1.1 Probabilistic Methods for Communication Systems

Alexander Hinsen and Benedikt Jahnel

The steadily increasing demand for fast and reliable data exchange in communications systems presents network operators worldwide with major challenges, but also opportunities. A very important aspect of this state of affairs is the strongly increasing use of connected machines as part of the internet of things (IoT) as well as smart devices such as mobile phones, tablets, or even self-driving cars; see Figure 1 (left).

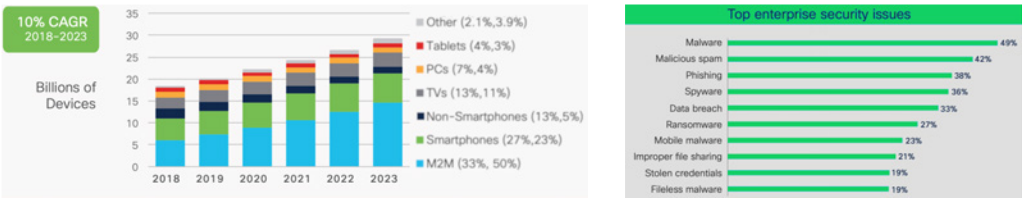

Fig. 1: *Left: Growth of communications driven by machine-to-machine and smartphone connections. Right: Network security breaches driven by malware attacks.* Cisco Annual Internet Report (2018–2023) White Paper

This situation is also reflected in the 5G (5th generation mobile network) specifications as well as in the negotiations for subsequent standards, which envisage faster connections, higher throughput, more capacity over enhanced mobile broadband, and highly reliable, low-latency communications to enable the system to support time-critical applications such as car-to-car communications as well as inter-machine connectivity.

In this context, *device-to-device (D2D) communications* is considered one of the key concepts that permeates a wide range of use cases. On the one hand, D2D systems have the potential to relieve today's cellular networks of at least some of the system pressure. On the other hand, D2D communications can provide, for example, faster and more robust connectivity. However, from an operator's perspective, D2D systems are much less controllable than traditional cellular networks, due to their dependence on individual user behavior. This lack of control is exacerbated when devices are *mobile* and the system is very dense due to the widespread use of connectable devices. Therefore, to correctly predict the performance and vulnerabilities of D2D systems, detailed and comprehensive modeling and analysis is essential. Here, a natural approach is to study the uncertainties of the system using *probabilistic methods*.

The starting point is the modeling of *random locations and movement* of the smart devices within their environment. Based on this information, the *transmission mechanism* between any pair of devices must be represented with an appropriate level of detail. Then, already the *connection times* in a sparse *cellular network* that is augmented by D2D communications, with a large but finite number of allowed D2D hops, is a significant performance indicator for a feasible D2D application; see below for more details in this direction.

A particularly relevant aspect of D2D networks is the *spread of malware* through the system. Due to the lack of central control, proximity-based sabotage software or viruses such as, for example, *Cabir, CommWarrior, or HummingBad* can potentially spread undetected in such networks. In this context, modeling and analyzing the unchecked infiltration of malware into D2D systems is already of immense relevance; see Figure 1 (right) and our results on limiting shapes of infected regions below. The natural next step then is to design and evaluate possible decentralized countermeasures against malware attacks, for which we present a short summary of our research at the end.

Stochastic geometry in telecommunications

In order to describe the system with all its uncertainties, we use methods from probability theory, more precisely from *stochastic geometry* [6]. In the first step, we associate to each device located at X_i an interaction zone given by an open disk $B_r(X_i)$ of radius r centered at X_i, where $2r$ is the range in which the device X_i can communicate in a peer-to-peer fashion. In other words, any pair of devices whose associated disks have nonempty intersection are connected by an edge, forming a random graph, the so-called *Gilbert graph* $\mathcal{G}_r(X)$, where $X = \{X_i\}_{i\in I}$ is the set of all device locations; see Figure 2 for an illustration. In absence of any refined statistical information on the spatial distribution of the devices, the null model is given by the stationary *Poisson point process*. This is a family of random point measures that enjoys strong spatial independence. It has one parameter, the spatial average density of the points. The associated *Poisson–Gilbert graph* is the fundamental object of the theory of *continuum percolation*, which investigates statistical properties of the *connected components* of the graph. In particular, it has been observed already in the early 1960s that there exists a phase transition in the density parameter for the almost-sure absence, respectively unique existence, of an infinite connected component $\mathcal{C}$ in $\mathcal{G}_r(X)$. This is the celebrated *phase transition of percolation*, which also has strong links to statistical physics (e.g., the probabilistic description of liquid-vapor phase transitions) and can be interpreted as an indicator for the D2D system to feature only local, respectively global, connectivity.

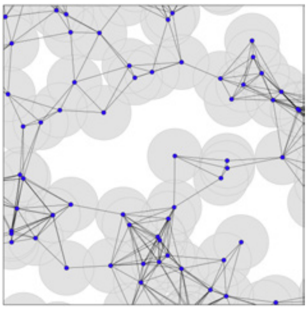

Fig. 2: *Realization of randomly placed devices (blue). Black edges are drawn whenever two devices have overlapping interaction zones (gray).*

The Poisson–Gilbert graph can serve as a model for a static pure D2D connectivity network, or as a snapshot for an otherwise *mobile system*. For example, considering the use case of car-to-car communications, the mobility of the nodes is of crucial importance for the system. Fortunately, *point-process theory* provides also a versatile framework for the modeling of mobile nodes, via the use of *markings*. More precisely, we can associate with each device location X_i an attribute, for example a trajectory Γ_i that represents the path of X_i. Then, the location of the device X_i at time $t \geq 0$ is given by $X_i(t) = X_i + \Gamma_i(t)$. Natural choices for these mobility models are, for example, independent and identically distributed (i.i.d.) *random walks* or *random-waypoint models* in continuous time and continuous space; see Figure 3 for an illustration. Mobile connectivity graphs now give rise to a new class of questions, for example, concerning the time at which a device makes contact with a large cluster, or the amount of time that it can communicate over large distances. Here is where the DYCOMNET group within WIAS makes contributions; see the following section for details.

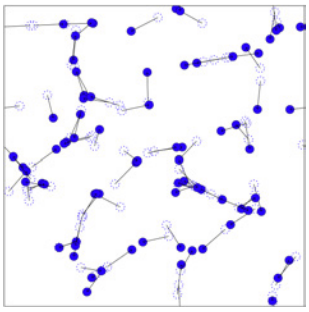

Fig. 3: *Realization of initial device positions (dotted blue) and their respective positions at a fixed positive time (blue), with arrows indicating the corresponding displacement.*

On the other hand, in connectivity graphs such as the Poisson–Gilbert graph, edges represent the possibility to transmit data from device to device, but the actual flow of messages is not represented. In order to incorporate this flow, a standard modeling approach is known under the name *first-passage percolation*, which plays also a big role in the probabilistic analysis of space-time epidemiological events. Here, a *passage time* τ_e is associated to every edge e in the graph. If, at time zero, a message is placed at a vertex X_i; then, the set of vertices $\mathcal{H}_t(X_i)$ that have received the message by some positive time $t > 0$, is given by

$$\mathcal{H}_t(X_i) = \Big\{X_j \in X\colon \exists \text{ path } \gamma \text{ in } \mathcal{G}_r(X) \text{ starting in } X_i \text{ and ending in } X_j, \text{ such that } \sum_{e\in\gamma} \tau_e \le t\Big\};$$

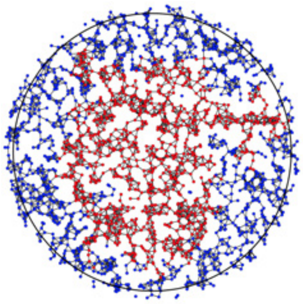

Fig. 4: Realization of first-passage percolation on a Gilbert graph at some positive time. Vertices in $\mathcal{H}_t(o)$ are indicated in red.

see Figure 4 for an illustration. Note that the set $\mathcal{H}_t(X_i)$ cannot be larger than the connected component of X_i in $\mathcal{G}_r(X)$. In the case that the connected component of X_i is infinite, then questions about the asymptotic speed of data propagation, the limiting geometry of the reachable device set, and properties of the shortest paths (geodesics) become highly relevant, but nontrivial. One particular use case is given by the propagation of malware in pure D2D systems with and without the presence of counter-measures. Also here, the DYCOMNET group performs research within WIAS, and we present details in the last section.

Connection intervals in mobile D2D networks

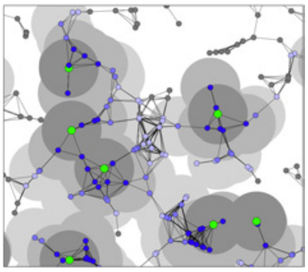

Fig. 5: Realization of infrastructure nodes (green) and devices (blue and grey). The grey areas indicate the k-hop coverage zones of the infrastructure where $k = 1$ is dark grey, $k = 2$ is grey, and $k = 3$ is light grey. Correspondingly, dark blue devices are connected to the infrastructure directly, blue devices need one intermediate hop, light blue ones need two hops, and grey devices need at least three hops.

Although pure D2D systems are already in use today, for example in sensor networks or disaster-rescue ad-hoc networks, in the foreseeable future, D2D systems will be mainly deployed as an extension to more traditional cellular networks. We thus consider a homogeneous Poisson point process $Y = \{Y_j\}_{j\in J}$ of *infrastructure nodes* with intensity $\lambda_S > 0$ in addition to the marked Poisson point process of *mobile nodes* $X(t)$ as presented above. We are interested in the times at which a typical device X_o is connected to the infrastructure in at most k hops

$$\Xi_k = \{t \in \mathbb{R}\colon X_o(t) \overset{k}{\rightsquigarrow}_t Y\},$$

where $\overset{k}{\rightsquigarrow}_t$ means that a connection is possible with at most k hops in $\mathcal{G}_r(X(t) \cup Y)$; see Figure 5 for an illustration. Our main interest lies in the distribution of *connection intervals* of the typical device. For this problem, we consider the length of maximal uninterrupted connection time intervals around a given time t

$$I(t, \Xi_k) = \sup_{a\le b\colon t\in[a,b]\subset\Xi_k} (b - a),$$

where $I(t, \Xi_k) = 0$ if $t \notin \Xi$. Using this formula, we define the k*-hop connection-interval measure*

$$\tau_T(\mathrm{d}\ell, \mathrm{d}t) = \frac{1}{T}\int_{[0,T]\cap\Xi_k} \delta_{(I(s,\Xi_k),s/T)}(\mathrm{d}\ell, \mathrm{d}t)\mathrm{d}s,$$

where δ is the Dirac measure. Here, T is the time horizon. Let us highlight that τ_T encodes a number of important network characteristics. For example, the integral $\tau_T(f)$ for $f(\ell, t) = 1$ is the *time-averaged connection time of the typical node*, and $T\tau_T(f)$ for $f(\ell, t) = 1\{\ell > 0\}/\ell$ is the *number of connection intervals* in $[0, T]$. We analyze τ_T in an asymptotic regime of large time horizons $T \uparrow \infty$, many hops $k \uparrow \infty$, and sparse infrastructure $\lambda_S \downarrow 0$, coupled as

$$\lambda_S(T)|B_{k/\mu}(o)| = c \qquad \text{and} \qquad \lambda_S(T) = T^{-\alpha}, \tag{1}$$

with parameters $\alpha, c > 0$, and where $\mu > 0$ is the so-called *stretch factor* of the Poisson–Gilbert graph in the supercritical percolation regime. This is the asymptotic quotient of the graph distance and the Euclidean distance between two distant sites in $\mathcal{C}$. Note that the constant c can be interpreted as the expected number of infrastructure nodes in the reachable region.

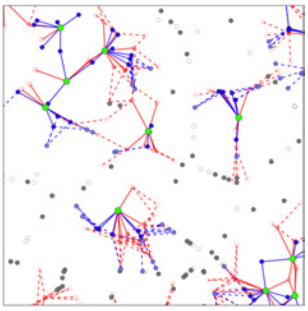

Fig. 6: *Realization of devices (transparent, grey, and blue) and infrastructure nodes (green). Devices at initial time (transparent) are either directly connected (solid red edges) or indirectly connected (dashed red edges) to infrastructure, or they are disconnected. At time* 1*, the devices have moved and their corresponding new locations offer either direct connections (solid blue edges) or indirect connections (dashed blue edges) to the infrastructure process, or no connections (grey).*

For the mobility scheme, we assume the trajectories Γ_i to be i.i.d. random walks, starting from zero, i.e., devices sequentially move with constant speed to random waypoints that are independently drawn from an isotropic probability measure $\kappa(\mathrm{d}x)$; see Figure 6 for an illustration. Now, our main results in [4, 5] establish distributional limits of τ_T, as $T \uparrow \infty$ under the scaling (1), for three different regimes that depend on α. First, for $\alpha < d/2$, we have

$$\tau_T(\mathrm{d}\ell, \mathrm{d}t) \xrightarrow{\mathrm{D}} \mathbb{E}[\delta_{I_o(N)}(\mathrm{d}\ell)]\mathrm{d}t,$$

where $I_o(N) = I(0, \Xi_\infty \cap (\cup_{j \le N} \Xi_\infty^{o,j}))$ with N an independent Poisson random variable with intensity c. Here, $(\Xi_\infty^{o,j})_{j \ge 1}$ is a family of i.i.d. copies of Ξ_∞^o where

$$\Xi_\infty^o = \{t \in \mathbb{R} \colon o \leftrightsquigarrow_t \infty\} \qquad \text{and} \qquad \Xi_\infty = \{t \in \mathbb{R} \colon X_o(t) \leftrightsquigarrow_t \infty\},$$

the events that the origin o, respectively the typical device, is part of the infinite cluster of $\mathcal{G}_r(X(t) \cup \{o\})$, respectively $\mathcal{G}_r(X(t))$. In words, in the regime of (relatively) dense infrastructure, the k-hop connection-interval measure approaches (in the spirit of a law of large numbers) a product measure that is given in terms of an expectation over interval lengths in which both, the typical device and at least one reachable infrastructure node, are part of the infinite cluster. On the other hand, for (relatively) sparse infrastructure, where $\alpha > d/2$, using the same definitions,

$$\tau_T(\mathrm{d}\ell, \mathrm{d}t) \xrightarrow{\mathrm{D}} \mathbb{E}[\delta_{I_o(N)}(\mathrm{d}\ell)|N]\mathrm{d}t.$$

Again in words, in this regime, there is less averaging, and the number of reachable infrastructure nodes remains random in the limit. Finally, in the critical regime $\alpha = d/2$, we see the emergence of a standard Brownian motion W_t, and

$$\tau_T(\mathrm{d}\ell, \mathrm{d}t) \xrightarrow{\mathrm{D}} \mathbb{E}[\delta_{I_o(Y'(B_{c'}(W_t)))}(\mathrm{d}\ell)|Y'(B_{c'}(W_t))]\mathrm{d}t,$$

where $c' = (c/|B_1(o)|)^{1/d}$, and Y' is a unit-intensity homogeneous Poisson point process. In this case, even the random number of reachable infrastructure nodes around the limiting trajectory of the typical random walker survives the limit.

Using various degrees of asymptotic decoupling in the highly detailed k-hop connection-interval measure τ_T, our results show that τ_T can be well approximated by much simpler connection-interval measures given in terms of expectations over percolation clusters.

Malware propagation in random geometric graphs

Let us now revisit the static Poisson–Gilbert graph and consider first-passage percolation as introduced above. We are interested in the behavior of $\mathcal{H}_t = \mathcal{H}_t(q(o))$, where $q(o)$ is the closest point of the origin in the unique infinitely large connected component $\mathcal{C}$. The version of first-passage percolation in which the passage times are i.i.d. exponential random variables with parameter $\rho > 0$ is called the *Richardson model* (on the Poisson–Gilbert graph), and we write $\mathcal{H}_t^{\lambda,\rho}$ to indicate both, the intensity λ of the underlying Poisson point process as well as the parameter ρ. Our first result from [1] establishes weak convergence of the paths $\mathcal{H}_{[0,t]}^{\alpha\lambda,\rho/\alpha}$ (with respect to the

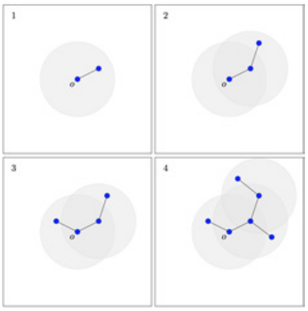

Fig. 7: *Illustration of $\mathcal{T}^{\lambda,\rho}$ at 4 increasing time steps. Gray disks indicate the area for possible offsprings.*

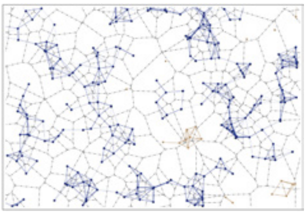

Fig. 8: *Realization of a Poisson–Voronoi tessellation with respect to the large connected component of a Poisson–Gilbert graph*

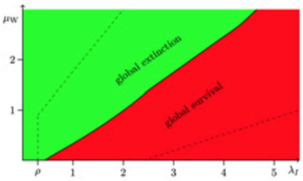

Fig. 9: *Phase diagram for global survival and extinction based on the infection rate $\lambda_{\rm I}$ and the white-knight intensity $\mu_{\rm W}$. The solid line is based on simulations, and the dashed lines indicate rigorous bounds. The constant ρ represents a threshold below which any positive white-knight intensity suffices to eliminate the malware on infinite clusters.*

Skorokhod topology based on the vague topology) towards a limiting branching process $\mathcal{T}^{\lambda,\rho}_{[0,t]}$, in the limit as α tends to infinity. This can be seen as high-density limit for devices with slow transitions. The limiting process $\mathcal{T}^{\lambda,\rho}_{[0,t]}$ has an initial device at the origin and then iteratively produces offsprings after independent exponential waiting times with parameter $|B_r(o)|\lambda\rho$. Here, the offsprings are uniformly distributed in the ball with radius r centered at the parent device; see Figure 7 for an illustration.

The main result in [1] is a *shape theorem* for $H_t = \bigcup_{x\in\mathcal{H}_t} V(x,\mathcal{C})$, the union of *Voronoi cells* $V(x,\mathcal{C})$ associated with points in $\mathcal{H}_t$, taken with respect to the infinite cluster $\mathcal{C}$; see Figure 8 for an illustration. The shape theorem can be understood as a spatial strong law of large numbers, i.e., almost surely,

$$\lim_{t\uparrow\infty} \frac{1}{t} H_t = B_\phi(o),$$

where $0 < \phi < \infty$ is a nontrivial speed constant. In words, the set of space points that are closest to a device that is reached at a time t by a message initially placed close to the origin, approaches a ball with radius given by $t\phi$. Note that, in order to avoid percolation already at time zero, it suffices to require that $\mathbb{P}(\tau = 0) < (\lambda|B_r(o)|)^{-1}$, the inverse of the expected degree of a typical node in the Poisson–Gilbert graph. On the other hand, in order to control fluctuations, we also have to assume that $\mathbb{E}[\tau^\eta] < \infty$ for some $\eta > 2d+2$, but otherwise the distribution of τ is arbitrary; however, note that it influences the speed ϕ. The main ingredients in the proof are a good control on the length of shortest paths in the graph and subadditivity arguments.

Let us finally report also on results for the propagation of malware in the supercritical Poisson–Gilbert graph (and more refined *Cox–Gilbert graphs*, i.e., Poisson–Gilbert graphs in random environments) in the presence of a counter-measure, as presented in [2, 3]. We consider the Richardson model on the Poisson–Gilbert graph in which a typical device carries a malware at initial time. In addition to regular susceptible devices, at initial time, there is also an independent Poisson point process of special devices called *white knights* in the system. Now, white knights carry a patch that eliminates the malware, but this patch can only be transferred to devices that are infected (due to privacy regulations) and not to susceptible devices. Once patched, the device also becomes a white knight and, on the long-term run, we observe a competition between an escaping malware spreading and a chasing patch. Allowing for different transmission rates for the malware and the patch, we see various behavioral regimes depending on four parameters, the different rates and different initial intensities of susceptible devices and white knights. Our main findings in [2] (based on rigorous arguments) and in [3] (mainly based on simulations) concern phase transitions of global and local survival and extinction of the malware as exemplified in Figure 9. In a nutshell, sufficiently large white-knight intensities or patching rates lead to global extinction of the malware, whereas sufficiently large intensities of susceptible devices or infection rates guarantee positive probabilities of a global survival of the malware.

References

[1] C. Coletti, L. de Lima, A. Hinsen, B. Jahnel, D. Valesin, *Limiting shape for first-passage percolation models on random geometric graphs*, arxiv Preprint no. 2109.07813, 2021.

[2] A. Hinsen, E. Cali, B. Jahnel, J.-P. Wary, *Phase transitions for chase-escape models on Poisson–Gilbert graphs*, Electron. Comm. Probab., **25**:25 (2020), pp. 1–14.

[3] ———, *Malware propagation in urban D2D networks*, in: IEEE 18th International Symposium on on Modeling and Optimization in Mobile, ad Hoc, and Wireless Networks, (WiOpt), Volos, Greece, Institute of Electrical and Electronics Engineers (IEEE), 2020, pp. 1–9.

[4] C. Hirsch, B. Jahnel, E. Cali, *Percolation and connection times in multi-scale dynamic networks*, arxiv Preprint no. 2103.03171, 2021.

[5] ———, *Connection intervals in multi-scale dynamic networks*, arxiv Preprint no. 2111.13140, 2021.

[6] B. Jahnel, W. König, *Probabilistic Methods for Telecommunications*, Compact Textbooks in Mathematics, D. Mazlum, ed., Birkhäuser Basel, 2020, 200 p.