Maximum Damage Malware Attack in Mobile Wireless Networks

M.H.R Khouzani, Saswati Sarkar, Eitan Altman

***Abstract*—Malware attacks constitute a serious security risk that threatens to slow down the large scale proliferation of wireless applications. As a first step towards thwarting this security threat, we seek to quantify the maximum damage inflicted on the system owing to such outbreaks and identify the most vicious attacks. We represent the propagation of malware in a battery-constrained mobile wireless network by an epidemic model in which the worm can dynamically control the rate at which it kills the infected node and also the transmission range and/or the media scanning rate. At each moment of time, the worm at each node faces the following trade-offs: (i) using larger transmission range and media scanning rate to accelerate its spread at the cost of exhausting the battery and thereby reducing the overall infection propagation rate in the long run or (ii) killing the node to inflict a large cost on the network, however at the expense of loosing the chance of infecting more susceptible nodes at later times. We mathematically formulate the decision problems and utilize Pontryagin Maximum Principle from optimal control theory to quantify the damage that the malware can inflict on the network by deploying optimum decision rules. Next, we establish structural properties of the optimal strategy of the attacker over time. Specifically, we prove that it is optimal for the attacker to defer killing of the infective nodes in the propagation phase for a certain time and then start the slaughter with maximum effort. We also show that in the optimal attack policy, the battery resources are used according to a decreasing function of time, i.e., mostly during the initial phase of the outbreak. Finally, our numerical investigations reveal a framework for identifying intelligent defense strategies that can limit the damage by appropriately selecting network parameters.**

I. Introduction

a) Motivation: Malicious self-replicating codes, known as malware, pose substantial threat to the wireless computing infrastructure. Malware can be used to launch attacks that vary from the less intrusive confidentiality or privacy attacks, such as traffic analysis and eavesdropping, to the more intrusive methods that either disrupt the nodes normal functions such as those in relaying data and establishing end-to-end routes (e.g., sinkhole attacks [1]), or even alter the network traffic and hence destroy the integrity of the information, such as unauthorized access and session hijacking attacks [2], [3]. Malware outbreaks like those of Slammer [4] and Code Red [5] worms in wired Internet have already inflicted expenses of billions of dollars in repair after the viruses rapidly infected thousands of hosts within few hours. New investments have increasingly been directed toward wireless infrastructure thanks to the rapid growth of consumer demands and advancements in wireless technologies. The economic viability of these investments is, however, contingent on the design of effective security countermeasures.

The first step in devising efficient countermeasures is to anticipate malware hazards, and understand the threats they pose, before they emerge in the hands of the attackers [6]. Recognizing the above, specific attacks such as the wormhole [7], sinkhole [1], and Sybil [8], that utilize vulnerabilities in the routing protocols in a wireless sensor network, and their counter-measures, have been investigated before they were actually launched. We pursue the complementary but closely related goals of (i) quantifying fundamental limits on the damages that the attackers can inflict by intelligently choosing their actions, and (ii) identifying the optimal actions that inflict the maximum damage on the network. Such quantification is motivated by the fact that while attackers can pose serious threats by exploiting the fundamental limitations of wireless network, such as limited energy, unreliable communication, constant changes in topology owing to mobility [9], their capabilities may well be limited by the above as well since they rely on the same network for propagating the malware. Finally, the answers will depend on the network parameters such as communication ranges of the nodes, mobility parameters, and also the counter-measure parameters such as the rates of updates of security patches, etc. This will in turn suggest appropriate counter-measures which minimize overall network costs that depend on the costs of the counter-measures and the damages inflicted by the malware.

b) Decision problems of the attackers: Worms spread during data or control message transmission from nodes that are infected (*infectives*) and those that are vulnerable, but not yet infected (*susceptibles*). We consider a pernicious worm that may (i) eavesdrop, (ii) analyze, (iii) alter or destroy traffic and (iv) disrupt the infective host's normal functions (such as relaying data or establishing routes), and even *kill* the host, that is, render it completely dysfunctional (*dead*). This killing process may be triggered by performing a code which inflicts irretrievable hardware damage. For instance, Chernobyl virus [10] could re-flesh the BIOS, corrupting the bootstrap program required to initialize the system. The worm can determine the time to kill, or equivalently the rate of killing the hosts, by regulating the rate at which it triggers such codes.

Counter-measures can be launched by installing security patches that 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. For instance, for SQL-Slammer worms [4], while StackGuard programs [11] immunize the susceptibles by removing the buffer overflow vulnerability that the worms exploit, specialized security patches [12] are required to re-

M. H. R Khouzani and Saswati Sarkar are at the Electrical and Systems Engineering Department of University of Pennsylvania. Their emails are {khouzani,swati} @seas.upenn.edu. Eitan Altman is at INRIA, Sophia Antipolis, France. His email is altman@sophia.inria.fr

move the worm from (and thereby heal) the infectives. Nodes that have been immunized or healed are denoted as *recovered*. Thus, depending on whether the worm kills the infective before it fetches a security-patch, the state of an infective changes to dead or recovered. States of susceptible nodes change to infective or recovered depending on whether they communicate with infectives before installing the security-patches. Note that the counter-measures incur costs, since the patches must be obtained through the bandwidth-limited wireless media involving energy-expensive communications, and different patches incur different costs depending on whether they treat susceptibles or infectives. Thus, such counter-measures must be resorted to, selectively and judiciously.

The goal of the attacker is to infect as many nodes as possible, and use the worms to disrupt the hosts as well as the network functions, while being cognizant of the counter-measures [13]. Killing an infective host sooner rather than later maximally disrupts its functions and thereby inflicts damage on the network right away, but also prevents it from propagating the infection in the network and eavesdropping, analyzing, altering or destroying network traffic. Deferral of killing, on the other hand, may allow the host to be healed of the infection before it can be killed or infect other hosts. It is therefore interesting to determine the instantaneous rate of killing that maximizes the damage inflicted by the worm. Another important decision of the worm pertains to its optimal use of the available energy of the infective nodes. The infectives can accelerate the rate of spread of the worm by increasing their contact rates with susceptibles by selecting higher transmission gains and media scanning rates. Such choice however depletes their energy reserves which are limited as those of any other nodes in wireless networks, which in turn limits the spread of the infection and also their other functionalities such as eavesdropping, traffic destruction, *etc.*

c) Contributions: The fundamental contributions of this paper are threefold. First, we construct a mathematical framework which cogently models the effect of the decisions of the attackers on the state dynamics and their resulting trade-offs through a combination of epidemic models and damage functions (Section II). Specifically, we assume that the damage inflicted by the worm is a cumulative function increasing in the number of infected and dead hosts, both of which change with time. We allow the function to be fairly general, in that it can be either linear or non-linear, and consider that the worm seeks to maximize the damage subject to satisfying certain constraints on the energy consumption of its hosts by dynamically selecting its killing rates and energy usages of its hosts while assuming full knowledge of the network parameters and the counter-measures. The maximum value of the damage function then quantifies the fundamental limits on the efficacy of the worm, particularly, since we assume that the worm has complete knowledge of all the contributing factors, and uses optimal dynamic strategies. The damage maximization problem turns out to be an elegant optimal control problem which can be solved numerically by applying Pontryagin's Maximum Principle [14]–[16] - an effective tool that so far has been rarely used in the context of network security (Section III).

Second, we seek to answer the natural next question of whether in practice the worm can indeed inflict the damage quantified above, or the above quantifications constitute only theoretical upper bounds. Specifically, if the optimal policies that inflict the above maximum damage are complex to execute, then the worm may not be able to execute them since they are limited by the capabilities of their resource constrained hosts as well. Towards this end, we investigate structures of the optimum policies for the worms. Our results are surprising and have negative connotations from the counter-measures point of view since we show that an attacker can inflict the maximum damage by using very simple decisions. We first investigate the case where the worm selects the killing rates dynamically and the energy consumption strategies statically (i.e., once at the beginning of network operation) (Section IV). We prove that the optimal killing rate has the following simple structure: until a certain time (which can be zero depending on the network and counter-measure parameters), the worm does not kill any host, and right after that, it annihilates its hosts at the maximum possible rate until the end of the optimization period (Theorem 1). Thus, the first phase is to *amass* the infectives and then arrives the *slaughter* time. The result carries a qualitative cautionary message for countermeasures as well: an apparently inoffensive malware with little to no disruptive behavior might well be stacking infective hosts for the imminent carnage. In optimal control terminology [14]–[16], we have proved that the optimal strategy has a *bang-bang* structure, that is, at any given time, the killing rate is either at its minimum or maximum possible values; in addition it has at most one jump which necessarily culminates at the maximum possible value. Optimality of this simple strategy for this nontrivial problem is in fact quite surprising.

We next investigate the complementary problem where the worm selects only the optimal energy consumption rate dynamically (Section V). We prove that when the energy consumption costs are convex the worm's optimal energy consumption rate is a decreasing function of time (Theorem 2). Thus, the worm seeks to infect as many hosts as possible early on by selecting the maximum possible values of the media scanning rates and transmission ranges, and thereafter starts to behave more conservatively so as to satisfy the energy consumption constraints. This inevitably slows the further spread of the worm towards the end of the optimization period, but then a large fraction of nodes have already been infected due to the choice of large values of these parameters early on. When the energy consumption costs are concave, the structure results are even more specific: the optimal media scanning rates and transmission ranges are not only decreasing functions of time, but also have a bang-bang nature with at most one jump from the maximum possible value to the minimum possible value. Our numerical computations reveal that when both the killing rates and energy usages are selected dynamically, the optimal strategies follow the above structures as well (Section VI).

Finally, we demonstrate how an understanding of the maximum value of the damage function can facilitate the design of suitable counter-measures. Our numerical computations affirm that as expected the damage can be reduced if the

nodes fetch the security patches at the maximum possible rate (Section VI). However, this incurs cost for the system owing to the energy-expensive communication of the patches through bandwidth-limited wireless media. We devise a framework for determining the above parameter so as to minimize the overall network cost which increases with the damage and the cost associated with installation rate of the security patches.

d) Related Works: Malware outbreaks in wireless networks constitute an emerging research topic (e.g., [17]–[21]), though, the research on spread of malware has traditionally focused on wired networks. Epidemic modeling based on the classic Kermack-Mckendrick model [22] has extensively been used to analyze the spread of malware in wired networks [5], [23], [24], *etc*, and more recently in wireless networks [25]. 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. Most of these however do not identify the optimal policies nor provide provable performance guarantees, but instead propose heuristic dynamic policies in different contexts, and evaluate through simulations the efficacies and various trade-offs of the policies they propose. For example, [24] proposes heuristics for dynamic quarantining of nodes in wired networks that appear suspicious through traffic analysis, and [26] introduces heuristic strategies for dynamically adjusting the transmission power of attacker nodes in wireless networks. We instead obtain attack policies that provably attain the maximum possible damage and consider a general model that incorporates healing, immunization and mortality of nodes.

Interestingly, tools from the optimal control theory such as the effective theorem of Pontryagin maximum Principle has rarely been used for analyzing network security - [27] and our previous work [28] constitute notable exceptions. The first formulates the trade-off for optimal treatment of the infective nodes in wired networks. However, in contrast to our work, the solution is based on numerical evaluations only and no structural property of the optimal policy is established. Our earlier work [28] proposes reduction of reception gain of wireless nodes as a counter-measure for slowing down the spread of malware in wireless networks. Our current work in contrast focuses on the attack viewpoint and considers the transmission range of the infective nodes and the rate of killing as dynamic parameters of the worm to inflict the maximum damage, and therefore invokes and answers a different set of questions using different analytical arguments. Also the model assumed here is more general than in [27], [28] in that the worm causes mortality and the counter-measures include both healing and immunization.

II. System Model

A. Dynamics of State Evolution

A **susceptible** node is a mobile wireless device 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**. Installation of appropriate security patches, by the respective users or the network operator, can *immunize* susceptibles to the recovered states and *heal* infectives to the recovered states. Different security patches may be required for immunization and healing as the first involves rectification of the vulnerability that rendered the susceptibles culpable to the attack, whereas the second involves both the removal of the worm and the vulnerability that the worm exploits.

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 denoted by $n_S(t), n_I(t), n_R(t)$ and $n_D(t)$, respectively, and the corresponding fractions be $S(t) = n_S(t)/N$, $I(t) = n_I(t)/N$, $R(t) = n_R(t)/N$, and $D(t) = n_D(t)/N$ respectively. Then, $S(t)+I(t)+R(t)+D(t) = 1$. We assume that at the time of the outbreak of the infection, that is at time zero, some but not all nodes are infected: $0 < I(0) = I_0 < 1$. For simplicity, we assume $R(0) = D(0) = 0$. Thus, $S(0) = 1 - I_0$.

We now model the dynamics of infection propagation. Nodes are assumed to roam in a vast 2-D region of area A with an average velocity v. An infective transmits a message to a susceptible with a given probability whenever the two are in *contact*, that is, the susceptible is in the transmission range of the infective. Now, this probability is a linear function of the rate at which the infective scans the media in search of susceptibles nearby, and the proportionality constant is determined by the message collision probability η_1. When the communication range of the nodes is small compared to A (which is usually the case in multihop networks), η_1 is essentially determined by the overall node density (N/A). Next, under mobility models such as random waypoint or random direction model [29], Groenevelt *et al.* [30] have shown that the time between consecutive contacts of a specific pair of nodes is nearly *exponentially* distributed, and the rate of this exponential process is linearly dependent[1] on the communication range of the nodes with a proportionality constant η_2 that depends only on v and A. Specifically, $\eta_2 \propto \frac{1}{A}$. Let $u(t)$ be the product of the infective's transmission range and its media scanning rate. Then, the worm is transmitted between a given infective-susceptible pair as per an exponential random process whose rate at any given time t is $\hat{\beta}u(t)$, where $\hat{\beta} = \eta_1\eta_2$. The worm regulates the spread of the infection by controlling $u(t)$ through appropriate choice of its transmission gain and media scanning rate.

We now model the dynamics of mortality, healing and immunization. The worm at an infective host kills the host after a random time which is exponentially distributed with rate $\nu(t)$ at any given time t. Here, the worm regulates the death process by appropriately choosing the instantaneous rate of killing $\nu(t)$ at t; this is accomplished by invoking and executing the code that kills the node at desired rates. The security patches are installed at an infective (susceptible, respectively) after

[1]The result has been proved when the communication range of the nodes is small compared to the total area of the region and v is sufficiently high. Numerical computations reveal that the result holds even otherwise.

exponentially distributed random times starting from when it is infected ($t = 0$, respectively). The delays account for the time required in detection of infection, and fetching the appropriate security patch, etc. The instantaneous rates of these exponential healing and immunization processes for any given infective at any given time t are $B\left(I(t)\right)$ and $Q\left(S(t)\right)$, respectively, where $B(.), Q(.)$ are arbitrary functions that satisfy the following mild assumptions: $\lim_{x\to 0} B(x), \lim_{x\to 0} Q(x)$ are finite, and for $0 < x < 1$, $B(x), Q(x)$ are positive and differentiable, $xB(x)$ is a concave non-decreasing function of x and $xQ(x)$ is a non-decreasing function of x. Note that the functions $B(.)$ and $Q(.)$ are likely to be constants (e.g., $B(x) = B_0, Q(x) = Q_0$ for all x), in practice[2], and any constant function satisfies all of the above properties. Nevertheless, we consider more general functions (such as $Q(x) = x^\alpha$ for $\alpha > -1$ and $B(x) = x^\alpha$ for $-1 < \alpha < 0$) so as to allow for more general scenarios.

Following the conditions assumed for the model, the number of nodes of each type evolves according to a pure jump Markov chain with state vector $(n_S(t), n_I(t), n_D(t))$ (since for all t, $n_S(t) + n_I(t) + n_R(t) + n_D(t) = N$, the state of the Markov chain is three dimensional). Let

$$\beta = \lim_{N\to\infty} N\hat{\beta}, \qquad q(S) = Q(S)S, \qquad b(I) = B(I)I.$$

Now[3] according to the results of [31], as N grows, $S(t)$, $I(t)$ and $D(t)$ converge to the solution of the following system of differential equations[4]:

$$\dot{S}(t) = -\beta u(t) I(t) S(t) - q\left(S(t)\right), \quad S(0) = 1 - I_0 \quad \text{(1a)}$$

$$\dot{I}(t) = \beta u I(t) S(t) - b\left(I(t)\right) - \nu(t) I(t), \quad I(0) = I_0 \quad \text{(1b)}$$

$$\dot{D}(t) = \nu(t) I(t), \quad D(0) = 0. \quad \text{(1c)}$$

and also satisfy the following constraints at all t:

$$0 \le S(t), I(t), D(t) \quad \text{(2a)}$$

$$S(t) + I(t) + D(t) \le 1. \quad \text{(2b)}$$

The convergence is in the following sense:

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

and likewise for $I(t)$ and $D(t)$.

Similar epidemic models have been validated through experiments as well as network simulations to provide an acceptable representation of the spread of malware in mobile wireless networks (see e.g. [32], [33]).

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

[2]This is because the users of infectives and susceptibles are likely to receive the security patches from software stores or servers distributed in the area A. In the first case, the rates are clearly constants. In the latter case, the reception rates of the patches depend on the host's reception gains, servers' transmission gains, collision probabilities etc, and none of the above depend on the infective and susceptible fractions (collision probability depends on the overall node density N/A).

[3]Note that since $\hat{\beta} = \eta_1\eta_2$, and η_1 depends only on the node density, and $\eta_2 \propto \frac{1}{A}$, the limit β exists as long as the node density $\lim_{N\to\infty} N/A$ exists for large N.

[4]Variables with dot marks (e.g., $\dot{S}(t)$) will represent their time derivatives (e.g., time derivative of $S(t)$) and the prime signs (e.g., $q'(S)$) designate their derivatives with respect to their argument (e.g., S).

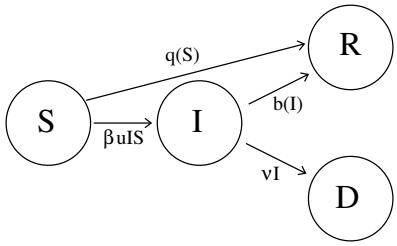

Fig. 1: Transitions. S, I, R, D respectively represent fraction of the susceptible, infective, recovered and dead.

Finally, owing to the technical assumptions we made on $B(.)$ and $Q(.)$, the functions $b(.)$, $q(.)$ exhibit the following properties: $b(0) = q(0) = 0$, and for $0 < I < 1, 0 < S < 1$ $b(I), q(S) > 0$, $b'(I) = db/\,dI \ge 0$, $q'(S) = dq/\,dS \ge 0$, and $b''(I) = d^2b/\,dI^2 \le 0$.

B. Maximum Damage Attack

We consider an attack that seeks to inflict the maximum possible damage in a time window $[0, T]$ of its choice. An attack can benefit over time from the infected hosts, by using the worms to (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. An attacker also benefits by inflicting a large death-toll by the end of the desired time window. These motivate the following damage function:

$$J = \kappa D(T) + \int_0^T f\left(I(t)\right)\, dt. \quad (3)$$

where κ is an arbitrary positive constant, and $f(.)$ is an arbitrary non-decreasing, convex function such that $f(0) = 0$. Note that the assumptions on $\kappa, f(.)$ are mild and natural,and a large class of functions, e.g., $f(I) = KI^\alpha$ for $\alpha \ge 1$ and $K \ge 0$, $f(I) = K(e^I - 1)$ for $K \ge 0$ satisfy them. Finally, an attacker that simply seeks to maximize the final tally of the dead without any other agenda is readily representable by taking $f \equiv 0$.

The attacker seeks to maximize the damage function by appropriately regulating its killing rate $\nu(t)$ and the product $u(t)$ of the transmission range and the scanning rate of the infective nodes[5] subject to:

$$0 \le \nu(t) \le \nu_{\max} \qquad 0 \le u_{\min} \le u(t) \le u_{\max} \quad \text{(4a)}$$

$$\int_0^T h\left(u(t)\right)\, dt \le C \quad \text{(4b)}$$

The bound on $\nu(t)$ is imposed by limitations on the worm's speed of killing an infective host. The bounds on $u(t)$ are dictated by the physical constraints of the transmitters and also for ensuring that the interference and hence collisions between simultaneous transmissions remain limited. The second constraint (*battery* constraint) arises because enhancing $u(t)$ depletes the infective's battery, and the worm needs to ensure

[5]The attacker does not control any other parameter such as the susceptible's reception gain, server's transmission gains, mobility patterns, immunization and healing rate functions, Q(.) and B(.) etc.

that the infective's battery lasts and it can continue to use it and to infect susceptibles for the time period of its operation $[0, T]$ (should it choose not to kill the host earlier). For appropriate functions, $h(.)$ (e.g., $h(u) = K_1 u^r$, for $r \geq 2$), $\int_0^T h(u(t))\, dt$ is the energy consumed by the host if it is infected at $t = 0$ and is not killed before $t = T$ - this is therefore an upper bound on the energy consumption of any infective while it remains infected. We assume that the susceptibles use their battery so as to last much longer than T, and therefore the energy consumed by a host before it is infected is relatively insignificant. Thus, the worm chooses $u(t)$ so that the above upper bound is less than its energy reserve.

It is natural to assume that $h(u)$ is non-decreasing and non-negative. We allow $h(u)$ to be either convex or concave for $u_{\min} \leq u \leq u_{\max}$. Note that when $h(u)$ represents power dissipation associated with u, $h(u)$ must be $K_1 u^r$, for $r \geq 2$ and some non-negative K_1, and is therefore convex. But, if $h(u)$ represents a cost associated with power dissipation, then it may be concave as well. Finally, without loss of generality, $h(u_{\min}) = 0$, as if $h(u_{\min}) > 0$, we can equivalently consider $h(u_{\min}) = 0$, and reduce the bound C appropriately. Any pair of piecewise continuous functions $(\nu, u) : [0, T] \to \mathbb{R}^2$ such that the left and right hand limits exist and that satisfy the above constraints belongs to the *control region* denoted by Ω.

We next show that for any $(\nu, u) \in \Omega$, the state constraints in (2) are automatically satisfied throughout $(0 \ldots T]$. Thus, we ignore (2) henceforth.

Lemma 1: For any $(\nu, u) \in \Omega$, the state functions $(S, I, D) : [0, T] \to \mathbb{R}^3$ that satisfy (1), also satisfy (2). Moreover, $S(t) \geq (1 - I_0)e^{-K_1 t} > 0$, $I(t) \geq I_0 e^{-K_2 t} > 0$ for $t \in [0, T]$ and some finite K_1, K_2.

The proof will reveal that $K_1 = \beta u_{\max} + \max_{0 \leq x \leq 1} q'(x), K_2 = \max_{0 \leq x \leq 1} b'(x)$. The proof is similar to that of lemma 1 in [28], and is relegated to our tech. report [34].

Once the control (ν, u) is selected, the system state vector (S, I, D) is specified at all t as a solution to (1) and hence the value of the damage function J is determined as well. Thus, the control (ν, u) is considered only as a function of time rather than that of the system states, and since the value of J is determined only by the selection of (ν, u), we will henceforth denote J as $J(\nu, u)$ instead.

The state and control functions pair $((S, I, D), (\nu, u))$ is called an *admissible pair* if (i) (ν, u) is in Ω, i.e. satisfies (4), (ii) (ν, u) is piecewise continuous such that the left and right hand limits exist at the points of discontinuity, and (iii) (1) hold. The function (ν, u) is then called an admissible control. Let $((S, I, D), (\nu, u))$ be an admissible pair. If

$$J(\nu, u) \geq J(\underline{\nu}, \underline{u}) \quad \text{for any admissible control } (\underline{\nu}, \underline{u})$$

then $((S, I, D), (\nu, u))$ is called an *optimal solution* and (ν, u) is called an *optimal control* of the problem.

In order to obtain fundamental bounds on the efficacy of the attack, we assume that the attacker computes its optimal control assuming full knowledge of the parameters of the system, such as the mobility pattern, the reception gain of the susceptibles and the healing and immunization rate functions $(B(.), Q(.))$. We also assume that the system selects the above parameters apriori and does not change them with time. The damage can only be equal or lower if the counter-measures are adaptive or the attacker does not know the above parameters.

III. Worm's optimal control

We now present a framework using which the worm can determine its *optimal control* functions (ν, u) and also compute the maximum value of the damage function.

The main challenge in computing the optimal control is that the differential equations (1) can be solved provided the functions (ν, u) are known. Thus, the only approach seems to be that of an exhaustive search on all functions (ν, u) in Ω. This will require the evaluation of the damage function $J(\nu, u)$ for each pair of such functions where the corresponding (I, D) functions required in evaluating $J(\nu, u)$ are obtained by solving (1) for each such pair. But, Ω consists of an uncountably infinite number of such pairs, which rules out an exhaustive search. *Pontryagin's Maximum Principle* however provides an elegant tool for solving this seemingly impossible problem, which we apply next.

First, we introduce a new state variable E to transform the constraint in (4b) to a more treatable one:

$$\dot{E}(t) = -h(u), \quad E(0) = 0, \tag{5}$$

with the final constraint:

$$E(T) \geq -C. \tag{6}$$

Now, note that (5) and (6) are together equivalent to (4b). Thus, the optimal control problem posed in section II can now be modified to augment (1) with (5) and (6), and omit (4b), without any alterations in the set of optimal solutions and in the maximum value of the damage function. We consider this version henceforth.

Let $((S, I, D), (\nu, u))$ be an optimal solution. Consider the *Hamiltonian* H, and *co-state* or *adjoint* functions $\lambda_1(t)$ to $\lambda_4(t)$, and a scalar $\lambda_0 \geq 0$ defined as follows:

$$\begin{aligned} H := \lambda_0 f(I) + (\lambda_2 - \lambda_1)\beta u I S - \lambda_1 q(S) - \lambda_2 b(I) \\ + (\lambda_3 - \lambda_2)\nu I - \lambda_4 h(u). \end{aligned} \tag{7}$$

$$\begin{aligned} \dot{\lambda}_1 &= -\frac{\partial H}{\partial S} = -(\lambda_2 - \lambda_1)\beta u I + \lambda_1 q' \\ \dot{\lambda}_2 &= -\frac{\partial H}{\partial I} = -\lambda_0 f' - (\lambda_2 - \lambda_1)\beta u S + \lambda_2 b' - (\lambda_3 - \lambda_2)\nu \\ \dot{\lambda}_3 &= -\frac{\partial H}{\partial D} = 0 \\ \dot{\lambda}_4 &= -\frac{\partial H}{\partial E} = 0. \end{aligned} \tag{8}$$

along with the transversality conditions:

$$\lambda_1(T) = 0, \quad \lambda_2(T) = 0, \quad \lambda_3(T) = \lambda_0 \kappa \tag{9a}$$

$$\lambda_4(T) \geq 0 \tag{9b}$$

$$\lambda_4(T)(E(T) + C) = 0. \tag{9c}$$

Then according to Pontryagin's Maximum Principle With Terminal Constraints ([14, P.111 theorem 3.14]), there exists continuous and piecewise continuously differentiable co-state functions $\lambda_1, \lambda_2, \lambda_3$ and λ_4, and constant $\lambda_0 \geq 0$ that at every

point $t \in [0 \dots T]$ where $(\nu(.), u(.))$ is continuous satisfy (8), and the transversality conditions (9), and we have:

$$\vec{\lambda} \not\equiv \vec{0} \tag{10a}$$

$$(\nu, u) \in \arg \max_{(\underline{\nu}, \underline{u}) \in \Omega} H(\vec{\lambda}, (S, I, D), (\underline{\nu}, \underline{u})) \tag{10b}$$

From (10b), (ν, u) must be selected to ensure that $(\lambda_3 - \lambda_2)\nu I$ is maximized, and $\frac{\partial H}{\partial u} = 0$. Since $I > 0$ (lemma 1),

$$\nu = \max(\lambda_3 - \lambda_2, 0)\nu_{\max} \tag{11a}$$

and for convex $h(u)$,

$$u = \begin{cases} u_{\min}, & \psi \leq h'(u_{\min}) \\ h'^{-1}(\psi), & h'(u_{\min}) < \psi \leq h'(u_{\max}) \\ u_{\max}, & h'(u_{\max}) < \psi. \end{cases} \tag{11b}$$

where $\psi := (\lambda_2 - \lambda_1)\beta IS/\lambda_4$. Combining (8), (11) and (9), we obtain a system of (non-linear) differential equations with final values specified that involve only the state and co-state functions (and not the control (ν, u)). Functions λ_1 to λ_4 and scalar λ_0 that satisfy the above differential equations and final values, can therefore be obtained using standard numerical procedures that solve differential equations [35]. Now, the optimal control (ν, u) can be obtained using the above solutions in (11).

Here, we obtain the following properties of the Hamiltonian, and system states, that we use later.

First, the system is *autonomous*, i.e., the Hamiltonian and the control region do not have an explicit dependency on the independent variable t.Thus, [15, P.236]

$$H(S(t), I(t), D(t), \nu(t), \lambda_1(t), \lambda_2(t), \lambda_3(t)) \equiv \text{constant}. \tag{12}$$

Second, we have the following lemma:.

Lemma 2: $(f'(I)I - f(I)) \geq 0$ and $(b(I) - b'(I)I) \geq 0$ for all $t \in [0 \dots T]$.

Proof: By lemma 1, I and S are nonnegative. Define $\xi(I) = f'(I)I - f(I)$. Since $f(0) = 0$, we have $\xi(0) = 0$. Also, $\frac{d}{dI}\xi(I) = \xi' = f''(I)I + f'(I) - f'(I) = f''(I)I$. Following lemma 1 and properties of f, we observe that $\xi' \geq 0$ for all $t \in [0 \dots T]$. Thus, since $\xi(0) = 0$, $\xi(I) = f'(I) - f(I)I \geq 0$ for all $t \in [0 \dots T]$. Likewise for b. ∎

IV. Optimal rate of killing

In this section, we consider the case in which the transmission range and media scanning rate in the infective nodes is selected a priori by the worm and is fixed throughout the $[0 \dots T]$ interval. Specifically, $u(t) = u_0 > 0$, for all $t \in [0 \dots T]$, and u_0 is chosen such that the constraint (4b) is satisfied, i.e., $h(u_0) \leq C/T$. Therefore, the state function E and thus, the co-state function λ_4 need not be introduced. Thus, without loss of generality, $\lambda_4 \equiv 0$ in (8).

We obtain structural results for the optimal killing rate $\nu(.)$ as a function of time, that maximizes the overall damage function in (3). Specifically, Theorem 1 shows that $\nu(.)$ is of *bang-bang* form, that is, it possesses only two possible values $\nu_{\max}$ and 0, and switches abruptly between them. It has at most one such jump, which necessarily culminates at $\nu_{\max}$.

Theorem 1: When $u(t) = u_0$ for all $t \in [0, T]$, such that $u_0 \in [u_{\min}, u_{\max}]$ and u_0 satisfies constraint (4b), the optimal $\nu(t)$ that maximizes the worm's damage function in (3) is characterized as follows: $\exists t_1 \in [0 \dots T)$ such that $\nu(t) = 0$ for $0 < t < t_1$ and $\nu(t) = \nu_{\max}$ for $t_1 < t < T$.

Proof: First, we assume, without loss of generality, $\lambda_0 = 1$. This is because if $\lambda_0 > 0$, then the Hamiltonian, H, can be can be re-scaled by $1/\lambda_0$, and by replacing λ_i/λ_0, $i = 1 \dots 4$ instead of $\lambda_1 \dots \lambda_4$, the conditions of Pontryagin Maximum Principle are satisfied for $\lambda_0 = 1$. On the other hand, if $\lambda_0 = 0$ then (7) constitutes a *linear autonomous ODE* with the final constraint of $\vec{\lambda}(T) = \vec{0}$ which, from vector space theory [35], has the unique solution of $(\lambda_1, \dots, \lambda_4) = \vec{0}$ for all $t \in [0 \dots T]$. This however contradicts the necessary condition of $\vec{\lambda} \not\equiv \vec{0}$ of (10a).

Let the switching function, φ, be defined as follows:

$$\varphi := (\lambda_3 - \lambda_2)I$$

which is a continuous and piecewise continuously differential function of time and referring to (9), has the following final value:

$$\varphi(T) = \kappa I(T) > 0. \tag{14}$$

where positivity comes from $\kappa > 0$, and $I > 0$ according to lemma 1. Introduction of φ, along with $\lambda_0 = 1$ and $\lambda_4 \equiv 0$, allow us to rewrite the Hamiltonian in (7) as follows:

$$H = f + (\lambda_2 - \lambda_1)\beta u_0 IS - \lambda_1 q - \lambda_2 b + \varphi\nu. \tag{15}$$

According to Pontryagin's Maximum Principle, we have:

$$H(S, I, D, \nu, \lambda_1, \lambda_2, \lambda_3) \geq H(S, I, D, \underline{\nu}, \lambda_1, \lambda_2, \lambda_3) \quad \text{over all admissible } \underline{\nu}. \tag{16}$$

Hence, the optimal ν satisfies $\varphi\nu \geq \varphi\underline{\nu}$, where $\underline{\nu}$ is any admissible controller, i.e., $\underline{\nu} \in [0 \dots \nu_{\max}]$. Thus, to find the optimal controller, one needs to maximize the linear function $\varphi\nu$ over the admissible set $\nu \in [0 \dots \nu_{\max}]$, which yields:

$$\nu = \begin{cases} 0, & \varphi < 0 \\ \nu_{\max}, & \varphi > 0, \end{cases} \tag{17}$$

hence, the name switching function. An immediate observation of the above property is the following important property:

$$\varphi\nu \geq 0. \tag{18}$$

Also note that according to the continuity of the φ and its final value (14) and following (17), we have $\nu = \nu_{\max}$ over an interval of nonzero length toward the end of $(0 \dots T)$ interval which extends until time T. Specifically, we have $\nu(T) = \nu_{\max}$ and ν at T is differentiable and $\dot{\nu}(T) = 0$.

Now, in order to establish the statement of the theorem, we will show that the switching function φ has at most one zero-crossing point. We show this by proving that the right side time derivative of φ at its potential zero-crossing points are necessarily (strictly) positive. Towards this end, we need to establish three lemmas first.

Let us begin by stating a simple real analysis property which we prove in Appendix A of our tech. report [34].

Property 1: Let $f(t)$ be a continuous and piecewise continuously differentiable function of t. Assume $f(t_0) > L$. Now

if $f(t_1) = L$ for the first time before t_0, i.e., $f(t_1) = L$ and $f(t) > L$ for all $t \in (t_1 \dots t_0]$, then $\dot{f}(t_1^+) \geq 0$.[6]

Lemma 3: $H = \text{constant} > 0$.

Proof: As we argued in section III, the system is *autonomous*, and thus the Hamiltonian is a constant. Therefore,

$$H = H(T) = f(I(T)) + \kappa\nu(T)I(T). \tag{19}$$

Following lemma 1, $I(T) > 0$; also $\nu(T) = \nu_{\max} > 0$, as we argued after (17). Thus $H(T) > 0$. ∎

Lemma 4: For all $t \in (0 \dots T)$, we have $\lambda_1 > 0$, $\lambda_2 > 0$ and $(\lambda_2 - \lambda_1) > 0$.

Proof: **Step-1.** Following (9), $\lambda_2(T) = (\lambda_2(T) - \lambda_1(T)) = 0$. From the discussion following inequality (18), ν is continuous at T. Thus, from (8) and (9), $\dot{\lambda}_2(T) = (\dot{\lambda}_2(T) - \dot{\lambda}_1(T)) = -f'(I(T)) - \kappa\nu_{\max}$, which is strictly negative due to lemma 1 and the discussion following inequality (18). Also, again from (8) and (9), $\lambda_1(T) = \dot{\lambda}_1(T) = 0$, and by taking the time derivative of (8) and using (9), we obtain $\ddot{\lambda}_1(T) = -\dot{\lambda}_2(T)\beta u_0 I(T) > 0$. Therefore, $\lambda_1(t), \lambda_2(t)$ and $(\lambda_2(t) - \lambda_1(t))$ are strictly positive over an interval of nonzero length towards the end of interval $(0 \dots T)$.

Step-2. Proof by contradiction. Let t^* be the last time at which (at least) one of these three non-negativity constraints is active, i.e., for $t^* < t < T$, we have:

$$\lambda_1(t) > 0, \quad \lambda_2(t) > 0, \quad (\lambda_2(t) - \lambda_1(t)) > 0. \quad \text{and:}$$
$$\lambda_1(t^*) = 0 \quad \text{OR} \quad \lambda_2(t^*) = 0 \quad \text{OR} \quad \lambda_2(t^*) - \lambda_1(t^*) = 0.$$

- Case 1: $\lambda_2(t^*) - \lambda_1(t^*) = 0$ and $\lambda_1(t^*) \geq 0$ and $\lambda_2(t^*) \geq 0$. Now:

$$\begin{aligned} &(\dot{\lambda}_2(t^{*+}) - \dot{\lambda}_1(t^{*+})) \\ &= -f' + \lambda_2 b' - (\lambda_3 - \lambda_2)\nu - \lambda_1 q' \quad [\because (8)] \\ &= -f' + \lambda_2 b' - (\lambda_3 - \lambda_2)\nu - \lambda_1 q' \\ &\quad - \frac{H}{I} + \frac{f}{I} - \frac{\lambda_1 q}{I} - \frac{\lambda_2 b}{I} + \frac{\varphi\nu}{I} \quad [\because (15)] \\ &= \frac{1}{I}[f - f'I] + \frac{\lambda_2}{I}[b'I - b] - \lambda_1 q' - \frac{\lambda_1 q}{I} - \frac{H}{I} \end{aligned} \tag{20}$$

  From lemma 2, $[f - f'I] \leq 0$ and $[b'I - b] \leq 0$. From the definition of t^*, $\lambda_1(t^{*+}) \geq 0$ and $\lambda_2(t^{*+}) \geq 0$. Now following Lemmas 1 and 3 and (20) and properties of $q(S)$, we observe that $[\frac{d}{dt}(\lambda_2 - \lambda_1)]|_{t^{*+}} < 0$. According to property 1, this is a contradiction. Thus, case 1 could not occur.
- Case 2: $\lambda_1(t^*) = 0$, $\lambda_2(t^*) \geq 0$ and $\lambda_2(t^*) - \lambda(t^*) > 0$, Then, from (8), $\dot{\lambda}_1(t^{*+}) = -(\lambda_2 - \lambda_1)\beta u_0 I$. Since in this case $(\lambda_2(t^*) - \lambda_1(t^*)) > 0$, thus $\dot{\lambda}_1(t^{*+}) < 0$ which is in contradiction with property 1. Hence case 2 is also impossible.
- Case 3: $\lambda_1(t^*) > 0$, $\lambda_2(t^*) - \lambda_1(t^*) > 0$ and $\lambda_2(t^*) = 0$. Thence, from (8),$\dot{\lambda}_2(t^{*+}) = -f' - (\lambda_2 - \lambda_1)\beta S - \frac{\varphi\nu}{I}$. For this case $(\lambda_2(t^*) - \lambda_1(t^*)) > 0$. These inequalities along with (18) and lemma 1, show $\dot{\lambda}_2(t^{*+}) < 0$. This is again in contradiction with property 1.

[6]For a general function $f(x)$, the notations $f(x_0^+)$ and $f(x_0^-)$ are defined as $\lim_{x \downarrow x_0} f(x)$ and $\lim_{x \uparrow x_0} f(x)$, respectively.

Therefore, none of the three cases could occur, which is a contradiction with existence of t^*. Hence, follows the lemma. ∎

Here, we state another general property of differentiable functions whose proof can be found in Appendix B of our tech. report [34].

Property 2: Assume $f(t)$ is a continuous and piecewise continuously differentiable function of t. Assume t_1 and t_2 to be its two consecutive L-crossing points, that is, $f(t_1) = f(t_2) = L$ and $f(t) \neq L$ for all $t_1 < t < t_2$. Now if $\dot{f}(t_1^+) \neq 0$ and $\dot{f}(t_2^-) \neq 0$, then $\dot{f}(t_1^+)$ and $\dot{f}(t_2^-)$ must have opposite signs.

Let us calculate the time derivative of the φ function wherever ν is continuous:

$$\begin{aligned} \dot{\varphi} &= (\dot{\lambda}_3 - \dot{\lambda}_2)I + \dot{I}\frac{\varphi}{I} \quad [\because (13)] \\ &= (f' + (\lambda_2 - \lambda_1)\beta u_0 S - \lambda_2 b' \\ &\quad + (\lambda_3 - \lambda_2)\nu)I + \dot{I}\frac{\varphi}{I} \quad [\because (8)] \\ &= f'I + (\lambda_2 - \lambda_1)\beta u_0 IS - \lambda_2 b'I + \varphi\nu + \dot{I}\frac{\varphi}{I} \\ &\quad + (H - f - (\lambda_2 - \lambda_1)\beta u_0 IS + \lambda_1 q \\ &\quad + \lambda_2 b - \varphi\nu) \quad [\because (15)] \\ &= H + \lambda_1 q + (f'I - f) + \lambda_2(b - b'I) + \dot{I}\frac{\varphi}{I}. \end{aligned} \tag{21}$$

Let a time at which $\varphi = 0$ be denoted by τ. From (21) we obtain:

$$\dot{\varphi}(\tau^+) = \dot{\varphi}(\tau^-) = H + \lambda_1 q + (f'I - f) + \lambda_2(b - b'I) \tag{22}$$

Equation (22) and Lemmas 1, 3, 2, 4 show that $\dot{\varphi}(\tau) > 0$.

Firstly, this shows that φ cannot be equal to zero over an interval of nonzero length, since that requires $\dot{\varphi} = 0$ over that interval, which is not possible. Thus, referring to (17), ν is bang-bang, i.e., $\nu \in \{0, \nu_{\max}\}$.

Secondly, referring to property 2, we conclude that φ has at most one zero-crossing point. Note that according to (17), ν can have jump only at zero-crossing points of φ. Now to find the direction of the jump, we note that according to (14), continuity of φ and (17), $\nu = \nu_{\max}$ for an interval of nonzero length towards the end of the $(0 \dots T)$. Thus, the Theorem follows.

∎

V. Dynamic Control of the Scanning Rate/Tx Range

In this section, we assume that the worm has selected a killing rate $\nu_0 \geq 0$ a priori and it is fixed throughout the optimization period and the attacker seeks to determine the optimum $u(.)$.

e) Convex $h(u)$: Recall that both $b(I)$ and $q(S)$ satisfy $b(0) = q(0) = 0$, and $b(I), q(S)$ are increasing functions of I, S for $I, S \in [0 \dots 1]$. Assume further that there exist constants $\hat{b}$ and $\hat{q}$ such that

$$\forall I, S \in [0 \dots 1], \quad b(I) \geq \hat{b}I \text{ and } q(S) \geq \hat{q}S. \tag{23}$$

Now, considering the supremum of such constants, we assume to have:

$$\hat{b} + \hat{q} \geq \beta u_{\max} \tag{24}$$

$\beta u_{\max}$ is the maximum rate of the spread of the infection, and intuitively, the above condition describes the scenario in which the recovery rate (healing + immunization) is larger than the rate of the spread of the infection. We present the structural characteristics of the optimal u in such *fast-healing* regime in Theorem 2. We show that the optimal transmission range times scanning rate of the infective nodes is a non-increasing function of time that necessarily ends at $u_{\min}$.

Theorem 2: Any optimal $u(t)$ that maximizes the worm's damage function in (3) for the case of static killing rate and convex $h(u)$, is constituted of the following *phases*:

1) $u = u_{\max}$ on $0 < t \leq t_0 < T$ for some $t_0 \geq 0$;
2) u strictly and continually decreases on $t_0 < t \leq t_1 < T$ for some $t_1 \geq t_0$;
3) $u = 0$ on $t_1 < t \leq T$.

f) Concave $h(u)$:

Theorem 3: Any optimal $u(t)$ that maximizes the worm's damage function in (3) for the case of static killing rate, and a concave $h(u)$, has the following characteristic: $\exists t_1 \in [0 \dots T)$ such that $u(t) = u_{\max}$ for $0 < t < t_1$ and $u(t) = u_{\min}$ for $t_1 < t < T$.

The proofs of Theorems 2 and 3 are similar to the proof of Theorem 1 and are transferred to our tech. report [28] due to the space constraint.

VI. Numerical Computations

Our numerical computations have been designed to complement our analysis in the previous two sections. We use the insights revealed by these computations in designing robust counter-measures.

We choose $T = 10, I_0 = 0.1, \beta = 0.6, u_{\max} = 1, \nu_{\max} = 1$ $h(u) = u^2$ (which is a convex function) and $C = 5$. We selected C such that $u(t) = u_{\max}$ for all $t \in [0, T]$ violates the constraint of (4b) Also, we assume that $Q(x) = B(x) = \gamma$ for all $x \in [0, 1]$, i.e., $q(S) = \gamma S$ and $b(I) = \gamma I$. The equal rates

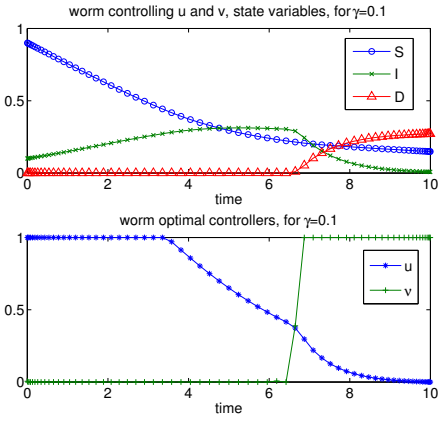

Fig. 2: Evaluation of the optimal controllers and the according states as functions of time. Specific parameters used are $f(I) = \frac{3}{2}I^2$, $\kappa = 4$, $\gamma = 0.1$

are justified if we assume that there is one type of security patch which successfully removes the infection, if any, and immunizes a node against future infection.

Our first observation is that for all the range of parameters that will follow in this section, the structural results of Theorems 1 and 2 for the optimal solution hold, although they were shown assuming that only one of the controllers is dynamic (i.e., only one is allowed to vary as a function of time and the other is chosen as a constant), whereas here, both $u(t)$ and $\nu(t)$ are chosen dynamically by the attacker (i.e., both u, ν are allowed to vary as functions of time). In addition, Theorem 2 was shown assuming a fast healing regime while we observe that the results are valid for cases that are not fast-healing as well. Owing to space constraints, we present only one corroborating figure, Fig. 2, which depicts the optimal controllers as well as the states as functions of time. Henceforth, we continue to consider the case in which the worm dynamically selects both u, ν. This reveals the full damage potential of the worm.

Next, we investigate the effect that changing γ causes on the optimal controllers. According to Fig. 3 and our other computation results, we observe that increasing the recovery rate generally (a) decreases the jump time in the ν; (b) extends the initial period during which $u = u_{\max}$ and makes the subsequent descent in u sharper. Intuitively, these phenomena can be explained in the following manner: In a system with large recovery rate, both the susceptible and infective nodes are recovered rapidly. Hence, the worm should use more of its power resources early on and also starts killing them earlier in order to not lose many nodes to the pool of recovered.

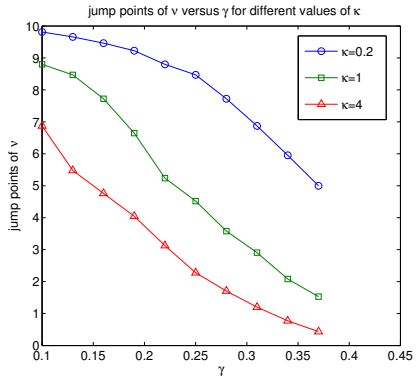

Fig. 3: Jump points of optimal ν versus γ for values of $\kappa = 0.2, 1, 4$. Here, $f(I) = \frac{3}{2}I^2$, and $\gamma = 0.1$.

Finally, we consider the problem of choosing the best parameters from the viewpoint of the system. Specifically, the system chooses the recovery rate a priori for a worst case scenario, which is when the attacker knows the parameters of the system (including the recovery rate) and chooses the optimal dynamic attack policy. As anticipated, our numerical computations reveal that higher the recovery rate (the sum of the immunization and healing rates which is 2γ in this case), the less is the damage due to the attack. For example, Fig. 4 depicts the damage inflicted by the worm versus γ (γ is varied between 0.10 to 0.37) for 3 different examples of damage functions: $f(I) = I$, $f(I) = 3/2I^2$, $f(I) = 2I^3$

and $f(I) = 0.5(e^I - 1)/(e - 2)$. The coefficients in $f(I)$ are chosen such that all of these functions have the same average for I from 0 to 1. But, increasing the recovery rate is achieved through greater usage of costly resources such as bandwidth and power, and thereby inflicts a recovery cost on the system. We consider the overall system cost as the sum of the damage caused by the worm and the expense of providing the immunization and healing rates of γ. The system faces a trade-off in choosing the least-costly recovery rate, which we resolve numerically. In the examples provided in this paper (Fig. 4), we have plotted the overall system cost assuming a simple linear recovery cost induced by γ (specifically 4γ), and the damage functions described above in this paragraph. In each case, the overall cost is minimized at a unique value of γ: $\gamma = 0.34, 0.25, 0.22, 0.19$ in the figures respectively.

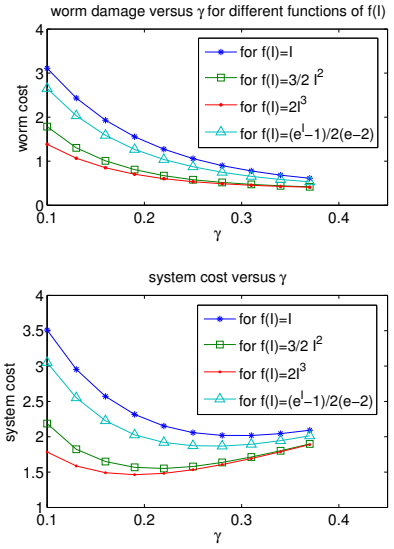

Fig. 4: Worm damage and system cost versus γ for different functions of $f(I)$. Here, $\kappa = 4$.

REFERENCES

[1] C. Karlof and D. Wagner, "Secure routing in wireless sensor networks: Attacks and countermeasures," *Ad Hoc Networks*, vol. 1, no. 2-3, pp. 293–315, 2003.

[2] D. Welch and S. Lathrop, "Wireless security threat taxonomy," in *Information Assurance Workshop, 2003. IEEE Systems, Man and Cybernetics Society*, pp. 76–83, 2003.

[3] A. Herzog, N. Shahmehri, and C. Duma, "An ontology of information security," *International Journal of Information Security and Privacy*, vol. 1, no. 4, pp. 1–23, 2007.

[4] D. Moore, V. Paxson, S. Savage, C. Shannon, S. Staniford, and N. Weaver, "Inside the slammer worm," *IEEE Security & Privacy*, vol. 1, no. 4, pp. 33–39, 2003.

[5] C. Zou, W. Gong, and D. Towsley, "Code red worm propagation modeling and analysis," in *Proceedings of the 9th ACM conference on Computer and communications security*, pp. 138–147, ACM New York, NY, USA, 2002.

[6] E. Filiol, M. Helenius, and S. Zanero, "Open problems in computer virology," *Journal in Computer Virology*, vol. 1, no. 3, pp. 55–66, 2006.

[7] Y. Hu, A. Perrig, and D. Johnson, "Packet leashes: a defense against wormhole attacks in wireless networks," in *IEEE INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications Societies*, vol. 3.

[8] J. Douceur, "The sybil attack," in *Peer-to-Peer Systems, First International Workshop, IPTPS 2002, Cambridge, MA, USA, March 7-8, 2002, Revised Papers*, pp. 251–260.

[9] J. Walters, Z. Liang, W. Shi, and V. Chaudhary, "Wireless sensor network security: A survey," *Security in Distributed, Grid, Mobile, and Pervasive Computing*, p. 367, 2007.

[10] F.-S. C. T. Page, "F-secure virus descriptions : Cih." http://www.f-secure.com/v-descs/cih.shtml.

[11] C. Cowan, C. Pu, D. Maier, H. Hinton, J. Walpole, P. Bakke, S. Beattie, A. Grier, P. Wagle, and Q. Zhang, "StackGuard: Automatic adaptive detection and prevention of buffer-overflow attacks," in *Proceedings of the 7th USENIX Security Conference*, vol. 78, San Antonio: USENIX Press, 1998.

[12] Symantec, "W32.sqlexp.worm," (02.13.2007).

[13] N. Weaver and V. Paxson, "A worst-case worm," in *Proc. Third Annual Workshop on Economics and Information Security (WEIS04)*, 2004.

[14] D. Grass, A. Vienna, J. Caulkins, and P. RAND, "Optimal Control of Nonlinear Processes," 2008.

[15] D. Kirk, *Optimal Control Theory: An Introduction*. Prentice Hall, 1970.

[16] A. Seierstad and K. Sydsaeter, *Optimal control theory with economic applications*. 1986.

[17] 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, Rio de Janeiro, Brazil*, 2009.

[18] B. Sun, G. Yan, Y. Xiao, and T. Andrew Yang, "Self-propagating mal-packets in wireless sensor networks: Dynamics and defense implications," *Ad Hoc Networks*, 2009.

[19] C. Fleizach, M. Liljenstam, P. Johansson, G. Voelker, and A. Mehes, "Can you infect me now?: malware propagation in mobile phone networks," in *Proceedings of the 2007 ACM workshop on Recurring malcode*, pp. 61–68, ACM New York, NY, USA, 2007.

[20] A. Bose, *Propagation, Detection and Containment of Mobile Malware*. PhD thesis, The University of Michigan, 2008.

[21] A. El Fawal, J. Le Boudec, and K. Salamatian, "Vulnerabilities in epidemic forwarding," in *IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks, 2007. WoWMoM 2007*, pp. 1–6, 2007.

[22] D. Daley and J. Gani, *Epidemic modelling: an introduction*. Cambridge Univ Pr, 2001.

[23] J. Kephart, S. White, I. Center, and Y. Heights, "Directed-graph epidemiological models of computer viruses," in *Research in Security and Privacy, 1991. Proceedings., 1991 IEEE Computer Society Symposium on*, pp. 343–359, 1991.

[24] C. Zou, W. Gong, and D. Towsley, "Worm propagation modeling and analysis under dynamic quarantine defense," in *Proceedings of the 2003 ACM workshop on rapid malcode*, pp. 51–60, ACM New York, NY, USA, 2003.

[25] S. Tanachaiwiwat and H. A., "Encounter-based worms: Analysis and defense," *Ad Hoc Networks, Elsevier JOURNAL*, 2009.

[26] V. Karyotis and S. Papavassiliou, "Risk-based attack strategies for mobile ad hoc networks under probabilistic attack modeling framework," *Computer Networks*, vol. 51, no. 9, pp. 2397–2410, 2007.

[27] X. Yan and Y. Zou, "Optimal Internet Worm Treatment Strategy Based on the Two-Factor Model," *ETRI JOURNAL*, vol. 30, no. 1, p. 81, 2008.

[28] M. Khouzani, E. Altman, and S. Sarkar, "Optimal Quarantining of Wireless Malware Through Power Control," in *Proceedings of the Fourth Symposium on Information Theory and Applications*, University of California at San Diego, 2009.

[29] C. Bettstetter, "Mobility modeling in wireless networks: categorization, smooth movement, and border effects," *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 5, no. 3, pp. 55–66, 2001.

[30] R. Groenevelt, P. Nain, and G. Koole, "The message delay in mobile ad hoc networks," *Performance Evaluation*, vol. 62, no. 1-4, pp. 210–228, 2005.

[31] T. Kurtz, "Solutions of ordinary differential equations as limits of pure jump Markov processes," *Journal of Applied Probability*, pp. 49–58, 1970.

[32] R. Cole, "Initial Studies on Worm Propagation in MANETS for Future Army Combat Systems," 2004.

[33] S. Tanachaiwiwat and A. Helmy, "VACCINE: War of the worms in wired and wireless networks," in *IEEE INFOCOM*, pp. 05–859, 2006.

[34] M. Khouzani, S. Sarkar, and E. Altman, "Maximum Damage Malware Attack in Mobile Wireless Networks," tech. rep., http://www.seas.upenn.edu/ swati/publication.htm, 2009.

[35] M. Hirsch and S. Smale, *Differential equations, dynamical systems, and linear algebra*. Academic Press Inc, 1974.