2, are conducive to implementation in resource constrained wireless devices. The threshold times can be computed by a central unit that estimates the system parameters β_0, β_1 and knows the damage coefficients $\kappa_I, \kappa_D, K_D, \kappa_r, \kappa_i$. This computation needs to be performed once, (at $t = 0$, i.e., when the central unit learns the presence of the worm in the system), and transmitted to all devices via a secure broadcast. Since this is a one-time transmission, such secure broadcasts can be afforded. The devices can subsequently execute the robust strategies without coordinating any further among themselves or with the central unit.

Note that t_1, t_2 can be determined by solving a system of 6 differential equations, as described in §III-B. Such systems can be solved very fast due to the existence of efficient numerical algorithms for solving differential equations, and the computation time is constant in that it does not depend on the number of nodes N . For example, we obtained computation times of 1 second using an 2.66 GHz Intel Xeon CPU X 5355. Given that many mobile devices have computing capabilities, and that this is a one time computation, it can even be executed at each mobile device once they have estimated and/or learned $\beta_0, \beta_1, \kappa_I, \kappa_D, K_D, \kappa_r, \kappa_i$.

In practice, due to drifts in local clocks, different nodes will increase (decrease, resp.) the reception (patching, resp.) rates to normal values at different times instead of exactly at t_1, t_2 . Our simulations presented in the next section reveal that the overall costs are robust to clock drifts.

V. PERFORMANCE EVALUATION

Epidemic models have been validated for several mobile wireless networks through experiments as well as network simulations (see e.g. [4], [20]). Our simulations for specific wireless networks such as DTNs and cellular networks also show a close match, even in cases where homogenous mixing assumption does not hold. Here, we compare the overall damages predicted by the epidemic differential equations (1) and obtained through simulations in three different scenarios.

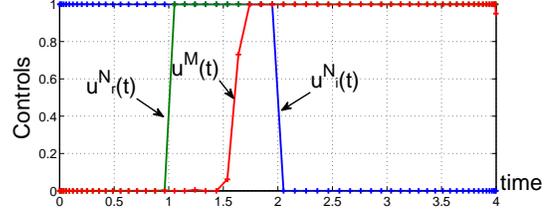


Fig. 5: Saddle-point defense and attack strategies for the game considered in §V. Here, $\kappa_I = 10$, $\kappa_D = 20$, $\kappa_u = 10$, $\kappa_r = 15$, $\beta_1 = \beta_0 = 0.109$, $\beta_2 = K_I = K_D = 0$, and initial fractions $I_0 = 0.1$, $R_0 = 0.1$, $D_0 = 0$, and $T = 4$ hours.

First, we consider a DTN with 41 nodes (the number 41 is in accordance with the experiment reported in [8]), where nodes communicate only when they move to communication range of each other, i.e., when they meet. We allow the nodes to move as per a uniform mobility model [6], with average speed $v = 15\text{km/h}$ and with communication range 50m . Defense and attack strategies are saddle-point strategies calculated based on the estimated β_0 and β_1 for each I_0 . We consider different initial fraction of the infectives, specifically $I_0 \in \{0.01, 0.02, 0.05, 0.10, 0.15, 0.20, 0.25\}$.

Fig. 3(a) reveals that the average of the state fractions ($S(t), I(t), D(t)$) over 20 runs of the simulation closely match those predicted by the epidemic model differential equations (1). Moreover, Fig. 4(a) reveals that the average of the total damage over 20 runs of the simulation with the above parameters, closely match those predicted by the epidemic model; also as expected the damage increases with increase in I_0 . Similar trends and matches can be observed for random waypoint and random direction mobility models (defined in [6]). Such close match is expected since homogenous mixing holds for these models [6]. We next consider the mobility pattern reported in [8] (based on measurements on human mobility during Infocom 2005) that does not satisfy homogenous mixing. Here, inter-contact times are power-law distributed, which arises since nodes which have just met are more likely to meet in near future than those who had met a long time ago. Nevertheless, the average of 20 runs shows that the overall damage follow similar trends (fig. 4(b)) as under the epidemic representations, with universally lower overall damages as compared with the calculated damage. This is intuitively because in the lack of homogenous mixing, the infection tends to stay local and less frequently reaches new (susceptible) nodes, which has a self-suppressing effect on the spread of the malware. This phenomenon can be better seen in fig. 3(b).

Finally, we consider a cellular network composed of 400 nodes and 8 base stations. Nodes follow uniform mobility and are associated with the nearest base-station. Infective nodes try to transmit the malware to randomly chosen IDs (cell phone number) - the communication proceeds through the base stations serving the node-pair. The security patches are distributed by base stations to the mobiles via control channels. The overall data (and control message) exchange bandwidth

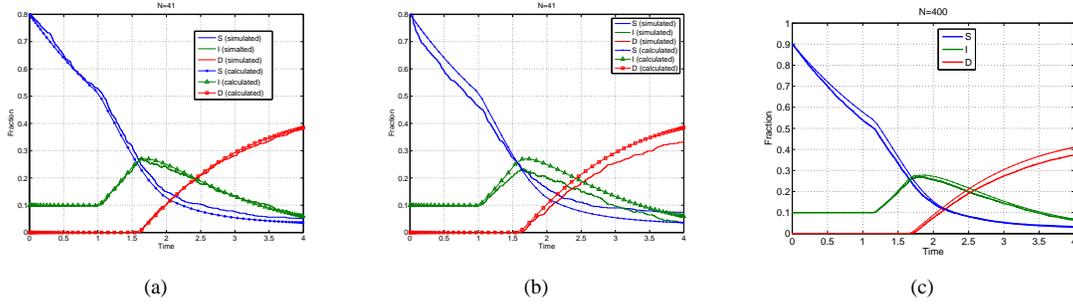


Fig. 3: Average of 20 different runs of the evolution of the states under their the saddle point strategies for DTNs with homogenous mixing (uniform mobility model) in (a), DTNs with non-homogenous mixing (power-law inter-meeting times) in (b), and for a cellular network in (c).

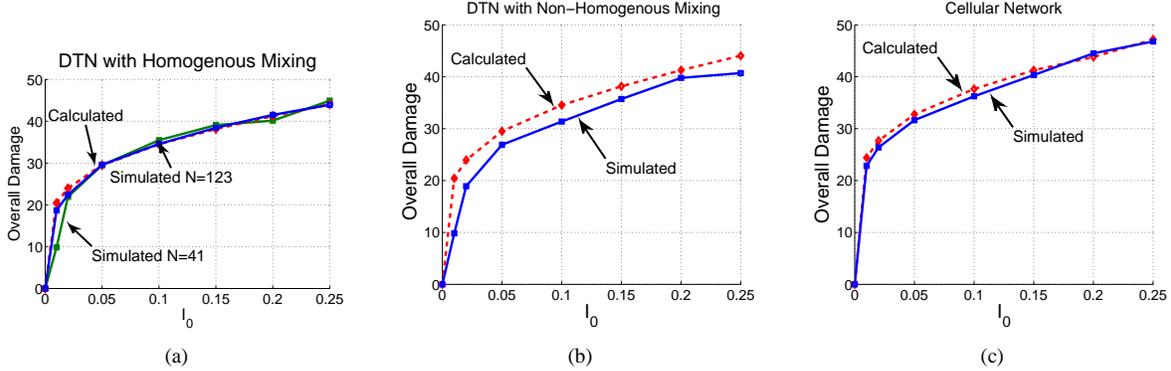


Fig. 4: Average of 20 different runs of the overall damage under their respective saddle point strategies for DTNs of 41 and 123 nodes with homogenous mixing (uniform mobility model) in (a), DTNs with 41 nodes with non-homogenous mixing (power-law inter-meeting times) in (b), and for a cellular network of 400 nodes and 8 BST's in (c). Fig. 4(a) also shows that the match improves with increasing N .

of each base-station is divided equally among the associated nodes. Fig. 3(c) and 4(c) show an acceptable match between our simulation and the epidemic model both for states and game values in the case of a cellular network as well.

We next evaluate the performance, i.e., the overall damage, when nodes' clocks drift from the global clocks by different amounts, and hence they choose different threshold times (optimal threshold time + individual drift). We consider the DTN setting with uniform mobility model, and clock drifts which are statistically independent and uniformly distributed between $-A$ and A . Fig. 6 depicts the overall damage as a function of A averaged over 100 simulation runs. Note that even for A as large as $T/2$ (i.e., 50% inaccuracy in the value of the threshold times) the increase in the overall damage is less than 9%.

At this step, using epidemic representations, we will assess the advantage of considering a dynamic game and implementing saddle-point strategies as robust defense against a dynamically optimizing malware. We now measure the gap between the maximum value of the incurred damage if the defense parameters, i.e., u^{N_r} and u^{N_i} , do not change with time, and that when saddle-point defense strategies are used. We will refer to the former as static strategies. Fig. 7 depicts the maximum damages incurred by the *best* static and dynamic saddle-point defense strategies for different values of the initial

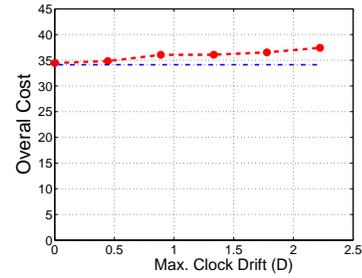


Fig. 6: Robustness of the saddle-point strategy with respect to clock drift. The increase in the overall cost is less than 9%.

fraction of infective nodes (i.e., I_0 is between 0.1 to 0.6) when the other parameters are the same as those reported in the caption of fig. 5. By best static, we mean the fixed reception rate, as well as the fixed dissemination rate of patches are those that achieve the least damage among all possible fixed choices. Saddle-point defense strategies result in a 220% to 270% reduction in the overall damage.

CONCLUSION

We have investigated strategic confrontations of malware attack and network defense in mobile wireless networks through dynamic choices of reception and patching rates (network's actions) and annihilation rate of the infectives (malware's

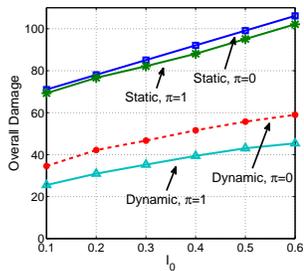


Fig. 7: Comparison of the maximum damage for the best static choice of defense parameters and dynamic saddle-point defense strategies.

action). Using a dynamic game formulation, we prove that the robust defense strategies have simple structures conducive to implementation in resource constrained wireless devices. Our performance evaluations based on simulations and numerical computations reveal that the performance (overall damage) is robust to clock drifts at nodes and is significantly better than when the reception and patching rates are fixed (i.e., are not allowed to vary with time). The analysis is directed towards capturing scenarios where neither the attack nor the defense has access to exact network state information, and the spread is homogenous; design of robust defense when node localities play significant role in the spread of malware constitutes directions for future research.

REFERENCES

- [1] T. Alpcan and T. Basar. A game theoretic analysis of intrusion detection in access control systems. In *Proc. of the 43rd IEEE Conf. on Decision and Control*, pages 1568–1573, 2004.
- [2] C. Bettstetter. Mobility modeling in wireless networks: categorization, smooth movement, and border effects. *ACM SIGMOBILE Mobile Computing and Communications Review*, 5(3):55–66, 2001.
- [3] Z. Chen, L. Gao, and K. Kwiat. Modeling the spread of active worms. In *IEEE INFOCOM 2003*, volume 3.
- [4] R. Cole. Initial Studies on Worm Propagation in MANETS for Future Army Combat Systems, 2004.
- [5] D. Daley and J. Gani. *Epidemic modelling: an introduction*. Cambridge Univ Pr, 2001.
- [6] R. Groenevelt, P. Nain, and G. Koole. The message delay in mobile ad hoc networks. *Performance Evaluation*, 62(1-4):210–228, 2005.
- [7] W. He, C. Xia, H. Wang, C. Zhang, and Y. Ji. A Game Theoretical Attack-Defense Model Oriented to Network Security Risk Assessment. In *Proc. of ICCSSE-Vol.3*, pages 1097–1103, 2008.
- [8] P. Hui, A. Chaintreau, J. Scott, R. Gass, J. Crowcroft, and C. Diot. Pocket switched networks and human mobility in conference environments. In *Proc. of the 2005 ACM SIGCOMM workshop on Delay-tolerant networking*, page 251. ACM, 2005.
- [9] W. Jiang, H. Zhang, Z. Tian, and X. Song. A Game Theoretic Method for Decision and Analysis of the Optimal Active Defense Strategy. In *Proceedings of the 2007 Intl. Conf. on Computational Intelligence and Security*, pages 819–823. IEEE Computer Society, 2007.
- [10] M. Khouzani, E. Altman, and S. Sarkar. Optimal Quarantining of Wireless Malware Through Power Control. In *Proc. of the Fourth Symp. on Inf. Theory and Applications*, 2009.
- [11] M. Khouzani and S. Sarkar. Dynamic Malware Attack in Energy-Constrained Mobile Wireless Networks. In *Proc. of the Fifth Symp. on Inf. Theory and Applications*, 2010.
- [12] M. Khouzani, S. Sarkar, and E. Altman. Maximum Damage Malware Attack in Mobile Wireless Networks. *INFOCOM 2010*.
- [13] M. Khouzani, S. Sarkar, and E. Altman. Dispatch then Stop: Optimal Dissemination of Security Patches in Mobile Wireless Networks. In *49th IEEE CDC*, 2010.
- [14] M. H. R. Khouzani, S. Sarkar, and E. Altman. Optimal propagation of security patches in mobile wireless networks: extended abstract. In *SIGMETRICS*, pages 355–356, 2010.

- [15] H. Kuhn and G. Szegő. *Differential games and related topics*. North-Holland Pub. Co., 1971.
- [16] T. Kurtz. Solutions of ordinary differential equations as limits of pure jump Markov processes. *Journal of Applied Probability*, pages 49–58, 1970.
- [17] P. Liu, W. Zang, and M. Yu. Incentive-based modeling and inference of attacker intent, objectives, and strategies. *ACM Transactions on Information and System Security (TISSEC)*, 8(1):118, 2005.
- [18] R. Núñez-Queija and B. Prabh. Scaling laws for file dissemination in P2P networks with random contacts. In *Proc. of IWQoS*, 2008.
- [19] S. Tanachaiwiwat and H. A. Encouter-based worms: Analysis and defense. *Ad Hoc Networks, Elsevier JOURNAL*, 2009.
- [20] S. Tanachaiwiwat and A. Helmy. VACCINE: War of the worms in wired and wireless networks. In *IEEE INFOCOM*, pages 05–859, 2006.
- [21] N. Weaver, V. Paxson, and S. Staniford. A worst-case worm. In *Proc. Third Annual Workshop on Economics and Information Security (WEIS04)*, 2004.
- [22] X. Yan and Y. Zou. Optimal Internet Worm Treatment Strategy Based on the Two-Factor Model. *ETRI JOURNAL*, 30(1):81, 2008.
- [23] Y. Zhang and W. Lee. Intrusion detection in wireless ad-hoc networks. In *Proc. of the 6th annual intl. conf. on Mobile computing and networking*, page 283. ACM, 2000.
- [24] Z. Zhu, G. Cao, S. Zhu, S. Ranjan, and A. Nucci. A social network based patching scheme for worm containment in cellular networks. *IEEE INFOCOM*, 2009.
- [25] C. Zou, W. Gong, and D. Towsley. Worm propagation modeling and analysis under dynamic quarantine defense. In *Proc. of the 2003 ACM workshop on rapid malware*, pages 51–60, 2003.

APPENDIX

APPENDIX-A: PROOF OF LEMMA 1

Statement: Any pair of strategies $(u^N(\cdot), u^M(\cdot))$ that satisfy the control constraints (5a), (5b), satisfy the state constraints (3) and ensure that $I(t) > 0$, $S(t) > 0$ for all $t \in [0, T]$.

Proof: All S , I and D , resulting from (1) (and thus any continuous functions of them) are continuous functions of time. Since $0 < I_0 < 1$, the initial conditions in (2) ensure that the state constraints $S > 0$ and $I > 0$ are strictly met at $t = 0$. The continuity of S and I functions ensure that there exists an interval of nonzero length starting at $t = 0$ on which both S and I are strictly positive. Thus, from (1c) and since $u^M(t) \geq 0$, $\dot{D} \geq 0$ in the above interval. Thus, since $D(0) = 0$, $0 \leq D$ in this interval as well. Since $\frac{d}{dt}(S + I + D)|_{t=0} = -\beta_1 u^{N_i}(0)R_0S(0) - \pi\beta_1 u^{N_i}(0)R_0I(0) \leq 0$ and $S(0) + I(0) + D(0) = 1$, there exists an interval after $t = 0$ over which the constraint of $S(t) + I(t) + D(t) \leq 1$ is met.

Now, suppose by contradiction that $t_0 \leq T$ be the first time after $t = 0$ at which, at least one of the constraints of $0 \leq S$, I and $S + I + D \leq 1$ becomes active, or $0 \leq D$ becomes violated right after it. That is, at t_0 , we have (1) $S = 0$ OR (2) $I = 0$ OR (3) $S + I + D = 1$ OR (4) there exists an $\epsilon > 0$ such that $D < 0$ on $(t_0 \dots t_0 + \epsilon)$; AND throughout $(0, t_0)$, we have $0 < S, I$ and $S + I + D < 1$ and $D \geq 0$. Hence, for $0 \leq t < t_0$ from (1a) we have $\dot{S} \geq -\beta_0 S - \beta_1 R_0$. Hence, $S(t) \geq S(0)e^{-(\beta_0 + \beta_1 R_0)t} \geq S(0)e^{-(\beta_0 + \beta_1 R_0)t_0}$ for all $0 \leq t < t_0$. Since S is continuous, $S(t_0) \geq S(0)e^{-(\beta_0 + \beta_1 R_0)t_0}$. Similarly, we can show that $I(t_0) \geq I(0)e^{-(\beta_0 + \pi\beta_1 R_0)t_0}$. Thus, since $S(0) > 0$, $I(0) > 0$, neither (1) nor (2) could have happened. Also, $\frac{d}{dt}(S + I + D) = -\beta_1 u^{N_i}(t)R_0S(t) - \pi\beta_1 u^{N_i}(t)R_0I(t) \leq \beta_1 R_0 S(0)e^{-(\beta_0 + \beta_1 R_0)t_0} - \pi\beta_1 R_0 I(0)e^{-(\beta_0 + \pi\beta_1 R_0)t_0} < 0$

throughout $[0 \dots t_0]$. Since $S(0) + I(0) + D(0) = 1$ we have $(S + I + D)|_{t=t_0} < 1$, showing that (3) is impossible. Moreover, from (1c), and since $I(t_0) > 0$, and I is continuous, there exists an ϵ' such that $\dot{D} \geq 0$ over $(t_0 \dots t_0 + \epsilon')$. From continuity of D , $D(t_0) \geq 0$. Thus, $0 \leq D$ over $(t_0 \dots t_0 + \epsilon')$, dismissing the possibility of (4). This negates the existence of t_0 and the lemma follows. ■

APPENDIX-B: PROOF FOR LEMMA 2

Statement: $\lambda_S > 0$ and $\lambda_I > \lambda_S$, $\lambda_D \geq 0 \forall t, 0 < t < T$.

Proof: Since $\lambda_D(T) = K_D \geq 0$ (from (8)), and $\frac{d}{dt}\lambda_D \leq 0$, $\lambda_D \geq 0$.

Now, for the rest, we argue in two steps.

Step 1: $\lambda_S(T) = 0$ and $\lambda_I(T) = K_I = 0$, also:

$$\dot{\lambda}_I(T) - \dot{\lambda}_S(T) = \dot{\lambda}_I(T) = -\kappa_I - K_D u^M(T) < 0$$

Therefore, $\exists \epsilon > 0$ s.t. on $(T - \epsilon \dots T)$ we have $\lambda_S > 0$ and $(\lambda_I - \lambda_S) > 0$.

Step 2: Proof by contradiction. Let τ be such that:

$$\begin{aligned} \lambda_S > 0, (\lambda_I - \lambda_S) > 0 \quad \text{on } (\tau \dots T) \quad \& \\ \lambda_S(\tau) = 0 \quad \text{OR} \quad \lambda_I(\tau) = \lambda_S(\tau) \end{aligned}$$

From the continuity of the co-state functions, $(\lambda_I(\tau) - \lambda_S(\tau)) \geq 0$, and $\lambda_S(\tau) \geq 0$.

We first prove that $(\lambda_I(\tau) - \lambda_S(\tau)) > 0$. Suppose not. Then, $\lambda_I(\tau) = \lambda_S(\tau)$. Thus:

$$\begin{aligned} \dot{\lambda}_I(\tau) - \dot{\lambda}_S(\tau) = \\ -\kappa_I + \lambda_I \beta_2 u^{Ni} - (\lambda_D - \lambda_I) u^M - \lambda_S \beta_1 u^{Ni} = \\ -\kappa_I - \lambda_S u^{Ni} (\beta_1 - \beta_2) - (\lambda_D - \lambda_I) u^M \end{aligned}$$

Here, (i) the first term is strictly negative¹⁶, (ii) the second term is negative because $\lambda_S(\tau) \geq 0$ and $\beta_2 \leq \beta_1$ and (iii) the third term is negative because of (13). Thus, $\dot{\lambda}_I(\tau) - \dot{\lambda}_S(\tau) > 0$. But, then both $\lambda_I(\tau) = \lambda_S(\tau)$, and $(\lambda_I - \lambda_S) > 0$ on $(\tau \dots T)$ can not happen. Thus, $(\lambda_I(\tau) - \lambda_S(\tau)) > 0$.

Now, suppose $\lambda_S(\tau) = 0$.

$$\dot{\lambda}_S(\tau) = -(\lambda_I - \lambda_S) \beta_0 u^{Nr} I|_{\tau} < 0.$$

The last inequality follows since $(\lambda_I(\tau) - \lambda_S(\tau)) > 0$, $\beta_0 > 0$, $u^{Nr} \geq u_{\min}^{Nr} > 0$ and $I(\tau) > 0$ (lemma 1). This again contradicts the assumptions that $\lambda_S(\tau) = 0$ and $\lambda_S > 0$ on $(\tau \dots T)$. Thus, $\lambda_S(\tau) \neq 0$, and hence $\lambda_S(\tau) > 0$. ■

APPENDIX-C: PROOF OF THEOREM 1

Statement: The minimax game defined above has a saddle-point pair of strategies.

Note that this theorem also implies the existence of the value.

Proof: This theorem directly follows from theorem 2 in page 91 of [15]. The necessary conditions of the theorem are readily satisfied in our game. Namely:

(i): the system function $f(t, x, u^P, u^E)$ in this game is continuous in states and controls and is moreover bounded. Note

that this is sufficient for the condition in page 83 of [15] to hold.

(ii): the instantaneous pay-off function $h(t, x, u^P, u^E)$ and the terminal pay-off function $g(x(T))$ are continuous in the states and controls (page 84 of [15]).

(iii): the system function $f(t, x, u^P, u^E)$ is linear in controls. the set defining controls are convex sets; and the instantaneous pay-off function $h(t, x, u^P, u^E)$ is linear in controls (page 91 of [15]). ■

¹⁶Negative in this proof is distinguished from strictly negative.