Stochastic Behavior of Random Constant Scanning Worms

Kurt R. Rohloff and Tamer Başar

***Abstract*— This paper discusses modeling and simulation issues associated with the stochastic behavior of a special type of a computer worm called a Random Constant Scanning (RCS) worm. Although these worms propagate by randomly scanning network addresses to find hosts that are susceptible to infection, traditional RCS worm models are fundamentally deterministic. A density-dependent Markov jump process model for RCS worms is presented and analyzed. Conditions are shown for when worm models can safely ignore some stochastic properties of RCS worm propagation. A computationally simple hybrid deterministic/stochastic model for the observed scanning behavior on a local network due to the global propagation of an RCS scanning worm is also presented and discussed.**

I. Introduction

A computer worm is a piece of malicious code that can spread automatically over a computer network without the need for human intervention. Due to this automatic propagation, worms can potentially spread on the Internet with staggering speed and cause damage on the order of billions of dollars [5, 6]. A special type of a worm, called a Random Constant Scanning worm (RCS worm for short) propagates by continually scanning randomly selected Internet addresses in attempts to infect other hosts. When an Internet address has been selected for scanning, the infected host attempts to transmit infectious packets to a host at the selected address. If the targeted host is susceptible to the infection, then, upon receiving the infectious packets, that host becomes infected and the scanning and infection process continues at both hosts. This branching behavior of the worm infection process causes the fantastic propagation speeds observed during attacks from RCS worms such as CodeRed1v2 and Slammer.

From data collected from previous worm attacks it has been found that the deterministic simple epidemic model ([4]) can effectively capture aspects of the behavior of an RCS worm epidemic's propagation [5, 6, 10]. However, the underlying propagation behavior of RCS worms is fundamentally stochastic in nature. It has even been noted in the literature that due to the random nature in which an RCS worm spreads, there could be variability between the overall propagation rates of RCS worm epidemics for worms with similar propagation properties [7, 12]. To the best knowledge of the authors there has been no work analyzing the stochastic properties of RCS worm epidemics despite their inherently stochastic behavior.

This paper presents an idealized stochastic propagation model for RCS worms taken from the literature of epidemiology and public health [1]. The large-scale propagation behavior of an RCS worm predicted by this model is compared to the large-scale behavior predicted by the standard deterministic simple epidemic model.

The deterministic simple epidemic model has been widely used in the literature as a basis for developing worm detection methods [9, 11–13]. It is the hope of the authors that from the analyses of the stochastic nature of RCS worm propagations discussed in this paper, more effective automatic detection methods for these worms might be developed. Related to this, conditions are also shown in this paper for when worm models can safely ignore some stochastic properties of RCS worm propagation. A hybrid deterministic/stochastic point process model for the observed scanning behavior on a local network due to the global propagation of an RCS scanning worm is also presented. Such a model has not been previously discussed in the literature.

The paper is organized as follows. Section II establishes the notation used throughout the paper and presents the well-known deterministic simple epidemic model. A density-dependent Markov jump model for worm propagation is introduced in Section III. Section IV presents a hybrid deterministic/stochastic point process model for a worm's scanning behavior as observed on a local network. The paper concludes with a discussion of the results and possible areas for future research in Section V.

This research was supported by the NSF grant CCR 00-85917 ITR.

K. Rohloff and T. Başar are with the Coordinated Science Laboratory, The University of Illinois, 1308 West Main St., Urbana, IL 61801, USA {krohloff,tbasar}@control.csl.uiuc.edu

II. Random Constant Scanning Worm Propagation

The underlying properties and notations for the modeling of RCS worm epidemics are now presented. It is assumed that in a given network (such as the Internet) there are n unique host addresses. Of these addresses, $n_s \leq n$ hosts could potentially become infected by the worm. The set of potential hosts is split into *infected* and *susceptible* subpopulations, where $I(t)$ is the number of hosts which are infected by the worm at time t, and $S(t)$ is the number of hosts which could become infected, but are not at the time. Note that due to the random scanning propagation behavior of RCS worms, $S(t)$ and $I(t)$ are inherently random variables. At time 0, $(S(0),I(0)) = (s_0,i_0)$ where $i_0 \geq 1$ is the initial infected population. Due to the fast dynamics of the epidemics modeled in this paper and for general simplicity in the model, it is assumed that infectious hosts are not removed from the general population, so that for all $t \geq 0$, $I(t)+S(t) = n_s$. However, the results of this paper can be generalized to the cases of host recovery and/or removal.

When an infected host attempts to spread the worm, the addresses selected for scanning are assumed to be selected with a uniform distribution for mathematical simplicity, but this assumption can be removed. Therefore, any one infectious packet sent at time t has probability $\frac{S(t)}{n}$ of being sent to a susceptible host. It is assumed that an infected host scans for susceptible hosts at a constant rate β which is called the *infection parameter*. Therefore $\frac{\beta}{n}S(t)$ is the rate at which an infected host transmits its infection to susceptible hosts and $\frac{\beta}{n}S(t)I(t)$ is the rate at which the infected population transmits the infection to susceptible hosts.

In order to motivate the discussion of properties of the stochastic epidemic models studied in this paper, the deterministic simple epidemic model is now presented where the variable $s(t)$ is used to represent the size of the susceptible population at time t, and the variable $i(t)$ is a similarly defined deterministic variable which represents the number of infected individuals at time t. Consequently, $\frac{\beta}{n}s(t)i(t)$ is the rate at which the $i(t)$ infected hosts propagate the epidemic, and

$$\frac{di}{dt} = \frac{\beta}{n}s(t)i(t) = \frac{\beta}{n}(n_s - i(t))i(t).$$

With $i(0) = i_0$,

$$i(t) = \frac{i_0 n_s}{i_0 + (n_s - i_0)e^{-\beta\frac{n_s}{n}t}}.$$

A plot of $i(t)$ vs. t, called the *infection curve*, for a model of the CodeRed1v2 epidemic can be seen in Figure 1 where it is assumed that $n = 2^{32}$ (the IP address space), $n_s = 350,000$ (an approximation of the size of the susceptible CodeRed population), $\beta = 10188$ (an approximation of the number of IP addresses scanned by an infected host scans per hour) and $i_0 = 1$ (the size of the initial infection).

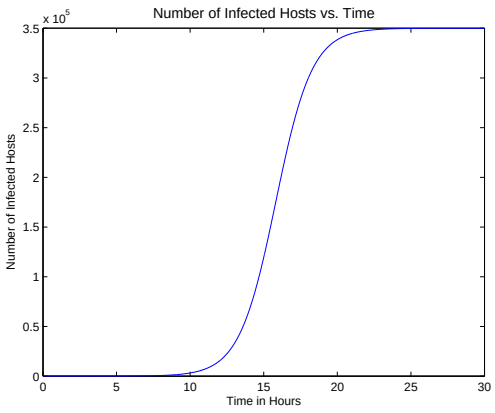

Fig. 1. Plot of a Deterministic CodeRed1v2 Propagation Simulation.

III. Stochastic Epidemiological Model for Scanning Worms

We now introduce a stochastic density-dependent Markov jump process propagation model for an RCS worm drawn from the field of epidemiology [1, 2]. When $(S(t),I(t)) = (s,i)$, the pair (s,i) can be thought of as the "state" of the epidemic. Note that due to the infection propagation, if the propagation process is at a state (s,i), then the next state must be $(s-1,i+1)$ and the next state after that must be $(s-2,i+2)$ and so on until state $(0,n_s)$ is reached. From $(0,n_s)$ no other state can be reached, so $(0,n_s)$ is an absorbing state and almost surely a time t^{fin} is eventually reached such that $(S(t^{fin}),I(t^{fin})) = (0,n_s)$. $\frac{\beta}{n}S(t)I(t)$ is the rate at which $(S(t),I(t)) = (s,i)$ goes to $(s-1,i+1)$. This process can be modeled as a jump process with a jump intensity:

$$q_{(s_a,i_a)(s_b,i_b)} = \begin{cases} \frac{\beta}{n}s_a i_a & \text{if } (s_b = s_a - 1) \wedge (i_b = i_a + 1) \\ 0 & \text{otherwise} \end{cases}.$$

This jump process is Markovian because at state $(S(t),I(t)) = (s,i)$, the current jump intensity depends only on (s,i) and is independent of the

previous states of the process. Consequently, this stochastic epidemic propagation process is by definition a *density-dependent Markov jump process* because the jump intensity at a state (s,i) depends on the "densities" of the number of susceptible hosts s and the number of infected hosts i. Several important aspects of this subclass of Markov jump processes are discussed in [1, 3].

Five simulations of a stochastic worm propagation model with growth parameters similar to that of the CodeRed1v2 worm and an initial infection of $i_0 = 1$ can be seen in Figure 2. Note that by visual inspection, the propagation curves in Figure 2 are approximately the same curve shifted in time.

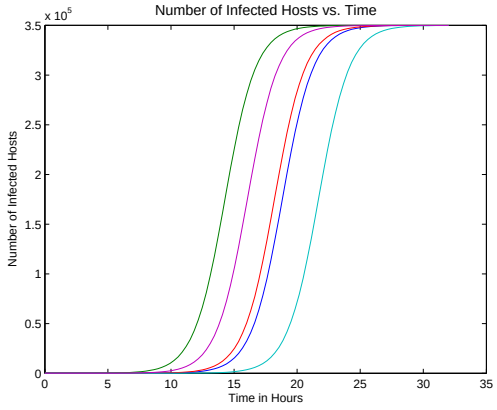

Fig. 2. Plot of Five Stochastic CodeRed1v2 Propagation Simulations.

It has been shown in [1] that the expected values of the susceptible and infected population sizes in the stochastic model at time t, $(E\{S(t)\}, E\{I(t)\})$, converge almost surely to the susceptible and infected population sizes predicted by the deterministic model $(s(t), i(t))$ as the size of the populations n_s and n increase. Also, the fluctuations of the susceptible and infected population sizes in the stochastic model around the deterministic solution are asymptotically Gaussian. However, as discussed in [1], it is difficult if not impossible to find a closed-form expression for the covariance of a density-dependent Markov jump process. Despite this, the covariances of worm propagation at various time intervals can be easily computed through simulation. The mean of 100 such CodeRed1v2 epidemic propagation simulations and the variance of these simulations at various instances of time can be seen in Figure 3 where initially one host is infected.

The plots in Figure 3 show that $(E\{S(t)\}, E\{I(t)\})$ is approximately the same as

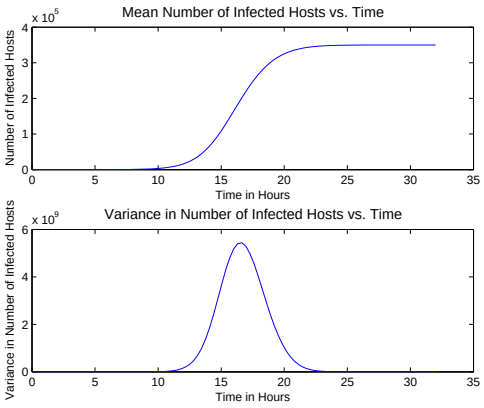

Fig. 3. Plots of the Mean and Variance of 100 Stochastic CodeRed1v2 Propagation Simulations.

the deterministic simulation $(s(t), i(t))$ as seen in Figure 1, but the variance of the stochastic epidemic simulations is potentially very large. With this in mind, consider the first plot of Figure 4 which shows five stochastic simulations of the propagation of a worm with growth parameters similar to that of the CodeRed1v2 worm where initially half of the susceptible population is infected. The variances computed from 100 such simulations can be seen in the second plot.

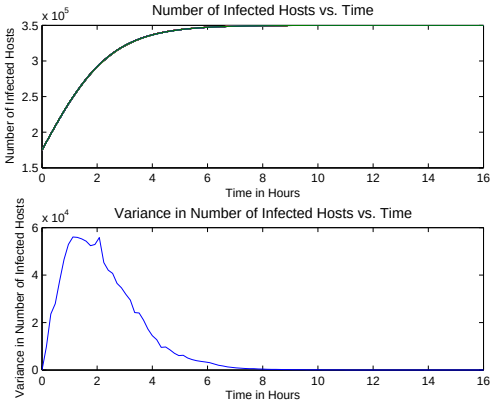

Fig. 4. Plots of 5 Stochastic CodeRed1v2 Propagation Simulations and the Variance of 100 Stochastic CodeRed1v2 Propagation Simulations.

As can be seen from the first plot of Figure 4, the various simulations of the epidemic propagations are effectively identical. This is also indicated in the variance plot of Figure 4. The maximum variance of these simulations where initially half of the susceptible population is infected is several orders of magnitude less than the maximum variance of

the simulations where the initial infection is one as in Figure 3. Although the propagation curves seen in Figure 2 are nearly identical, there can be differences on the order of hours in the amount of time it takes a worm epidemic to infect half of the susceptible population.

Define T_{i_j} to be a random variable which represents the amount of time the infection process is in state (s_j, i_j). T_{i_j} is exponentially distributed with mean $\frac{n}{\beta(n_s - i_j)i_j}$ and variance $(\frac{n}{\beta(n_s-i_j)i_j})^2$. Note that $\max_{i_j}(Var(T_{i_j})) = \left(\frac{n}{\beta}\right)^2 \left(\frac{1}{(n_s-1)^2}\right)$ and $\min_{i_j}(Var(T_{i_j})) = \left(\frac{n}{\beta}\right)^2 \left(\frac{4}{n_s^4}\right)$.

Let $i_a, i_b \in \{1, \ldots, n_s\}$ be such that $i_b > i_a$ and let $T_{i_a i_b}$ be a random variable that represents the amount of time it takes a stochastic infection propagation process to go from state (s_a, i_a) to state (s_b, i_b). By definition, $T_{i_a i_b} = \sum_{i_j=i_a}^{i_b-1} T_{i_j}$. $T_{i_a i_b}$ is not necessarily exponentially distributed, but a closed-form expression for its probability distribution function exists. Relevant to the discussions in this paper,

$$E\{T_{i_a i_b}\} = \sum_{i_j=i_a}^{i_b-1} E\{T_{i_j}\} = \frac{n}{\beta} \sum_{i_j=i_a}^{i_b-1} \frac{1}{(n_s - i_j)i_j}$$

and

$$\begin{aligned} Var(T_{i_a i_b}) &= \sum_{i_j=i_a}^{i_b-1} Var(T_{i_j}) \\ &= \left(\frac{n}{\beta}\right)^2 \sum_{i_j=i_a}^{i_b-1} \left(\frac{1}{(n_s - i_j)i_j}\right)^2. \end{aligned}$$

Although it is not demonstrated here for the sake of brevity, it can be shown that $E\{T_{1\frac{n_s}{2}}\} = \frac{n}{\beta n_s}[C + \ln(n_s - 1) + f(n)]$ where $C \neq 0$ is a constant and $f(n) \in O(1/n)$. Note that this closed-form expression for $E\{T_{1\frac{n_s}{2}}\}$ is not equal to the expression $\frac{n}{\beta n_s}\ln(n_s - 1)$, the same value predicted by the deterministic model in Section II. More quantitatively, for the epidemic parameters used to generate the simulations of the CodeRed1v2 worm in the plots shown above, $E\{T_{1\frac{n_s}{2}}\}$ differs from $\frac{n}{\beta n_s}\ln\frac{n_s-1}{1}$ by over half an hour. This is due to the fact that both the deterministic and stochastic models are different abstractions of the same underlying process where the deterministic model assumes a continuous state process, which is not a completely accurate reflection of the underlying behavior of a worm's propagation. However, it can be shown through the application of l'Hôpital's rule that as n_s increases, then $E\{T_{1\frac{n_s}{2}}\}$ approaches $\frac{n}{\beta n_s}\ln(n_s - 1)$.

Now consider Figure 5, which shows plots of $E\{T_{i_j}\}$ vs. i_j and $Var(T_{i_j})$ vs. i_j on semilog scales. Note that both of the plots in Figure 5 are bowl-shaped. Note also that the bowl-shaped curve of $Var(T_{i_j})$ vs. i_j in Figure 5 has a relatively flat bottom with steep sides. At initialization, when $I(t)$ is close to 0, there is potentially large variation in the evolution of $I(t)$ as indicated in Figure 5 (hence the observed time-shifting.) However, an epidemic's growth pattern is relatively stable once the epidemic has been established on a large enough portion of the susceptible population and there is still a relatively large number of susceptible hosts left to infect.

For an RCS worm epidemic to be a threat to the Internet, n_s, the size of the susceptible population, should be relatively large (on the order of several thousand or more hosts). If this holds then n_s^4 will be very large compared to $(n_s - 1)^2$ and hence $\min_{i_j}(Var(T_{i_j}))$ will be much smaller than $\max_{i_j}(Var(T_{i_j}))$. This indicates that $Var_{T_{i_a i_b}}$ will be relatively small when $0 << i_a$ and $i_b << n_s$ compared to $Var_{T_{i_a i_b}}$ when i_a is close to 0 or i_b is close to n_s. This explains why the propagation simulation curves in Figure 2 appear to be the same curve shifted in time.

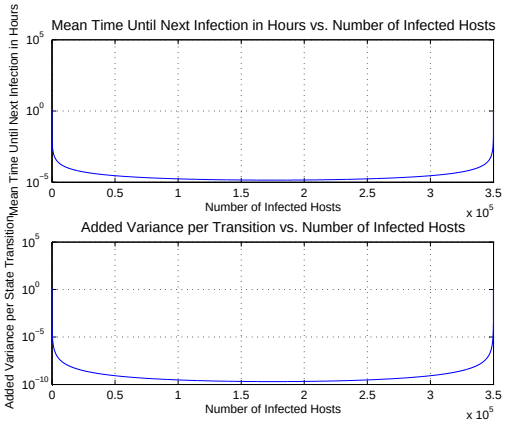

Fig. 5. Plot of the mean and variance of the inter-state jump intervals for a CodeRed1v2 worm.

This analysis of $Var_{T_{i_a i_b}}$ indicates that excluding the early and late growth stages of an epidemic and the "time-shifting" effects due to the probability distribution on $T_{i_a i_b}$, simulations of a worm's propagation using the deterministic and stochastic models are effectively equivalent when n_s is sufficiently large. Therefore, the possible variability in the stochastic propagation of RCS worm epidemics mentioned in [7, 12] is surprisingly minor if these effects can be ignored. We believe that this distinction for when the deterministic model can be used in place of the stochastic model has not been

previously noted in the Internet worm literature.

IV. Observed Local Scanning Model

We now introduce a new hybrid model for the scanning behavior of a global RCS worm epidemic observed on a local network. Let random variables $\tau_1, \tau_2, \tau_3, \ldots$ represent the times at which scans due to the worm are observed on the local network. For computational simplicity this hybrid model uses the deterministic simple epidemic model to describe the large-scale global propagation behavior, but a stochastic model to generate $\tau_1, \tau_2, \tau_3, \ldots$ based on the current state of the large-scale propagation. Below it is described how and when this particular combination of deterministic and stochastic behavior can be justified as an accurate representation of the observed worm scanning behavior.

Let n_t be the number of unique addresses in the local network. Because it is assumed that the addresses selected for scanning by infected hosts are assumed to be selected with a uniform distribution, every scanning attempt by an infected host during the epidemic is a Bernoulli trial with probability n_t/n of successfully scanning a host in the set of n_t local addresses. Let $P_1, P_2, P_3, \ldots$ be random variables such that for $i \in \{1, 2, \ldots\}$, P_i represents the global number of scans which have occurred due to the worm up until the time of the ith successful scan of the local network. By definition, $P_1, P_2 - P_1, P_3 - P_2, \ldots$ are independent and geometrically distributed random variables with equal mean n/n_t. Because n_t is small compared to n, $P_1, P_2 - P_1, P_3 - P_2, \ldots$ can be approximated as exponentially distributed with mean n/n_t.

It is now shown how $P_1, P_2 - P_1, P_3 - P_2, \ldots$ can be used to find $\tau_1, \tau_2, \tau_3, \ldots$, the times at which scans due to the worm are observed on the local network. Let $P(t)$ be the random variable which represents the total number of infection attempts that have been made due to the epidemic up to time t using the stochastic model, and let $p(t)$ be its deterministic approximation. Then,

$$\frac{dp}{dt} = \beta i.$$

It is already known from the dynamics of the deterministic epidemic model that

$$\frac{di}{dt} = \frac{\beta}{n} si$$

where $s = n_s - i$. Therefore, by basic mathematical manipulation,

$$\frac{n}{(n_s - i)} di = dp.$$

Consequently,

$$\int_{i_0}^{i(p)} \frac{n}{(n_s - i)} di = \int_0^p dp.$$

From here it is easy to show that

$$i(p) = n_s - \frac{n_s - i_0}{e^{p/n}}.$$

Therefore,

$$\frac{dt}{dp} = \frac{1}{\beta \left[n_s - (n_s - i_0) e^{-p/n} \right]}$$

and

$$t(p) = \frac{n}{\beta n_s} \left[\frac{p}{n} + \log \left(\frac{n_s - (n_s - i_0) e^{-p/n}}{i_0} \right) \right].$$

If p scans have occurred due to the worm globally, then $t(p)$ represents the time at which the pth scan occurred under the assumption of epidemic propagation dynamics modeled by the simple epidemic model.

With $t(p)$, a model can now be presented to generate the times $\tau_1, \tau_2, \tau_3, \ldots$ at which scans due to the worm are observed on the local network. Recall that $P_1, P_2 - P_1, P_3 - P_2, \ldots$ are independently exponentially distributed and can therefore be generated using standard methods. Therefore, values for $P_1, P_2 - P_1, P_3 - P_2, \ldots$ can be used to generate $P_1, P_2, P_3, \ldots$ and then, using the $t(p)$ function, $\tau_1 = t(P_1), \tau_2 = t(P_2), \tau_3 = t(P_3), \ldots$. An example of $\tau_1, \tau_2, \tau_3, \ldots$ values generated for the CodeRed1v2 model with $n_t = 15$ can be seen in Figure 6.

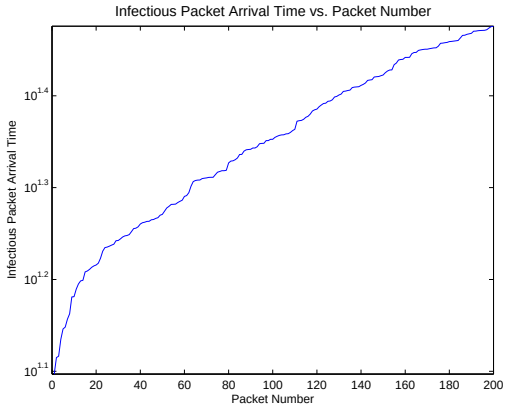

Fig. 6. Plot of Infectious Packet Arrival Times for a CodeRed1v2 worm with $n_t = 15$.

We describe next how this hybrid model can be justified as an accurate representation of the locally observed behavior due to a worm epidemic when

the variability in the initial propagation of a worm described in Section III can be ignored.

If it is assumed that $n_s >> i_0$, then

$$i(p) = n_s\left(1 - \frac{1}{e^{p/n}}\right).$$

It was shown above that $E\{P_1\} = n/n_t$. This therefore shows that

$$i(E\{P_1\}) \quad = \quad n_s\left(1 - \frac{1}{e^{1/n_t}}\right).$$

Note that $T_{i(E\{P_1\})}$ is the amount of time it would take a stochastic epidemic process to transition to the next state from $i(E\{P_1\})$, the size of the global infection when the first infectious packet is expected to be observed locally. Substituting $i(E\{P_1\})$ into the equation for $Var(T_{i_j})$,

$$Var(T_{i(E\{P_1\})}) = \left(\frac{n}{\beta}\right)^2 \frac{1}{n_s^4\left(e^{-1/n_t}(1-e^{-1/n_t})\right)^2}$$

Recall that $Var(T_{i(E\{P_1\})})$ represents the variance in the amount of time it takes the stochastic epidemic propagation process to jump from state $i(E\{P_1\})$ to $i(E\{P_1\})+1$. Therefore, $Var(T_{i(E\{P_1\})})$ gives an indication of the stability of the epidemic's propagation when the first scan due to the worm is observed on the local network. If $Var(T_{i(E\{P_1\})})$ is relatively small, then the epidemic has been sufficiently established on the global network such that the deterministic simple epidemic model will accurately represent its future large-scale growth. Recall that $\max_{i_j}(Var(T_{i_j})) = \left(\frac{n}{\beta}\right)^2\left(\frac{1}{(n_s-1)^2}\right)$ and $\min_{i_j}(Var(T_{i_j})) = \left(\frac{n}{\beta}\right)^2\left(\frac{4}{n_s^4}\right)$. Therefore, if n_s is relatively large or n_t is relatively small, as is generally assumed, then $Var(T_{i(E\{P_1\})})$ is on the same order of magnitude as $\min_{i_j}(Var(T_{i_j}))$. Hence, if n_s is relatively large or n_t is relatively small, then the first infectious packet is expected to be sent to an addresses on the local network when the underlying stochastic worm epidemic process is in a relatively stable region of growth with little variance in the growth dynamics. This also means that there is unlikely to be any observations of scanning behavior on the local network when the deterministic model is not as good a representation of the worm's growth dynamics. Therefore, if the exact starting time of the worm propagation is not known, then the deterministic epidemic model is an effective representation of the stochastic epidemic process under these conditions, and the hybrid observation model would likewise be an effective representation of the local observation process $\tau_1, \tau_2, \tau_3, \ldots$ when n_t is small and n_s is large.

V. Discussion

This paper has discussed issues associated with the idealized stochastic properties of RCS worm epidemics. A density-dependent Markov jump process model for the large-scale propagation behavior of these worms has been presented, which is something that has not previously been covered in the worm literature. Several commonly satisfied conditions are presented, under which the variability in the stochastic propagation of RCS worm epidemics predicted in [7, 12] can be ignored. A hybrid deterministic/stochastic model for the observations of a worm's scanning behavior on a local network has also been presented and discussed.

It is the hope of the authors that the worm modelling methods contained herein will aid in the development of better automated worm detection methods. For instance, the hybird worm propagation model presented in this paper is used in [8] to analyze the optimal RCS worm detection methods under idealized conditions.

References

[1] H. Andersson and T. Britton. *Stochastic Epidemic Models and Their Statistical Analysis*. Number 151 in Lecture Notes in Statistics. Springer-Verlag, New York, 2000.

[2] D.J. Daley and J. Gani. *Epidemic Modelling: An Introduction*. Cambridge University Press, Cambridge, 1999.

[3] S.N. Ethier and T.G. Kurtz. *Markov Processes, Characterization and Convergence*. Wiley Series in Probability and Mathematical Statistics. John Wiley and Sons, New York, 1986.

[4] W.O. Kermack and A.G. McKendrick. A contribution to the mathematical theory of epidemics. *Royal Society of London Proceedings Series A*, 115:700–721, August 1927.

[5] D. Moore, V. Paxson, S. Savage, C. Shannon, S. Staniford, and N. Weaver. Inside the slammer worm. *IEEE Security and Privacy*, 1(4):33–39, 2003.

[6] D. Moore, C. Shannon, and J. Brown. Code-red: A case study on the spread and victims of an Internet worm. In *Proceedings of the Internet Measurement Workshop (IMW)*, 2002.

[7] D. Moore, C. Shannon, G.M. Voelker, and S. Savage. Internet quarantine: Requirements for containing self-propagating code. In *INFOCOM*, 2003.

[8] K. Rohloff and T. Başar. The detection of RCS worm epidemics. Preprint.

[9] S. Staniford. Containment of scanning worms in enterprise networks, 2003. Silicon Defense White Paper.

[10] S. Staniford, V. Paxson, and N. Weaver. How to 0wn the Internet in your spare time. In *Proceedings of the 11th USENIX Security Symposium (Security '02)*, 2002.

[11] C. Wong, C. Wang, D. Song, S. Bielski, and G.R. Granger. Dynamic quarantine of Internet worms. In *Proceedings of the International Conference on Dependable Systems and Networks (DSN-2004)*, 2004.

[12] C.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*, pages 51–60. ACM Press, 2003.

[13] C.C. Zou, D. Towsley, and W. Gong. A fi rewall network system for worm defense in enterprise networks. Technical Report TR-04-CSE-01, Department of Computer Science and Engineering, University of Massachusetts, 2004.