

# A Topologically-Aware Worm Propagation Model for Wireless Sensor Networks<sup>\*</sup>

Syed A. Khayam and Hayder Radha  
Department of Electrical & Computer Engineering / 2120 Engineering Building  
Michigan State University  
East Lansing, MI 48824 USA  
{khayamsy, radha}@egr.msu.edu

## Abstract

*Internet worms have repeatedly revealed the susceptibility of network hosts to malicious intrusions. Recent studies have proposed to employ the underlying principles of worm propagation to disseminate security-critical information in a network. Wireless sensor networks can benefit from a thorough understanding of worm propagation over sensor networks to defend from worms and to efficiently disseminate security-critical information. In this paper, we develop a topologically-aware worm propagation model (TWPM) for wireless sensor networks. In addition to simultaneously capturing both time and space propagation dynamics, the TWPM also incorporates physical, MAC and network layer considerations of practical sensor networks. Simulation results show that the proposed model follows actual propagation dynamics quite closely.*

## 1. Introduction

Computer worms have recently emerged as one of the most imminent and effective threats against information confidentiality, integrity and service availability. Internet worms have repeatedly revealed the susceptibility of Internet hosts to malicious intrusions by compromising millions of vulnerable hosts at an extremely fast pace [1]–[3]. Some recent studies (see for example [3]) have shown that *anti-worms* (also referred to as *good worms*) can serve as an effective counterattack tool by spreading disinfection and immunization information in the same way as malicious worms. The anti-worms can hence be employed to dissem-

inate security-critical information at a very fast pace.

An accurate propagation model is instrumental to understand the automated spread of a worm [1]–[3]. Worm propagation models also facilitate the design of real-time detection strategies. The battery-constrained, time-critical and military-oriented natures of many sensor networks necessitate a robust security framework. Design of secure sensor networks should therefore consider real-time monitoring, detection and mitigation of malicious worms<sup>1</sup>. Further, good worms can efficiently disseminate security-critical information to the nodes in a sensor network. An accurate model is necessary to characterize and evaluate propagation of worms over sensor networks. In this paper, we analytically derive a novel and accurate worm propagation model for wireless sensor networks. The proposed model, referred to as a *topologically-aware*<sup>2</sup> *worm propagation model* (TWPM), simultaneously captures the time and space dynamics of worms spreading over a sensor network.

In Section 2, we define worm propagation characteristics that are specific to sensor networks. We parameterize the effects of physical channel conditions, medium access control (MAC) layer contention, network layer routing and transport layer protocol on worm propagation in sensor networks. Section 3 incorporates these parameters in the TWPM which borrows its basic formulation from models of epidemic diseases [5]. The advanced model parameters and the mathematical treatment following the formulation are then developed specifically for sensor networks. The basic model formulation results in a partial differential equation which is solved in the frequency domain to yield a closed-form solution for the TWPM. For performance evaluation, in Section 4 we simulate the spread of a worm over a sensor

<sup>\*</sup>This material is based upon work supported by the National Science Foundation under NSF CyberTrust Grant No. 0430436. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

<sup>1</sup>This fact motivated the authors' prior work [4] which investigated the spread of active worms over vehicular ad hoc networks.

<sup>2</sup>The term *topologically-aware scanning* was introduced by Staniford *et al.* in their seminal study to refer to worms "which use information available on a victim's machine to select new targets" [1].

network. The simulated and TWPM-predicted worm propagation dynamics are then compared to evaluate the accuracy of the model. We show that TWPM predicts the worm propagation dynamics very accurately. Section 5 summarizes key conclusions of this paper.

## 2. Sensor Network Propagation Parameters and Assumptions

In this section, we define new worm propagation parameters which arise due to the inherent attributes of a sensor network. Further, in this section we also state the assumptions made in this work.

### 2.1. System Description

We consider a network composed of  $N$  stationary and identical sensors which are placed on a rectangular (two-dimensional) grid. The sensors are equipped with omnidirectional antennas which have a maximum transmission range of  $r$  meters. The horizontal and vertical axes are represented by  $\xi$  and  $\eta$ , respectively. To simplify analysis, we uniformly (and logically) sample both axes and treat  $\xi$  and  $\eta$  as discrete variables. Each discrete  $(\xi, \eta)$  position is referred to as a *segment*. Let  $l$  and  $h$  denote the length and height of a segment, respectively. Here, it should be emphasized that this segmentation is only logical with just one constraint:  $r \ll l \times h$ . Due to this constraint, a sensor in segment  $(\xi, \eta)$  can receive traffic from sensors in segment  $(\xi, \eta)$  or (at maximum) from sensors in neighboring segments of  $(\xi, \eta)$ . The neighboring segments of segment  $(\xi, \eta)$  are shown in Figure 1.

The distribution of sensors on the grid is governed by a two-dimensional, discrete-time random process  $\mathfrak{S}(\xi, \eta)$ . Each constituent random variable  $D(\xi, \eta)$  of  $\mathfrak{S}(\xi, \eta)$  describes the *number of sensors* in the  $(\xi, \eta)$  segment. The random variables of  $\mathfrak{S}(\xi, \eta)$  are assumed to be independent and identically distributed (IID). Thus on-average we have

$$\begin{aligned} E\{D(\xi, \eta)\} &= E\{D\} \\ &= \text{average number of nodes per segment.} \end{aligned} \quad (1)$$

We assume that the sensors in a segment are distributed (located) uniformly within the boundaries of the segment.

Figure 1 outlines that due to the  $r \ll l \times h$  constraint, sensors on the edges of a segment can communicate with sensor in parts of the neighboring segments. In essence, a sensor in segment  $(\xi, \eta)$  can only communicate with sensors inside the thick broken-line of Figure 1. For instance, the sensor located at the corner of the  $(\xi, \eta)$  segment can at most send/receive traffic to/from sensors within its transmission range as represented by the circle in Figure 1.

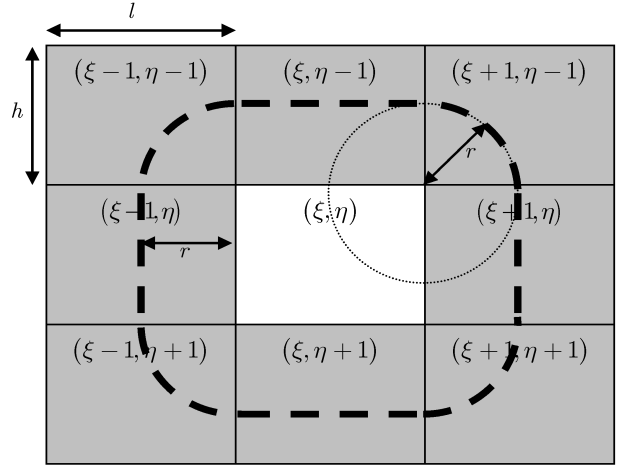


Figure 1. Neighbors of segment  $(\xi, \eta)$ .

### 2.2. Physical Layer Parameters

In order to simultaneously capture distance and fading based attenuations in the wireless medium, we employ a log-normal shadow fading model [6] to define the probability that a transmitted packet is successfully received in a sensor network. Previous studies (see for example [7] and [8]) have illustrated the efficacy of this channel model in defining ad hoc network topologies. We assume that channel conditions do not change drastically during transmission of a given infectious packet. The conditions can, however, change for different packets.

A packet transmission between two nodes  $u$  and  $v$  at a distance  $d(u, v)$  from each other is successful if the received signal power,  $p_r$ , is greater than or equal to a certain threshold power  $p_{r,th}$ . In other words, given a receiver sensitivity,  $p_r \geq p_{r,th}$ , a packet transmission from  $u$  to  $v$  is successful if the signal attenuation between  $u$  and  $v$  is constrained as  $\beta(u, v) \leq \beta_{th}$ , where the threshold attenuation is  $\beta_{th} = 10 \log \left( \frac{p_t}{p_{r,th}} \right)$  dB. The probability of a successful packet transmission between sensors  $u$  and  $v$  in the presence of shadow fading can then be expressed as [6]

$$\begin{aligned} p &= \Pr \{ \beta(u, v) \leq \beta_{th} \} \\ &= \frac{1}{2} + \frac{1}{2} \operatorname{erf} \left( \frac{\beta_{th} - \alpha 10 \log(d(u, v))}{\sqrt{2}\sigma} \right). \end{aligned} \quad (2)$$

### 2.3. MAC Layer Parameters

As opposed to the Internet, worms over sensor networks will face channel contention which should in theory reduce the overall rate of spread. Depending on the node density and MAC layer fairness, the highest achievable probe rate

might be significantly lower than the Internet. A similar trend has been observed for Internet worms, where after the initial fast spread phase, worm traffic causes severe congestion at routers and hence the spread rate decreases. Staniford *et al.* [1] emphasize that future worms will employ better scanning techniques to achieve high virulence. In view of the added constraint of MAC contention, sensor network worms (good or malicious) should be more bandwidth- and contention-aware than Internet worms.

Gupta and Kumar [9] showed that a packet transmission between nodes  $u$  and  $v$  is successful if:

1. The distance between  $u$  and  $v$  is not greater than  $r$ ,  $d(u, v) \leq r$
2. For every other node  $x$  simultaneously receiving,  $d(u, x) > r$
3. For every other node  $y$  simultaneously transmitting,  $d(y, v) > r$

The above considerations are incorporated in the simulations of subsequent sections. We assume that, to avoid unsuccessful transmissions, the sensors employ a CSMA/CA mechanism with handshaking. Thus, while we account for channel contention, it is assumed that packet transmissions are collision-free. Let there be  $N(\xi, \eta)$  nodes in segment  $(\xi, \eta)$ . Then  $N'(\xi, \eta) \leq N(\xi, \eta)$  nodes can be granted simultaneous channel access in the segment.

## 2.4. Network and Transport Layer Considerations

Most contemporary Internet worms uniformly scan the IP address space, that is, every IP address in the  $2^{32}$  IPv4 space has an equal probability of being probed. This results in many “missed scans” due to two reasons: (a) the unused IP address space; (b) many of the uniformly scanned computers are already patched. Staniford *et al.* [1] discussed strategies that can increase the virulence of an Internet worm. One such strategy that has been effectively employed by many recent worms (e.g., CodeRed v2) is *localized scanning*. The local scanning worms after compromising a host scan the nearby hosts (e.g., machines in the same subnet) with a higher probability. This strategy has proven to be quite effective since presence of a single vulnerable host implies that with high probability other hosts on the same network will also be vulnerable. This method increases virulence while reducing the outgoing network traffic. Nevertheless, even localized scanning suffers from unused IP address scans.

In the localized scanning context, a sensor network worm has an invaluable resource available to it in the form of its next-hop neighbor list. We assume that neighbor list is maintained at each node. This can be achieved by an ad

hoc routing algorithm such as [10] and [11]. An infectious sensor can spread the infection quite effectively by communicating it only to its next-hop neighbors. This strategy, which we refer to as *next-hop scanning*, will provide effective worm propagation with minimal channel contention delays. It should be emphasized here that despite the negative meaning associated with the term *infectious* (and consequently the term *infected*), in this paper *infectious/infected sensor* refers to a node which has received the worm payload and is actively participating in spreading the payload. Thus, no assumption is made about the intent of the *infectious* payload. Throughout this paper we assume that the worm employs the next-hop infection strategy. Since the worm under consideration employs (next-hop) information from a host to infect other hosts, we refer to it as a *topologically-aware worm* [1].

Recent highly virulent worms<sup>3</sup> are employing datagram communications due to the low protocol overhead and the consequent lower bandwidth consumption. In this paper, we also assume that infections are transmitted using the user datagram protocol (UDP). We assume that the worm cannot determine that a host is already infected. Due to lack of such knowledge and due in part to the use of UDP, a node can receive multiple infectious packets from different transmitters. We assume that a node will be infected when it receives its first infectious packet. An infected node will send the packet to its neighbors only once and all subsequent duplicate packets received by the infected node will be dropped.

## 2.5. Worm Properties

We focus on *unknown worms* which have also been referred to as zero-day worms and novel worms in previous literature. For the malicious worm case, we assume that high virulence and unknown nature of the present next-hop worm renders immunization ineffective. The unknown/novel payload assumption is obviously true for the good worms. Although recent Internet worms have exhibited probabilistic scanning behavior, the infection process of most known worms is still largely deterministic. This results in constant infection rates as shown in [1]. We also assume a constant infection rate for the next-hop scanning sensor network worm.

## 3. The Topologically-Aware Worm Propagation Model

Using the parameters defined in the last section, we now describe the *topologically-aware worm propagation model* (TWPM).

<sup>3</sup>The Witty worm [2], which has the fastest spread rate among all worms to date, had a UDP payload.

### 3.1. TWMP Formulation

We focus solely on propagation dynamics of unknown worms and therefore a node can be in one of two possible states: *Susceptible* or *Infected*. We emphasize again that the term *infected* simply refers to a node which has received the worm payload, without any assumption about the (good or bad) intent of the payload. A susceptible node becomes infected as soon as it is contacted by an infectious node. Immediately after getting infected, a node starts spreading the worm. Let the total number of susceptible and infectious nodes in segment  $(\xi, \eta)$  at time  $t$  be denoted by  $S(\xi, t)$  and  $I(\xi, t)$ , respectively. Using *average* statistics, the sum of nodes in both states should be

$$S(\xi, t) + I(\xi, t) = E\{D\}, \quad (3)$$

where  $E\{D\}$  represents the average number of sensor in a segment as defined in (1). This model is referred to as the classical SI model of epidemic diseases [5]. The rate of change of susceptible population with respect to time can then be expressed as [5]

$$\frac{\partial S(\xi, t)}{\partial t} = -\beta S(\xi, t) I(\xi, t), \quad (4)$$

where  $0 < \beta \leq 1$  represents the constant infection rate. We assume that the total population of initially susceptible nodes is large enough so that during the initial stages of the worm spread, the susceptible population is approximately constant. More specifically, an infectious node can infect  $\beta S(\xi, t)$  susceptible nodes in one unit of time. Thus,  $I(\xi, t)$  infectious nodes can create a total of  $\beta S(\xi, t) I(\xi, t)$  new infections in each time unit. However, in accordance with our discussion in Sections 2.2 and 2.3, channel conditions and contention will reduce the virulence of the worm. Specifically,  $I(\xi, t)$  infectious nodes will create a total of  $p\beta N'(\xi, t) I(\xi, t)$  new infections in each time unit, where  $p$  is the probability of successful packet transmission, and  $N'(\xi, \eta)$  is the number of nodes in segment  $(\xi, \eta)$  which can receive a packet (despite channel contention) in one time unit; note that  $N'(\xi, \eta) \leq E\{D\}$ . Let us denote the rate of infectious contacts received from neighboring segments of  $(\xi, \eta)$  as  $\phi$ , where  $\phi \leq \beta$  since all of the infectious contacts from a neighboring segment are not targeted at segment  $(\xi, \eta)$ .

A closer look at Figure 1 shows that if a sensor is located exactly at one of the corners of the  $(\xi, \eta)$  segment, then at maximum it can receive an infectious contact from a node which is at distance  $r$  from it (shown by the circle). For instance, at most infected nodes from segment  $(\xi - 1, \eta - 1)$  that are within  $\pi r^2/4$  area of the corner of  $(\xi, \eta)$  can spread infection to nodes in segment  $(\xi, \eta)$ . Since the total area of a segment is  $l \times h$  and nodes are uniformly distributed inside a segment, a total of  $\phi p \frac{\pi r^2}{4lh} N'(\xi - 1, \eta - 1) I(\xi - 1, \eta - 1)$

infectious contacts are received by segment  $(\xi, \eta)$  from the neighboring segment  $(\xi - 1, \eta - 1)$ . By similar logic, infected nodes of segment  $(\xi, \eta - 1)$  will transmit  $\phi p \frac{r}{h} N'(\xi, \eta - 1) I(\xi, \eta - 1)$  to the  $(\xi, \eta)$  segment. Infections from the remaining segments can be expressed similarly.

Thus the rate of change in the infectious population is

$$\begin{aligned} \frac{\partial I(\xi, \eta, t)}{\partial t} = & \beta p N'(\xi, \eta) I(\xi, \eta, t) \\ & + \phi p \frac{\pi r^2}{4lh} N'(\xi, \eta) \left[ \begin{array}{l} I(\xi - 1, \eta - 1, t) + \\ I(\xi + 1, \eta - 1, t) + \\ I(\xi - 1, \eta + 1, t) + \\ I(\xi + 1, \eta + 1, t) \end{array} \right] \\ & + \phi p \frac{r}{h} N'(\xi, \eta) \left[ \begin{array}{l} I(\xi - 1, \eta, t) + \\ I(\xi + 1, \eta, t) + \\ I(\xi, \eta + 1, t) + \\ I(\xi, \eta - 1, t) \end{array} \right]. \end{aligned} \quad (5)$$

Now that we have defined the fundamental equations, we focus on obtaining a closed-form solution for the above model. Previous studies of Internet worm epidemics have outlined that the spread is exponential during the initial stages [1], [2]. We are, therefore, particularly interested in ascertaining the solution for  $I(\xi, \eta, t)$  during initial stages of the worm outbreak. The next section derives the closed-form solution.

### 3.2. Closed-Form Solution

The expression for TWPM given in (5) is somewhat convoluted. To simplify this expression, let us rewrite (5) as

$$\begin{aligned} \frac{\partial I(\xi, \eta, t)}{\partial t} = & AI(\xi, \eta, t) \\ & + \frac{B}{2} \left[ \begin{array}{l} I(\xi - 1, \eta - 1, t) + \\ I(\xi + 1, \eta - 1, t) + \\ I(\xi - 1, \eta + 1, t) + \\ I(\xi + 1, \eta + 1, t) \end{array} \right] \\ & + \frac{C}{2} \left[ \begin{array}{l} I(\xi - 1, \eta, t) + \\ I(\xi + 1, \eta, t) + \\ I(\xi, \eta + 1, t) + \\ I(\xi, \eta - 1, t) \end{array} \right], \end{aligned}$$

where  $A = \beta p N'(\xi, \eta)$ ,  $B = \phi p \frac{\pi r^2}{2lh} N'(\xi, \eta)$  and  $C = 2\phi p \frac{r}{h} N'(\xi, \eta)$ . In order to solve this partial differential equation, we take a two-dimensional DTFT along the  $\xi$  and  $\eta$  axes. Using  $M(\omega, \theta, t)$  to denote the DTFT of  $I(\xi, \eta, t)$ , we obtain

$$\begin{aligned} \frac{\partial M(\omega, \theta, t)}{\partial t} = & AM(\omega, \theta, t) \\ & + \frac{B}{2} M(\omega, \theta, t) \left( \begin{array}{l} e^{\omega+\theta} + e^{-\omega+\theta} \\ + e^{\omega-\theta} + e^{-\omega-\theta} \end{array} \right) \\ & + \frac{C}{2} M(\omega, \theta, t) \left( e^\theta + e^\omega + e^{-\omega} + e^{-\theta} \right), \end{aligned}$$

which can be expressed as

$$\frac{\partial M(\omega, \theta, t)}{\partial t} = \begin{bmatrix} A + C \begin{pmatrix} \cos(\omega) + \\ \cos(\theta) \end{pmatrix} \\ + B \begin{pmatrix} \cos(\omega + \theta) + \\ \cos(\omega - \theta) \end{pmatrix} \end{bmatrix} M(\omega, \theta, t).$$

Assuming that the infection starts with a single infectious node (the initial condition), the solution for the above differential equation is [12]

$$M(\omega, \theta, t) = \exp \left\{ \begin{array}{l} At + Ct(\cos(\omega) + \cos(\theta)) \\ + Bt(\cos(\omega + \theta) + \cos(\omega - \theta)) \end{array} \right\}. \quad (6)$$

The exponent in the above expression is mathematically cumbersome and hence we employ the Taylor series approximation of the cosine function which is given by  $\cos(\omega) = 1 - \frac{\omega^2}{2!} + \frac{\omega^4}{4!} - \frac{\omega^6}{6!} + \dots$ . Using the first two terms of the above expansion, an approximation of  $M(\omega, \theta, t)$  can be written as

$$\begin{aligned} M(\omega, \theta, t) &\approx \exp \left\{ \begin{array}{l} At + Ct \left( 1 - \frac{\omega^2}{2} + 1 - \frac{\theta^2}{2} \right) + \\ Bt \left( 1 - \frac{(\omega - \theta)^2}{2} + 1 - \frac{(\omega + \theta)^2}{2} \right) \end{array} \right\} \\ &\approx \exp \left\{ \begin{array}{l} At + Ct \left( 2 - \frac{\omega^2 + \theta^2}{2} \right) \\ + Bt \left( 2 - \frac{2\omega^2 + 2\theta^2 + 2\omega\theta - 2\omega\theta}{2} \right) \end{array} \right\} \\ &\approx \exp \left\{ \begin{array}{l} At + \frac{Ct}{2} (4 - \omega^2 - \theta^2) \\ + Bt (2 - \omega^2 - \theta^2) \end{array} \right\} \\ &\approx \exp \left\{ \begin{array}{l} t(A + 2B + 2C) \\ - \omega^2 t \left( B + \frac{C}{2} \right) - \theta^2 t \left( B + \frac{C}{2} \right) \end{array} \right\}. \end{aligned}$$

Let  $\frac{1}{2F} = t \left( B + \frac{C}{2} \right)$  and the above expression becomes

$$M(\omega, \theta, t) \approx \exp \{ t(A + 2B + 2C) \} \exp \left\{ - \frac{\omega^2 + \theta^2}{2F} \right\}.$$

Taking the inverse DTFT gives

$$\begin{aligned} I(\xi, \eta, t) &\approx e^{t(A+2B+2C)} \int \int e^{-\frac{\omega^2 + \theta^2}{2F}} e^{j\xi\omega + j\eta\theta} d\omega d\theta \\ &\approx e^{t(A+2B+2C)} \int e^{-\frac{\omega^2}{2F}} e^{j\xi\omega} d\omega \int e^{-\frac{\theta^2}{2F}} e^{j\eta\theta} d\theta. \end{aligned}$$

The above expression denotes the inverse Fourier transform of a Gaussian function. The forward Fourier transform of the Gaussian function  $e^{-\frac{\xi^2}{2F}}$  is given by  $\sqrt{F}e^{-F\omega^2/2}$  [13]. By duality we obtain  $\sqrt{F}e^{-F\xi^2/2} \xrightarrow{\text{DTFT}} e^{-\frac{\omega^2}{2F}}$ . Using this expression for inverse DTFT we get

$$I(\xi, \eta, t) \approx F e^{t(A+2B+2C)} e^{-F\frac{\xi^2}{2}} e^{-F\frac{\eta^2}{2}}.$$

Plugging in the values of  $A, B, C$  and  $F$  renders the final (approximate) closed-form expression as

$$\begin{aligned} I(\xi, \eta, t) &\approx \frac{1}{t\phi p N'(\xi, \eta) \left( \frac{\pi r^2}{lh} + 2\frac{r}{h} \right)} \times \\ &\exp \left\{ t p N'(\xi, \eta) \left( \beta + \phi \frac{\pi r^2}{lh} + 4\phi \frac{r}{h} \right) \right\} \times \\ &\exp \left\{ \frac{-\xi^2 - \eta^2}{2t p \phi N'(\xi, \eta) \left( \frac{\pi r^2}{lh} + 2\frac{r}{h} \right)} \right\}. \end{aligned} \quad (7)$$

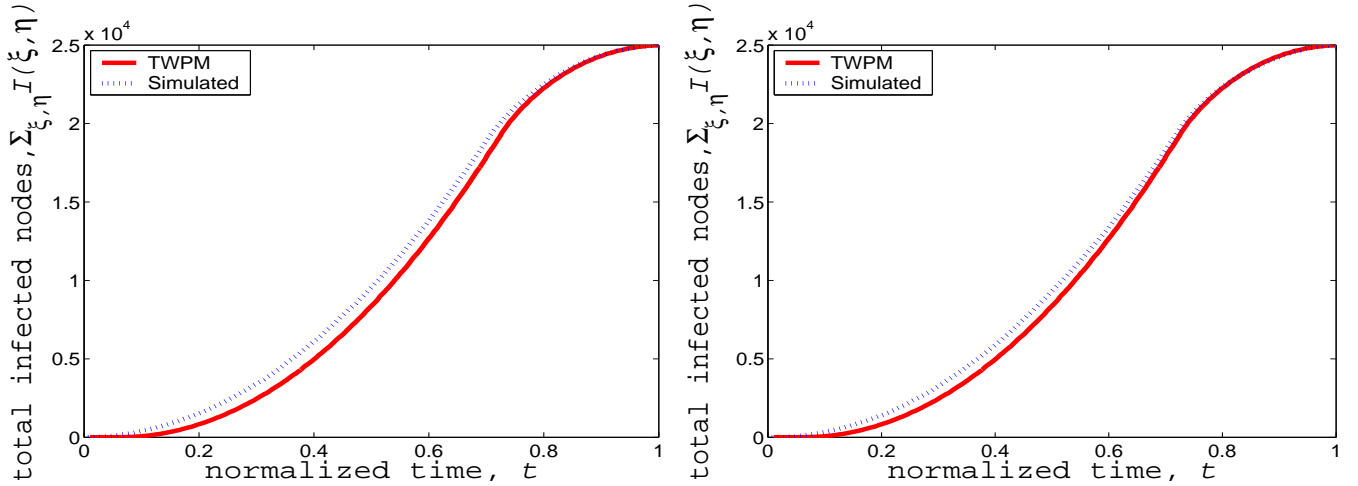
The above expression gives a closed-form solution for the TWPM model. The first exponential term shows that the initial spread is an exponential function of the infection rate,  $\beta$ , and the channel contention, represented by  $N'(\xi, \eta)$ , in the current segment. The  $e^{\phi t}$  terms in the first exponent emphasize that the number of infectious contacts,  $\phi$ , received from neighboring segments further expedite the infection process in the current segment. The second exponent in (7) exponentially decreases with an increase in  $\xi$  or  $\eta$ . This result is intuitive since nodes which are spatially far away from the infectious concentration are much less likely to contract infections. Thus, the number of infectious nodes in a segment  $\xi$  is a function of its distance from the infectious concentration.

## 4. Simulation Results

We developed a simulator which can abstractly simulate worm traffic over a sensor network. Given the total number of nodes, a two-dimensional grid size and a random distribution, the simulator placed the nodes on the grid using the random distribution as the constituent distribution of a two-dimensional IID process. Once transmission range of each node is specified, the simulator calculated the next-hop neighbors using the Euclidean distance measure. The following parameters comprise the input of the simulator:

1. infection rate,  $\beta$
2. maximum nodes in a segment that can access the channel in a given time unit,  $N'(\xi, \eta)$
3. the threshold attenuation and the path-loss exponent

Furthermore, the MAC layer interference considerations discussed in Section 2.3 were incorporated in the simulator. At each time instance, every infected node communicated the infection to  $\beta$  fraction of its neighbors. The receiver node simulated the fading effects by generating a Gaussian random variable. This random variable was generated by performing the Box-Muller transformation [14] on a random variable generated using the R250 simulator [15]–[17]. A transmitted packet was dropped or received at the receiver on the basis of the level of (simulated) channel attenuation. Some of the nodes received multiple infections through different neighbors. The simulator generated worm



**Figure 2. Total number of infections given by TWPM and simulation: (left)  $\beta = 0.2$ , (right)  $\beta = 0.5$ .**

propagation traces for total number of infected sensors in the grid.

We performed many experiments with varying parameters. It was observed that the model followed the simulation results quite closely. As an example, in Figure 2 we show results from a simulation on a  $250 \times 250$  m<sup>2</sup> grid with  $N = 25000$  sensors. Other parameters are:  $r = 3$ m,  $p = 0.95$ ,  $\alpha = 2$ ,  $l = h = 10$ m,  $N'(\xi, \eta) = E\{D\} = 40$ .

Total number of infected nodes at different time instances is shown for two infection rates,  $\beta = \phi = 0.2$  and  $\beta = \phi = 0.5$  in Figure 2. The results in Figure 2 are plotted against normalized times of the TWPM-predicted and simulated worm propagations. It can be seen that the TWPM follows the simulation results quite accurately, especially during the initial and the final stages of the infection. Even during the intermediate stages, the TWPM performance is quite close to the simulation results. Thus, we conclude that the TWPM provides an accurate model for worm propagation in a sensor network.

Figure 2 also reveals that the TWPM is quite similar to the spread of Internet worms, i.e., an exponential initial spread followed by a linear increase and finally a slow spread. This similarity between the worm spread dynamics over a two sensor network and the Internet can be explained as follows. The exponential initial spread of Internet worms is due to the availability of large numbers of vulnerable hosts. Since we are modeling an unknown (zero-day) worm, even in the sensor network case, the initial size of the susceptible population is quite large which results in a fast initial increase. Similar to the Internet, as time progresses more and more susceptible sensors are infected and therefore the curve assumes a linear increase. The slow final spread in the Internet was attributed to the fact that only few vulnerable hosts remain and it takes more time to search

out these vulnerable hosts. The explanation for the slow final spread of sensor networks has precisely the same explanation, that is, in the last stages of the infection almost all sensors are surrounded by neighbors which are already infected thereby resulting in a slow spread.

## 5. Conclusions

In this paper, we proposed and evaluated a novel worm propagation model for wireless sensor networks, namely the topologically-aware worm propagation model (TWPM). We derived a closed-form expression for the model and verified its correctness through simulations. We demonstrated that the TWPM provides an effective and accurate worm propagation model for sensor networks.

## 6. Acknowledgments

The authors thank the National Science Foundation (NSF) for supporting this project. The authors also thank Wajahat Ali Syed for his helpful comments on the original manuscript of this paper.

## 7. References

- [1] S. Staniford, V. Paxson, and N. Weaver, "How to Own the Internet in Your Spare Time," *Usenix Security Symposium*, 2002.
- [2] C. Shannon and D. Moore, "The Spread of the Witty Worm," *IEEE Security & Privacy*, vol. 2, no. 4, July/August 2004.
- [3] F. Castaeda, E. C. Sezer, and J. Xu, "WORM vs. WORM: A Preliminary Study of an Active Counter-attack Mechanism," *ACM International Workshop on Rapid Malcode (WORM)*, October 2004.

- [4] S. A. Khayam and H. Radha, "Analyzing the Spread of Active Worms over VANET," *ACM International Workshop on Vehicular Ad Hoc Networks (VANET)*, October 2004.
- [5] N. T. J. Bailey, "The Mathematical Theory of Infectious Diseases and Its Applications," Charles Griffin & Co. Ltd.: London, 1975.
- [6] T. S. Rappaport, "Wireless Communications: Principles and Practice," Prentice-Hall, 2nd ed., December 2001.
- [7] C. Bettstetter and C. Hartmann, "Connectivity of Wireless Multihop Networks in a Shadow Fading Environment," *ACM International Workshop on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWiM)*, September 2003.
- [8] R. Hekmat and P. van Mieghem, "Study of Connectivity in Wireless Ad-Hoc Networks with an Improved Radio Model," *IEEE International Symposium on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt)*, March 2004.
- [9] P. Gupta and P. R. Kumar, "The Capacity of Wireless Networks," *IEEE Transactions on Information Theory*, vol. 46, March 2000.
- [10] C. E. Perkins, E. M. Royer, and S. Das, "Ad Hoc On-Demand Distance Vector (AODV) Routing," RFC 3561, February 1999.
- [11] J. Broch and D. B. Johnson, "The Dynamic Source Routing Protocol for Mobile Ad Hoc Networks," IETF Internet Draft, July 2004.
- [12] E. Kreyszig, "Advanced Engineering Mathematics," Wiley: New York, 1998.
- [13] B. P. Lathi, "Signal Processing and Linear Systems," Berkeley-Cambridge Press, 1998.
- [14] G. E. P. Box and M. E. Muller, "A Note on the Generation of Random Normal Deviates," *Annals Math. Stat.*, vol. 29, pp. 610–611, 1958.
- [15] R250 Random Number Generator Webpage, <http://www.taygeta.com/random.xml>.
- [16] N. Zierler and J. Brillhart, "On Primitive Trinomials (mod 2)," *Information and Control*, vol. 13, no. 6, pp. 541–554, December 1968.
- [17] N. Zierler and J. Brillhart, "On Primitive Trinomials (mod 2) II," *Information and Control*, vol. 14, no. 6, pp. 566–569, June 1969.