

Analyzing the Spread of Active Worms over VANET

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

Interactive communications among nodes in Vehicular Ad Hoc Networks (VANET) and the safety-oriented nature of many VANET applications necessitate a robust security framework. An active worm over VANET can, in addition to the well-known threats to information confidentiality, integrity and service availability, pose a whole new class of traffic-related threats (ranging from congestion to large-scale accidents). This paper investigates the parameters governing the spread of active worms over VANET. To this end, we first define the average effective distance between two VANET vehicles using parameters of freeway traffic (such as velocity, time lag, number of lanes and traffic density). This effective distance measure is then used to describe the behavior of a VANET link as a log-normal shadow fading channel. The channel model is employed to define the VANET topology as a geometric random graph. We derive an analytic expression describing the average node degree of the VANET graph. The spread of a worm over VANET is modeled using a stochastic model of infectious diseases, namely the *standard Susceptible, Infected, Removed (SIR) epidemic model*. We run the stochastic SIR epidemic model on the VANET graph. For both congested and low-density traffic scenarios, we derive expressions for the rate of worm spread as a function of the average degree of the graph and the patching process. Analysis is provided for: 1) preemptive patching, where the number of patched VANET nodes remains constant; 2) interactive patching, where real-time patching is performed during a worm outbreak. We demonstrate that the latter can effectively curb the spread of a VANET worm in both congested and low-density traffic scenarios.

1. INTRODUCTION

The ease of deployment and inherently scalable nature of wireless ad hoc networks makes them a natural choice for vehicle-to-vehicle and roadside-to-vehicle communications in vehicular networks. With a vision to “increase traveler safety, reduce fuel consumption and pollution, and continue to advance the nations economy” [1], the Federal Communications Commission (FCC) has recently allocated 75 MHz of spectrum at 5.9 GHz for Dedicated Short-Range Communications (DSRC) [1], [2], [3]. A wide spectrum of services, including yet not limited to safety applications, real-time traffic management, traveler information, on-the-road entertainment and mobile Internet access, will

be offered by the DSRC Vehicular Ad Hoc Networks (VANET).

Mobile ad hoc networks have received considerable research attention during the last decade. However, the high data rates of VANET network traffic, the critical nature of VANET safety applications, time-varying vehicle densities, one-dimensionality of real-life traffic movements and high speeds of freeway vehicles distinguish the design of VANET from other forms of ad hoc networks. In view of these design considerations, research at different layers of the protocol stack is being conducted to build a robust, scalable and secure DSRC system [2]–[9].

Active computer worms, which spread over a network without human intervention, have recently emerged as one of the most imminent and effective threats against information confidentiality, integrity and service availability. An active worm over VANET can, in addition to the well-known threats, pose a whole new class of traffic-related threats ranging from congestion to large-scale accidents. Design of secure VANET applications can benefit from a thorough understanding of the spread of computer worms over a typical VANET¹.

This paper investigates the parameters governing the spread of active worms over VANET. To this end, we define an *average effective distance* between two VANET vehicles using parameters of freeway traffic (such as velocity, time lag, number of lanes and traffic density) [10]. This effective distance measure is then used to describe the behavior of a VANET link as a *log-normal shadow fading channel* [11]. The shadow fading channel model is employed to define the VANET topology as a *geometric random graph* [12]. We derive an analytic expression describing the

¹ Studies of the spread of Internet worms [13]–[17] have ascertained the efficacy of such analysis in identifying vulnerabilities, understanding threats and designing effective countermeasures.

analytic expression describing the average node degree of the VANET graph. The spread of a worm over VANET is modeled using a stochastic model of infectious diseases, namely the *standard Susceptible, Infected, Removed (SIR) epidemic model* [18]–[20]. We run the stochastic SIR epidemic model on the geometric random VANET graph. For both congested and low-density traffic scenarios, we derive expressions for the rate of worm spread as a function of the average degree of the graph and the rate at which VANET nodes are being patched. Analysis is provided for two cases of worm spread: 1) preemptive patching, where the number of patched VANET nodes remains constant; 2) interactive patching, where real-time patching is performed during a worm outbreak.

The rest of this paper is organized as follows. Section 2 develops a realistic VANET channel model. Section 3 defines the VANET topology as a geometric random graph and derives the average node degree of the VANET graph. The SIR epidemic model is applied to the graph in Section 4 and simulation results are discussed. Section 5 summarizes key conclusions of this work.

2. VANET CHANNEL MODEL

Analyzing the spread of active worms over VANET requires sound and realistic models for the level of connectivity, form of topology, and underlying wireless channels among VANET nodes. In this section we define a realistic channel model for VANET which takes into account freeway traffic parameters, such as velocity, time lag, number of lanes and traffic density. These parameters are in turn plugged into a log-normal shadow fading model to describe the probability of a link between two VANET nodes. This link probability is then used to provide a measure for the level of connectivity among VANET nodes.

2.1 Traffic Density

Ad hoc networking studies typically assume a rectangular area on which mobile nodes are placed according to some probability distribution. The time-varying traffic densities, high speeds of freeway vehicles and one-dimensionality of real-life traffic movements stipulate a different distribution for VANET nodes. From a high-level traffic density (vehicles/km) perspective, freeway traffic can be broadly classified as either *low-density* or *congested* [10].

In the case of a low traffic density, vehicles generally travel at a speed which is close to (and bounded by) a certain speed limit restriction, a phenomenon known as “free speed” in traffic flow models [10]. Since on average all vehicles will be traveling at (or around) the speed limit, a uniform (low) traffic density can be assumed here. On the contrary, congested vehicles move in a “bumper-to-bumper” manner with very high traffic densities. Nevertheless, the (high) traffic density can again be characterized as being uniform and approaching the maximum traffic density with increase in congestion. Hence, throughout this paper we assume a uniform traffic density, ρ_a , where large values of this variable represent congestion and small values represent low traffic densities.

2.2 Average Effective Vehicle Distance

The Fleetnet group defined the average distance between two vehicles as [21]

$$d_a = \frac{v_a \tau}{L} \quad (1)$$

where, v_a is the average velocity of vehicles; τ is the average time lag between two vehicles (sometimes referred to as *time headway*); and L is the number of lanes in the freeway

We use this definition of effective distance for our derivations later in the paper. In addition to the velocity and time parameters, this definition introduces the number of lanes in the distance measure. Hence, this definition captures the fact that the average distance between vehicles decreases with an increase in the number of lanes, i.e., more lanes allow vehicles to approach free speed.

2.3 Channel Model

Most ad hoc networking studies define the topology by connecting nodes that are within a certain threshold distance of each other (see references of [22], [23]). Such a geometric construction does not take the effect of interference and transmission/reception power into consideration. VANET nodes may have varying number and types of transmit/receive antennas. Similarly, signal strength between VANET nodes is largely dependent on weather conditions and roadside structures (e.g., trees, mountains, buildings etc.) A realistic VANET channel model should therefore have a (deterministic) average effective distance com-

ponent and a (random) fading/interference component.

In order to simultaneously capture distance and fading based attenuations, we employ a log-normal shadow fading model [11] to define the probability of a link between two VANET nodes. Previous studies (see for example [22] and [23]) have illustrated the efficacy of this channel model in defining ad hoc network topologies.

2.3.1 Log-normal Shadow Fading Model

Consider two VANET nodes u and v at an average effective distance d_a from each other, where d_a is defined in (1). The signal attenuation in decibels (dB) is then expressed as

$$\beta(u, v) = 10 \log \left(\frac{p_t}{p_r} \right) \text{dB},$$

where p_t and p_r represent the transmitted and received powers, and $p_r \leq p_t$. These powers are in turn a function of the number and types of transmit/receive antennas installed on the VANET nodes, the distance between them, and the overall environment surrounding these nodes.

In a shadow fading environment, $\beta(u, v)$ comprises of two additive components [11],

$$\beta(u, v) = \beta_1(u, v) + \beta_2.$$

The first component is a deterministic distance dependent variable defined as

$$\beta_1(u, v) = \alpha 10 \log(d(u, v)) \text{dB},$$

where α is a path loss exponent that depends on the environment (generally $2 \leq \alpha \leq 5$). Using the average effective distance described in (1) we obtain

$$\beta_1(u, v) = \beta_a = \alpha 10 \log(d_a) = \alpha 10 \log \left(\frac{v_a \tau}{L} \right) \quad (2)$$

which gives an on average (distance-dependent) signal attenuation in a freeway VANET.

The second component, β_2 , captures the fading effects and is defined as a normal random variable with zero mean and variance σ^2 ,

$$f_{\beta_2}(\beta_2) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left(-\frac{\beta_2^2}{2\sigma^2} \right) \quad (3)$$

Nodes u and v have a bi-directional link between them if the received signal power, p_r , is greater than or equal to a certain threshold power $p_{r,th}$.

2.3.2 Link between Two VANET Nodes

Given a receiver sensitivity, $p_r \geq p_{r,th}$, a link exists between nodes u and v if $\beta(u, v) \leq \beta_{th}$, where the threshold attenuation is

$$\beta_{th} = 10 \log \left(\frac{p_t}{p_{r,th}} \right) \text{dB}$$

Thus, the probability a link between nodes u and v is

$$\begin{aligned} \Pr \{ \text{link between } u \text{ and } v \} &= \Pr \{ \beta(u, v) \leq \beta_{th} \} \\ &= \int_{-\infty}^{\beta_{th} - \beta_a} \frac{1}{\sqrt{2\pi}\sigma} \exp \left(-\frac{\beta_2^2}{2\sigma^2} \right) d\beta_2 \end{aligned}$$

Let $\frac{\beta_2}{\sqrt{2}\sigma} = x$ which gives $d\beta_2 = \sqrt{2}\sigma dx$ and the upper integral limit becomes $x = \frac{\beta_{th} - \beta_a}{\sqrt{2}\sigma}$. Using

x as the variable of integration and employing the average effective distance from (1) yields a probability of link between two nodes as

$$p_{link} = \int_{-\infty}^{\frac{\beta_{th} - \beta_a}{\sqrt{2}\sigma}} \frac{1}{\sqrt{\pi}} \exp(-x^2) dx$$

Symmetry of the normal distribution gives

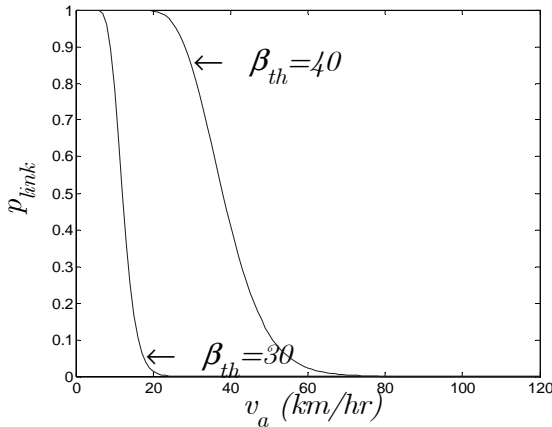
$$p_{link} = \frac{1}{2} + \int_0^{\frac{\beta_{th} - \beta_a}{\sqrt{2}\sigma}} \frac{1}{\sqrt{\pi}} \exp(-x^2) dx.$$

Expressing the probability using the error function and plugging in the value of β_a from (2), we get

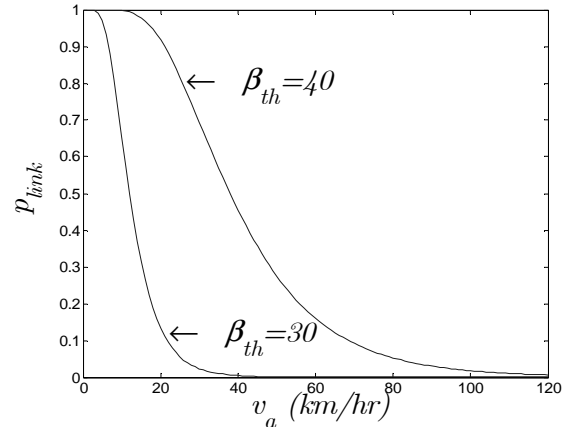
$$p_{link} = \frac{1}{2} + \frac{1}{2} \operatorname{erf} \left(\frac{\beta_{th} - \alpha 10 \log \left(\frac{v_a \tau}{L} \right)}{\sqrt{2}\sigma} \right) \quad (4)$$

The above expression defines the (on average) probability that a link exists between two nodes using freeway VANET parameters (v_a , τ and L), fading effects (σ) and receiver characteristics² (β_{th}). The link probabilities for different values of velocity and a 10 dB perturbation in thresh-

² The threshold attenuation (β_{th}) can be related to the percentage (circular) transmission range using a closed-form solution given in [11] (pg. 141–143).



(a) Fading coefficient $\sigma = 2$ dB



(b) Fading coefficient $\sigma = 4$ dB

Figure 1. Probability of a link between two VANET nodes for different values of average velocity.

old attenuation are shown in Figure 1. The number of freeway lanes, time lag between vehicles and the path loss exponent are kept constant; $L = 2$, $\tau = 5$ seconds and $\alpha = 2$.

It is clear from Figure 1 that the link probability is maximum when the vehicles are not moving ($v_a = 0$) which represents a congestion on the freeway. In this case, the effective distance between vehicles reduces to zero and two nodes establish a link with probability one. As the average velocity starts increasing, the probability of a link between two nodes starts decreasing. This substantiates the fact that quality of a link between vehicles is a function of the velocity at which they are traveling. This result is congruent with [24] and [25] which show that if the sender and receiver are within transmission range of each other then the throughput decreases due to channel interference and contention. In other words, mobility increases the capacity of ad hoc networks only when there are multiple hops between the sender and receiver.

The probability of link between mobile vehicles should also be a function of the transmission/reception range of the vehicles, which is captured by the threshold attenuation component of the model. It can be observed that a larger value of threshold attenuation allows a higher link probability at all velocities. For a fading coefficient of $\sigma = 4$ dB and a threshold attenuation of $\beta_{th} = 40$ dB, there is a non-negligible link probability even at a high velocity of 120 km/hr.

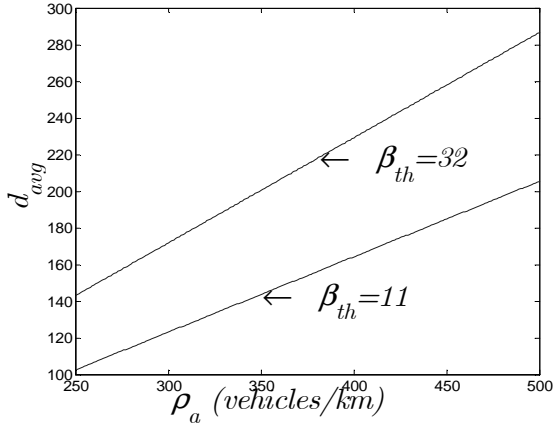
At this point the (on-average) probability of link has been defined for a VANET. The graph-theoretic interpretation of the stochastic SIR epidemic model requires that the average number of neighbors for each node is identified [18]. Therefore, we derive the average degree of a VANET graph in the following section.

3. VANET TOPOLOGY AND AVERAGE DEGREE

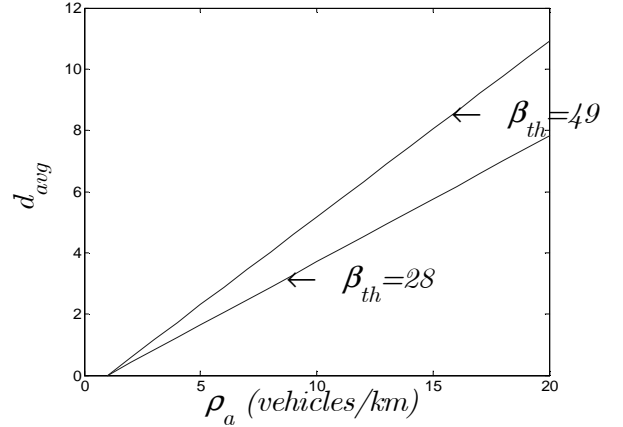
Previous studies have extensively analyzed the connectivity and minimum node degree properties of ad hoc networks (see [22], [23] and their references). For the purpose of analyzing the spread of worms over VANET, we focus on the “average node degree” of an underlying VANET channel. The derivation of an average node degree also yields the probability that a VANET node has k neighbors for two different traffic scenarios: 1) *low-density* and 2) *congested* (high density) traffic.

It was emphasized in Section 2.1 that for both low and very high traffic densities, vehicles on a freeway can be considered as uniformly distributed over a unit region (e.g., a km). The uniform vehicle density, ρ_a , then gives the average number of vehicles per unit region. Let the event that a VANET node has a link with k of the total $\rho_a - 1$ vehicles in the unit region be denoted by $\Lambda(k)$. The (on average) probability of $\Lambda(k)$ is

$$\Pr\{\Lambda(k)\} = \binom{\rho_a - 1}{k} (p_{link})^k (1 - p_{link})^{\rho_a - 1 - k} \quad (5)$$



(a) Congested traffic, $v_a = 15$ km/hr



(b) Low-density traffic, $v_a = 100$ km/hr

Figure 2. Average degree of a VANET graph in (a) congested and (b) low-density traffic scenarios.

The average degree of each node then becomes,

$$d_{avg} = (\rho_a - 1) p_{link}.$$

Using the value of p_{link} from (4), we get

$$d_{avg} = \frac{1}{2} (\rho_a - 1) \left(1 + \operatorname{erf} \left(\frac{\beta_{th} - \alpha 10 \log(v_a \tau / L)}{\sqrt{2} \sigma} \right) \right) \quad (6)$$

The above expression is intuitively convincing since it shows that the average node degree is determined by the average traffic density, the transmitted and received powers (β_{th}), the effective distance between the vehicles and the variance of the fading coefficient. As mentioned above, the β_{th} parameter is dependent on the number and types of antennas installed on the transmitter and receiver VANET nodes and the environment surrounding these nodes. Further, the effective distance is a function of the velocity and time lag of the vehicles, and the number of freeway lanes. The two traffic density scenarios under consideration are low-density and congested.

For congested traffic, $\rho_a - 1$ will take large values and therefore (5) can be approximated by a Poisson distribution as

$$\Pr \{ \Lambda(k) \} = \left(\frac{(\rho_a - 1)^k (p_{link})^k}{k!} \right) e^{-(1-\rho_a)(p_{link})} \quad (7)$$

Note that the average degree is still the same as given in (6). Hence, while (5) and (7) give the probability that a VANET node has k neighbors in low-density and congested scenarios, the expression for average node degree remains unaltered.

3.1 Simulation Results

The average node degree for a VANET in congested and low-density traffic scenarios is shown in Figure 2. As explained earlier, the average velocity of freeway vehicles decreases with an increase in traffic congestion whereas a low-density of traffic allows vehicles to travel at “free speed”. Thus, for the results in Figure 2 we used an average velocity of 15 km/hr and 100 km/hr for the congested traffic and low-density scenarios. The time lag, number of lanes, path loss coefficient and fading coefficient are fixed; $L = 2$, $\tau = 5$ seconds, $\alpha = 2$ and $\sigma = 4$.

It is clear from Figure 2 (a) that, since the effective distance between vehicles decreases in congestion, the average degree of the VANET graph increases with an increase in the traffic density. The second parameter governing the average degree is the threshold attenuation; larger β_{th} implies a larger average degree for the same traffic density. It can be seen that, for 500 vehicles/km and a threshold attenuation of 32 dB, a congested VANET node is on-average connected to about 60% of other vehicles in the unit region.

The characteristics of very low-density traffic (Figure 2 (b)) are somewhat different from the

congested scenario. It is intuitively obvious that if there is only one vehicle per unit region then the degree should be zero (assuming that the transmission range is not large enough to contact vehicles in the next unit region). Nevertheless, even for low-density traffic, the average degree increases with an increase in traffic density. However, to achieve the same (60%) level of connectivity as in Figure 2 (a), for the 20 vehicles/km traffic density a threshold attenuation of 49 dB is required, as opposed to 32dB for the congested case.

The average node degree expression of the VANET topology is now used in the stochastic SIR epidemic model to characterize the spread of an active worm over VANET.

4. VANET WORM SPREAD MODEL

The last few years have witnessed a dramatic increase in malicious network traffic. Computer worms and viruses have repeatedly revealed the susceptibility of Internet hosts to malicious intrusions. In particular, the automated self-propagating nature of active worms allows them to compromise vulnerable hosts at an extremely fast pace thereby eluding human countermeasures. Therefore, there has been a recent interest to understand the anatomy and spread dynamics of Internet worms [13]–[17]. This analysis facilitates the identification of vulnerabilities, understanding of threats and design of effective countermeasures.

A worm over VANET can compromise safety applications and hence can, in addition to other well-known threats, cause a new class of traffic-related threats, ranging from traffic congestion to large-scale accidents. It is therefore imperative that VANET applications, and in particular safety applications, are robust against computer worms. Design of these applications can benefit from a thorough understanding of the active worm spread patterns over VANET.

In this section, we employ the *stochastic standard Susceptible, Infectious, Removed (SIR) model* [18]–[20] from the spread of infectious diseases to model active worms over VANET. Some previous studies have described the spread of Internet worms using epidemic models, see for example *Kephart et al.* [13] and *Zou et al.* [16]. However, both these studies use deterministic epidemic models and we argue that the spread of worms over VANET is inherently stochastic in nature be-

cause: (1) deterministic models operate on the premise of mass action (relying on the law of large numbers) whereas the community size in VANET is ascertained by the traffic density and may not follow the law of large numbers; and (2) some VANET nodes might have been patched/immunized which necessitates that there be a probability associated with infection transmission, rather than stating certainly whether or not an infection transmission will occur. In view of the above considerations, recent studies on the spread of Internet worms are also adopting the stochastic modeling approach (see *Chen et al.* and *Garetto et al.* [17].)

4.1 The Stochastic Standard SIR Epidemic Model for VANET

An individual (i.e., a vehicle in the VANET case) can be in one of three possible states: Susceptible, Infected or Removed. Assume that in a population with a traffic density of N vehicles/unit region, there are initially n susceptible and m infectious vehicles ($N = n + m$). The *infectious period* of an infected vehicle is an exponential distribution, I , with rate γ , i.e., once infected a vehicle remains in the infected state for a mean time of $1/\gamma$ after which it transits to the removed state. During its infectious period, an infected node makes contacts with a given vehicle according to a time homogeneous Poisson process, C , with rate λ , i.e., an infected vehicle makes λ contacts with another (given) VANET node in one infectious period (λ infectious contacts/infectious period). We henceforth refer to C as the *infection process*. Thus, the time between contacts is exponentially distributed with mean $1/\lambda$. If the contacted vehicle is still susceptible then it becomes infectious and immediately starts spreading the worm. An infected node after the end of its infectious period becomes removed and plays no further part in the worm spread. From a VANET perspective, a removed vehicle has either been patched or is no longer a part of the VANET. The spread of the worm ceases when there are no infected vehicles left in the population. The SIR model [20] defines all Poisson processes to be independent of each other; they are also independent of the infectious period.

4.2 The SIR Epidemic Model on the Random VANET Graph

We apply the standard stochastic SIR model on the random geometric VANET graph whose average degree was computed in Section 3. Generally, a worm attains its peak spread rate during the initial stages. Hence, in this paper we focus on quantifying the rate of infection in the initial stages of the worm spread. To this end, we define a measure, referred to as the *spread factor*³, which is the average number of infections (spread by a newly infected vehicle) normalized by the total size of the population. Mathematically,

$$\text{spread factor} = \frac{E\{\text{infections spread by an infected node}\}}{N}$$

where, $E\{\cdot\}$ denotes an expectation operation, and as before N is the total size of the VANET population (i.e., susceptible + infected). Thus, the spread factor gives a normalized average number of vehicles that will be infected by a given (infectious) vehicle during its infectious period.

We focus our attention on two worm spread scenarios:

1. *Preemptive Patching*, where a fixed number of nodes, μ , have been patched before the start of the worm outbreak. This is a typical Internet environment where a security vulnerability is generally discovered and publicly announced by a digital security firm, such as eEye [26], TrueSecure [27], CVE [28] etc. A patch is also released for this vulnerability. A worm exploiting the announced vulnerability typically starts spreading after the release of the patch and can infect only those systems that are still unpatched (i.e., susceptible).
2. *Interactive Patching*, where vehicles are being patched in real-time while the worm is spreading. In this case, we model the *patching process* as a homogeneous Poisson process with patching rate η .

³ The definition of *spread factor* is closely related to the *basic reproduction number* which is used extensively in epidemic literature. While the spread factor does not have the asymptotic properties of the basic reproduction number, it renders a normalized (on-average) measure to quantify the rate of infection without any assumption about the susceptible population size.

Before proceeding, we derive the spread factor of a special case of the “preemptive” and “interactive” worm outbreaks.

4.3 Spread Factor of a Novel Worm

We consider the special case where the worm is exploiting a novel (unknown) security vulnerability and therefore vehicles are neither initially patched nor being patched in real-time, i.e., the case where $\mu = 0$ for preemptive patching and $\eta = 0$ for interactive patching.

Since a susceptible vehicle is infected the first time it is contacted by an infected vehicle, the worm spreading Poisson process, C , can be thought of as the time it takes for an infected vehicle to make the first contact with a given susceptible vehicle. As explained before, the mean time between contacts is exponentially distributed with mean $1/\lambda$. Let us denote this exponential random variable, which captures the time taken for the first contact as \tilde{C} . Thus, the expected value of this exponential random variable is $1/\lambda$. In other words, \tilde{C} is an exponential random variable, which is derived from C , and has *infection rate* λ .

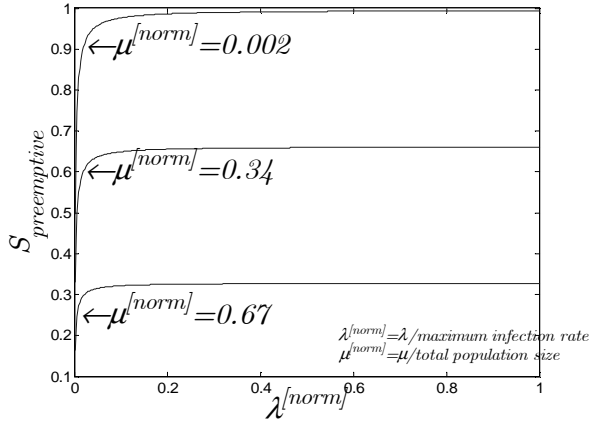
A given susceptible is infected if only if it is contacted by an infected vehicle before the end of the infectious period. Let us express the probability of this event (i.e., the probability of infection spread) as

$$\begin{aligned} \Pr\{\tilde{C} < I\} &= \Pr\{\text{time for first contact} < \text{infection period}\} \\ &= \Pr\{\text{infection spread}\} \end{aligned}$$

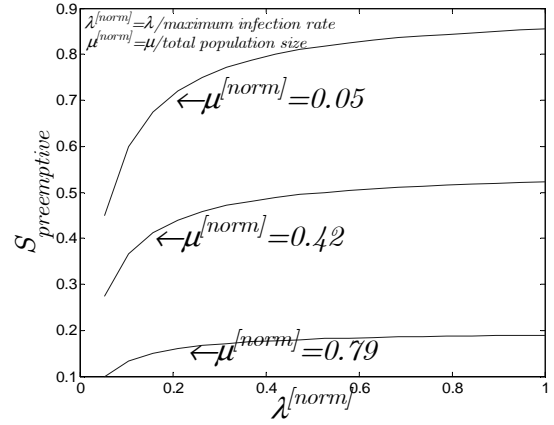
Recall that the contact process, \tilde{C} , and infectious periods process, I , are independent of each other. For two independent exponentially distributed random variables, the above probability is the probability of *competing exponentials*,

$$\Pr\{\text{infection spread}\} = \Pr\{\tilde{C} < I\} = \frac{\lambda}{\lambda + \gamma} \quad (8)$$

From (6) we know the average number of nodes, d_{avg} , connected with an (infected) VANET node. Thus, an infected VANET node will infect each of its d_{avg} neighbors with probability $\Pr\{\text{infection spread}\}$. Employing the average traffic density ρ_a as the total population, the



(a) Congested traffic, $\rho_a = 500$ vehicles/km, $v_a = 15$ km/hr



(b) Low-density traffic, $\rho_a = 20$ vehicles/km, $v_a = 100$ km/hr

Figure 3. Spread factor of an active worm over a VANET with preemptive patching.

spread factor for novel vulnerability, S_{novel} , can be expressed as,

$$S_{novel} = \frac{1}{\rho_a} \left(\frac{\lambda}{\lambda + \gamma} \right) d_{avg} \quad (9)$$

In the following sections, we provide simple extensions of the novel spread factor expression given in (9) for the general cases of preemptive and interactive patching.

4.4 Preemptive Patching

In the preemptive patching scenario, the number of patched vehicles is a non-negative number, $\mu \geq 0$. Thus, the probability of infection spread is,

$$\Pr\{\text{infection spread}\} = \Pr\left\{ \begin{array}{l} (\text{time for first contact} < \text{infection period}) \\ \cap (\text{node is unpatched}) \end{array} \right\}$$

Let the two events be independent of each other. Since μ vehicles out of a population of $\rho_a - 1$ vehicles are patched, the probability of an unpatched vehicle is

$$\Pr\{\text{node is unpatched}\} = 1 - \frac{\mu}{\rho_a - 1} \quad (10)$$

The spread factor for a worm exploiting a known vulnerability, $S_{preemptive}$, can then be expressed as

$$S_{preemptive} = S_{novel} \Pr\{\text{node is unpatched}\}$$

Plugging in values from (6), (9) and (10) yields,

$$S_{preemptive} = \frac{1}{2} \left(\frac{\rho_a - 1 - \mu}{\rho_a} \right) \left(\frac{\lambda}{\lambda + \gamma} \right) \times \left(1 + \operatorname{erf} \left(\frac{\beta_{th} - \alpha 10 \log(v_a \tau / L)}{\sqrt{2} \sigma} \right) \right) \quad (11)$$

which is the spread factor of a known worm over VANET. Comparison with (9) readily outlines that, as expected, using $\mu = 0$ in (11) gives the spread factor of a novel worm over VANET.

4.4.1 Simulation Results

Figure 3 illustrates the spread factor for the preemptive patching case in congested and low-density traffic scenarios. All other parameters are fixed; $L = 2$, $\tau = 5$ seconds, $\alpha = 2$, $\sigma = 4$ dB, $\beta_{th} = 150$ dB and $\gamma = 1$ is used as a normalized time unit.

It is evident from Figure 3 that, in accordance with the level of preemptive patching, a certain number of nodes are immune to the spread of the worm. The spread factor therefore never reaches one. However, due to the high average degree in congested scenarios, for all levels of preemptive patching the worm spreads extremely quickly in the susceptible (unpatched) population (see Figure 3 (a)). For example, in all three cases shown in Figure 3 (a) more than 90 % of the ‘‘susceptible’’ vehicles in the unit region are compromised by a worm with a infection rate (λ) as low as 0.1.

Figure 3 (b) shows that the worm outbreak in the ‘‘susceptible’’ population is much more controlled for low-density traffic due to the smaller average

degree as compared to the congested scenario. A moments thought substantiates the intuition of this result; worm spread is slower since, in the low-density case, each VANET node has (on average) lesser number of neighbors as compared to the congested case. A small average degree ensures that some infected nodes are completely surrounded by patched neighbors thereby mitigating the propagation of infection in certain graph segments. Thus in the low traffic density case, even for a very small number of preemptively patched nodes, such as 5%, a very high infection rate is required by a worm to compromise all the VANET nodes in the unit region.

4.5 Interactive Patching

Now we consider the case where vehicles are patched in real-time while the worm is spreading. These patches can be downloaded into the VANET nodes using roadside-to-vehicle communications. We model the *patching process* as a time-homogeneous Poisson process, T , which, for a given infectious period⁴, contacts a vehicle with rate η , i.e., the patching process contacts a vehicle η times in one infectious period (η patching contacts/infectious period). A susceptible vehicle is patched the first time it is contacted by the patching process. Thus, the patching Poisson process can be thought of as the time it takes to patch a given susceptible. Since the time between patching contacts in exponentially distributed, on average the time for the first contact is $1/\eta$. Let us represent this exponential random variable as \tilde{T} , which captures the time taken for the first contact by the patching process. Since the mean time is $1/\eta$, the rate of \tilde{T} is given by η , and we refer to it as the *patching rate*. The probability of an unpatched (susceptible) node is given by,

$$\begin{aligned} \Pr \{\text{node is unpatched}\} &= \Pr \{\tilde{C} < \tilde{T}\} \\ &= \Pr \{\text{time for first infection} < \text{time for first patch}\} \end{aligned}$$

Assuming that the patching and infection processes are independent, we get

$$\Pr \{\text{node is unpatched}\} = \Pr \{\tilde{C} < \tilde{T}\} = \frac{\lambda}{\lambda + \eta}$$

⁴ *Infectious period* is used as a unit of time here in order to draw a fair comparison with the infection spread process, C .

The spread factor for interactive patching, $S_{\text{interactive}}$, can hence be written as

$$S_{\text{interact}} = \frac{1}{2} \left(\frac{\lambda}{\lambda + \eta} \right) \left(\frac{\lambda}{\lambda + \gamma} \right) \left(\frac{\rho_a - 1}{\rho_a} \right) \times \left(1 + \operatorname{erf} \left(\frac{\beta_{th} - \alpha 10 \log(v_a \tau / L)}{\sqrt{2} \sigma} \right) \right) \quad (12)$$

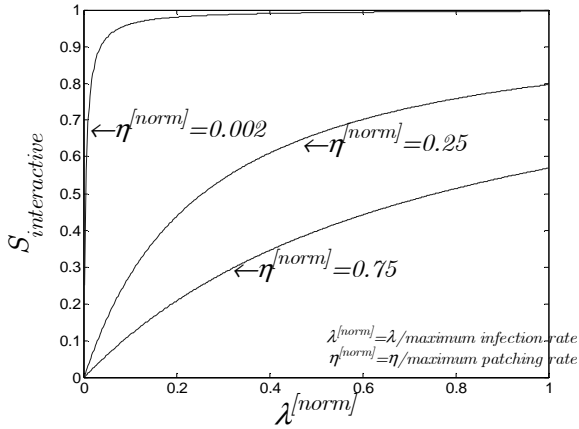
It can be observed that the new term, involving η , introduced in the above expression signifies the interplay between the patching and infection processes.

4.5.1 Simulation Results

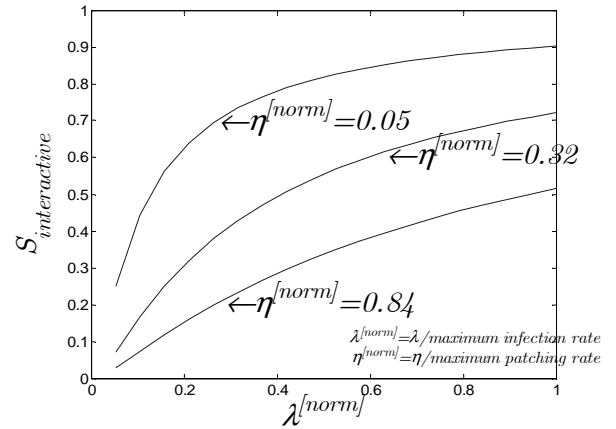
Figure 4 outlines the interplay between a real-time patching process and the infection process for congested and low-density traffic scenarios. All other parameters are fixed with same values as Section 4.4.1. It is clear from Figure 4 (a) that, due to the high average degree of the VANET graph in congested scenarios, the patching rate plays a crucial role in curbing a worm outbreak. For instance, in the case of a very low patching rate $\eta = 0.2\%$, 100% of the nodes are infected quite rapidly even for very low infection rates. Nevertheless, as the patching rate increases, the spread of the worm is mitigated quite effectively, for example, in the case of $\eta = 75\%$ and the highest possible infection rate, less than 60% of the susceptible nodes are eventually infected. The spread factor for the low-density traffic (see Figure 4 (b)) is even smaller than the congested scenario for all values of η . For a patching rate of $\eta = 32\%$, approximately 30% of the population does not contract the infection.

5. CONCLUSIONS

In this paper, we defined a log-normal shadow fading VANET channel model using freeway traffic parameters. This model was employed to define the VANET topology as a geometric random graph. Average degree of the graph was computed for low-density and congested traffic scenarios. The SIR epidemic model was then applied to evaluate a worm outbreak over the VANET graph. Two scenarios which curb the worm outbreak were analyzed: preemptive patching and interactive patching. We conclude that in preemptive patching, although some vehicles are patched be-



(a) Congested traffic, $\rho_a = 500$ vehicles/km, $v_a = 15$ km/hr



(b) Low-density traffic, $\rho_a = 20$ vehicles/km, $v_a = 100$ km/hr

Figure 4. Spread factor of an active worm over VANET with interactive patching.

fore the start of the outbreak, during traffic congestion all the susceptible vehicles can be compromised very rapidly. Due to the smaller average degree of the VANET graph in low-density scenarios, the worm spread rate for the preemptive case is lower as compared to the congested case. A worm with a high spread rate can however compromise all the susceptible population. Moreover we deduce that interactive patching, in both congested and low-density traffic scenarios, can very effectively curb the spread of a VANET worm.

6. REFERENCES

- [1] FCC Report and Order 03-024, released February 10, 2004.
- [2] Dedicated Short Range Communications Home, (www.leearmstrong.com/dsrc/dsrchomeset.htm.)
- [3] R. Sengupta and Q. Xu "DSRC for Safety Systems", *Intellimotion*, vol. 10, no. 4, 2004.
- [4] Q. Xu, T. Mak, J. Ko, and R. Sengupta, "MAC Protocol Design for Vehicle Safety Communications in Dedicated Short Range Communications Spectrum," submitted to *IEEE ITSC*, (<http://path.berkeley.edu/dsrc/public.htm>.)
- [5] Q. Xu, R. Sengupta, and D. Jiang, "Design and Analysis of Highway Safety Communication Protocol in 5.9 GHz Dedicated Short Range Communication Spectrum," *IEEE VTC*, Spring 2003.
- [6] R. Jain, A. Puri, and R. Sengupta, "Geographical Routing Using Partial Information for Wireless Ad Hoc Networks," *IEEE Personal Communications*, vol. 8, no. 1, pp. 48–57, February 2001.
- [7] Q. Xu and R. Sengupta, "Simulation, Analysis, and Comparison of ACC and CACC in Highway Merging Control," *IEEE Intelligent Vehicle Symposium*, 2003.
- [8] D. Lee, R. Attias, A. Puri, R. Sengupta, S. Tripakis, and P. Varaiya, "A Wireless Token Ring Protocol For Intelligent Transportation Systems," *ITSC*, August 2001.
- [9] M. Zennaro and J. Misener, "A State-map Architecture for Safe Intelligent Intersection," *ITSA*, 2003.
- [10] W. Leuzbach, "Introduction to the Theory of Traffic Flow," Springer: Verlag, 1988.
- [11] T. S. Rappaport, "Wireless Communications: Principles and Practice," Prentice-Hall, 2nd ed., December 2001.
- [12] B. Bollobas, "Random Graphs," Cambridge University Press, 2nd edition, January 2001.
- [13] J. O. Kephart and S. R. White, "Directed-Graph Epidemiological Models of Computer Viruses," *IEEE ISSP*, 1991.
- [14] *Communications of the ACM*, vol. 32, no. 6, pp. 678–698, 1989 and vol. 25, no. 3, pp. 172–180, 1982.
- [15] D. Moore, C. Shannon, and D. Brown, "Code-Red: A Case Study on the Spread and Victims of an Internet Worm," *ACM Sigcomm*, 2002.
- [16] Proceedings of *ACM WORM 2003*.
- [17] Network Security Session, *IEEE Infocom*, 2003.

- [18] H. Andersson and T. Britton, "Stochastic Epidemic Models and Their Statistical Analysis," Springer: Verlag, 2000.
- [19] W. O Kermack and A. G. McKendrick, "A Contribution to the Mathematical Theory of Epidemics," *Proc. Roy. Soc. Lond.*, A 115, pp. 700–721, 1927.
- [20] M. S. Bartlett, "Some Evolutionary Stochastic Processes," *J. Roy. Statistic. Soc.*, B 11, pp. 211–229, 1949.
- [21] H. Hartenstein, B. Bochow, A. Ebner, M. Lott, M. Radimirsch, and D. Vollmer, "Position-Aware Ad Hoc Wireless Networks for Inter-Vehicle Communications: the Fleetnet Project," *ACM MobiHoc*, October 2001.
- [22] C. Bettstetter and C. Hartmann, "Connectivity of Wireless Multihop Networks in a Shadow Fading Environment," *ACM MSWiM*, September 2003.
- [23] R. Hekmat and P. Van Mieghem, "Study of Connectivity in Wireless Ad-Hoc Networks with an Improved Radio Model," *WiOpt*, March, 2004.
- [24] M. Grossglauser and D. N. C. Tse, "Mobility Increases the Capacity of Ad Hoc Wireless Networks," *IEEE/ACM Trans. on Networking*, vol. 10, no 4, pp. 477–486, August 2002.
- [25] P. Gupta and P. Kumar, "The Capacity of Wireless Networks," *IEEE Trans. on Information Theory*, vol. 46, pp. 388–404, March 2000.
- [26] eEye Digital Security, <http://www.eeye.com/html/>.
- [27] Trusecure, <http://www.trusecure.com/>.
- [28] Common Vulnerabilities and Exposures (CVE), <http://www.cve.mitre.org>.