Chapter 9
Leveraging Topological and Temporal Structure of Hospital Referral Networks for Epidemic Control

Vitaly Belik, André Karch, Philipp Hövel, and Rafael Mikolajczyk

Abstract Antimicrobial-resistant pathogens constitute a major threat for health care systems worldwide. The hospital-related pathway is a key mechanism of their spread. Contrary to intra-hospital transmission data that requires sophisticated contact tracing technologies, data on inter-hospital transmission is collected on a regular basis. We investigate the dataset of patient referrals between hospitals in a large region of Germany. This dataset contains approximately one million patients over a 3-year period. The dataset is used to build a dynamic network of hospitals where nodes are hospitals and edges represent movements of patients between them. We consider the worst-case scenario of a highly contagious disease corresponding to deterministic infection dynamics. Furthermore, we investigate the impact on epidemic processes of the correction to the temporal network due to home (or community) visits of possibly contagious patients returning to hospitals. Moreover, we implement an extensive stochastic agent-based computational model of epidemics on this network. By leveraging the topological and temporal network structure for epidemic control, we propose intervention schemes able to hinder spread. Our approach can be used to design optimal control strategies for containment of nosocomial diseases in health-care networks.

V. Belik (✉)
System Modeling Group, Institute for Veterinary Epidemiology and Biostatistics, Freie Unversität Berlin, Königsweg 67, 14163, Berlin, Germany
e-mail: vitaly.belik@fu-berlin.de

A. Karch • R. Mikolajczyk
Helmholtz Center for Infection Research, Inhoffenstraße 7, 38124, Braunschweig, Germany
e-mail: andre.karch@helmholtz-hzi.de; rafael.mikolajczyk@helmholtz-hzi.de

P. Hövel
Institut für Theoretische Physik, Technische Universität Berlin, Hardenbergstraße 36, 10623, Berlin, Germany
e-mail: phoevel@physik.tu-berlin.de

N. Masuda, P. Holme (eds.), *Temporal Network Epidemiology*,
Theoretical Biology, DOI 10.1007/978-981-10-5287-3_9

9.1 Introduction

Nosocomial or healthcare-associated infections are a significant mortality and morbidity factor in Europe and across the globe [1, 2]. As an additional challenge, many of these are caused by pathogens which are drug-resistant. Annually, up to 700.000 deaths worldwide can be attributed to antimicrobial resistances [2]. The administrative data on patient hospital admission and discharge constitute referral patterns and is routinely collected by healthcare providers. This data could be used to identify factors facilitating the spread of nosocomial diseases in a healthcare system. For this purpose it is important to consider patient-resolved data, because the identity and causal order of movement events have significant implications for the spreading dynamics [3–6]. From the data the underlying hospital referral network could be reconstructed, with vertices being hospitals and edges being movements of patients between hospitals. The patient movements occur only on some days and thus the edges of the referral network appear and disappear on daily basis. Such networks changing form one time instance to another are known as temporal or dynamic [7, 8].

Note that some aspects of epidemics on hospital referral networks from different countries were considered in Refs. [9–13]. One crucial assumption in such models concerns the disease-free status of patients discharged from hospitals, which may not hold. If such patients return to hospitals, still carrying the pathogen, they may facilitate the further spread of pathogens. Another frequently neglected aspect are temporal and topological correlations [14–16]. In Germany only one small regional network of hospitals was considered on a descriptive level without modeling epidemic spread on it [17]. In the present study we consider referral patterns in a big region in Germany with almost one million patients over 3 years and investigate epidemics on the corresponding network.

In this chapter, we begin by discussing structural properties of the static and the temporal representations of the referral network of hospitals. Then, we present results on the analysis of generic deterministic worst-case spreading phenomena (SI and SIR epidemics) in the static and the temporal frameworks. Finally, we present results of extensive numeric simulations of an endemic disease (modeled as SIS epidemics) and evaluate the effect of various control measures.

9.2 Dataset on Hospital Referrals

From patient referral data we extract a network with vertices being hospitals and edges between hospitals corresponding to direct relocations of patients between two hospitals. One should note, that direct relocations between hospitals correspond only to approximately 3% of referrals in the system. The rest are relocations from or into the community (non hospital whereabouts of patients). We consider only relocations

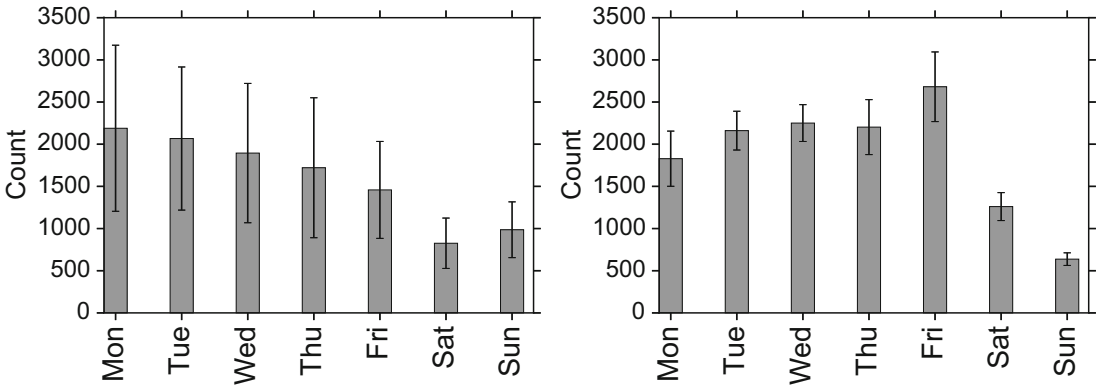

Fig. 9.1 The average number of referral events for every day of the week. *Left panel*: admission frequency. *Right panel*: discharge frequency. Error bars correspond to the standard deviation

between two hospitals on a single day, which is the majority of cases.[1] The dataset was preprocessed to exclude overlapping stays in different hospitals and overlapping stays in a single hospitals were merged together.

9.2.1 Referral Patterns

The dataset spans from the 1st of January 2009 until the 31st of December 2011. Each data record corresponds to a hospital stay and includes the day of admission t^{in}, the day of discharge t^{out}, the anonymized hospital ID and the anonymized patient ID. In our dataset we have 2,037,460 records for 917,834 individuals. The dataset contains patients that were in the system on the 1st of January 2009 (9,874 individuals). The first admission date mentioned in the dataset is the 30th of November 2005. Admission patterns manifest strong temporal regularities (Fig. 9.1). The number of admissions is maximal on Monday (around 2200) and continuously decreases until it reaches a minimum around 700 on Saturday. The number of discharges is minimal on Sunday (around 600), increases until Wednesday, has a small decrease on Thursday and reaches a maximum on Friday (around 2500).[2]

There were 1654 hospital IDs in the data. However because the data comes from a major insurance provider of the federal state under consideration (Lower Saxony) it makes sense to restrict ourselves only to hospitals located in Lower Saxony. This is done by considering only hospitals with a maximal number of patients per hospital per day larger than 30 (as estimated from the data). This results in 185 vertices (Fig. 9.2, see also Ref. [18] for justification of the procedure).

[1] However there were around 300 patients which were apparently transferred between 3 hospitals in one single day. We exclude those from the network reconstruction.

[2] This weekly dynamics will be also reflected in the epidemic dynamics (Fig. 9.6).

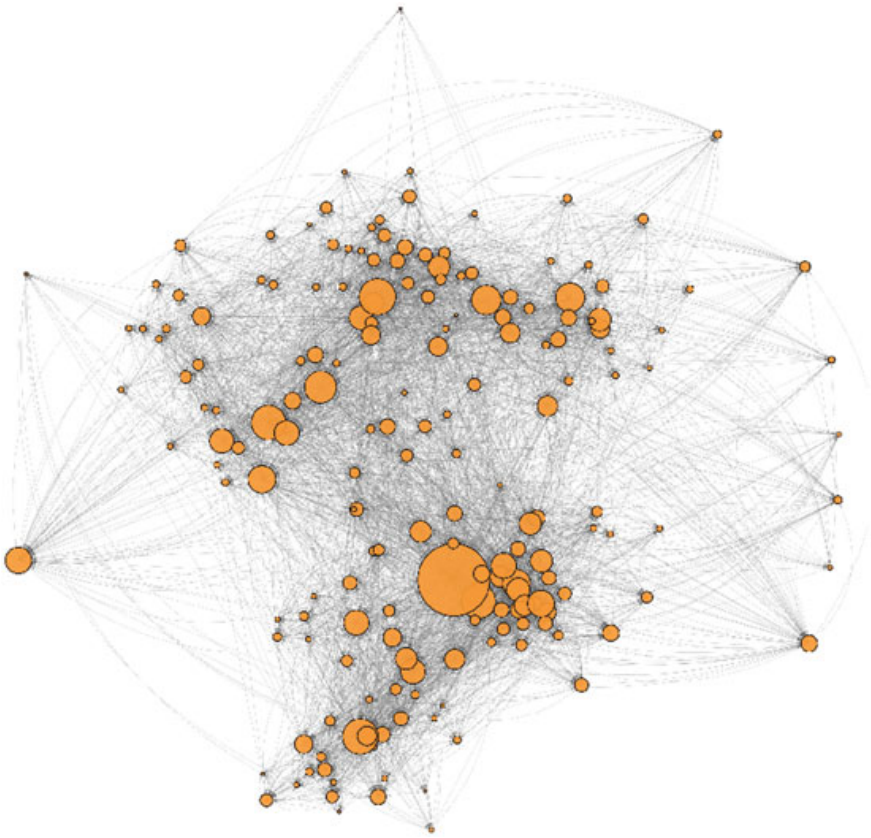

Fig. 9.2 Visualization of the aggregated network of hospitals obtained by considering direct movements of patients between hospitals as edges. The size of vertices is proportional to their in-degree

9.2.2 Network Properties

If we consider the whole available time span and aggregate the temporal edges, we obtain an unweighted static network. In this case, a static edge between two vertices is present, if a temporal edge occurred at least once. In the resulting static network there are 4,949 directed edges. The average in/out-degree is 27 ($\pm$17 and $\pm$13 respectively, $\pm$ denotes standard deviation) with the diameter equal to 3. In Fig. 9.3 the distributions of in- and out-degrees are presented. Note that the in-degrees show a less heterogeneous distribution, than the out-degrees. The chance to be taken out of a hospital into another one is more less heterogeneouly distributed than the chance of being admitted to a hospital from another hospital – there are just a few major hospitals admitting patients from other hospitals. Note that degree distributions resemble an exponential distribution characteristic of random Erdős-Rényi networks (in the limit of a large network size).

So far we considered the aggregated network of hospitals, where we were not concerned with the temporal order of edges. Now we consider a temporal network and review its basic temporal properties.

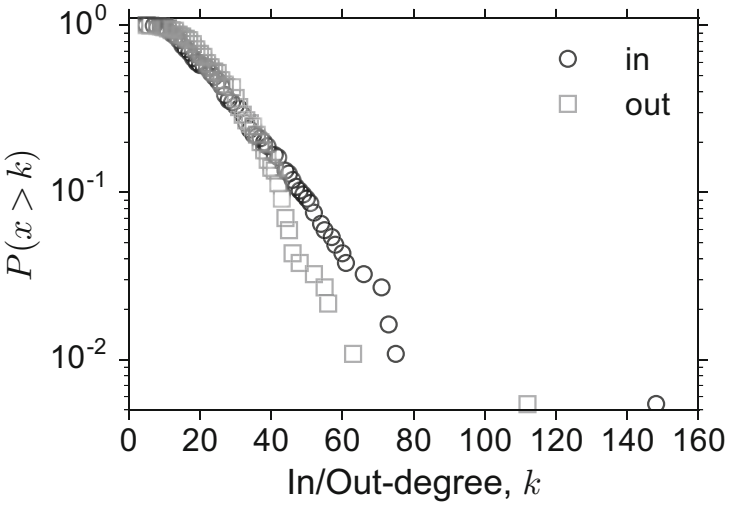

Fig. 9.3 The complementary cumulative probability distribution function of in-degrees (*circles*) and out-degrees (*squares*). Note that the in-degrees are more broadly distributed than the out-degrees. This reveals a hierarchy of hospitals in the healthcare network with a few large broadly specialized clinics receiving the majority of patients

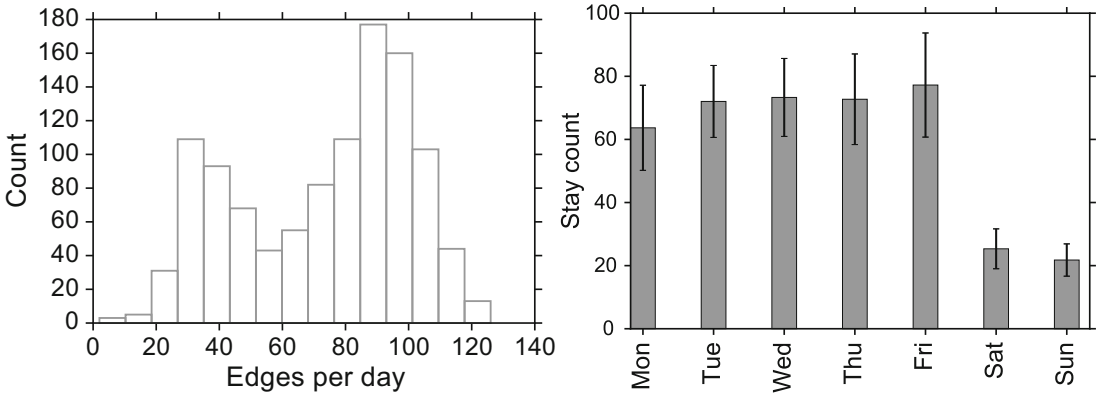

Fig. 9.4 Daily properties of the temporal referral network of hospitals. Histogram of the number of edges per day (*left*). The bimodal distribution corresponds to the small number of edges (around 30) on weekends and large number of edges (around 90) during the rest of the week. This becomes clear from the plot of the average number of edges versus day of the week (*right*). Error bars correspond to standard deviations

After data preprocessing we obtain about 67,000 temporal edges (edges with the corresponding timestamps of their occurrence) for 1,099 days. On average there were 61 ± 22 edges per day. However, in the distribution of the daily number of edges there are two peaks attributable to particular week days (Fig. 9.4).

We define the activity of an edge $j \to i$ as the number of its occurrences ω_{ij} in a temporal network. In a directed graph an edge could be an incoming edge for a recipient node or an outgoing edge for a donor node. The number of occurrences (or activity) of incoming/outgoing edges for a node reads $a_i^{\text{in}} = \sum_j \omega_{ij}$ and $a_i^{\text{out}} = \sum_j \omega_{ji}$ respectively. Distributions of ω_{ij}, a_i^{in}, and a_i^{out} over the whole time span are presented in Fig. 9.5. Concerning activity of incoming/outgoing edges, we see the picture similar to the in/out-degree distribution (Fig. 9.3) – there are less vertices (hospitals) with the high recipient activity. Activity of outgoing edges is distributed more homogeneously. As it could be seen from the semi-logarithmic plot, the activity distributions resemble the exponential distribution.

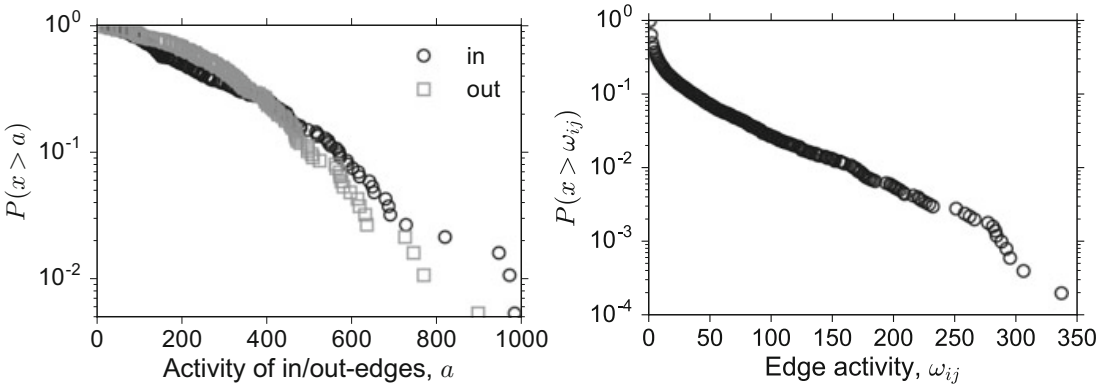

Fig. 9.5 *Left*: The complementary cumulative distribution of the node activity a_i^{in} and a_i^{out} over the whole time span of outgoing ("out") and incoming ("in") edges of a node respectively. *Right*: The complementary cumulative distribution of the activity of directed edges ω_{ij} over the whole time span of a node

9.3 Epidemic Dynamics

In the next sections we investigate on epidemic dynamics on the network of hospitals. First, we examine the network of hospitals as a directed contact network, considering a hospital as a single unit being in one of the susceptible, infected or recovered states. Second, we take into account the detailed referral patterns of single individuals, but assume the well-mixed approximation for the stochastic infection dynamics within a single hospital.

9.3.1 Deterministic SI Model

First of all we consider a deterministic (corresponding to the worst-case scenario) disease with no recovery or an infinite infectious period. Such an SI (susceptible-infected) process has the kinetics

$$S + I \to 2I,$$

where a susceptible node becomes immediately infected upon contact (via temporal edge) with an infectious node.

The spread of a deterministic SI process during t days in a static network could be described by the reachability or accessibility matrix

$$\boldsymbol{P}_t = \bigcup_{n=1}^{t} A^n, \tag{9.1}$$

where A denotes the adjacency matrix[3] of the static network and $\cup$ denotes the Boolean operator OR (a Boolean analog of the matrix multiplication). Elements of this matrix are non zero if there is a path from the vertex i to the vertex j — a connected sequence of edges, where the target vertex of the previous edge ($i \to j$, $i \neq j$) is the start vertex of the next one ($j \to k$, $j \neq k$). Equation (9.1) is easy to understand if we recall, that all possible paths up to the length t are given by $\sum_{n=1}^{t} A^n$. Equation (9.1) is just the Boolean version of the last relation. Analogously the spread of a deterministic SI process on a temporal network could be described by the temporal accessibility matrix

$$\mathcal{P}_t = \bigcap_{n=1}^{t} (\mathbb{1} \cup A_n),$$

where $\cup$ is the Boolean operator AND (a Boolean analog of the matrix addition), $\mathbb{1}$ is the identity matrix and A_n is the adjacency matrix for the n-th snapshot (on the n-th day) of the temporal network [19]. The temporal accessibility contains all possible spreading paths of duration less or equal to t of a deterministic infection with the infinite infectious period started at all vertices. The elements of the temporal accessibility matrix are non-zero, when there is a time respecting path from the vertex i to the vertex j — a connected sequence of temporal edges, where the target vertex of the previous edge ($n : i \to j$, $i \neq j$) is the start vertex of the next one ($n + \tau : j \to k, j \neq k, 0 < \tau \leq t$).

The density $\rho(\mathcal{P}_t)$ (fraction of non-zero elements) of the accessibility matrix gives a cumulative distribution F_t of the number of shortest paths of duration less or equal than t between any two nodes [19]. Thus the difference between two successive values of the cumulative distribution $F_t - F_{t-1}$, $F_0 \equiv 0$ gives us the probability distribution of shortest paths of duration t. It is shown in Fig. 9.6. As it is clear from Fig. 9.6, the characteristic time scale of the spread in this network is around 80 day (corresponding to the peak position). The total duration of the spreading activity is around 100 days.

Not all paths possible in a static aggregated network are present in the temporal one. The causal fidelity gives the fraction of paths present in the temporal network relative to the paths in the corresponding static (aggregated) network [19]

$$c(t) = \rho(\mathcal{P}_t)/\rho(\boldsymbol{P}_t), \tag{9.2}$$

where $\boldsymbol{P}_t$ is the accessibility in the static aggregated case with all daily network snapshots being the same. It tells us how important the temporal resolution is compared with the aggregated network. The dependence of the causal fidelity on time is depicted in Fig. 9.7. As it can be clearly seen, if the time scale (e.g. infectious period) of a dynamic process on the network is larger than approximately 300 days, we can consider the aggregated network as a static network.

[3]Its element $a_{ij} = 1$ if there is the edge $j \to i$ and zero otherwise.

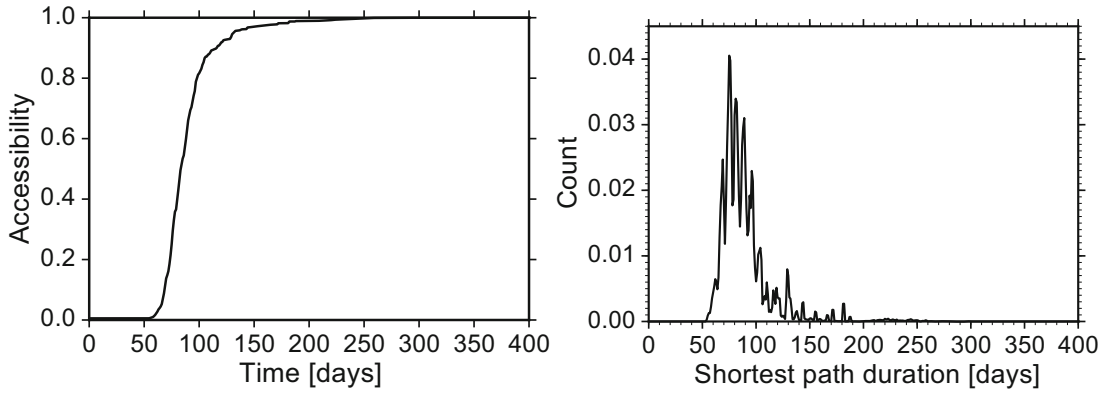

Fig. 9.6 *Left*: the density (fraction of non-zero elements) of the accessibility matrix $\rho(\mathcal{P}_t)$ of the temporal referral network of hospitals. *Right*: the difference between successive values of the accessibility density $\rho(\mathcal{P}_t)$ corresponding to the distribution of shortest path durations [19]

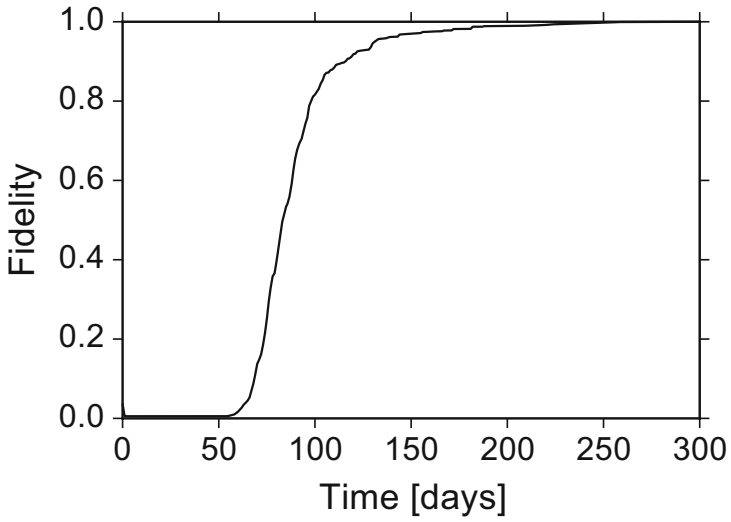

Fig. 9.7 The causal fidelity is given by the fraction of time respecting causal paths present in the aggregated static network, Eq. (9.2)

9.3.2 Deterministic SIR Model

So far we considered a deterministic infection without recovery. However the causal chain of contacts of a temporal network manifests itself even stronger in the case of a disease with a finite fixed infectious period. If vertices become immune to the disease after infection, the following kinetics can be used

$$S + I \to 2I$$
$$I \to R.$$

Note that we again consider the deterministic case, and thus upon contact with an infected, a susceptible individual becomes infected for sure. For a general analysis of deterministic SIR epidemics on temporal networks see [20]. As a quantity of interest we consider an out-component of the vertex i – a set of vertices which

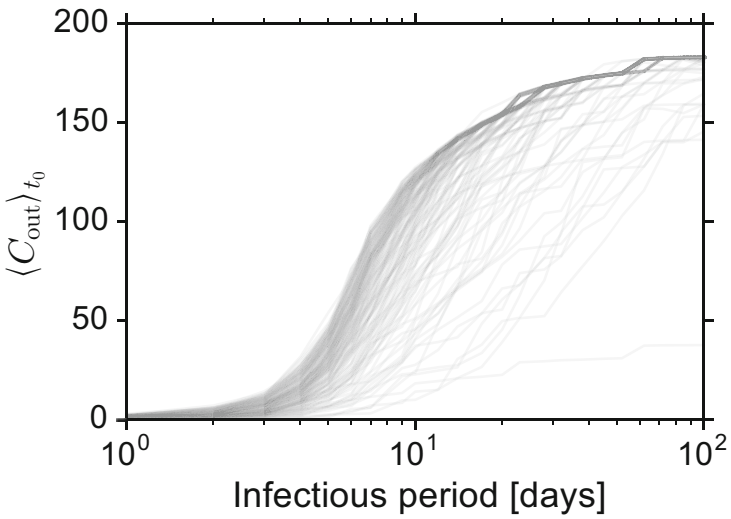

Fig. 9.8 Sizes of the out-components $\langle C_{out} \rangle_{t_0}$ for all of the 185 nodes in the referral network of hospitals in dependence on the infectious period k. Averaging was performed over initial time of infection t_0

could be reached by an epidemic with the infectious period k started from the given node i at the initial time t_0 by following all temporal edges respecting their time order [21]. We denote the size of the out-component by $C_{out}(i, k, t_0)$ which can be also considered as an upper bound for the number of infected nodes in the case of a stochastic disease transmission.

In Fig. 9.8 the out-components sizes averaged over initial times t_0 are presented for all nodes in the referral network of hospitals in dependence on the infectious period. We observe, that except for a few nodes, the majority of out-components reaches the size of the whole network for a disease with the infectious period of $k \sim 100$ days. Note that in the case of MRSA (Methicillin-resistant Staphylococcus aureus), the carriage of the pathogens could be even longer than 100 days. Figure 9.8 could not immediately reveal how long it would take to actually reach all the nodes in the out-component. Furthermore, we rank 185 hospitals according to their out-component size (Fig. 9.9, left panel) for different infectious periods. We observe a strong heterogeneity in the rank due to changes in tied values of the out-components for small (for $k \sim 3$ days there is a maximum number of ranks) and intermediate values of the infectious period. For high values of the infectious period the rank becomes very similar for all nodes. Small values of infectious period lead to highly fluctuating rank due to the high importance of the precise timing of outgoing links from the given node and small sizes of the resulting out-components.

We also average the ranks over all considered infectious periods (Fig. 9.9, right panel) and observe a strongly heterogeneous distribution – only a few nodes have a small rank, the majority has high rank values.

To analyze the robustness of the rank we considered the entropy of the rank distribution values for different infectious periods

$$\mathcal{H} = \sum_i p_i \ln p_i, \tag{9.3}$$

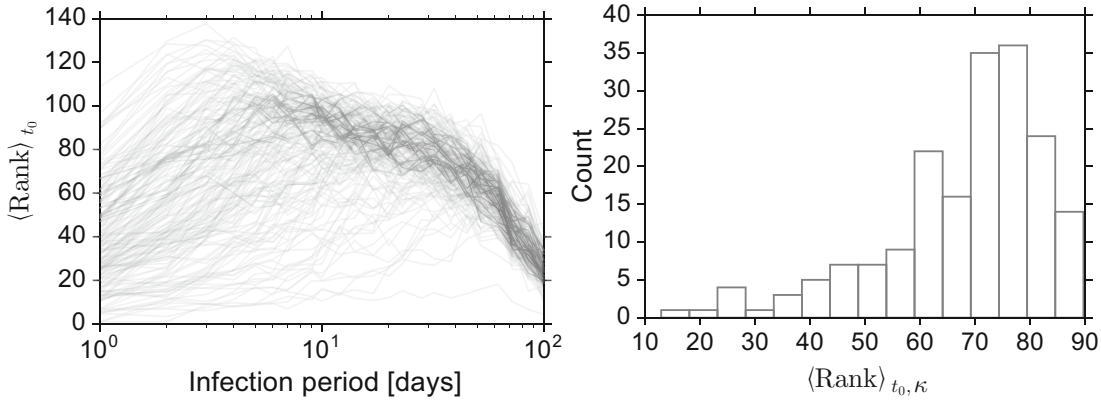

Fig. 9.9 *Left*: The node ranking according to the size of their out-component, averaged over initial times t_0. In the case of tied values the minimal rank values are taken. *Right*: Node ranks averaged over both initial time t_0 and infectious period κ

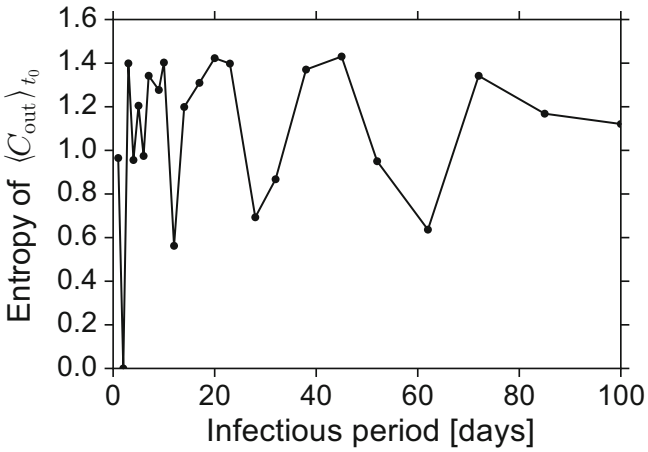

Fig. 9.10 The entropy Eq.(9.3) of the out-component rank distribution averaged over initial time t_0 versus the infectious period

where p_i is the probability of the out-component $\langle C_{out}\rangle_{t_0}$ to have rank i (Fig. 9.10). High entropy values correspond to strong heterogeneity. We observe a peculiar behavior – the entropy oscillates with the infectious period, especially for very low and very high values of k. It may be due to some "resonance" effect. Some values of the infectious period lead to similar out-components corresponding to the low entropy.

9.3.3 Network Correction Due to Community Stays

Until now we neglected the community stays and considered only direct transfers of patients between hospitals as edges in a dynamic network. This holds under the assumption of the complete recovery of a patient after a hospital visit. This approach was also adopted e.g. in Ref. [10]. If patients are not pathogen-free upon discharge from the hospital n, they could still carry pathogens when they return to the hospital m from the community. This happens, if the time spent in the community

is less than the infectious period of the pathogen: $t_m^{\text{in}} - t_n^{\text{out}} < k$, where subsripts of discharge and admission times denote the corresponding hospital. Now we relax this assumption and allow patients to carry the pathogens also in the community (without transmission) and adjust the corresponding effective dynamic network and quantify the impact of the pathogen carriage during community stays.

We call the network of direct patients transfers between two hospitals considered above an H-network. The H-network with additional edges due to possible transmission events after stays in the community is a C-network. In the C-network we include the edge between a hospital visited before the community stay and a hospital visited afterwards, if the duration of the community stay is less than the infectious period. Out-components $\tilde{C}_{\text{out}}$ in the C-network are larger than the out-components C_{out} in the H-network, because the H-network is a subset of the C-network (Fig. 9.11), left panel). In Fig. 9.11, right panel, the difference between $\langle \tilde{C}_{\text{out}} \rangle_{t_0,h}$ and $\langle C_{\text{out}} \rangle_{t_0,h}$ is shown. We see a pronounced peak around the infectious period of $k \approx 7$ days. For this infectious period, using only the H-network we underestimate the actual out-component by 20%.

This difference could be clarified if we look on the Jaccard coefficient $\Theta(k)$ between the edges of H- and C-networks given by

$$\Theta(k) = \frac{|E_{\text{H}} \cap E_{\text{C}}|}{|E_{\text{H}} \cup E_{\text{C}}|}, \tag{9.4}$$

where $|\cdot|$ denotes the cardinality of a set (number of its elements) and E_H and E_C are set of edges in H- and C-networks respectively. The Jaccard coefficient quantifies the relative overlap of two sets. It is maximal (equal to one), if two sets coincide and is zero if two sets are disjoint. In Fig. 9.12, left panel, we observe that maximal overlap is reached at the value of the infectious period $k \sim 7$ days. Additional increase of the infectious period does not lead to higher Jaccard index. This effect could be explained if we look on the denominator and the numerator in Eq. (9.4) separately (Fig. 9.12, right panel). Both the total number of edges in the C-network $|E_{\text{H}} \cap E_{\text{C}}|$ and the number of edges common in both C- and H-networks $|E_{\text{H}} \cup E_{\text{C}}|$ increases with the infectious period k. However, the former increases first sub-linearly but after the value of $k \sim 7$ days it increases in a super-linear manner. The latter increases first in a super-linear way but after the value of $k \sim 7$ days it increases sub-linearly. Together this leads to the peak in the Jaccard coefficient.

So far, in the epidemic analysis we considered the network extracted from the available dataset as a ground truth. However, we actually observe only around 50% of individuals [18] due to the market share limitations. This could lead to the underestimation of epidemic effects, such as timescales of the spread and the number of affected nodes. On the other hand, because we considered deterministic epidemics, we overestimated the epidemic effects. To account for both deficiencies – missing data and stochasticity of the transmission events, in the next section we consider a fully stochastic, discrete-event, and agent-based computational model of a disease with total recovery but no immunity and implement control measures.

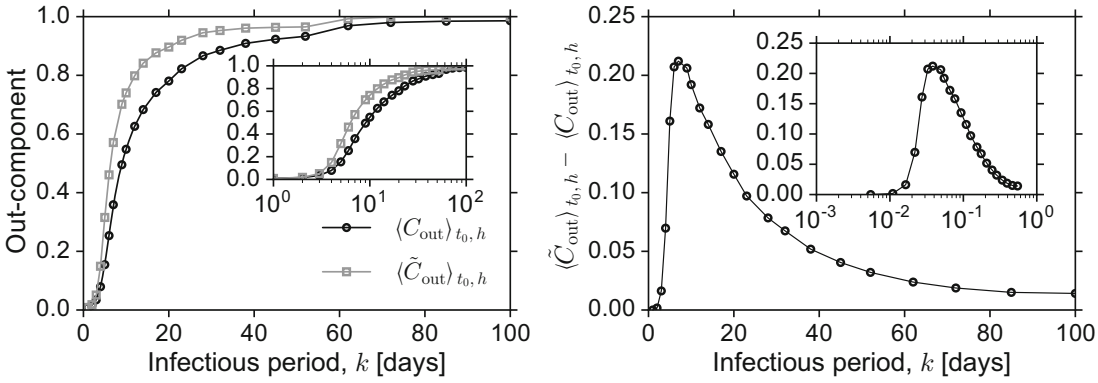

Fig. 9.11 *Left*: The effect of community stays on the out-component averaged over different initial times t_0 and over different hospitals h. $\tilde{C}_{\rm out}$ and $C_{\rm out}$ denote the out-components of the C and H-network respectively. *Right*: The difference between $\langle \tilde{C}_{\rm out}\rangle_{t_0,h}$ and the out-component $\langle C_{\rm out}\rangle_{t_0,h}$. Note that we consider the whole dataset (with initial time starting from the beginning of the data). *Insets* show the same dependences as the main plot in semi-logarithmic coordinates

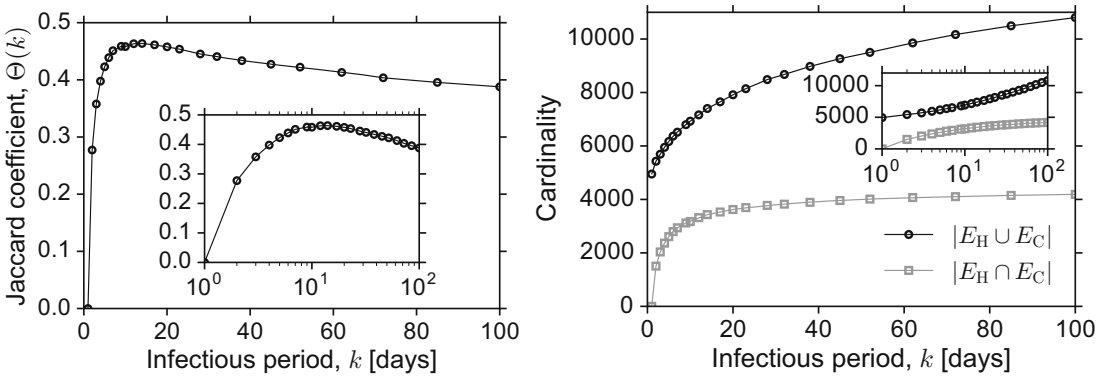

Fig. 9.12 The effect of community stays. *Left*: The relative overlap of the aggregated network is measured by the Jaccard coefficient $\Theta(k)$, Eq. (9.4) versus the infectious period k. *Right*: the total number of edges in C-network and the number of edges common in both C- and H-networks versus the infectious period k. *Insets* show the same dependences as the main plots in semi-logarithmic coordinates

9.3.4 Agent-Based Computational Model

In this section we introduce the computational framework for modeling the disease spread in a network of hospitals. In a single hospital we assume a randomly mixed situation, i.e. every patient could encounter every other patient. Information on healthcare workers was not available and they are assumed not to contribute to epidemic dynamics. We consider an endemic disease modeled by the standard SIS kinetics

$$S + I \xrightarrow{\alpha} 2I \tag{9.5}$$

$$I \xrightarrow{\beta} S.$$

Here the number of infected individuals I in a single hospital increases due to encounters with susceptibles S at the per capita rate α. The infecteds could become susceptible again at the rate β. The spread of some nosocomial diseases such as resistant pathogens reached an extent of an endemic with the prevalence for the MRSA around 4% in hospitals [22]. To ensure the prevalence of 4%, we use the following parameter values: $\alpha = 0.023\,\mathrm{day}^{-1}$ and $\beta = 0.0027\,\mathrm{day}^{-1}$ corresponding to half a year of carriage of the pathogens before recovery.

For our computational model, we need an artificial or a surrogate population due to the incompleteness of the data (only around 50% of the total population is included in the data), privacy reasons and to make projection into the future. Generation of a surrogate population is a non-trivial task, because the data is highly spatially and temporally correlated which is usually neglected. E.g. in the study by Donker and colleagues [10], hospital stays, separated by a community stay, were considered uncorrelated which makes sense only for patients healthy at discharge from a hospital. See also Ref. [23] for related issues in intra-hospital contact tracing. To produce the surrogate population and to keep the correlations present in the original dataset, we use the following bootstrapping procedure.

For every surrogate patient, we randomly choose a patient from the original dataset considering only the middle year $[T, 2T]$ (out of three) for bootstrapping, where $T = 365$ days as counted from the beginning of the dataset (the 1st of January 2009). We randomly from a uniform distribution choose the day $t \in (T, 2T]$ of the first appearance of the surrogate patient within the interval $(T, 2T]$ and replicate individual referral history periodically, n times.[4] To account for vital (birth/death) dynamics we need M – the total number of patients staying in hospitals during the year – for every annual period $[(k-1)T, kT]$, $k \leq n$. We choose randomly μM patients at a random time point within the current k-th time interval $((k-1)T, kT]$, exchange their IDs by new ones and reset their infection status to susceptible for the next hospital stay in the future. Here $\mu = 0.25$ is the turnover rate of the population (disappearance and appearance of patients) in the censored period of time.

The simulation was implemented using a modified stochastic Gillespie algorithm [24] for epidemic dynamics within a hospital, combined together with explicitly scheduled patient transfer events. All relevant events were implemented as a priority queue data structure [25]. The health status of all the patients in the system was tracked. We let the system equilibrate to a stationary level before starting the epidemic simulation.

Using the above computational model we can easily access the effect of different intervention measures based on the topological and temporal properties of the patient referral network. We choose two exemplary intervention scenarios. In both of them the resources for random screening of 50% of incoming patients and subsequent decolonization (this reduces the rest recovery time 3-fold) and isolation were allocated to the selected 10% of hospitals with the highest ranks. The ranking

[4] Since hospital stays could fall outside the interval $(T, 2T]$, we choose to cut the stays to fit into the interval.

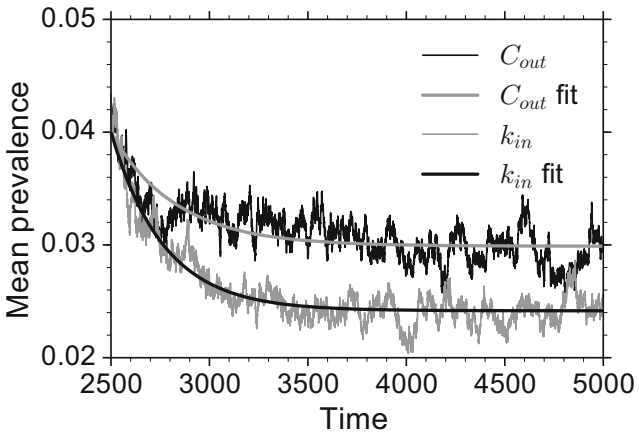

Fig. 9.13 Impact of interventions. Results of the numerical simulations of an agent based model, Eq. (9.5) with interventions – allocation of resources for screening and subsequent decontamination and isolation to 10% of hospitals (prioritized according to the out-component $\langle C_{out} \rangle_{t_0,k}$ of nodes (*blue*) and according to the aggregated in-degree k_{in} of a node (*red*)). For the fit an exponential function Eq. (9.6) was chosen. The in-degree ranking seems to be more efficient to reduce the prevalence than the time-averaged out-component ranking

was performed (i) according to the deterministic out-component $\langle C_{out} \rangle_{t_0}$ of a node and (ii) according to the in-degree of a node in the aggregated network. In Fig. 9.13 the time course of the prevalence, averaged over all hospitals, after intervention were applied, is presented for both scenarios. The comparison shows, that the in-degree ranking is more appropriate for the prevalence reduction than the out-component. The deviation in the final prevalences is up to 20% $\pm$ 10%. As an estimate for the final prevalence y_0 we took the limit $t \to \infty$ of the fit function to the prevalence time series $y(t)$ of the form

$$y = y_0 - \left(\pi^* - y_0\right) e^{-a(t-t_0)}, \tag{9.6}$$

where $\pi^* = 0.4$ is the baseline prevalence. This curve describes an abrupt prevalence decrease after intervention, afterwards slowly approaching the endemic prevalence level. We obtain for the ranking by the out-component $y_0 = 0.30$ and for the in-degree $y_0 = 0.24$. Note also that allocation of limited resources only to some hospitals usually leads to reduction of the prevalence but not to (almost) eradication of the disease, which could be achieved if the intervention measures were applied at all the hospitals. This supports the hypothesis, that in the network of hospitals the prevalence reduction is achieved immediately in the hospitals where control measures are implemented (high in-degree selects hospitals with a lot of incoming patients) and not due to the indirect effects of the reduction of disease transmission to other hospitals (which corresponds to the high out-component).

9.4 Conclusion

We have investigated patient referral patterns in a large federal state in Germany. We extracted the underlying network of hospitals and investigated its properties with respect to pathogen spread. For the worst case scenario of a highly contagious pathogen with and without recovery we examined the time scales and, in the presence of recovery, sizes of an outbreak dependent on the infectious period. We investigated the role of patients returning to hospitals but still carrying the pathogen acquired during previous hospital visits. We showed that this results in a deviation (underestimate) of the size of the out-components (or outbreak size) up to 20% for the values of the infectious period of one week without patient returns. By deploying an agent- and discrete event-based computational model of an endemic disease (MRSA), we assessed the impact of intervention strategies based on the out-component size and the in-degree resource allocation. Our analysis showed the advantage of the in-degree based allocation.

There is still the need for generic models of contacts patterns, to account for the missing data as well as to project the results into the future. Recently some additional epidemic control strategies based on the temporal aspects of the network were proposed as a promising direction for the future research [26, 27]. Prioritization scheme according to the risk of disease introduction should be used based on novel network distance measures [28, 29]. Our approach helps to understand the spread of infections in a network of hospitals and could be used to plan preventive measures as well as to design informed clinical studies.

Acknowledgements All the authors acknowledge the courtesy of the AOK Niedersachsen for providing the anonymized data on patient referrals. VB and PH acknowledge funding by the Deutsche Forschungsgemeinschaft in the framework of Collaborative Research Center 910. At the early stage of this study VB was financially supported by the fellowship "Computational Sciences" of the VolkswagenStiftung.

References

1. Cassini, A., Plachouras, D., Eckmanns, T., Sin, M.A., Blank, H.P., Ducomble, T., Haller, S., Harder, T., Klingeberg, A., Sixtensson, M., et al.: PLoS Med. **13**(10), e1002150 (2016)
2. O'Neill, J.: The review on antimicrobial resistance. Tackling drug-resistant infections globally: final report and recommendations. http://amr-review.org [WebCite Cache ID 6jI5znBnd] (2016). Accessed 26 July 2016
3. Keeling, M.J., Danon, L., Vernon, M.C., House, T.A.: Proc. Natl. Acad. Sci. **107**(19), 8866 (2010)
4. Belik, V., Geisel, T., Brockmann, D.: Phys. Rev. X **1**(1), 011001 (2011)
5. Rosvall, M., Esquivel, A.V., Lancichinetti, A., West, J.D., Lambiotte, R.: Nat. Commun. **5**, 4630 (2014)
6. Scholtes, I., Wider, N., Pfitzner, R., Garas, A., Tessone, C.J., Schweitzer, F.: Nat. Commun. **5**, 5024 (2014)
7. Holme, P., Saramäki, J.: Phys. Rep. **519**(3), 97 (2012)

8. Casteigts, A., Flocchini, P., Quattrociocchi, W., Santoro, N.: Int. J. Parallel Emergent Distrib. Syst. **27**(5), 387 (2012)
9. Fernandez-Gracia, J., Onnela, J.P., Barnett, M., Eguiluz, V.M., Christakis, N.A.: Spread of pathogens in the patient transfer network of US hospitals. arXiv preprint arXiv:1504.08343 (2015)
10. Donker, T., Wallinga, J., Slack, R., Grundmann, H.: PLoS One **7**(4), 1 (2012)
11. Ohst, J., Liljeros, F., Stenhem, M., Holme, P.: EPJ Data Sci. **3**(1), 1 (2014)
12. Rocha, L.E., Singh, V., Esch, M., Lenaerts, T., Stenhem, M., Liljeros, F., Thorson, A.: arXiv preprint arXiv:1611.06784 (2016)
13. Karkada, U.H., Adamic, L., Kahn, J.M., Iwashyna, T.J.: Intensive Care Med. **37**(10), 1633 (2011)
14. Schneider, C.M., Belik, V., Couronné, T., Smoreda, Z., González, M.C.: J. R. Soc. Interface **10**(84), 20130246 (2013)
15. Kovanen, L., Karsai, M., Kaski, K., Kertész, J., Saramäki, J.: Temporal motifs. In: Temporal Networks, pp. 119–133. Springer, Berlin/Heidelberg (2013)
16. Lentz, H.H.K., Koher, A., Hövel, P., Gethmann, J., Sauter-Louis, C., Selhorst, T., Conraths, F.: PLoS One **11**(5), e0155196 (2016)
17. Wieler, L.H., Ewers, C., Guenther, S., Walther, B., Lübke-Becker, A.: Int. J. Med. Microbiol. **303**(6–7), 380 (2013)
18. Belik, V., Hövel, P., Mikolajczyk, R.: Control of epidemics on hospital networks. In: Control of Self-Organizing Nonlinear Systems, pp. 431–440. Springer International Publishing, Cham (2016)
19. Lentz, H., Selhorst, T., Sokolov, I.M.: Phys. Rev. Lett. **110**(11), 118701 (2013)
20. Koher, A., Lentz, H.H.K., Hövel, P., Sokolov, I.: PLoS One **11**(4), e0151209 (2016)
21. Konschake, M., Lentz, H.H.K., Conraths, F.J., Hövel, P., Selhorst, T.: PLoS One **8**(2), e55223 (2013)
22. Marschall, J., Mühlemann, K.: Infect. Control **27**(11), 1206 (2006)
23. Génois, M., Vestergaard, C.L., Cattuto, C., Barrat, A.: Nat. Commun. **6**, 8860 (2015)
24. Gillespie, D.T.: J. Phys. Chem. **81**(25), 2340 (1977)
25. Cormen, T.H.: Introduction to Algorithms. MIT Press, Cambridge (2009)
26. Liu, S., Perra, N., Karsai, M., Vespignani, A.: Phys. Rev. Lett. **112**(11), 118702 (2014)
27. Belik, V., Fengler, A., Fiebig, F., Lentz, H.H.K., Hövel, P.: arXiv preprint arXiv:1509.04054 (2016)
28. Brockmann, D., Helbing, D.: Science **342**(6164), 1337 (2013)
29. Iannelli, F., Koher, A., Brockmann, D., Hövel, P., Sokolov, I.M.: Phys. Rev. E **97**, 012313 (2017)