Weierstraß-Institut
für Angewandte Analysis und Stochastik

Leibniz-Institut im Forschungsverbund Berlin e. V.

Preprint ISSN 2198-5855

Malware propagation in urban D2D networks

Alexander Hinsen[1], Benedikt Jahnel[1], Eli Cali[2], Jean-Philippe Wary[2]

submitted: January 14, 2020

[1] Weierstraß-Institut
Mohrenstr. 39
10117 Berlin
Germany
E-Mail: alexander.hinsen@wias-berlin.de
benedikt.jahnel@wias-berlin.de

[2] Orange SA
44 Avenue de la République
92326 Châtillon
France
E-Mail: elie.cali@orange.com
jeanphilippe.wary@orange.com

No. 2674
Berlin 2020

2010 *Mathematics Subject Classification.* 60J25, 60K35, 60K37.

Key words and phrases. Random environment, Cox–Gilbert graph, Poisson–Voronoi tessellation, interacting particle system, ad-hoc network, data propagation, white knight, speed of propagation, survival, extinction.

This research was supported by Orange S.A. grant CRE I07263 as well as the German Research Foundation under Germany's Excellence Strategy MATH+: The Berlin Mathematics Research Center, EXC-2046/1 project ID: 390685689. We thank A. Boubaya for his contributions to the text.

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Leibniz-Institut im Forschungsverbund Berlin e. V.
Mohrenstraße 39
10117 Berlin
Germany

Fax: +49 30 20372-303
E-Mail: preprint@wias-Cberlin.de
World Wide Web: http://www.wias-Cberlin.de/

Malware propagation in urban D2D networks

Alexander Hinsen, Benedikt Jahnel, Eli Cali, Jean-Philippe Wary

Abstract

We introduce and analyze models for the propagation of malware in pure D2D networks given via stationary Cox–Gilbert graphs. Here, the devices form a Poisson point process with random intensity measure $\lambda\Lambda$, where Λ is stationary and given, for example, by the edge-length measure of a realization of a Poisson–Voronoi tessellation that represents an urban street system. We assume that, at initial time, a typical device at the center of the network carries a malware and starts to infect neighboring devices after random waiting times. Here we focus on Markovian models, where the waiting times are exponential random variables, and non-Markovian models, where the waiting times feature strictly positive minimal and finite maximal waiting times. We present numerical results for the speed of propagation depending on the system parameters.

In a second step, we introduce and analyze a counter measure for the malware propagation given by special devices called white knights, which have the ability, once attacked, to eliminate the malware from infected devices and turn them into white knights. Based on simulations, we isolate parameter regimes in which the malware survives or is eliminated, both in the Markovian and non-Markovian setting.

1 Introduction

Present day telecommunication networks are ill equipped for the rapidly growing demand for mobile data transfers. With the fifth generation of mobile networks, paradigmatic shifts in the design of the network are on the agenda. A critical aspect here is the rôle of infrastructure. Multilayered cellular networks with possible incorporation of relaying mechanisms are under investigation not only in the scientific community [1–4], but also in industry [5]. All these new designs have in common a rapid increase of degrees of freedom in the system. The central rôle of base stations is reduced in favor of an increasingly important rôle of relays. In particular, also the users of the system will be attached a relay functionality in the system. As a result, the network becomes more and more decentralized. Exploring the possible benefits of such new architectures is in full swing in the academic and the industrial research. For a survey on device-to-device (D2D) communication in cellular networks see for example [6].

Of crucial importance in all these future scenarios with less centralized architectures is a good understanding of vulnerability and security, in particular of the way in which malware (e.g., proximity-based propagation sabotage software or computer killing viruses like Cabir or CommWarrior) spreads in such networks [7]. For usual networks, a number of strategies have been exploited by operators in order to restrict the spread of the malware and to keep the functionality of the network available [8]. For security in D2D communication networks see the review [9]. However, the new challenges accompanying the system decentralization also include the question how successful these defense strategies can be in such systems, in particular since the spread of malware (more generally of any information) follows a different set of rules than in centralized networks.

DOI 10.20347/WIAS.PREPRINT.2674 Berlin 2020

In this paper, we introduce a framework for the modeling of malware spreading in an urban D2D scenario. One of the main parts of this framework is that the high complexity of the real-world system is reflected by a probabilistic perspective. Configurations of devices in the cities appear as random configurations of points or vertices, and their connectivity structure is represented by edges between these vertices. Similarly, the times at which data is transmitted along edges to neighboring devices are also assumed random, which results in a stochastic process of interacting devices. In order to find a balance between describing these complexities and being able to interpret the theoretical and numerical findings, we will mainly focus on models that can be described by a set of parameters of reasonable size.

We analyze numerically the speed of the infection spread based on the system parameters, for which we introduce two different updating mechanisms. We also evaluate the possibility of tackling the malware by the introduction of *white knights* into the system. The success of this counter measure is then measured via its effect on the survival of the malware over long times and large areas. Competition models of this type have been considered in the literature for various types of fixed network models and in a Markovian setting [10–15]. However, the present paper is devoted to the study of such models on particularly realistic and random network models, also going substantially beyond the Markovian setting for the propagation model.

2 Model

2.1 Street-system based network models

For the distribution of devices in space we adopt the stochastic geometry model of *Cox point processes* and later base our simulations on particular choices of environment measure. More precisely, a Cox point process in $\mathbb{R}^2$ is a Poisson Point process with random intensity measure $\lambda\Lambda$, where the directing measure Λ represents a stationary ergodic random environment. For example, starting from a homogeneous Poisson–Voronoi tessellation S, we define the (random) measure $\Lambda(A) = |S \cap A|$ which assigns to any measurable subset $A \subset \mathbb{R}^2$ the street length in A, see red cells in Figure 1. This modeling ansatz is supported by some statistical analysis performed in [16], at least for European cities. (Non-European cities with more rectangular street systems have been analyzed for example in [17].) In this example, since the underlying homogeneous Poisson point process that is used to create the Poisson–Voronoi tessellations can be parametrized by one parameter, the density of streets can also be described by one parameter, $\gamma = \mathbb{E}|S \cap [-1/2, 1/2]^2|$, the *expected street length in a unit area*. In our applications, the quantities of interest will usually only depend on the parameter γ, rather than on the whole distribution of the street system.

To finish the description of a Cox point process, for a given street system S, we assume that the devices form a Poisson point process $X = \{X_i\}_{i\geq 1}$ with intensity measure $\lambda|S \cap \mathrm{d}x|$, see blue dots in Figure 1. Here, $\lambda > 0$ is another system parameter describing the *linear intensity of devices* on the streets.

For the network model, we use the classical *Gilbert graph* $g_r(X)$ *with connection radius* $r > 0$. That is, devices $X_i, X_j \in X$ are connected by an edge in $g_r(X)$ if and only if $|X_i - X_j| \leq r$, see thin black lines in Figure 1.

The network model thus depends on the three central parameters λ, γ and r, which correspond to fundamental characteristics of real-world wireless networks and obey certain scaling relations. For

example, the product $\lambda\gamma$ describes the spatial intensity of devices, i.e., the expected number of devices per unit area. Not all parameters are equally accessible from an operator's perspective. For example, the connection radius r depends very much on technology (e.g., LTE, Bluetooth, or WIFI). On the other hand, the street intensity γ is given, but it is different from one city to the other. Hence, results should cover a certain range of values. Similarly, the linear intensity λ depends on the number of devices participating in the system and thus might vary substantially.

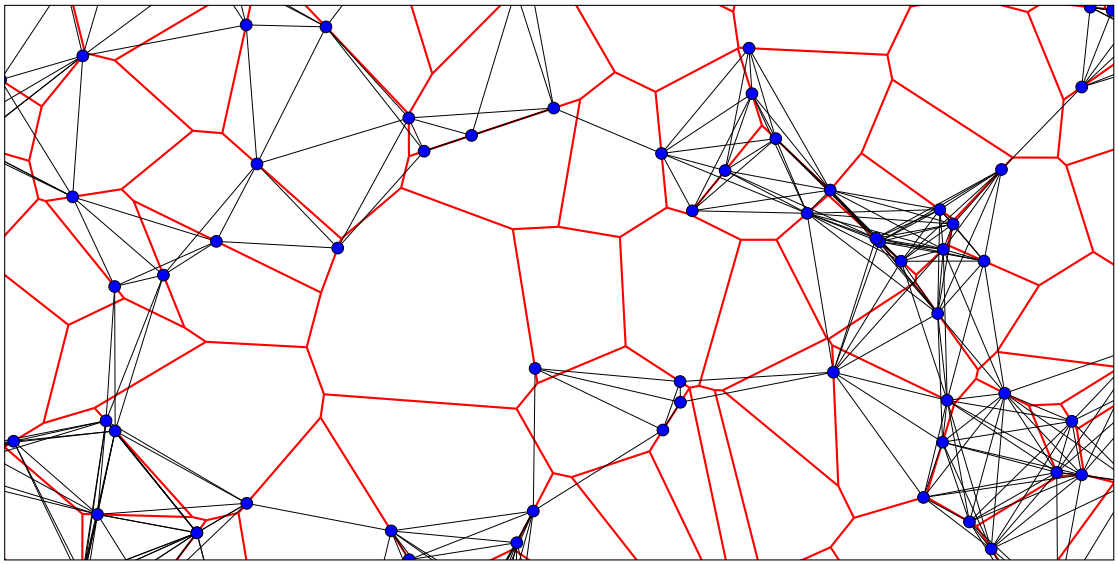

Figure 1: Realization of randomly placed devices (blue) confined to a street system given by a Poisson–Voronoi tessellation (red). Edges (black) are drawn whenever two devices are at sufficiently close proximity of each other.

Next, we introduce malware to the system and describe how its propagation can be formulated.

2.2 Malware propagation models

So far, devices X_i in the system carry no information about being infected or not, and there is no time component involved. We start by labelling each device $X_i \in X$ and write $\xi(t, X_i)$ for the state of the device X_i at time $t \geq 0$. In the first step, we only want to distinguish two possible states, namely S and I, for *susceptible* and *infected*, i.e., $\xi(t, X_i) \in \{\mathrm{S}, \mathrm{I}\}$:

$$\xi(t, X_i) = \begin{cases} \mathrm{S}: \text{device } X_i \text{ is susceptible at time } t \\ \mathrm{I}: \text{device } X_i \text{ carries malware at time } t. \end{cases}$$

We then define the set of infected devices at time t by

$$I(t) = \{X_i \in X : \xi(t, X_i) = \mathrm{I}\}.$$

We are interested in the collective behavior of a large number of interacting devices given by the vertices of $g_r(X)$ as presented above. Our general modeling assumption is that a change of status of $X_i \in X$ happens at random times due to fluctuations in the system, and these times depend on the configuration in the vicinity of X_i. Systems of this type are often called *interacting particle system* and a large body of literature is available typically for non-random particle positions and exponential waiting times, see for example [18–21].

In view of the application to malware propagation in realistic D2D networks, in this paper, we use an extended framework where instead of fixed geometries, the underlying geometric model is random and given by $g_r(X)$. Further, besides the usual exponential waiting times that we introduce next, we also study models with non-exponential waiting times, at least via simulations, see Section 3.

In order to introduce interaction into our system, we define the jump rate of a device X_i to jump from a state S to a state I as $\lambda_{\mathrm{I}} > 0$ times the number of infected neighbors of X_i in $g_r(X)$. At this point there are no defense mechanisms at work and thus we assume that infected devices cannot become susceptible again. Using the power expansion of the exponential and the Landau notation, we define

$$\mathbb{P}(\xi(t+h, X_i) = \mathrm{I} \mid \xi(t, X_i) = \mathrm{S}) = h\lambda_{\mathrm{I}} \#\{X_j \in X : |X_j - X_i| < r \text{ and } \xi(t, X_j) = \mathrm{I}\} + o(h).$$

Here, the scalar parameter λ_{I}, the *infection rate*, can be used to adjust the speed of the microscopic updates. We call this model the *Markovian* SI*-model*. In the literature it is also referred to as the *Richardson model*, see [22]. We have the following well definedness result.

Lemma 2.1. *For almost-all Gilbert graphs $g_r(X)$ we have that $\#I(0) < \infty$ implies $\#I(t) < \infty$ for all $t \geq 0$.*

Proof. Let $\tau_n = \inf\{s \geq 0 \colon I(s) \not\subset B_n\}$, where $B_n = [-nr, nr]^2$, then it suffices to show that $\tau_n \uparrow \infty$ almost surely. The proof rests on coupling arguments for the propagation process and ergodicity of the random environment. First, note that

$$\tau_n = \sum_{i=1}^{n} \tau_i - \tau_{i-1} \geq \sum_{i=1}^{n} \tau_i',$$

where $\tau_i' \sim \mathrm{Exp}(\lambda_{\mathrm{I}} \sum_{X_j \in X \cap B_i \setminus B_{i-1}} \deg(X_j))$ represents the fastest possible crossing of the annulus $B_i \setminus B_{i-1}$. Here $\deg(X_j)$ denotes the degree of X_j in $g_r(X)$.

Next, due to ergodicity, there exist $a, b > 0$ and $m = m(g_r(X))$ such that for all $n > m$ we have

$$\sum_{X_j \in X \cap B_i \setminus B_{i-1}} \deg(X_j) < a|B_i \setminus B_{i-1}| \leq 4ari$$

for at least bn-many annuli i. There, $\mathbb{E}[\tau_i'] > \frac{1}{\lambda 4ari}$ holds. This implies that

$$\lim_{n\uparrow\infty} \mathbb{E}[\tau_n] \geq \lim_{n\uparrow\infty} \sum_{i=0}^{n} \mathbb{E}[\tau_i'] = \infty,$$

Using similar arguments we can show for the variance that $\lim_{n\uparrow\infty} \mathbb{V}[\sum_{i=0}^{n} \tau_i'] < \infty$ and hence $\tau_n \uparrow \infty$ almost surely. □

Although the Markov property has substantial advantages for the analytical approach to malware propagation, the implied description via exponential waiting times is a restriction and also does not reflect some of the details of a realistic transmission process. Since, from a simulation perspective, the Markovian approach is not a substantial advantage, in this section, we develop an extended framework for interacting particle systems based on more general waiting-time distributions. The distributions that we have in mind feature a deterministic initial time step where transmission is impossible and a maximal waiting time until which transmissions must be finished. Further, here statistical input can be incorporated, which might propose certain bell-like shapes for the density of the waiting times. In the simplest case, we can work with a distribution where the infection attempt happens at times which are

uniformly distributed in $[c_1, c_2]$ with $0 < c_1 < c_2$, i.e., we have a density $f(t) = (c_2 - c_1)^{-1}\chi_{[c_1,c_2]}(t)$ for the iid renewal times.

We call this model the *non-Markovian* SI*-model*. We have the following well-definedness result.

Lemma 2.2. *For the propagation model with iid renewal times that have a density with support of the form* $[c_1, c_2]$ *for some* $0 < c_1 < c_2$*, we have that for almost-all Gilbert graphs* $g_r(X)$*,* $\#I(0) < \infty$ *implies* $\#I(t) < \infty$ *for all* $t \geq 0$*.*

Proof. Since in any finite time window, there can only be at most t/c_1 consecutive jumps and hence the infection can reach at most distance rt/c_1. □

In Figure 2 we present pairs of snapshots for the SI-model with exponential waiting times and uniform waiting times on $[c, d]$, all in a finite ball. The simulations are stopped at the stopping times at which a certain radius is reached.

2.3 White-knights counter measure

Let us introduce a counter measure based on the presence of another type of software, which permanently eliminates malware. For this we introduce a new state G and thus increase the local state space to $\{\mathrm{S}, \mathrm{I}, \mathrm{G}\}$. Updates can only be performed in this order

$$\mathrm{S} \to \mathrm{I} \to \mathrm{G}.$$

In words, susceptible devices, once infected, cannot become susceptible again, and only infected devices can become patched. The reason why the software patch cannot be transmitted to susceptible devices at once is the following. Due to privacy regulations, only after an infection attempt, the infection becomes known to other devices and only then retaliation is authorized without explicit consent of the infected device. The software patch corresponds to an immunization of the device and thus G is an absorbing state. We call devices in the state G *white knights*.

In order to have a Markovian structure, we assume exponential waiting times for the patch installation with parameter $\lambda_{\mathrm{G}} > 0$, the *patch rate*. This leads to the following description via generators

$$\mathbb{P}(\xi(t+h, X_i) = \mathrm{I} \mid \xi(t, X_i) = \mathrm{S}) = h\lambda_{\mathrm{I}} \#\{X_j \in X : |X_j - X_i| < r \text{ and } \xi(t, X_j) = \mathrm{I}\} + o(h),$$
$$\mathbb{P}(\xi(t+h, X_i) = \mathrm{G} \mid \xi(t, X_i) = \mathrm{I}) = h\lambda_{\mathrm{G}} \#\{X_j \in X : |X_j - X_i| < r \text{ and } \xi(t, X_j) = \mathrm{G}\} + o(h).$$

We call this process the *Markovian* SIG*-model*.

Typically, in our analysis, we will adopt a setting where initially the infection is present at one typical node in the network, which we assume to be the origin. This reflects the idea that the malware is brought into the system by one device and we start our clock at that time. On the other hand, white knights have to be present in the system already at time zero. Their locations are random, more precisely, white knights are assumed to have the same spatial distribution as any other device. In other words, at initial time, we will assume that white knights form a Poisson point process $Y = \{Y_i\}_{i \geq 1}$ on the street system S, independent of the other devices $X = \{X_i\}_{i \geq 1}$. Its intensity measure is then given by $\rho |S \cap \mathrm{d}u|$, where $\rho > 0$ is another system parameter describing the *linear intensity of white knights* on the streets. We note that alternatively, a Cox process with joint intensity $(\rho + \lambda)|S \cap \mathrm{d}u|$ can be realized and in a second step, devices can be labelled as white knights with iid probability $\rho/(\rho + \lambda)$.

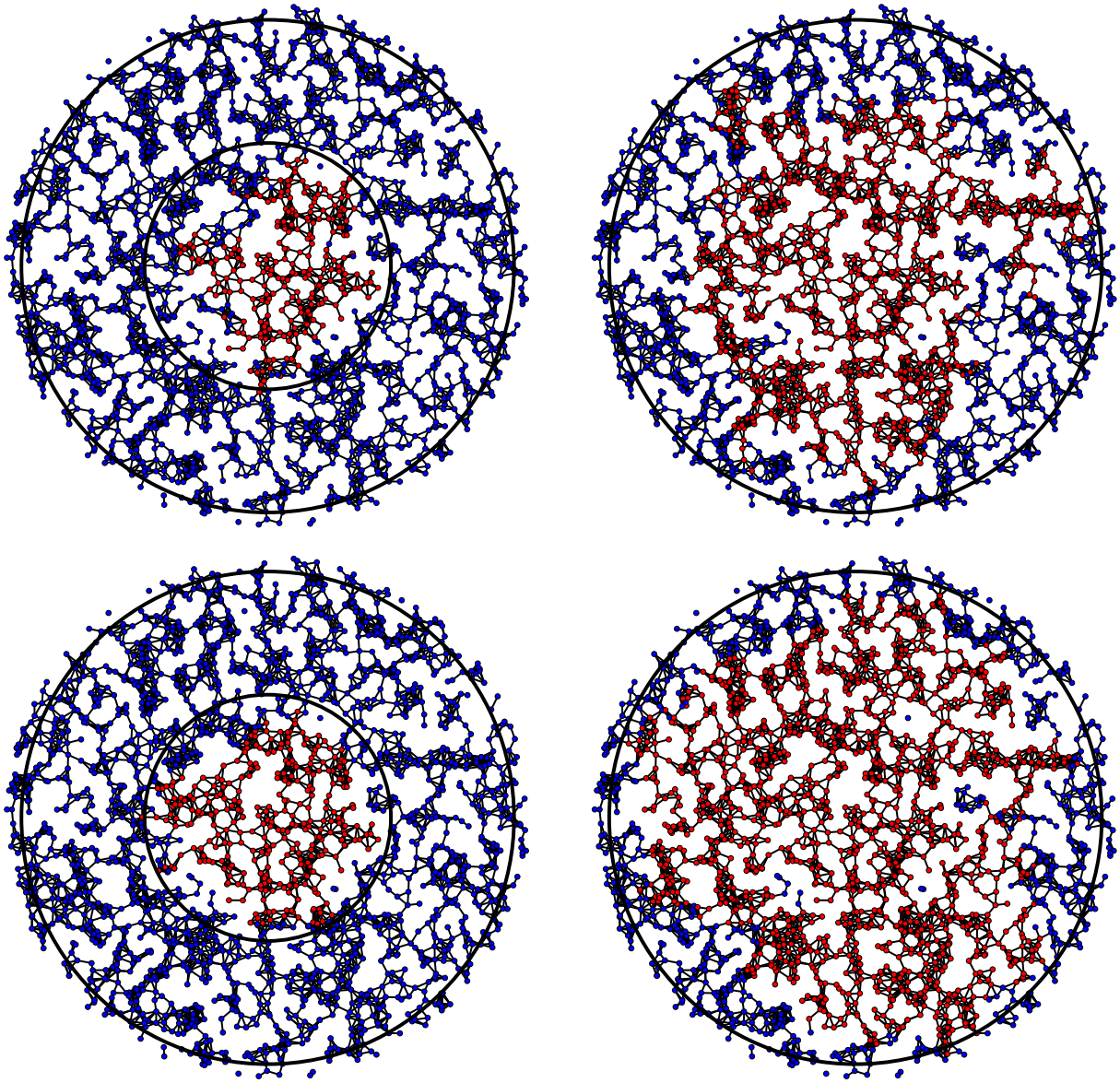

Figure 2: Realization of randomly placed devices on a street system of Poisson–Voronoi tessellation type with $\gamma = 20\,\mathrm{km}^{-1}$, $\lambda = 1.2\,\mathrm{devices/km}$ and $r = 300\,\mathrm{m}$. Upper row: The Markovian SI-model with parameter $\lambda_{\mathrm{I}} = 1$ stopped at the time at which the malware has reached the radius $u = 2.5\,\mathrm{km}$ (left) and $u = 5\,\mathrm{km}$ (right), indicated in black. Lower row: The non-Markovian SI-model with uniform waiting times on $[40\,\mathrm{sec}, 120\,\mathrm{sec}]$ stopped at the time at which the malware has reached the radius $u = 2.5\,\mathrm{km}$ (left) and $u = 5\,\mathrm{km}$ (right), indicated in black.

Like in the previous sections, also the white-knight process can be formulated where exponential waiting times are replaced by more general renewal processes, which we call the *non-Markovian* SIG-*model*, see Figures 3 for illustrations.

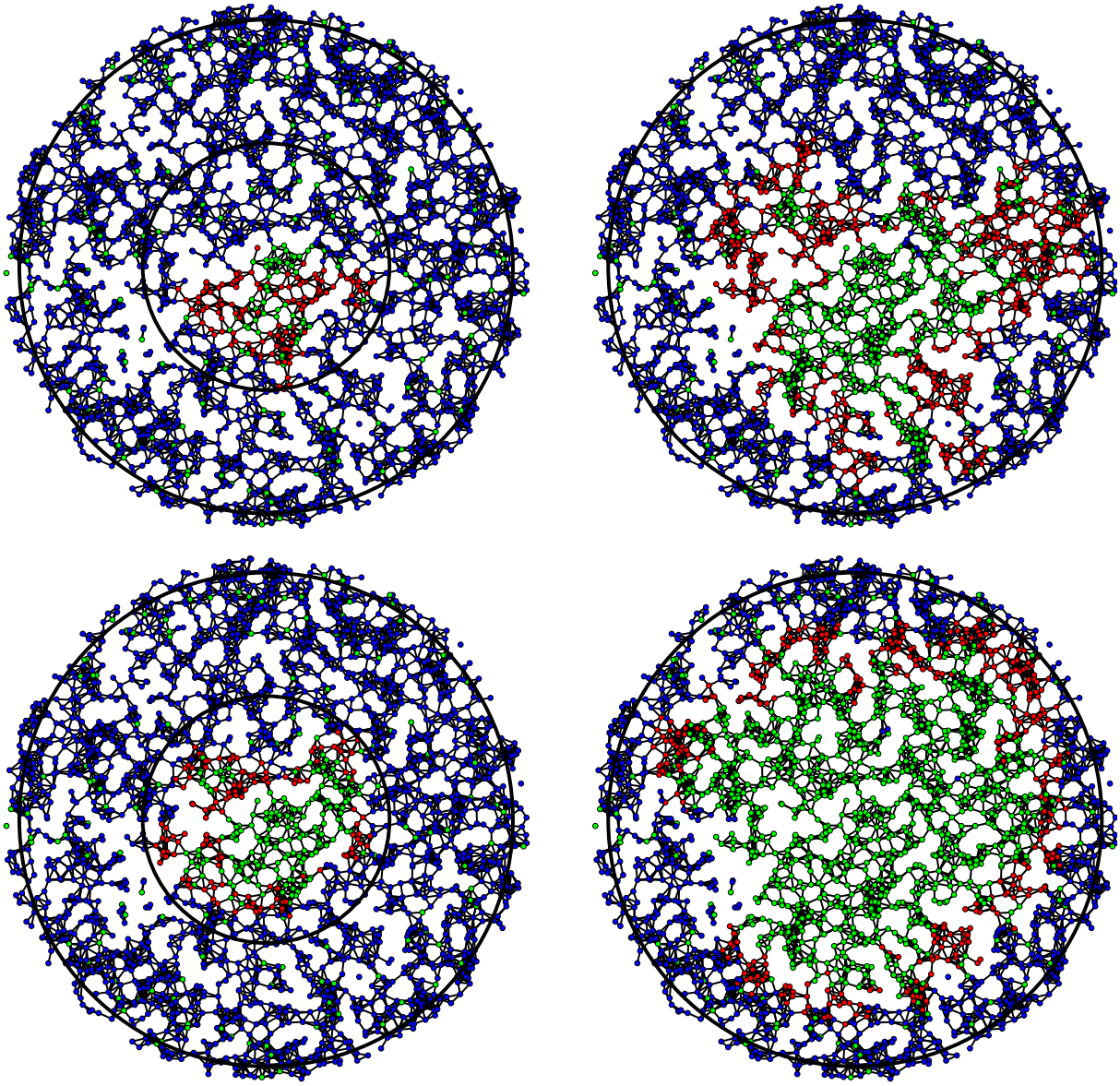

Figure 3: Realization of randomly placed devices on a street system of Poisson–Voronoi tessellation type with $\gamma = 20\,\mathrm{km}^{-1}$, $\lambda = 1.3\,\mathrm{devices/km}$, $\rho = 0.1\,\mathrm{devices/km}$ and $r = 300\,\mathrm{m}$. Upper row: The Markovian SIG-model with parameter $\lambda_{\mathrm{I}} = 1$ stopped at the time the malware has reached the radius $u = 2.5\,\mathrm{km}$ (left) and $u = 5\,\mathrm{km}$ (right), indicated in black. Lower row: The non-Markovian SIG-model with uniform waiting times on $[40\,\mathrm{sec}, 120\,\mathrm{sec}]$ for both infected and immune devices, stopped at the time the malware has reached the radius $u = 2.5\,\mathrm{km}$ (left) and $u = 5\,\mathrm{km}$ (right), indicated in black.

After having set up our models, in the next section we initiate the investigation of the models with respect to a decisive quantity, the speed of propagation.

2.4 The speed of malware propagation

One of the most important characteristics of any malware propagation model is the speed with which the malware spreads in space. More precisely, we want to assume that the malware is present only at the origin at initial time and consider balls $B_u(o)$ of radius $u \geq 0$ centered at the origin o. Then, we define the *speed* of infection spreading as

$$\limsup_{u\uparrow\infty} u\mathbb{E}^o[1/\tau_u] = \alpha,$$

where $\tau_u = \inf\{t > 0 : I(t) \not\subset B_u(o)\}$ is the hitting time of the infection at the distance u. Here, $\mathbb{P}^o$ is the Palm distribution of the Cox point process based on the street system S given by a Poisson–Voronoi tessellation with edge-length intensity γ, see for example [1] for background on Palm distributions.

3 Results

3.1 Simulations for the SI models

We will approximate α by $\alpha_u = u\mathbb{E}^o[1/\tau_u]$ for some fixed radius u. In order to determine how fast the convergence is as u tends to infinity, we estimate α_u for different values of u. As mentioned earlier, it is reasonable to expect that $u \mapsto \alpha_u$ is decreasing, since boundaries of larger balls are more easily disconnected from the initially infected node in the network model. Another aspect that works in the same direction is that the spread of the malware should behave almost independently in different directions. But, for the speed of the malware, only the maximal distance in one of the directions is taken into account. While for larger radii the law of large numbers averages the growth over all directions, for small radii, the different speeds have a higher variance. Therefore, the expected maximum is larger for small radii, although the expected growth speed in a given direction is independent of the direction. In Figure 4 we present estimated curves $\lambda \mapsto \alpha_u(\lambda)$ for different values of u, which indeed show the expected behavior. In particular, we can observe indications for convergence since the graphs start to become closer for linearly-growing radii and linear dependence of α_u on λ.

Next we present the associated estimates for the non-Markovian dynamics, where the exponential waiting times with rate λ_{I} are replaced by uniform waiting times in the interval $[c, d]$ with $c = 40\,\mathrm{sec}$ and $d = 120\,\mathrm{sec}$. The results are presented in Figure 5. We can clearly observe a different behavior of the curves compared to the Markovian case. First, the convergence with respect to growing balls is faster, so that estimating α via α_u for large u seems to be less important. Second, and more interesting, there is no linear dependence on λ any more but rather a saturation effect for large λ. This saturation effect can be understood via the following observation. For denser and denser graphs, the malware is more and more able to achieve its maximal updating speed c. This c, properly scaled by r, then serves as an upper bound for the achievable speed. More precisely, the infection cannot spread faster than one hop every $40\,\mathrm{sec}$. With a connection radius of $300\,\mathrm{m}$, an upper bound for the maximum speed is given by $450\,\mathrm{m/min}$, when we ignore the additional distance given by the street system. Let us note that this additional distance can be related to the *stretch factor* of the Cox point process, see [23] or [24]. There are two additional effects that lead to the fact that this maximum speed is not achieved for finite λ. First, the maximum speed can only be obtained along a straight line. On the graph however, the stretch factor encodes the shortest distance in the graph towards a large radius, asymptotically. As the number of devices increases, the difference between euclidean and

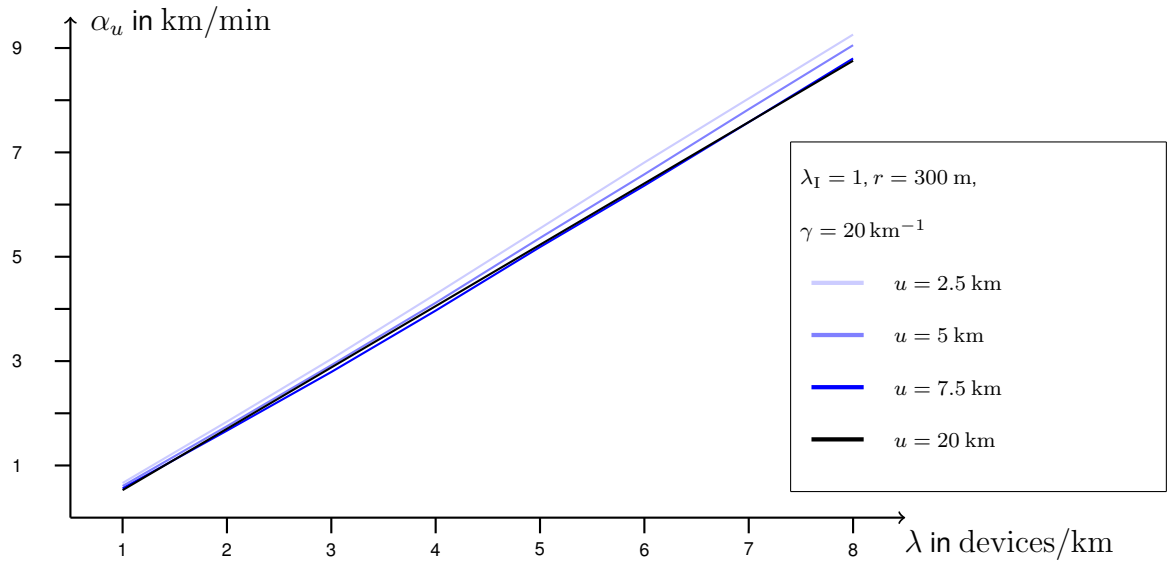

Figure 4: Simulations for the speed of malware propagation for the Markovian SI-model for growing observation windows as a function of λ.

graph distance encoded in the stretch factor becomes smaller. The second effect however is probably the dominating one for the shape of the curves in Figure 5. The maximal speed is achieved if at every step the minimal infection time is used. As the number of devices increase, so does the number of possible edges between devices (even quadratically). More available edges lead to a higher likeliness to sample an edge with a fast infection time.

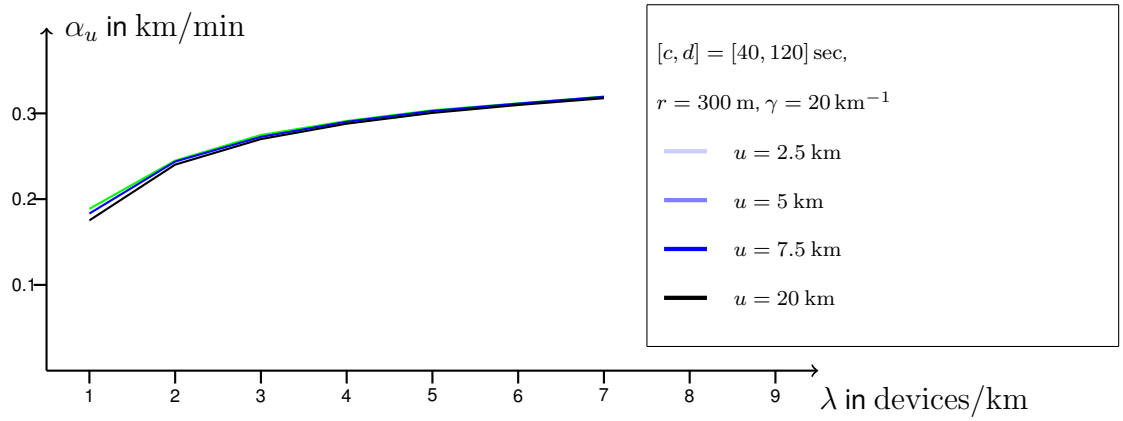

Figure 5: Simulations for the speed of malware propagation for the non-Markovian SI-model with uniform waiting times, for growing observation windows as a function of λ.

In order to further evaluate the convergence properties of our estimates, we track the associated approximate variance

$$\mathbb{V}^o[1/\tau_u] \approx n^{-2} \sum_{i=1}^{n} (1/\tau_u^i)^2 - \left(n^{-1} \sum_{i=1}^{n} 1/\tau_u^i\right)^2.$$

The corresponding curves for the Markovian setting are plotted in Figure 6 in blue. We observe a substantial decrease in variance for larger windows. This is not surprising since larger windows provide better averaging with respect to the waiting times. In other words, in smaller observation windows, less updating is involved which consequently creates more variance. At the same time, in larger observation windows it is even more unlikely to see clusters which contain the initial node where the malware starts and which are also large but do not reach the boundary of the window. Essentially, in large windows only the infinite cluster reaches the boundary of the window.

However, for larger values of λ the absolute variance increases, which is simply due to the fact that the variance measures the L^2 distance to larger expected values for larger λ. In order to see that larger values of λ lead to less deviation from the mean, in Figure 6 we also plot the relative deviation from the mean in red. The relative deviation is formally defined by

$$\mathbb{D}^o[1/\tau_u] \approx \frac{\sqrt{n^{-2}\sum_{i=1}^n (1/\tau_u^i)^2 - \left(n^{-1}\sum_{i=1}^n 1/\tau_u^i\right)^2}}{n^{-1}\sum_{i=1}^n 1/\tau_u^i}.$$

This decrease in the relative deviation is due to the fact that more nodes, encoded in larger values of

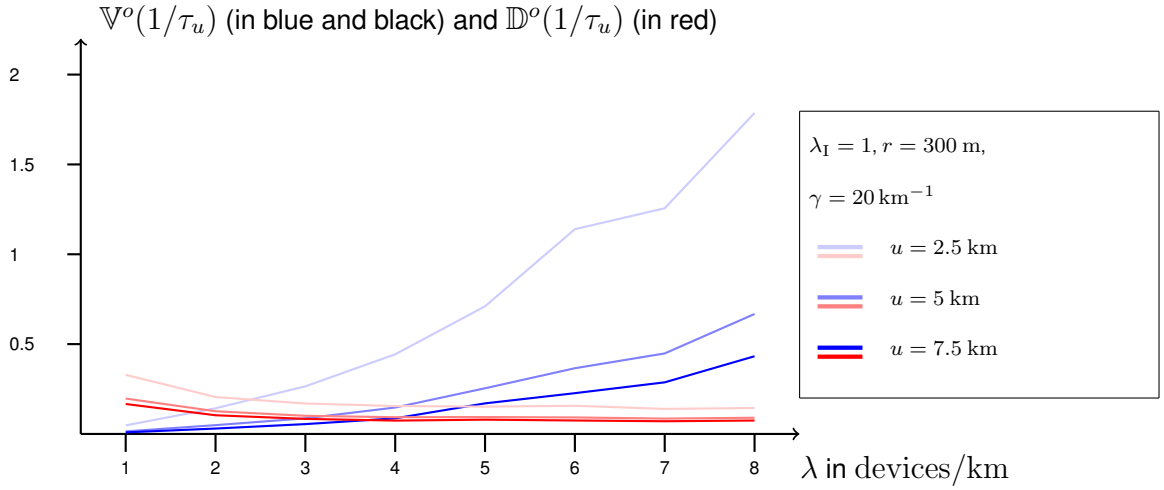

Figure 6: Simulations for the variances (in blue and black) and the relative deviations (in red) in the simulations of the speed of malware propagation for the Markovian SI-model for growing observation windows.

λ, provide more updates per unit volume, which results in more averaging, i.e., less deviation. Plus, larger λ refers to being deeper in the supercritical percolation regime for the network model, which also boosts the effect that we just described, namely that in large windows, for the boundary to be connected to the initial node by clusters other than the (unique) infinite one is highly unlikely.

Let us finally note that for the non-Markovian model the variances in all distances are already very small for λ not too close to the critical one, therefore we omit the associated plots.

3.2 Simulations for the SIG models

In this section we provide simulation results for the SIG-models, where the counter measure is given by white knights that are able to install patches exclusively on neighboring devices that carry the malware.

Our main objective is to present simulations for the phase diagram of survival and extinction of the malware. Here, we determine the critical intensity of white knights needed in order to eliminate the malware in the system for varying rates of the infection transmission. Let us again fix the parameters $\gamma = 20\,\mathrm{km}^{-1}$, $\lambda = 2\,\mathrm{devices/km}$ and $r = 300\,\mathrm{m}$.

3.2.1 The Markovian models

Let us start by giving an account on the Markovian models. Recall that here we can fix $\lambda_{\mathrm{G}} = 1$, the rate at which patches are installed on previously infected devices, due to time rescaling. In Figure 3 we illustrate the system stopped at two random times at which the malware has first reached the boundary of a certain prescribed centered ball.

Note that we can observe that in general the spreading of the malware can be divided into two main phases. In the first phase, as the white knights are random and maybe rare, the process has not discovered any white knights and spreads unobstructed like the SI-models. In the second phase, the first white knight has been discovered and a chase-escape dynamic unfolds in which the patch starts to spread on the previously infected nodes.

In the infinite volume, it is reasonable to assume, although not rigorously proven in our setting, that, for any value of intensity of the white knights ρ, there exists a *unique* critical infection rate $\lambda_{\mathrm{I}}^{\mathrm{c}}$ such that for $\lambda_{\mathrm{I}} > \lambda_{\mathrm{I}}^{\mathrm{c}}$ the infection survives for all times and infects infinitely-many devices, and for $\lambda_{\mathrm{I}} < \lambda_{\mathrm{I}}^{\mathrm{c}}$, the infection goes extinct on any infinite network component. Let us note that this result is available for d-regular trees, see [10]. In [15], for the case where Λ is the Lebesgue measure, we could rigorously determine areas in the phase diagram in which extinction, respectively survival, is present, but had to leave open the precise shape of the critical phase-separation line, and in fact also an area around that suspected line, due to lack of monotonicity. To determine this line via simulations in a much more realistic setting is one of the main results of this paper.

In order to determine the critical infection rate $\lambda_{\mathrm{I}}^{\mathrm{c}}$ for one given ρ, we have to cope with the fact that we simulate in bounded observation windows. More precisely, in the infinite domain we could simply define $\lambda_{\mathrm{I}}^{\mathrm{c}} = \inf\{\lambda_{\mathrm{I}} \colon \tilde{\alpha} > 0\}$, where $\tilde{\alpha}$ is defined as α but now under the measure $\mathbb{P}^o$ conditioned on connectedness of the origin to infinity. However, in finite windows survival is checked only within the window and thus survival inside the window could in principle still mean extinction in any larger domain. In Figure 7 we present estimates for the probability that the infection reaches varying distances u in finite time.

As could be anticipated, the curves become more pronounced in larger observation windows, since there we see the general principle that it is highly unlikely to survive a long time and then die out. In order to have a good balance between computational complexity and accuracy of the estimates, to find the critical intensity $\lambda_{\mathrm{I}}^{\mathrm{c}}$, we simulate large windows and check for which λ_{I} the proportion of realizations in which the malware reaches the boundary of the window exceeds 60%. For this note that it is computationally cheaper to run simulations in the extinction regime than in the survival regime. This is simply because in the extinction regime the infection dies out faster and hence our stopping criterion applies earlier. Therefore, in order to obtain the critical values, we run 1000 simulations (10 realizations of the propagation model on 100 realizations of the D2D-network model) until the probability to survive (in this case for the infection to reach a distance more than $17.5\,\mathrm{km}$ from the initially infected node) is larger than 60%. Note that the accuracy of the estimate of the critical value depends on the chosen discretization in the infection speed λ_{I} and the number of different white-knight intensities ρ that are simulated.

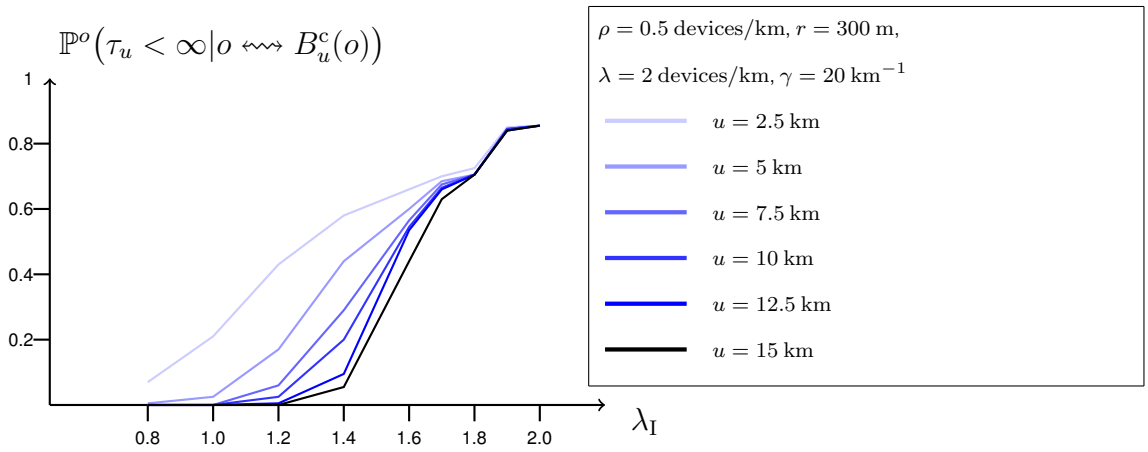

Figure 7: Estimates for the survival probability, conditioned to the event that a cluster exists connecting the initially infected node at the origin to the boundary of the observation window, for varying radii and in the Markovian setting.

Performing this step for multiple choices of ρ, we arrive at the simulated phase diagram presented in Figure 8. Let us mention that we simulated the critical infection speed for white-knight intensities

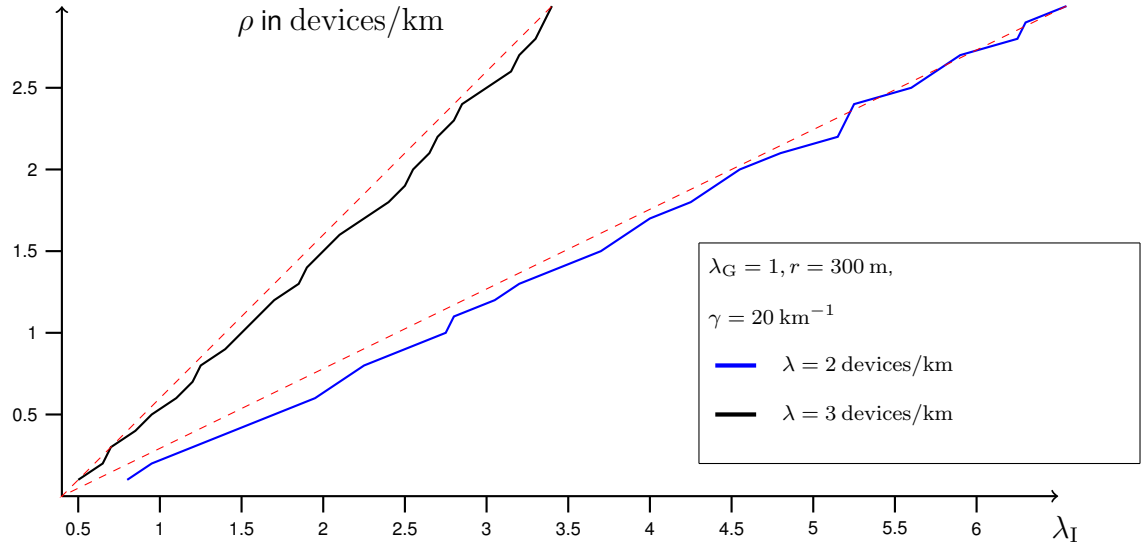

Figure 8: Estimated phase-separating curve for the survival of the malware in the Markovian SIG-model for two values of the linear device intensity. The dashed red lines indicate possible linear relationships. Below the line the malware is in the survival regime. Above the line, the malware becomes extinct.

ranging from 0.1 to 3 in 0.1 increments, i.e., 30 values of ρ and increased values of λ_{I} in steps of size 0.05 from 0 to 7. This would in principle lead to around 30×140 data points. However, we can use the property that for an increasing number of white knights, the critical infection rate has to increase as it becomes harder to spread. Therefore, new simulations for increasing ρ can use the previous critical value as a lower bound. This cuts down the number of simulations to roughly $30 + 140$.

The simulated phase-separating line resembles a straight line, indicating a linear dependence in the

function $\rho \mapsto \lambda_{\mathrm{I}}(\rho)$. One way to understand this effect may be to consider an infected device surrounded by white knights. In order for the malware to escape before the infected device receives the patch from the surrounding white knights, it has to be faster than the minimal patching event coming from the white knights. But for twice as many white knights, the first patch attempt is twice as fast and thus the malware should also be twice as fast in order to escape. This idea may serve as a rough guideline for the understanding of the linear relationship.

3.2.2 The non-Markovian models

The difference to the previous section is that now infection and immunization attempts are performed after non-exponential waiting times. As mentioned before we choose for the infection the uniform distribution $[c, d]$ with $c = 40\,\mathrm{sec}$ and $d = 120\,\mathrm{sec}$. Similarly, the waiting times for the patches are given by $[c, g]$ with the same $c = 40\,\mathrm{sec}$ and a $g > c$, which is potentially different from d. This allows us to incorporate in our simulations the situation in which the patches can be installed faster or slower than the malware transmits. In Figure 3 we illustrate the system stopped at two random times at which the malware has first reached the boundary of a certain prescribed centered ball, where we choose $g = d$.

As in the Markovian case, in order to simulate the phase diagram, we start by estimating the system for fixed white-knight intensity. Note that our propagation model has now the three parameters c, d, g instead of the two parameters $\lambda_{\mathrm{I}}, \lambda_{\mathrm{G}}$. By invariance with respect to time rescaling we can eliminate one of the parameters. In order to create a setting which makes it easier to compare the non-Markovian case to the Markovian case, we decided to keep c and d fixed and vary g, since then again, for larger g, the probability for survival grows. As in Figure 7, also in the non-Markovian case, for larger u, the curves become more pronounced.

Finally using the same approach as for the Markovian model we obtain the estimated phase diagram for the non-Markovian case, which we present in Figure 9.

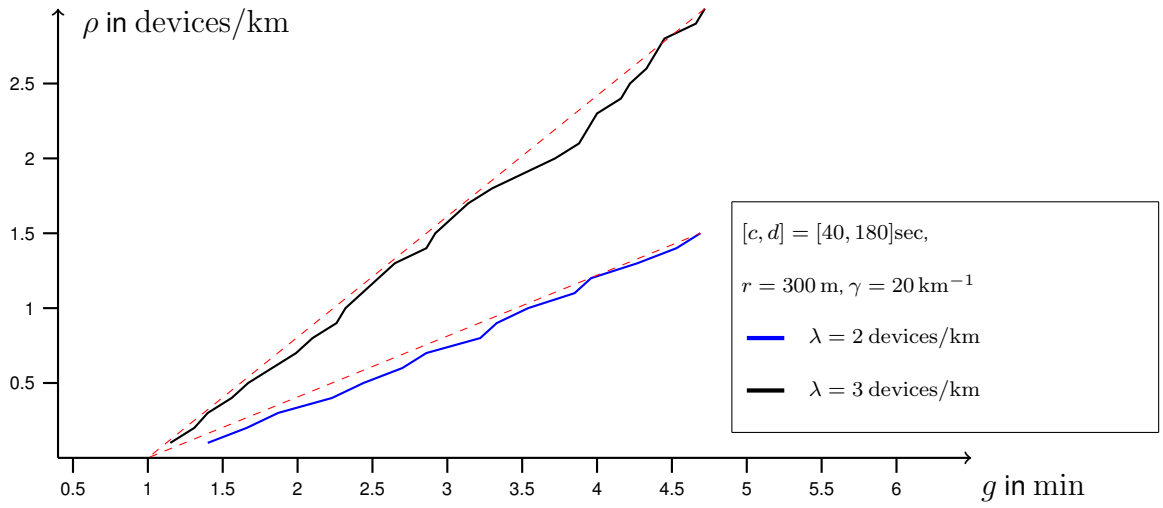

Figure 9: Estimated phase diagram for the survival of the malware in the non-Markovian SIG-model for two values of the linear device intensity. The dashed red lines indicate possible linear relationships. Above the line the malware is in the survival regime. Below the line, the malware becomes extinct.

Again, the simulated phase-separating line resembles a straight line, indicating a linear dependence

in the function $\rho \mapsto g(\rho)$. This prediction needs to be further investigated in future research.

3.2.3 The dependence on the street system

In this section we analyze the influence of the intensity of the street system on the critical infection speed $\lambda_{\mathrm{I}}^{\mathrm{c}}$ with respect to the intensity of white knights ρ. In order to better compare the critical behavior of the system, we keep the spatial intensity of devices fixed, i.e., as γ varies, we set $\lambda = c/\gamma$. Recall that as $\gamma \to \infty$ and $\lambda\gamma = c$, the distribution of devices converges towards a Poisson point process with intensity c, see for example [25].

In Figure 10 we present critical curves in the λ_{I}-ρ-plane corresponding to different values of γ where we fix λ such that $\lambda\gamma = 40\ \mathrm{devices/km}^2$. As in Figure 9, it is reasonable to believe that the curves are indeed linear and perturbations are due to simulation errors. We can observe a clear monotonicity, where for larger values of γ, for fixed λ_{I}, the critical intensity of white knights decreases. Note however that there is no reason to believe that the decrease in the gradient of the curves is linear in γ. We

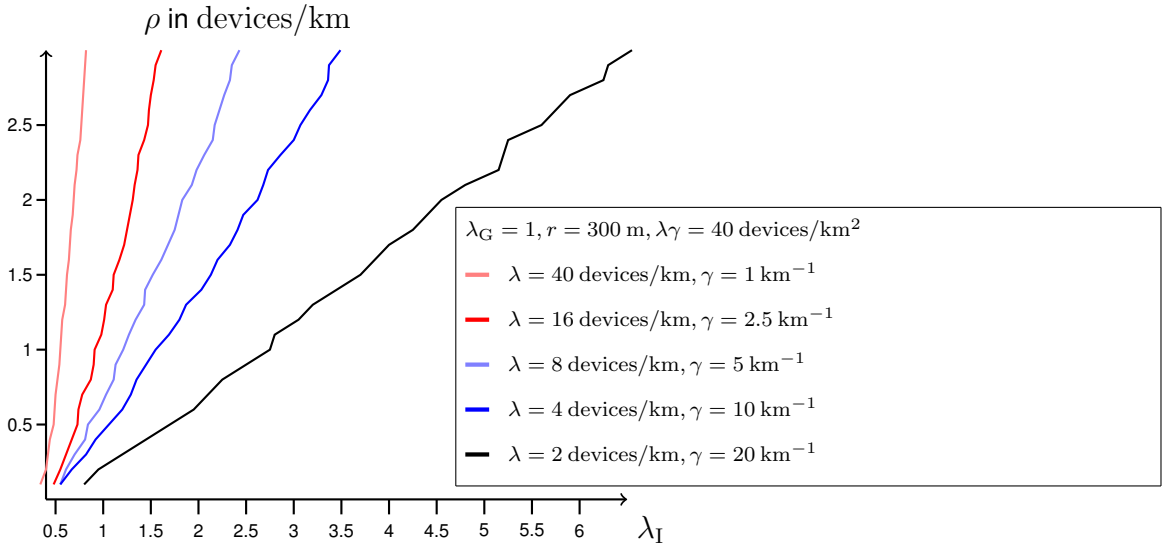

Figure 10: Estimated phase-separating curve for the survival of the malware in the Markovian SIG-model for two values of the linear device intensity. Below the line the malware is in the survival regime. Above the line, the malware becomes extinct.

believe that the average degree of the devices in the connection graph is the crucial factor for the gradient of the critical curves. If this assertion is valid, then it can help to explain the decreasing gradient of the critical curves. Indeed, note that for sparse street systems and a fixed spatial intensity of devices, the degree of the system is large, since many devices must be positioned on few streets, with empty space around them. On the other hand, for dense streets, the degree is similar to the degree of the planar Poisson point process. In this context, let us mention that, if we fix the connectivity threshold r and the linear intensity of devices λ and decrease the street intensity, the number of devices and therefore the average degree of device decreases. For $\gamma \to 0$, the limiting rural environment, this would even lead to disconnected graphs. This is the main reason why we couple λ to γ.

4 Conclusion

Based on a random configuration of devices placed in an urban street system of Poisson–Voronoi type, we first studied the spreading of some malware through the system without restriction. The malware is initially at the origin and infects neighboring devices after iid random waiting times. In the Markovian setting of exponential waiting times, we can clearly observe a linear dependence of the propagation speed on the device intensity. Here, the speed is measured in terms of the malware reaching large distances. We support our numerical findings by tests on the convergence in growing observation windows and estimates on the deviations. In the non-Markovian setting, where the waiting times are uniform in a compact time-interval strictly bounded away from zero, the picture changes substantially. The set of infected devices becomes much more ball-like in space and the linear dependence of the infection speed on the device intensity is replaced by a saturation effect, where more devices do not increase the speed any more.

Finally, we augmented the system by introducing a random set of white knights with intensity ρ at initial time. We simulated the critical phase-separating line of survival and extinction of the malware in the plane of ρ versus the infection rate λ_{I}, for different device intensities. Again, we support our findings by tests on the convergence in growing sampling windows. The main observation is that there seems to be a linear relation between λ_{I} and ρ, both in the Markovian and in the non-Markovian setting.

This leads to a number of open questions, subject to future investigations. For example: How fast is the convergence of the malware propagation speed for growing sampling windows? Is there a shape theorem for the set of infected devices? Can we determine precisely the phase diagram of survival and extinction? What is the dependence of the critical curve of survival and extinction on other model parameters, for example the device intensity?

References

[1] F. Baccelli and B. Błaszczyszyn, *Stochastic Geometry and Wireless Networks: Volume 1: Theory*. Now Publishers Inc, 2009.

[2] ——, *Stochastic Geometry and Wireless Networks: Volume 2: Application*. Now Publishers Inc, 2009.

[3] Y. Wu, W. Guo, H. Yuan, L. Li, S. Wang, X. Chu, and J. Zhang, "Device-to-device meets LTE-unlicensed," *IEEE Communications Magazine*, vol. 54, no. 5, pp. 154–159, 2016.

[4] B. Jahnel and W. König, *Probabilistic Methods for Telecommunications*, ser. Compact Textbooks in Mathematics. Birkhäuser, 2020.

[5] 3GPP, "Relay architectures for E-UTRA (LTE-Advanced)," *Evolved Universal Terrestrial Radio Access (E-UTRA)*, vol. TR 36.806, 2010.

[6] A. Asadi, Q. Wang, and V. Mancuso, "A survey on device-to-device communication in cellular networks," *Communications Surveys & Tutorials, IEEE*, vol. 16, no. 4, pp. 1801–1819, 2014.

[7] N. Valler, B. Prakash, H. Tong, M. Faloutsos, and C. Faloutsos, *Epidemic Spread in Mobile Ad Hoc Networks: Determining the Tipping Point*. Springer Berlin Heidelberg, 2011, pp. 266–280.

[8] L. Xie, X. Zhang, A. Chaugule, T. Jaeger, and S. Zhu, "Designing system-level defenses against cellphone malware," in *28th IEEE International Symposium on Reliable Distributed Systems*, 2009, pp. 83–90.

[9] M. Wang and Z. Yan, "Security in D2D communications: A review," *Proceedings of the 2015 IEEE Trustcom/BigDataSE/ISPA - Volume 01*, vol. 1, no. 1, pp. 1199–1204, 2014.

[10] G. Kordzakhia, "The escape model on a homogeneous tree," *Electron. Commun. Probab.*, vol. 10, pp. 113–124, 2005.

[11] C. Bordenave, "Extinction probability and total progeny of predator-prey dynamics on infinite trees," *Electron. Commun. Probab.*, vol. 19, 2014.

[12] I. Kortchemski, "A predator-prey sir type dynamics on large complete graphs with three phase transitions," *Stoch. Process. Their Appl.*, vol. 125, no. 3, pp. 886–917, 2015.

[13] ——, "Predator-prey dynamics on infinite trees: a branching random walk approach," *J. Theor. Probab.*, vol. 29, no. 3, pp. 1027–1046, 2016.

[14] R. Durrett, M. Junge, and S. Tang, "Coexistence in chase-escape," *arXiv preprint arXiv:1807.05594*, 2018.

[15] A. Hinsen, B. Jahnel, E. Cali, and J.-P. Wary, "Phase transitions for chase-escape models on Gilbert graphs," *arXiv preprint arXiv:1911.02622*, 2019.

[16] T. Courtat, "Promenade dans les cartes de villes-phénoménologie mathématique et physique de la ville-une approche géométrique," *These de doctorat de l'Université Paris-Diderot*, 2012.

[17] A. Hinsen, C. Hirsch, B. Jahnel, and E. Cali, "The typical cell in anisotropic tessellations," *arXiv preprint arXiv:1811.09221*, 2018.

[18] T. Liggett, *Interacting Particle Systems*, ser. Grundlehren der mathematischen Wissenschaften. Springer-Verlag, 1985.

[19] ——, *Stochastic Interacting Systems: Contact, Voter and Exclusion Processes*, ser. Grundlehren der mathematischen Wissenschaften. Springer Berlin Heidelberg, 2013.

[20] H. Kesten and R. Durrett, *Random Walks, Brownian Motion, and Interacting Particle Systems: A Festschrift in Honor of Frank Spitzer*, ser. Progress in Probability. Birkhäuser Boston, 2012.

[21] F. Spitzer, *Random fields and interacting particle systems: Notes on lectures given at the 1971 MAA Summer Seminar, Williams College, Williamstown, Massachusetts*, ser. Random Fields and Interacting Particle Systems. Mathematical Association of America, 1971.

[22] D. Richardson, "Random growth in a tessellation," in *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 74, no. 3, 1973, pp. 515–528.

[23] C.-L. Yao, G. Chen, and T.-D. Guo, "Large deviations for the graph distance in supercritical continuum percolation," *J. Appl. Probab.*, vol. 48, no. 1, pp. 154–172, 2011.

[24] E. Cali, N. N. Gafur, C. Hirsch, B. Jahnel, T. En-Najjary, and R. I. A. Patterson, “Percolation for D2D networks on street systems,” in *WiOpt*, 2018, pp. 1–6.

[25] C. Hirsch, B. Jahnel, and E. Cali, “Continuum percolation for Cox point processes,” *Stoch. Process. Their Appl.*, vol. 129, no. 10, pp. 3941–3966, 2019.