

## **DYNAMIC MULTIPLEX SOCIAL NETWORK MODELS ON MULTIPLE TIME SCALES FOR SIMULATING CONTACT FORMATION AND PATTERNS IN EPIDEMIC SPREAD**

Günter Schneckeneither

Niki Popper

Institute for Analysis and Scientific Computing  
Vienna University of Technology  
Wiedner Hauptstraße 8-10  
1040 Wien, AUSTRIA

DEXHELPP  
Decision Support for Health Policy and Planning  
Neustiftgasse 57-59  
1070 Wien, AUSTRIA

### **ABSTRACT**

This contribution presents a model for dynamic networks of physical contacts among large populations and their application for reproducing complex patterns in epidemic spread. The networks are constructed from statistical data on demography, geography, organizational structure and contact behavior. Due to the heterogeneous nature of the data and by construction, rich topological characteristics such as overlapping communities, layering along multiple dimensions and multiplex dynamics on different time scales can be observed. The generated dynamic networks can furthermore be regarded as subgraphs or derivatives of latent social networks. General results and observations from social network theory apply naturally and are used for explaining dynamic effects in epidemics. An exemplaric analysis investigates the impact of weak ties and effects of communities with decreased immunization on epidemic spread. Optimized implementation and visualization techniques turn out to be a key asset for dynamic simulation of contacts within large populations.

### **1 INTRODUCTION**

This paper presents ongoing work on dynamic multiplex social and contact networks which operate on multiple time scales and their application for reproducing artifacts and dynamic phenomena in epidemic spread.

A visual outline of the construction and layout of the network model can be found in Figure 1. The population (a) is sampled from statistical data with geographic locations and various social attributes. As part of a social structuring individuals are assigned to households, school classes and workplaces honoring statistical data on age-distribution, size and location of such “organizational units” (b). In combination with external survey data, typical patterns of physical contacts can be assigned to individuals according to their attributes and their group memberships. Once pairs of contact patterns have been matched into links of a multi-layer social network (c, d), time-varying and dynamic contact behavior of the population can be simulated (e). As a consequence this model incorporates a slow moving or static social network which is heavily characterized by overlapping communities, an intermediate meta network which primarily describes periodic contact patterns and a derived partially periodic fast-changing network of actual contacts.

In combination with a basic model for spread of infectious diseases effects of inhomogeneous patterns in the population on epidemic characteristics can be simulated. For example the existence of cohorts with lower than average immunization dramatically increases speed of spread. Conforming with the general theory on social networks, inter-community contacts can be identified as key parameters for the dynamics of epidemics.

In general, a detailed realistic model of physical contacts offers greater insight into complex processes and phenomena of disease propagation within real populations than aggregated or artificial models can deliver. Furthermore as an adaptable framework the developed model allows to assess the effect of a broad

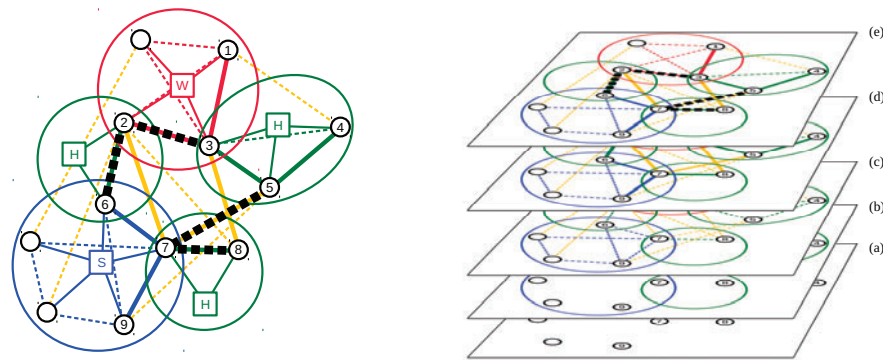


Figure 1: Visualization of the basic network layout. Households (H), workplaces (W) and school classes (S) and their memberships are depicted as green, red and blue blocks (b) of nodes (a). Social ties are indicated by dashed lines (c), derived contact patterns are shown as thick lines (d). Inter-block connections are distinguished by yellow lines. Instantiated temporal contacts are visualized as an overlaid black dashed pattern (e).

range of modifications in the structure and connectivity of the contact network as well as dynamic changes in individual contact behavior on various epidemic outcomes.

The sampling, network generation and simulation processes were implemented in Python and C++/Boost using geographic information processing, R-trees and graph libraries and are aimed at efficient simulation of large populations.

## 2 CHARACTERISTICS AND DYNAMICS OF SOCIAL AND CONTACT NETWORKS

From statistical data, observations and studies we know that for *social dimensions* (Vynnycky and White 2010) like age, gender, income/wealth, education and also geographic or administrative attribution of residence very characteristic patterns can be found in social relations and interactions. Besides link patterns that depend on various nodal attributes or social dimensions, social networks also feature intrinsic characteristics such as assortativity of degree, small-world phenomena, clustering, communities, subgraph patterns, etc. (Granovetter 1973, Wasserman and Faust 1994, Newman 2010).

### 2.1 Patterns and Network Formation

To a certain degree both types of link patterns – with respect to “microscopic” person-centric attributes or as “mesoscopic” natural properties of social networks (compare the analogous network formation approaches discussed by Anthony and desJardins (2008) and Toivonen et al. (2009)) – can be credited on the organizational structuring of societies into communities like households, school classes and workplaces.

One particular algorithmic approach to this concept is known as the *block model* which assumes that nodes – probably depending on their attributes – are assigned to blocks for which specific connectivities apply to the nodes within a block and between different blocks (Holland, Laskey, and Leinhardt 1983, Wasserman and Faust 1994). Often communities or blocks are assumed to be organized in a hierarchical and overlapping fashion in order to better explain particular characteristics of social networks (Ravasz and Barabási 2003, Clauset, Moore, and Newman 2008).

Nevertheless social communities like a group of close friends or other features also emerge randomly or from existing links without any underlying organizational block structure. On a technical level this concept is related to generative mechanisms which are based on the iterative modification of links and addition and deletion of nodes and often incorporate some kind of *triadic closure*. The challenge in this approach lies in finding algorithms or agent behavior that leads to the desired network structure (Barabási

1999, Newman, Watts, and Strogatz 2002, Toivonen et al. 2006, Kumpula et al. 2007, Klimek and Thurner 2013, Kertesz et al. 2016). Despite knowing the underlying mechanics for social relations very well, it proves to be a difficult task to define agent behavior that produces or maintains a certain social network structure with all its high-dimensional covariates (Toivonen et al. 2009).

## 2.2 Multiple Time Scales

Theoretically a model of a dynamic population with social relations and physical contacts as discussed here can be separated into different connected submodels. A demographic model (compare Bicher et al. 2015) for simulating births, migration and the evolution of individual properties of persons; a model for handling the block structure and assignments to blocks; mechanisms for creation and deletion of social ties based on the block structure and nodal attributes; and finally a model for simulating contacts between individuals.

It is important to note that a social network is not necessarily a network of physical contacts. In fact, social relations alter on a much slower time scale than interactions between individuals happen. Nevertheless it is often social relations that lead to physical contacts. In that context we can assume that contacts among individuals induced by particular social relations show specific patterns concerning periodicity and duration that may also depend on additional personal attributes of the involved individuals (edge and node attributes of the social graph). Again an agent-based approach seems suitable for generating contact patterns based on the existing social network structure.

Table 1 displays different network mechanisms operating on different time scales and layers.

Table 1: Different time scales for social ties and physical contacts.

Time scale	Layer	Mechanisms	Simulation of
life-time	all	node creation and deletion	births, deaths, migration
decades	social structure	creation and allocation of new blocks	household, family
years	social structure	creation of new blocks, block swapping	workpl., school cl.
months/weeks	social ties	additional link creation, alteration, deletion	other social acquaintances
days/hours	contacts	temporary links or instantiations	physical contacts

However, depending on application and simulation timeframe, a static population and social model is often sufficient. Still not straight-forward, a momentary and constant social structure can be sampled from survey data in combination with a block layout in a much simpler way without the necessity of fitting and modeling the temporal evolution of social ties.

## 3 CONCEPTUAL AND TECHNICAL MODEL DESCRIPTION

As suggested before, the idealized concept of a fast moving dynamic contact network completely based on a slow moving social network (Table 1) is difficult to implement and parameterize. This is especially due to the lack of extensive statistical data on social relations within a population and on the instantiation of contacts based on social ties.

A conceptually equal but technically slightly modified approach for simulating contacts in a population was applied in this work. Instead of inferring contact patterns from social links, here survey data, nodal attributes and block-memberships were used to sample contact patterns for individuals directly (compare Figure 1 (c) and (d)). The population and block structure is based on publicly available statistical data (Statistics Austria 2014). A dataset from the POLYMOD Project (funded by the European Commission, FP6, SSP22-CT-2004-502084, Mossong et al. 2008) on the contact behaviors in several European countries was used for sampling the contact patterns. Note that in their paper Mossong et al. (2008) define a contact

as “a two-way conversation with three or more words”. The term “physical contact” is used exclusively if two persons physically touched each other.

Once contacts have been established based on the sampled patterns, indirectly also a (latent) social network of the population is generated (compare (1) in Section 4.1). Of course random contacts must be excluded and social relations without any contact are neglected.

This section outlines the developed concept (Figure 1) and methods and gives insight into the technical implementation.

### 3.1 Sampling of Base Population and Block Structure

In a first step a static population and block structure with the following properties is sampled from basic statistical data. **person:** municipality, age, sex; **household:** district, location, size, age distribution, income; **school class:** district, location, size, age distribution, school-type; **workplace:** district, location, size, age distribution.

In multiple occasions multivariate and indirect sampling techniques have to be applied. In case of school classes for example in a first step, size, type and administrative affiliation are sampled. Based on size and school-type a random mean age is chosen according to the age distribution within classes of this type of school. The target age distribution within the class is then defined by a maximum allowed deviation from the mean age. This procedure ensures assortative mixing with respect to age for contacts within school classes. For workplaces on the contrary the age distribution is based on the age distribution of the employed population etc.

Sampling was implemented using marginal approaches as well as Markov Monte Carlo methods. Locations are sampled uniformly within administrative boundaries based on geographic data. Due to the computational effort associated with the allocation algorithms (see following sections) at the time of writing usually a subset (e.g. an arbitrary administrative district or state) of the Austrian population, which in total is about 8 million, was sampled.

### 3.2 Allocation of Blocks

The key challenge in allocating persons into predefined blocks consists in meeting the predefined target age distribution within blocks as close as possible while simultaneously obtaining realistic geographic distances between the location of the block and the location of the individual. For workplaces the latter should be distributed according to typical commuting distances (compare Figure 2 left). For school classes especially in lower grades shorter distances seem realistic. The situation for households (which are allocated at first) is slightly different: Once assigned, persons get a new geographic location in the vicinity of their household. Of course a great number of the involved parameters had to be estimated and the defined strategies include some heuristic assumptions.

Since global optimization of this allocation problem was not feasible an iterative approach based on search trees (R-trees) was implemented in C++/Boost. The previously insinuated strategies were translated into nearest neighbor statements and bounding boxes.

### 3.3 Sampling of Contact Patterns and Matching

The contact patterns available in the dataset are organized in a person-centric fashion and describe the contact behavior of a person during one day. This means that for a person with given age and sex a set of multiple contacts is recorded. Every recorded contact includes among other parameters (e.g. day of the week) the age of the “contact owner”, the age of the contacted person, the *frequency* with which the contact happens and the *regime* of the contact. The regime can be one of household, school class, workplace, leisure or other. Allowed frequencies or periodicities are daily, weekly and random. We will refer to these one-sided contact patterns as *contact requests*.

Once persons have been equipped with contact requests from the dataset (again randomized search strategies are applied), a similar optimization problem as for block allocation occurs when contact requests should be joined into pairs, which then correspond to links in the meta network of contact patterns. In contrast to the previous section where a bipartite one-slot-to-one-person network structure was generated, here a person-to-person network is generated. In this context contact requests can be regarded as unconnected/open directed links which are joined into undirected edges in the pattern layer of the network.

This procedure was again implemented using high-dimensional R-trees and graphs (C++/Boost) in combination with different modular configurable search strategies.

**exact matching of target attributes:** Since a contact between two individuals can be regarded as an undirected process, the regime and frequency or periodicity of joined contact patterns should match exactly.

**asymmetric allocation:** For two contact requests from different persons the contact-owner-age and the age of the contacted person should match in an asymmetric fashion.

**disallow multiple links:** A matched pair of contact requests is immediately stored as a link (with the contact pattern as edge properties) in a graph data structure. Before a pair of contact requests is joined, the graph is queried for an already existing link in order to prevent multiple links (i.e. contact patterns) between two persons.

**intra-block allocation:** Contact request matching can be restricted to individuals within the same block. This is obviously wanted for generating contacts within households, school classes or workplaces (the block model). Consequently intra-block matching only joins contact requests from the same regime. This strategy is especially suitable for high frequency contacts, which supposedly also in reality are likely to take place within communities.

**bounding boxes:** In order to restrict deviation of certain attributes from their target value it is possible to define maxima in arbitrary dimensions. This strategy can be extended with iterative relaxation of bounding criteria.

**assortative matching:** For assortative matching of attributes like income and geographic position also bounding boxes can be used.

**triadic closure:** Based on already established links in the graph new links which close open triangles can be preferred. This approach can also take into account nodal or link attributes like requiring that both existing links of an open triangle are part of the same regime (for example workplace contacts). This technique is required for reproducing clustering effects which are not covered by the block model.

**nearest neighbor:** Despite tricky in certain circumstances nearest neighbor search can be used if other strategies fail to deliver enough matched pairs.

**random contact allocation:** For daily contacts within households, random or complete graphs can be an alternative model. This approach is based on the Erdős-Réni model and mainly implemented for testing purposes.

An interface to a spreadsheet document was implemented in order to compile detailed configurations (50 and more rows of approximately 25 configuration parameters) at runtime into a sequential and partially parallel composition of combined allocation algorithms operating on different subsets of the population and the corresponding contact requests. This allows to quickly build and evaluate models for generating contact networks in a simulator-like environment.

### **3.4 Dynamic Simulation of Contacts**

A fully sampled contact pattern layer with daily, weekly and random contact patterns serves as initial condition for a simulation run. The basic layout of a simulation run is characterized by the following three steps which are applied at different intervals.



1. At the beginning of every day all random-periodicity contact requests are reconnected in a randomized fashion as temporary contact patterns.
2. Afterwards all contact patterns are scheduled for instantiation during the day.
3. Finally all vertices and their contacts are processed in an iterative manner.

Different approaches for scheduling contacts on basis of the contact pattern network are possible. Contacts can either be regarded as atomic events that happen at a particular time during the day (with duration as a virtual attribute) or handled as time periods according to their duration. At the time of writing a simple atomic-event approach was used. Based on the regime of the contact a random time is chosen. Contacts at work are scheduled during usual business hours. Random contacts are scheduled with a uniform distribution throughout the day etc.

When the contacts of a node are processed the disease state of the node can be updated according to arbitrary disease dynamics ((5) in Section 4.1). On the other hand also effects from an altered node state on all outgoing contact requests of a person can be configured, for example for simulating reduced contacts in case of infection ((3) in Section 4.1).

### 3.5 Visualization

Since it is the aim to simulate contact patterns in large populations with up to 8 million nodes (with fixed geographic locations), which implies an even larger number of edges, decent visualization techniques are required. Throughout development of the model geographic information systems turned out to provide capable rendering techniques which can be exploited for network visualization.

Different strategies with varying implementation effort and also benefit are listed below.

1. Open-source libraries like GDAL provide interfaces to many programming languages and can be used for rastering (i.e. rendering) vector data like nodes and edges of a graph (Figure 2 left).
2. Vector data can also be exported to ESRI Shapefiles which in turn can be loaded into geo-information systems such as QGIS for visual editing and rendering (Figure 2 center). Once a graphical representation has been set up, a QGIS configuration file can be stored and reused at a later time.
3. The QGIS framework provides APIs in Python and C++ which (in combination with configuration files) can be used for automated rendering at runtime. This is especially useful for generating animation frames (Figure 2 right) of a dynamic simulation run.

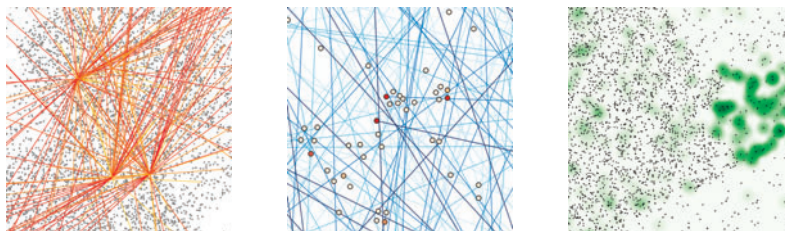


Figure 2: Geographic visualizations. On the left (static block structure): Four random workplaces with their block-memberships visualized as connections in a bipartite network (block-person links). Block assignments with a large geographic distance are indicated by red lines close geographic distances are yellow. Center (contact patterns): A small geographic section with persons as circles and (static) contact patterns as lines. The color intensity encodes the time of infection (red color scale for persons) and transmission of the disease (blue color scale for contacts). Right hand side (spread of disease): A larger geographic section with persons as black dots and the density of infection within the population on a green color scale. The initially infected community is visible as a highly infectious area in the right half of the image.

## 4 NETWORK ANALYSIS AND SIMULATION OF EPIDEMIC SPREAD

This section recapitulates the multiplex layout of the model and highlights some important aspects of the topology of the contact network. Subsequently effects of some topological characteristics on basic epidemic spread are analyzed.

### 4.1 Multiplex Dynamics

Resulting from the distinction of multiple time scales, links in the same network can simultaneously exist as social ties, contact patterns or instantiated contacts (Figure 1). Since instantiated contacts are based on contact patterns which in turn – neglecting random contacts (\*) – are based on social relations, we can establish the formal equation

$$A_{ij}^{(\text{contact})}(t) \neq 0 \implies A_{ij}^{(\text{pattern})}(t) \neq 0 \xrightarrow{(*)} A_{ij}^{(\text{social})}(t) \neq 0, \quad (1)$$

where  $A_{ij}^{(\text{layer})}(t)$  denotes the attributes and existence of a (symmetric) connection between node  $i$  and node  $j$  in a specific network layer at time  $t$ .

With  $v_i(t)$  as the attributes of node  $i$ , a formal mathematical representation of the multiplex network dynamics can be written as

$$\frac{d}{dt}A_{ij}^{(\text{social})}(t) = 0 \quad (2)$$

$$\frac{d}{dt}A_{ij}^{(\text{pattern})}(t) = G^{(\text{pattern})}(A_{ij}^{(\text{social})}(t), v_i(t), v_j(t)) + G^{(\text{random})}(v_i(t), v_j(t)) \quad (3)$$

$$\frac{d}{dt}A_{ij}^{(\text{contact})}(t) = G^{(\text{contact})}(A_{ij}^{(\text{pattern})}(t)) \quad (4)$$

$$\frac{d}{dt}v_i(t) = F(A_{ij}^{(\text{contact})}(t), v_i(t), v_j(t)). \quad (5)$$

Additional disjunct sub-layers result for example from the separation of contacts which happen in different regimes (at home, in school, at work, etc.) or with different frequencies (daily, weekly, random),

$$A_{ij}^{(\text{contact})}(t) = A_{ij}^{(\text{daily contacts})}(t) + A_{ij}^{(\text{weekly contacts})}(t) + A_{ij}^{(\text{random contacts})}(t).$$

Each layer has its distinct topological features which play a central role the following investigations.

### 4.2 Network Topology

We can assume that high frequency contacts (patterns) are based on strong social ties and therefore are suitable for allocation with the block model or triadic closure approach. Low frequency or periodicity contacts on the other hand can be allocated with methods that do not enforce a community structure or clustering on an algorithmic level. Figure 3 shows the degree and clustering coefficient distributions for daily, weekly and random contacts which were scheduled for instantiation during a random day. As a consequence of the configuration of allocation algorithms clustering in the random periodicity layer is much lower.

Note that the overall degree distribution in the network of instantiated contacts is basically exponential. Although exponential cutoff is totally plausible for high frequency contacts which are based on social relations (Dunbar 2010), a large number of random or low frequency contacts can for example occur in the tertiary sector, at ticket sales or during simulation conferences. The absence of nodes with a very large number of random contacts can be explained by limitations – or incorrect processing – of the data sample.

### 4.3 Model and a Basic Scenario for Epidemic Spread

In a first stage a prototype SIR model with characteristics similar to measles was integrated into the dynamic contact simulation (Vynnycky and White 2010, World Health Organization 2017). The model was

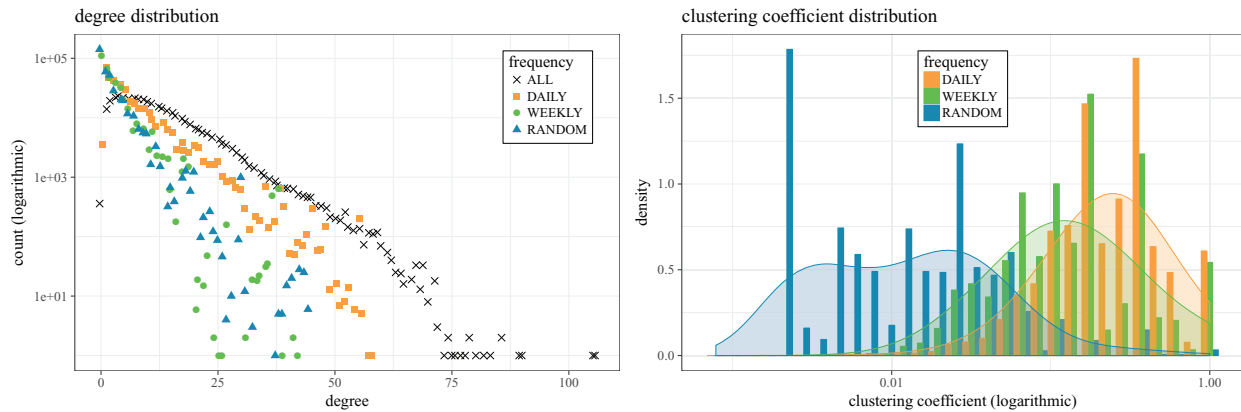


Figure 3: Distribution of degree and clustering coefficient in different frequency layers of the contact network.

configured to display an average latent period of 12 days and an average infectious period of 16 days. In case of contact the disease is highly contagious. Different heuristic penalties for decreasing the likelihood of infection based on nodal and contact link attributes like the age of the susceptible person, the duration of the contact, the environment in which the contact took place (e.g. at home) etc. were implemented resulting in an average infection probability of 0.6.

All simulations referred to by this paper were conducted with a population of about 400,000 and an immunization rate of 95%.

#### 4.4 The Strength of Weak Ties in Epidemic Spread

From Section 4.2 and Figure 3 we can expect that low-frequency contacts serve as bridges and weak ties in the contact network. According to Granovetter (1973) these weak ties are likely to have an important contribution to the spread of an infectious disease. To assess this assumption let us define the following partially overlapping concepts of weakness.

**social (frequency/regime):** The contact does not emerge from a close social relation and therefore has a low frequency and happens not within a household, school class or workplace.

**algorithmic:** Closely related to the previous concept. The contact was allocated neither as intra-block contact nor using the triadic closure algorithm.

**epidemiologic:** A contact is weak if it is associated with a low probability of disease transmission. This is the case for contacts with a short duration, between persons not in the high-risk age group, in an environment for which transmission is not likely etc.

**topologic:** The contact is a bridge or happens between two persons with a low combined clustering coefficient.

We now comparatively remove an equal amount of arbitrary, weak or strong (inverse definitions) contacts and observe the difference in the effects on the dynamics of an epidemic. Figure 4 shows that for all types of weakness the removal of weak contacts yields a lower epidemic incidence when compared to the removal of strong contacts.

In the visualization of the incidence in Figure 4 we can furthermore see that in an initial phase (200 hours) all strategies behave the same way until a pool of clustered susceptibles is depleted and spread of the disease relies on weak contacts in order to reach distant clusters or blocks. Of course when only a small number of weak links is removed this effect is often not clearly recognizable and can only be observed at a



smaller scale. This is partially due to the overlapping block structure and application of a combination of allocation algorithms, which perturbs an otherwise clear community structure (comp. Kertesz et al. 2016).

These results also play an important role in combination with network percolation and herd immunity which is addressed in the next section.

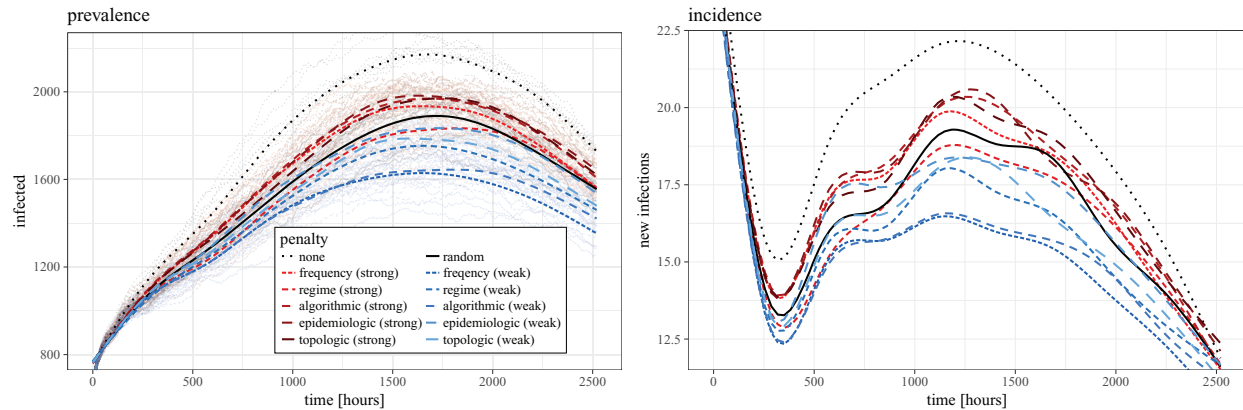


Figure 4: Disease dynamics for different strategies for removing strong (red) or weak (blue) contacts. As initial condition a group of 700 individuals out of the population of a random municipality was set to infectious. Each strategy was averaged over 10 simulation runs. For every tie removal strategy 1 of about 10 million contact patterns were removed before simulation. For comparison the default configuration without tie removal (none) and a random tie removal strategy (random) are shown.

#### 4.5 Secondary Infections and Herd Immunity

The *basic reproduction number*  $R_0$  is defined as the “average number of secondary infectious persons resulting from one infectious person following their introduction into a totally susceptible population” (Vynnycky and White 2010). For measles this number is about 12-18. The *effective reproduction number*  $R_n = R_0 s$  adjusts for a population with a fraction  $s$  of susceptibles. Accordingly the *herd immunity threshold* defines the fraction of susceptibles  $s$  or immune  $i = 1 - s$  necessary for  $R_n < 1$ . For a highly infectious disease this calculation shows that an immunization rate of over 95% can be necessary (Vynnycky and White 2010).

We investigate heterogeneous distributions of the susceptible population (formally  $s = s(x)$ ) with different strategies for selecting susceptible nodes (20,000) in the network. **block:** All members of selected blocks in the school class layer with some requirements on class size and school-type are susceptible. **regime:** Arbitrary individuals who are members of a school class with some requirements on age can be susceptible. **cluster:** Random nodes and their neighbors are susceptible. **random:** Random nodes are susceptible.

In Figure 5 (left) we see that for random placement the dynamics conform to the previous statements about the herd immunity threshold. With an immunization rate of about 95% the incidence constantly remains at a low level.

For the block and regime placement strategies spread of the disease is much faster because a single infection entering a block or connected community (students) with an overproportionate fraction of susceptible individuals leads to total infection of that community. Since the allocation algorithm was configured to additionally connect students from different school classes (this is necessary to represent the correct number of contacts from the data and also suits to simulate schools) the school contact layer consists of large connected components with school classes represented as tight clusters. This characteristic explains the similarity in the behavior of the block and regime strategy.

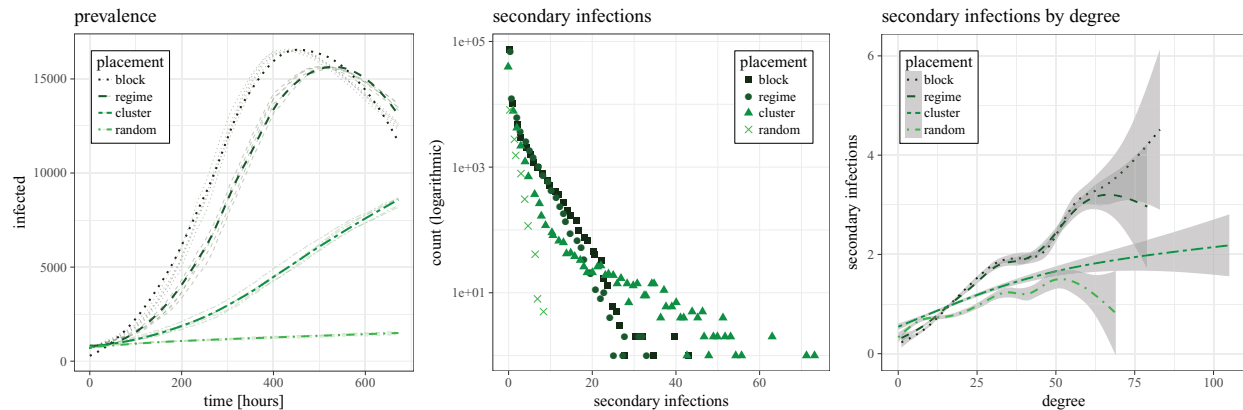


Figure 5: Different placement strategies for susceptible individuals. Initially a local group of about 700 individuals was infectious. The susceptible population (20,000) was distributed according to different placement strategies. Results were averaged over 5 simulation runs.

The different slopes of the averaged measurements in Figure 5 (right) indicate that in case of tightly clustered susceptibles an infectious nodes can make “better use” of their degree in spreading the infection. With the cluster strategy the selection of infectious nodes is not limited to students and clusters are not necessarily strongly connected. As a consequence spread is slower than in the block/regime approach but *infectious hubs* with a very large number of secondary infections emerge (power law behavior in Figure 5 center).

## 5 CONCLUSIONS AND DISCUSSION

The previous sections confirm the obvious finding that the usual estimation of the herd immunity threshold is not valid for inhomogeneously susceptible populations. A possible countermeasure when susceptible individuals are clustered could be prevention of disease transmission along weak ties. These conclusions of course need further and thorough assessment. On the one hand validation and improvement of the contact model and its heuristic assumptions is a necessary step for making quantitative examinations in the future. Furthermore the stability of qualitative results with respect to different, more detailed and realistic diseases or infection parameters needs inspection.

The exemplaric results presented show that detailed simulation of contacts within large populations is essential for investigating complex dynamics in disease propagation. In current literature on epidemic spread large contact networks are often generated in a non-person-centric fashion relying on aggregated “macroscopic” data. Although some important phenomena can be reproduced with fully artificial networks, detailed simulation of dynamic contacts on a large scale can presumably reveal additional effects in disease propagation.

For example integration of geographic information or heterogeneity of a population in general can be a crucial improvement to simulation models especially in epidemiology and health system research (Schneckenreither et al. 2008, Romstorfer and Schneckenreither 2012). Spatial distribution of nodes is also an important concept for explaining topological phenomena in networks. Kertesz et al. (2016) conclude that a geographic embedding and inter-layer correlation in terms of link overlap are essential for “Granovetterian properties [in combination with] overlapping communities”. Furthermore it can be argued that a geographic embedding can be interpreted as a hierarchical community structure (with perturbations and in addition to an organizational block structure), which in turn – according to the theory (Ravasz and Barabási 2003, Clauset, Moore, and Newman 2008) – also promotes the scale-free and clustering property. For assessment of the influence of a geographic embedding on the clustering property, networks allocated with different geographic assortativity conditions should be compared.

Due to the computational effort in sampling and matching of contact patterns on an individual basis, the possibility of an equivalent stochastic approach to contact network formation should be investigated and could be based on inference of stochastic block models (Nowicki and Snijders 2001, Hoff, Raftery, and Handcock 2002, Airoldi et al. 2008, Peixoto 2014) and exponential random graph models (ERGMs) (Robins et al. 2007, Fronczak, Fronczak, and Bujok 2013). Also here a thorough analysis of multiplex dynamics and topological characteristics in combination with link and community overlap (Szell, Lambiotte, and Thurner 2010, Hristova et al. 2016, Klimek et al. 2016, Kertesz et al. 2016) is a precondition for further investigations.

## ACKNOWLEDGMENTS

The work presented partially emerged from a lecture on complex systems attended at Medical University of Vienna and is connected to the authors work on his PhD thesis. The contribution was also supported by DEXHELPP (Decision Support for Health Policy and Planning, K-Project), which is funded by the Austrian Ministries for Innovation, Technology, Science, Research and Economy as well as the state of Vienna via COMET (Competence Centers for Excellent Technologies). The COMET program is processed by FFG (Austrian Research Promotion Agency).

## REFERENCES

- Airoldi, E. M., D. M. Blei, S. E. Fienberg, and E. P. Xing. 2008. “Mixed Membership Stochastic Block-models”. *The Journal of Machine Learning Research* 9 (1): 1981–2014.
- Anthony, A., and M. desJardins. 2008. “Generative Models for Clustering: The Next Generation”. In *Social Information Processing, Papers from the 2008 AAAI Spring Symposium, Technical Report SS-08-06*, 7–10. Stanford, USA.
- Barabási, A. 1999. “Emergence of Scaling in Random Networks”. *Science* 286 (5439): 509–512.
- Bicher, M., B. Glock, F. Miksch, G. Schneckenreither, and N. Popper. 2015. “Definition, Validation and Comparison of Two Population Models for Austria”. In *Proceedings of 4th UBT Annual International Conference on Business, Technology and Innovation*. Durres, Albania: UBT.
- Clauset, A., C. Moore, and M. E. J. Newman. 2008. “Hierarchical Structure and the Prediction of Missing Links in Networks”. *Nature* 453 (7191): 98–101.
- Dunbar, R. I. M. 2010. “Constraints on the Evolution of Social Institutions and their Implications for Information Flow”. *Journal of Institutional Economics* 7 (03): 345–371.
- Fronczak, P., A. Fronczak, and M. Bujok. 2013. “Exponential Random Graph Models for Networks with Community Structure”. *Physical Review E* 88 (3): 032810.
- Granovetter, M. S. 1973. “The Strength of Weak Ties”. *American Journal of Sociology* 78 (6): 1360–1380.
- Hoff, P. D., A. E. Raftery, and M. S. Handcock. 2002. “Latent Space Approaches to Social Network Analysis”. *Journal of the American Statistical Association* 97 (460): 1090–1098.
- Holland, P. W., K. B. Laskey, and S. Leinhardt. 1983. “Stochastic Blockmodels: First Steps”. *Social Networks* 5 (2): 109–137.
- Hristova, D., A. Noulas, C. Brown, M. Musolesi, and C. Mascolo. 2016. “A Multilayer Approach to Multiplexity and Link Prediction in Online Geo-Social Networks”. *EPJ Data Science* 5 (1).
- Kertesz, J., J. Torok, Y. Murase, H.-H. Jo, and K. Kaski. 2016. “Multiplex Modeling of the Society”. *arXiv:1609.08381*.
- Klimek, P., M. Diakonova, V. M. Eguíluz, M. S. Miguel, and S. Thurner. 2016. “Dynamical Origins of the Community Structure of an Online Multi-Layer Society”. *New Journal of Physics* 18 (8): 083045.
- Klimek, P., and S. Thurner. 2013. “Triadic Closure Dynamics Drives Scaling Laws in Social Multiplex Networks”. *New Journal of Physics* 15 (6): 063008.
- Kumpula, J. M., J.-P. Onnela, J. Saramki, K. Kaski, and J. Kertész. 2007. “Emergence of Communities in Weighted Networks”. *Physical Review Letters* 99 (22): 228701.

- Mossong, J., N. Hens, M. Jit, P. Beutels, K. Auranen, R. Mikolajczyk, M. Massari, S. Salmaso, G. S. Tomba, J. Wallinga, J. Heijne, M. Sadkowska-Todys, M. Rosinska, and W. J. Edmunds. 2008. “Social Contacts and Mixing Patterns Relevant to the Spread of Infectious Diseases”. *PLoS Medicine* 5 (3): 381–391.
- Newman, M. 2010. *Networks: An Introduction*. Oxford University Press, USA.
- Newman, M. E. J., D. J. Watts, and S. H. Strogatz. 2002. “Random Graph Models of Social Networks”. *Proceedings of the National Academy of Sciences* 99 (Supplement 1): 2566–2572.
- Nowicki, K., and T. A. B. Snijders. 2001. “Estimation and Prediction for Stochastic Blockstructures”. *Journal of the American Statistical Association* 96 (455): 1077–1087.
- Peixoto, T. P. 2014. “Efficient Monte Carlo and Greedy Heuristic for the Inference of Stochastic Block Models”. *Physical Review E* 89 (1): 012804.
- Ravasz, E., and A.-L. Barabási. 2003. “Hierarchical Organization in Complex Networks”. *Physical Review E* 67 (2): 026112.
- Robins, G., P. Pattison, Y. Kalish, and D. Lusher. 2007. “An Introduction to Exponential Random Graph ( $p^*$ ) Models for Social Networks”. *Social Networks* 29 (2): 173–191.
- Romstorfer, G., and G. Schneckenreither. 2012. “Using Open-Source Geo-Data in Agent-Based Models of Health Care Utilization”. *IFAC Proceedings Volumes* 45 (2): 1308–1312.
- Schneckenreither, G., N. Popper, and F. Breiteneker. 2008. “Modelling SIR-type Epidemics by ODEs, PDEs, Difference Equations and Cellular Automata – A Comparative Study”. *Simulation Modelling Practice and Theory* 16 (8): 1014–1023.
- Statistics Austria 2014. “Online Resources: Demographic, Economic, Educational and Geographic Data”. [www.statistik.at](http://www.statistik.at) (accessed between 2014-06 and 2017-03).
- Szell, M., R. Lambiotte, and S. Thurner. 2010. “Multirelational Organization of Large-Scale Social Networks in an Online World”. *Proceedings of the National Academy of Sciences* 107 (31): 13636–13641.
- Toivonen, R., L. Kovanen, M. Kivelä, J.-P. Onnela, J. Saramäki, and K. Kaski. 2009. “A Comparative Study of Social Network Models: Network Evolution Models and Nodal Attribute Models”. *Social Networks* 31 (4): 240–254.
- Toivonen, R., J.-P. Onnela, J. Saramäki, J. Hyvönen, and K. Kaski. 2006. “A Model for Social Networks”. *Physica A: Statistical Mechanics and its Applications* 371 (2): 851–860.
- Vynnycky, E., and R. G. White. 2010. *An Introduction to Infectious Disease Modelling*. Oxford University Press.
- Wasserman, S., and K. Faust. 1994. *Social Network Analysis: Methods and Applications. Structural Analysis in the Social Sciences*. Cambridge University Press.
- World Health Organization 2017. “Online Ressource: Measles Fact Sheets”. Available via [www.who.int](http://www.who.int) (accessed 2017-04).

## AUTHOR BIOGRAPHIES

**GÜNTER SCHNECKENREITHER** is a PhD candidate at the Institute for Analysis and Scientific Computing at Vienna University of Technology. His research interests include stochastic simulation, heterogeneous dynamical systems, stochastic partial differential equations, dynamic networks and complex systems. He is currently working in an applied research project on modeling of dynamic processes in health system research and health policy planning. His e-mail address is [guenter.schneckenreither@tuwien.ac.at](mailto:guenter.schneckenreither@tuwien.ac.at).

**NIKI POPPER** is the coordinator of COCOS “Centre for Computational Complex Systems” at TU Wien as well as chairman of DEXHELPP “Decision Support for Health Policy and Planning” (COMET K-Project), which is dedicated to the development of models and to efficient and safe use of data for decision making in health systems. His main research interests include theory and applications of modeling and simulation of dynamic and complex systems. His e-mail address is [niki.popper@dexhelpp.at](mailto:niki.popper@dexhelpp.at).