Spreading of rodent infestations through a city

Dalton Brooks and John Matta[*]

*Correspondence: jmatta@siue.edu
Southern Illinois University
Edwardsville, Edwardsville, IL, USA

Abstract

This paper details a network-based analysis of the spreading of rodent infestations through a city under varying conditions. Models of two very different cities, Tulsa, OK, USA and Providence, RI, USA, are created as geometric graphs using publicly available map data. The SIR epidemic model is applied with varying parameters, and the resulting infestation rates are measured. Results indicate that the density of a city, the size of the largest connected component (from the point of view of a rodent migrating to a new location), and the distance and probability of migration all play a role in determining the resulting rate of infestation. Results obtained are consistent with infestation patterns and rates reported by studies of other similarly dense city environments. For Providence a targeted attack simulation is run, which shows that carefully targeted extermination interventions could reduce infestation in the city by approximately 40%.

Keywords: City, Rodent infestation, Spreading, Random geometric graph

Introduction

Spreading algorithms have been studied with many different networks and contexts, such as diseases spread through human contacts (Newman 2002; Schneeberger et al. 2004; Bearman et al. 2002), information spread through email chains (Ebel et al. 2002) and mobile cellular networks (Onnela et al. 2007), and memes spread through marketing and promotion (Aral and Walker 2011; Weng et al. 2012). Similarly, rodents in a city will spread through a network of buildings. The spreading is typically not welcome and is referred to as an infestation. However, the rodents enjoy the surplus of food and habitat-constructing resources, easy accessibility, and warmth that urban environments provide. The spreading and location of rodents has been studied statistically (French et al. 1968; Feng and Himsworth 2014), but few network-based studies exist. Humans have numerous reasons to study the behavior of rodents, including management of rodent wastes, mitigation of damage that rodents cause, and prevention of diseases that are transmitted from rodents to humans (Mills and Childs 1998; Würbel 2001; Lund 2015).

The migratory patterns of rodents follow relatively short distances between hospitable environments. In an urban setting this likely means moving along a two-dimensional path from one building to another. One generative model for these migrations is a Random Geometric Graph (RGG), in which a radius r represents a rodent's migratory distance, n nodes representing buildings are distributed over a unit square $[0,1]^2$, and two nodes n_1 and n_2 are connected by an edge if their euclidean distance

$d(n_1, n_2) < r$. Similarly, Geometric Graphs, where the nodes are not placed randomly, but according to the locations of real buildings, can be used to model existing cities. As rodents spread through these networks, their pattern of infestation can be described by epidemic models such as Susceptible-Infected-Recovered (SIR) (Antulov-Fantulin et al. 2012). We examine their spreading by simulating infestation of two cities with very different characteristics, Providence, Rhode Island, USA and Tulsa, Oklahoma, USA.

It is probable that cities are not well represented by standard random network models, but in fact have topologies that are determined by natural factors such as bodies of water, properties of the terrain, and human factors such as urban planning. Residential and commercial buildings are generally built in blocks, not randomly, and cities typically have regulations that must be followed as new buildings are built. Since cities are often defined by their neighborhoods, we expect that the graph of a city will be dense, or will have communities that are dense. Highly populated cities are likely to be highly susceptible to infestation. Sparse and open cities may not be as susceptible, but even sparse cities have potential for large connected components that will be susceptible.

Many aspects of rodent spreading are not known with certainty. In this work we use a real city based geometric map model to study infestations with respect to a rodent's migration distance as well as the probability of spreading. By comparing simulations across different migration distance assumptions with known infestation data, we are able to determine bounds on true spreading rates.

Related Work
Random Geometric Graphs
RGGs in general have been extensively studied (Penrose 2003) and have been used to model wireless sensor networks (Kenniche and Ravelomananana 2010; Jia 2004), radio broadcasting (Elsässer et al. 2008) and vehicular ad hoc networks (Zhang et al. 2014). Many theoretical properties of RGGs are known. For example, the expected degree of $G(n, r)$ is $n\pi r^2$, and the critical radius at which connectivity is attained with high probability is $r_c = \sqrt{\frac{\log n \pm O(1)}{\pi n}}$. It is shown in (Dall and Christensen 2002) that toroidal (continuous) boundary conditions give different theoretical results than models with open boundary conditions. Here, we view the models as having boundaries, as exist with cities.

Variations on the RGG include a dynamic model where each vertex follows a random walk over time. Diaz et al. study the changing connectivity of this type of network (Díaz et al. 2009). The Gilbert model assumes the nodes locations are picked according to a Poisson process with density D points per unit area (Gilbert 1961). We note that the k-nearest neighbor model, where each node is connected to some number k of neighbors the closest Euclidean distance, is similar to a RGG (Balister et al. 2008). A preferential attachment model for geometric networks that combines aspects of RGGs and preferential attachment graphs is given in (Flaxman et al. 2006).

This work studies spreading behavior in RGGs. Preciado and Jadbabaie use spectral methods to develop a theoretical framework to describe spreading in RGGs in (Preciado and Jadbabaie 2009). The cover and mixing time of RGGs is studied in (Avin and Ercal 2007).

Network Spreading and SIR

Spreading is a classic problem in Network Science, to which the SIR model (where the states of nodes transition between Susceptible, Infected and Recovered) can be applied. This model has been applied to many problems, such as the spread of Salmonella in the pork supply chain (van der Gaag et al. 2004), the spread of the ecologically invasive South American tomato pinworm (Biondi et al. 2018), the social network spread of Charles Darwin's correspondence (Floyd 2019), identifying influential nodes (Chen et al. 2012), and tracking the spread of rumors through media (Zhao et al. 2013).

   Examples of using this model to study the spread of epidemic diseases are numerous (Newman 2002), with an abundant amount of literature on the spread of sexually transmitted diseases (Rocha et al. 2011; Sloot et al. 2008) and the effects of vaccination (Ruan et al. 2012). Many studies such as (Dean et al. 2018) and (Molefi 2001) look at the spread of plague, which is indirectly a study of the spread of rodents. The spreading of disease in mice contact networks is studied in (Lopes et al. 2016), although not the spreading of the mice themselves. One review paper calls network models "an underutilized tool in wildlife epidemiology" (Craft and Caillaud 2011). We could not find any studies describing the movement of rodents with a network approach. Hence, the novelty of this work.

Rodents

The nesting and migration patterns of rodents are interesting both in the contexts of conservation and pest control. Of particular interest are studies of rodent behavior in cities. Cavia et al. (2009) studied rodent communities in Buenos Aires in the diverse habitats of cities, shanty towns, parklands and natural reserves. Rodents caught in the city environment were mostly of the non-native species *R.Rattus* and *R. norvegicus*. Generally it was found that species diversity decreased in an urban environment. A study of rodent populations in urban Salzburg, Austria (Traweger et al. 2006) found that rats were most attracted to garden areas, and areas with nearby running or standing water sources. The study, which measured rat populations by trapping, showed that roughly 35% of the 71 discrete patches of the city in which traps were located contained rat populations. This density was well below estimates from government sources and pest control companies. It was also shown that rat populations were "distributed throughout the city area in patches" and not evenly as previously thought. A similar study in Sao Paulo, Brazil (Masi et al. 2010) which sampled 23,512 premises found an infestation rate of 23.1%. Conditions leading to rodent infestation in buildings and structures were found to be easy access, which is enabled by structural deficiencies, and food sources, particularly human food and garbage. It is concluded that unrepaired buildings and trash-littering are more common in low income areas, making rodent infestation to an extent representative of the socioeconomic conditions of a neighborhood.

   A study in Madrid, Spain (Ayyad et al. 2018), using the locations of 470 government-reported rat sightings applied various statistical models to determine the likelihood of rodent infestations at different locations. It found that "water source points, cat feeding stations, and green zones are closely related to rat proliferation providing a significantly direct relationship with the presence of near by rats" (Ayyad et al. 2018). Other related studies, such as Baker et al. (2003) discuss the factors affecting the distribution of rodents in an urban area, some of which include the prevalence of cats and the distance to food sources. In (Childs et al. 1998) the epidemiology of rodent bites in New York City is

studied to determine the locations of infestation which are then examined against control locations to predict attractiveness to rodents. It is found that 22% of city blocks are high risk for rodent bites, implying serious infestation, while fully 50% of blocks are rated low risk (although some rodent populations were found even in low-risk areas).

The problems associated with rodent habitation in cities are widely known. Rodents such as rats have single pairs of sharp upper and lower incisors that are not only used to break up food, but also to gnaw through and remove "non-food particles that may intervene between a rodent and its food" (Drummond 2001). This means that, for humans, the cost of rodent infestation includes not only damage to and loss of food supplies, but also gnawing damage to containers in which food is stored (Lund 2015), as well as to personal items and electrical wiring. In addition, there are hygienic reasons to control rodent populations, such as avoiding contact with their waste. Associated with food are also costs of food industry closures due to infestations (Meehan 1984), increased costs of hygienic food-storage procedures and economic loss of business and reputation when rodents are discovered. In addition to the direct losses, it was claimed by US insurance companies in the 1940s that "about 25% of fires for which there was no apparent cause may have been caused by rodents" (Battersby 2004).

A final important issue is rodent involvement in human health problems. In addition to the bites mentioned above, rodents are associated with transmission of human diseases. Plague is the most well-known rodent-involved illness, with over 2000 cases still reported annually (Titball and Leary 1998). Other diseases caused by direct or indirect rodent contact include hantavirus pulmonary syndrome, leptospirosis (Gubler et al. 2001), hemorrhagic fever (Tsai 1987), lassa fever (Ter Meulen et al. 1996), and the gastrointestinal disease cryptosporidiosis (Quy et al. 1999).

Methods

The research method followed in this work includes four steps, as shown in Fig. 1. Data for cities (in this case Tulsa and Providence) are collected from an open source mapping software, and are cleaned, normalized, and converted in format. RGG networks are generated from this data. Spreading simulations are run assuming different input parameters. If promising, the city data is further analyzed for important nodes using the immunization test.

Fig. 1 Data Driven Research and Analysis Method
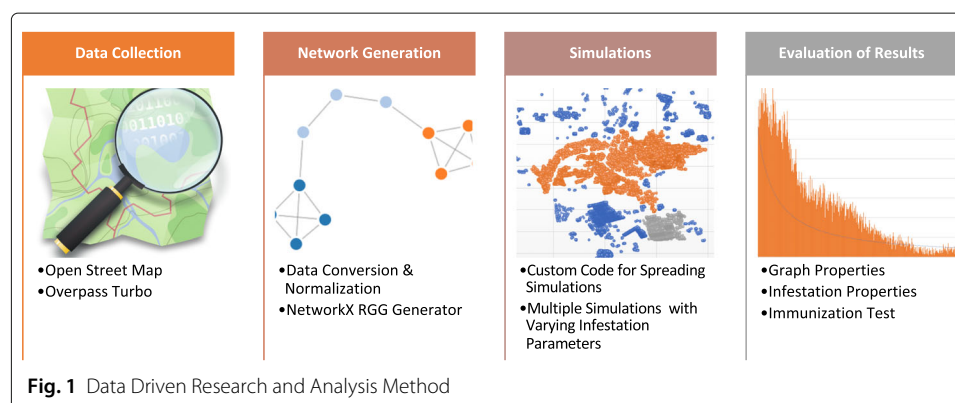
Data Collection

Providence is one of the oldest cities in the United States. It was founded in 1636, is a capital city, and is densely covered with over 2100 buildings per square mile. Providence was formerly heavily industrialized, has many factory buildings and warehouses, and is bordered to the east by the Providence River, which rodents cannot cross. Tulsa is a much newer city, founded between 1828 and 1836, and is much less densely covered with only 472 buildings per square mile. Tulsa is located on two sides of the Arkansas River and has a hilly terrain.[1] We used the OpenStreetMap API (Bennett 2010) to obtain data about both cities.

The OpenStreetMap API is an open-source mapping software comprised of data entered by its users. The information consists of over 5 billion nodes contributed by over 1 million contributors. The correctness of the OpenStreetMap data has been widely analyzed and was found to be accurate within about 6 meters (Haklay 2010). This paper uses building footprint data from OpenStreetMap. A study of such data in Munich found that OpenStreetMap footprints exhibit high completeness and semantic accuracy, and that "there is an offset of about four meters on average in terms of position accuracy" (Fan et al. 2014).

Overpass Turbo (Raifer 2018) is a web application that facilitates querying Open-StreetMap data. We query Overpass Turbo for all buildings in a selected city. An example query for Providence is shown in Listing 1, and a map of the results is shown in Fig. 2a.

```
[out:json];

{{geocodeArea:Providence}}->.searchArea;
(
node["admin_level"](area.searchArea);
way["admin_level"](area.searchArea);
relation["admin_level"](area.searchArea);
);

(
way["building"](area.searchArea);
relation["building"](area.searchArea);
);

out geom;
>;
out skel qt;
```
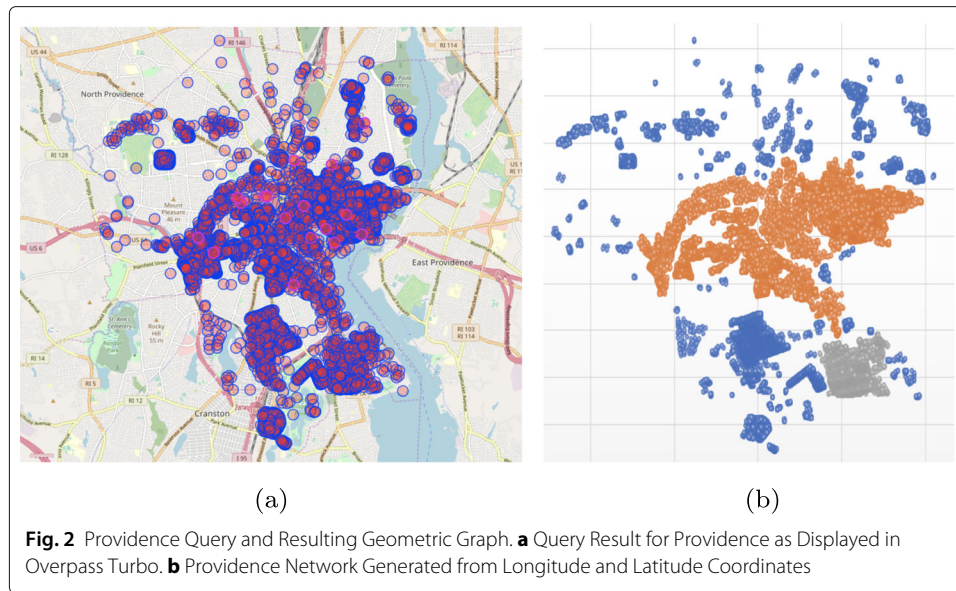
Listing 1 Overpass-Turbo Query

The query returns raw data which consists of longitude and latitude coordinates for each building, as well as unneeded information. The information returned is carefully cleaned and parsed such that it contains only the coordinates of each building. Buildings will become nodes in our network. The coordinates of each node are then stored in a Python dictionary with the node as the key and coordinates as the value, and the dictionary is then saved to a file.

Network Generation

NetworkX (Hagberg et al. 2008) is used for network generation and further analysis. We read the Python dictionary containing building coordinate data for a chosen city and use

---

[1]The density calculations are derived from our results using OpenStreetMap data.

(a)   (b)

Fig. 2 Providence Query and Resulting Geometric Graph. **a** Query Result for Providence as Displayed in Overpass Turbo. **b** Providence Network Generated from Longitude and Latitude Coordinates

it to generate a geometric graph network. The resulting network for Providence is shown in Fig. 2(b). The raw coordinate data must be converted into a format that can be loaded into NetworkX, such as GML (Graph Modeling Language) (Himsolt 2000). To generate a GML file for the chosen city, we use the NetworkX Random Geometric Graph function, which allows passing the dictionary of positions as a parameter. To facilitate computation, the coordinate dictionary is normalized using min-max normalization. With the coordinates normalized, the geometric graph radius (or the migration radius of a traveling rodent) must be converted from feet to degrees of latitude. Each degree of latitude is approximately 69 miles. Using this information we can convert a radius of 450 feet into approximately 0.00124 degrees latitude. However, this still must be normalized to match the coordinates specific to each city.

After the raw data is formatted, we use NetworkX's Random Geometric Graph function with the parameters of city size, normalized radius, dimensions, and normalized dictionary of building positions. This gives a NetworkX graph object on which that we can then perform simulations.

Simulations

In this work simulations are run on geometric graphs created from real data on the cities of Tulsa and Providence. The simulations examine changes in infestation rates that result from different parameter assumptions. Important parameters in our simulations include a rodent's migration distance r, and the spreading probability β of a building becoming infested if its neighbor is infested.

Two studies exist that help determine the migration distance. Feng and Himsworth (2014) states that rodents have a home range of about a city block, and French et al. (1968) studied many species of rats and their migration patterns. These studies help to define a radius for the networks. For migration distance, we start with one of French's results for the *dipodomys microps* species. Although this is not the typical city rat, it gives a starting point for a rodent's migration pattern. French tracked the movement of *dipodomys*

microps, and determined that they stayed within their original grid but moved 450 feet. Based on the research of block sizes in (Siksna 1997) and (Stangl 2015), and Feng's suggestion that rodents travel about a city block, 450 feet is a good estimate of distance for our migration radius. Because this distance is not definitively known and can vary among rodent species and geographic conditions, we test 4 different values of r: 400, 450, 500 and 550 feet.

The spreading probability β of a building becoming infested is unknown and relies on factors that vary between buildings, such as accessibility, amount of resources, and number of deterrents. Thus, we test multiple probabilities in order to get a range of results. For the spreading parameter β we considered 20 values ranging from 0.5 to 40%. At $\beta = 40\%$ an infestation spreads quickly to almost 100% of nodes, so larger values were not evaluated. For each β value 100 simulations were run on the Tulsa network and 10 simulations were run on the Providence network. For each simulation we keep a master list identifying whether each node was infested. The set statistics give an average infestation rate, and the most frequently infested nodes across all simulations can be determined from the master list.

Simulations are run using the NetworkX graph object. The simulation function is based on the SIR epidemic algorithm from Antulov-Fantulin et al. (2012). The algorithm takes as parameters a (NetworkX) Graph, a probability of infestation β, a probability for recovery, a queue of initially infested nodes, and a list of susceptible nodes. This algorithm was modified slightly to fit our needs. In order to make each node equally susceptible, the list of susceptible nodes was removed. We assume that once an infestation has cycled through a node's neighbors, the node will be removed from the queue of infested nodes and is thus fully recovered. This implies a recovery rate of 100 percent. Besides full recovery of each infested node, we also assume that once a node is infested it can not be infested again by its neighbors. Initially, experiments were run with the assumption of non-complete recovery rates and re-infestation. It was observed that, although the infestation rate grew more quickly, the difference in results did not warrant the greatly increased execution time. These modifications are analogous to a situation in which rodents migrate and do not return to a previous location. This is a realistic assumption, as rodents migrate due to changing conditions, such as removal of food sources or the presence of cats.

The Immunization Test

Where appropriate, the robustness of the networks to targeted attacks is examined using the immunization test. This test, as described in (Holme et al. 2002), involves repeatedly removing high importance nodes and measuring the damage caused, as indicated by the size of the largest connected component. A quick drop in the size of the largest component with relatively few nodes removed indicates a susceptibility to attack. We used betweenness centrality, approximated via a GPU algorithm (McLaughlin and Bader 2014; Matta et al. 2019), as the measure of node importance. In this paper, the removal of a node (which represents a geographical place like a building) occurs when circumstances cause the extermination of rodents from that node. Examples would include interventions such as human residents getting a cat, setting traps, or patching holes in walls that allow rodents to enter.

Results

During data collection the locations of 88,260 buildings in Tulsa and 43,662 buildings in Providence were gathered, and four separate networks were created for each city based on rodents' migration radius r. Information for all networks created is summarized in Table 1. These two cities vary in density and size. Tulsa has an area of 186.8 square miles and Providence's area is 20.6 square miles. Providence has over 2100 buildings per square mile, whereas Tulsa has only 472 buildings per square mile.

These density differences affect the corresponding networks. The average degree is much higher for Providence than for Tulsa. For both networks increasing r by 50 feet adds approximately 1,000,00 edges. Because of Providence's dense nature, it has a much larger largest connected component than Tulsa. The size of its largest connected component ranges from 23,263 to 26,165 nodes depending on r. Therefore, in all cases more than half of Providence's buildings are connected within a geometric network. The geographical structure of Providence is also reflected in its largest connected component. It has three distinct regions that are difficult for rodents to travel between. One result of this is that the size of the largest connected component remains the same when r is increased from 450 to 500 feet. Tulsa's largest component varies from 7909 to 12,420 nodes. It is apparent that Tulsa has much more variance in largest component size than Providence. Tulsa's low density results in many disconnected components, while Providence has relatively few.

Tulsa

Network Structure

Tulsa is the larger of the two city networks studied. The Tulsa network had many clusters that could facilitate the spreading of infestation. However, these clusters were isolated from each other, and thus the infestation could only spread within the starting node's cluster. This prevented a full outbreak and contained the infestation to the cluster. Since our simulations started with a single random node, the random node could only infest a subgraph of the overall network, implying that the greatest amount of nodes infested would be the largest component. A visualization of the Tulsa clusters is shown in Fig. 3a.

Simulation Results

Tulsa citywide infestation rates with varying values of r and β are shown in Fig. 3b. Note that, while higher values of r imply higher infestation rates, because of the relatively small size of the largest connected component, the citywide infestation rates stay within a narrow range of values and do not rise above 4.5%. Over the range of spreading probabilities,

Table 1 Networks analyzed

| City	r (feet)	Nodes	Edges	Clusters	Largest component	Average degree
Tulsa	400	88260	5004030	669	7907	113.4
Tulsa	450	88260	5991094	587	7936	135.8
Tulsa	500	88260	7025832	546	8499	159.2
Tulsa	550	88260	8081604	506	12420	183.1
Providence	400	43662	4182319	102	23263	191.8
Providence	450	43662	5089332	93	25891	233.1
Providence	500	43662	6062430	84	25891	277.7
Providence	550	43662	7108720	74	26165	325.6
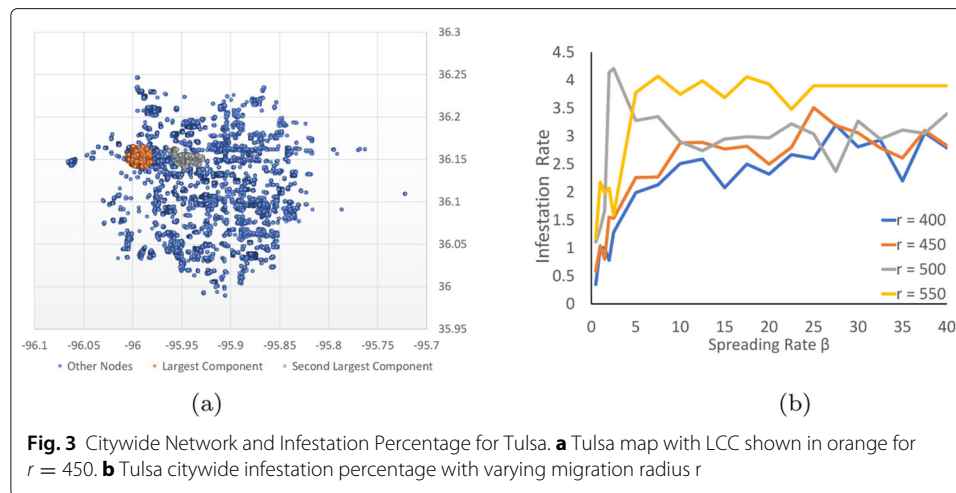
Fig. 3 Citywide Network and Infestation Percentage for Tulsa. **a** Tulsa map with LCC shown in orange for $r = 450$. **b** Tulsa citywide infestation percentage with varying migration radius r

the highest infestation was with the $r = 500, \beta = 0.025$ network, with 4.21% of the network infested. At β values over 5%, the $r = 550$ networks were consistently more infested than the others, although the infestation rate generally did not rise above 4%. If one of the random initially infested nodes was within Tulsa's largest connected component, then a substantial amount of the largest component became infested. Because infestation studies are often conducted within the largest connected component, it is useful to explore its infestation properties further.
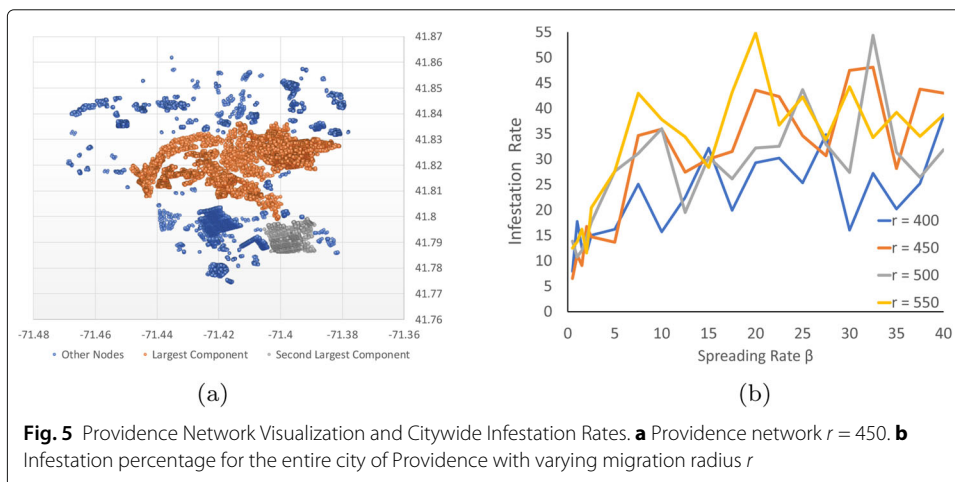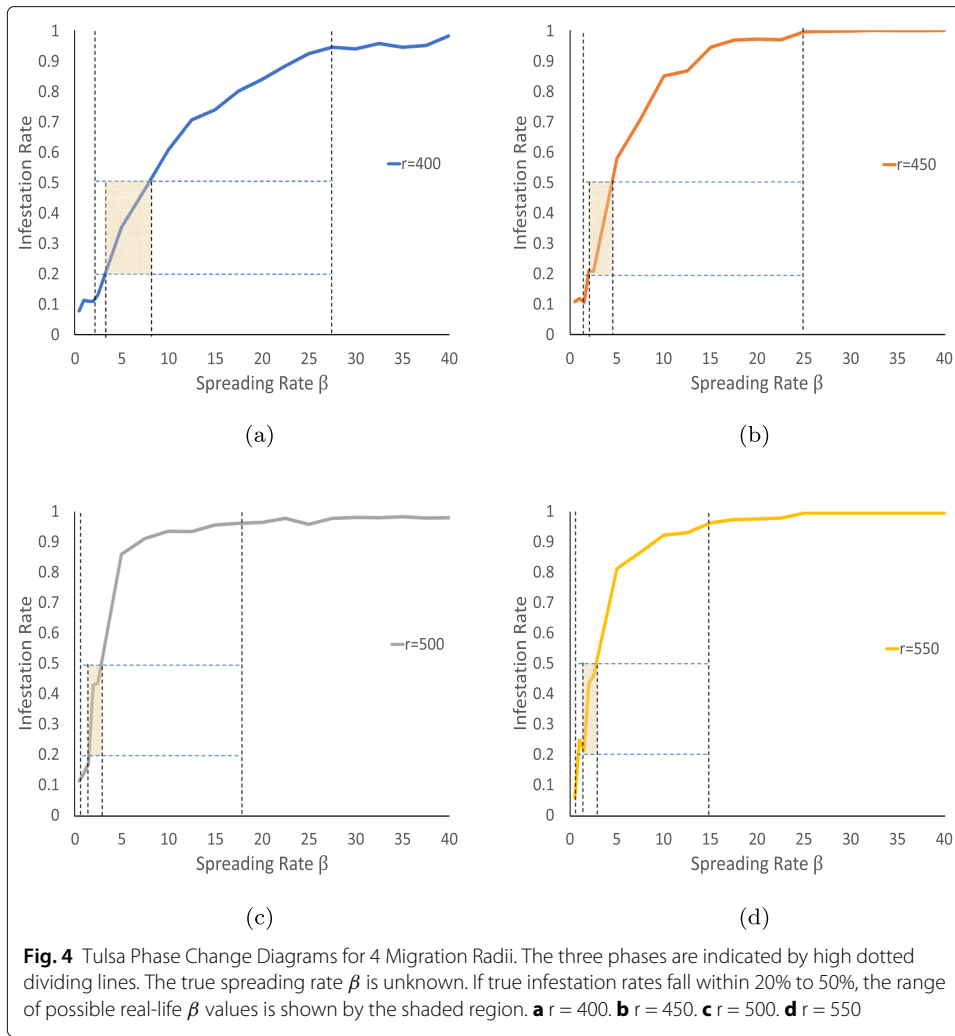
Phase transition diagrams showing infestation rates for the largest component over a range of r values (400, 450, 500 and 550) as well as a series of β values (0.5% to 40%) are shown in Fig. 4. For all 4 values of r, the infestation moves through three phases, which are delineated in the charts by high dotted lines. There is an initial phase, in which infestation rates are small (approximately 10%), followed by a phase of rapid growth, then a saturation phase. We consider the saturation phase to begin at the β which results in an infestation over 96%. It is noted that as r increases, the rapid growth phase and the saturation phase both start at lower values of β, and that the rapid growth region becomes smaller.

Previous work has estimated infestation rates at between 20% (Masi et al. 2010) and 50% (Childs et al. 1998). These rates are demarcated by horizontal lines in Fig. 4. The true spreading rate is unknown, but it must occur in the range of plausible infestation rates. The possible range of true spreading parameter values is shown as a shaded region. For example, the true β for Tulsa must be between approximately 3.5% and 8%, if $r = 400$. For $r = 450$, the range is reduced to between 2% and 5%. Results are quite similar for $r = 500$ and $r = 550$, with a range between approximately 1.5% and 3%.
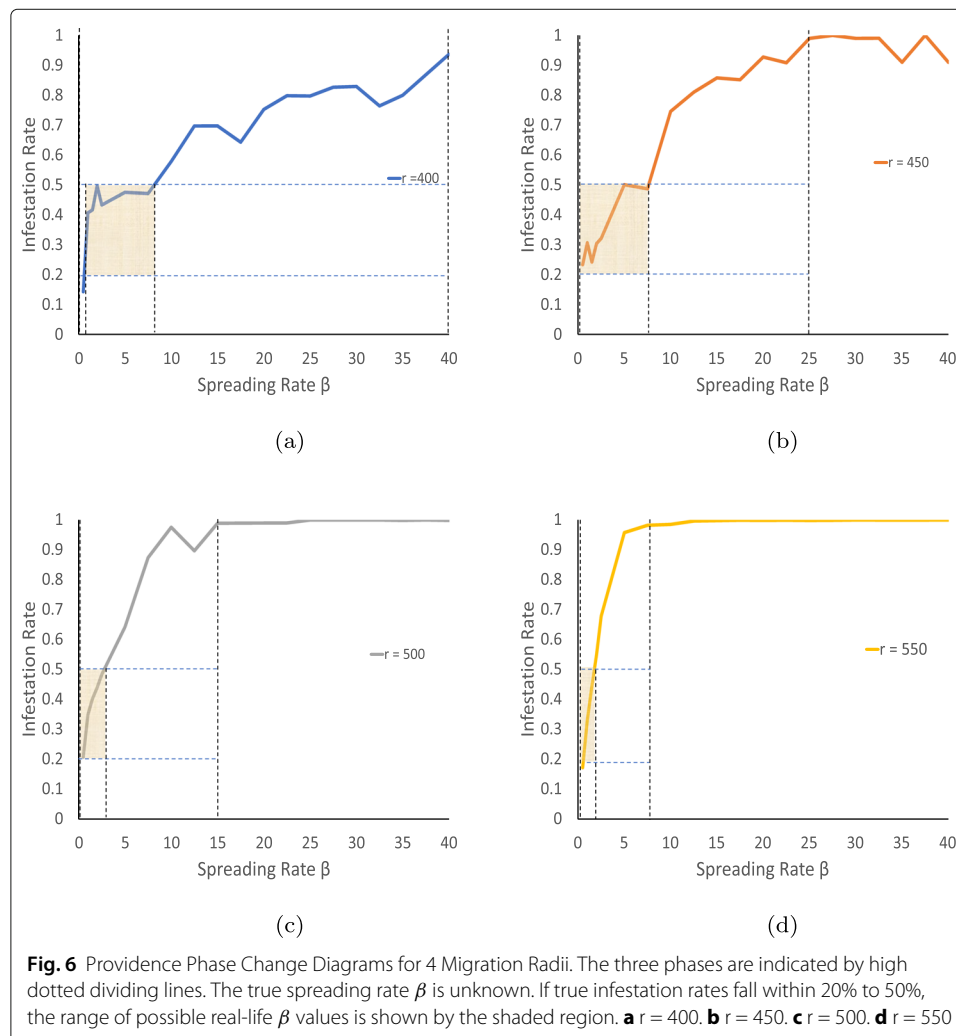
Providence

Network Structure

Figure 5a shows the Providence network for $r = 450$. Providence's structure is different from Tulsa's. Since the area is more dense, there are fewer clusters. In the case shown, the largest component of the city contains 25,891 nodes and makes up almost 60% of the network. With 60% of the network being connected, there are many more opportunities for the infestation to spread and a higher chance that a random starting node will be within the large component.

Fig. 4 Tulsa Phase Change Diagrams for 4 Migration Radii. The three phases are indicated by high dotted dividing lines. The true spreading rate β is unknown. If true infestation rates fall within 20% to 50%, the range of possible real-life β values is shown by the shaded region. **a** r = 400. **b** r = 450. **c** r = 500. **d** r = 550

Fig. 5 Providence Network Visualization and Citywide Infestation Rates. **a** Providence network r = 450. **b** Infestation percentage for the entire city of Providence with varying migration radius r

Simulation Results

As shown in Fig. 5b, the Providence network's infestation rate was much higher than that of Tulsa. As an example, with $r = 450$ and spreading probability β of 0.10, about 25% of nodes were infested on average. This is substantially higher than Tulsa's maximum of 4.2%. The infestation rate increased to as high as 54.8% in the $r = 550$ network. Overall, the citywide infestation rate tended to increase with r, although generally the infestation rate stayed between 15 and 50%. These infestation rates are consistent with infestation rates found in similarly dense cities such as Salzburg, Austria (35%) (Traweger et al. 2006), Sao Paulo, Brazil (23.1%) (Masi et al. 2010), and New York City, USA (between 22% and 50%) (Childs et al. 1998).

Phase transition diagrams for several different r values showing infestation rates for the Providence largest component over a series of β values are given in Fig. 6. These results are different in nature from the Tulsa phase diagrams. First, because of the density of Providence, the initial stage ends very rapidly. In our simulations, except where $r = 400$, the initial stage occurred somewhere below $\beta = 0.5\%$. At $\beta = 0.5$ the rapid growth phase had already begun, and the infestation rate was generally 20% or above. At $r = 400$ there is a relatively wide range of values for β, from 0.5% to 8%. Interestingly, for $r = 400$ a

Fig. 6 Providence Phase Change Diagrams for 4 Migration Radii. The three phases are indicated by high dotted dividing lines. The true spreading rate β is unknown. If true infestation rates fall within 20% to 50%, the range of possible real-life β values is shown by the shaded region. **a** r = 400. **b** r = 450. **c** r = 500. **d** r = 550

saturation phase is never achieved, even with spreading parameter as large as 40%. With $r = 450$ the range for true β is decreased to a maximum of to 7%. As r is increased to 500 and 550, the upper bound on the true β is further decreased, to 4% and 2%.

Immunization

We have shown that the largest component is influential in spreading simulations. Small, isolated clusters produce lower infestation rates, but the Providence networks had at least 60% of nodes in a connected component. Therefore, Providence is more vulnerable to infestation. One network-based way to contain an infestation is to follow an attack strategy of removing nodes with high betweenness centrality (Holme et al. 2002). This breaks up the largest connected component and limits the maximum number of nodes that can become infested in an outbreak.

Performing the immunization algorithm on the real network of Providence ($r = 450$) provides interesting results, as shown in Table 2 and Fig. 7. Upon removing the 10 nodes with highest betweenness centrality, the size of the largest component drops by 7374, from 25,891 nodes to 18,517 nodes. The separation of the largest component is demonstrated in Fig. 7b. The next large change occurs when 2 additional nodes are removed. With removal of just 11 nodes, three components emerge. The size of the largest component drops from 18,517 to 15,888 nodes as the third component becomes isolated. This is depicted in Fig. 7c. The final result is 3 components with 18,516 nodes, 7374 nodes, and 2628 nodes respectively.

The nodes are based on real building coordinates, which means that the addresses of each building immunized during the split up of the largest component can be geographically pinpointed. For Step 9, node 2491 is located at (41.81830, -71.41670), which equates to 47 Franklin St., Providence, RI 02903. The building at 47 Franklin St. is the United Church of Christ. Therefore, removing the church from the network, perhaps via an extermination intervention, along with the 9 nodes before it, splits the largest component and would be a major deterrent to the further spread of rodents. The next major change occurs at step 11 when node 43590 is removed. Node 43590 is located at (41.83249, -71.41329), which is an office building at 1 State St., Providence, RI 02908. The building hosts a Retail Decisions gift shop, American Heart Association, and Strategies for Wealth Financial Institute. Again, an intervention at this location would cause a major disruption to spreading in the network.

Table 2 Providence immunization

Step	Node removed	Node coordinates	Clusters	Largest component
0	18767	(41.83062, -71.41322)	93	25890
1	1682	(41.82955, -71.41733)	93	25889
2	34200	(41.81773, -71.41769)	93	25888
3	18768	(41.83048, -71.41327)	93	25887
4	38988	(41.82790, -71.41555)	93	25886
5	38989	(41.82790, -71.41551)	93	25885
6	2495	(41.81841, -71.41692)	93	25884
7	34199	(41.81762, -71.41756)	93	25883
8	1339	(41.82075, -71.41455)	93	25882
9	2491	(41.818302, -71.41670)	94	18517
10	42890	(41.83125, -71.40982)	94	18516
11	43590	(41.83249, -71.41329)	95	15888

Discussion

Providence provided a dense network that contained relatively few clusters and a single large connected component. Tulsa provided a sparse network with many smaller clusters and a few large components. Networks were generated based on varying migration radius values. For both cities, these networks had an average degree ranging from approximately 100 to 300. In practical terms, even a relatively low average degree such as 113.4 for the Tulsa $r = 400$ network implies that an average rodent has many options for migration to surrounding buildings. Increases in r have a substantial influence on average degree. For a rodent in Providence, increasing its mobility from 400 to 550 feet increases its migration options from almost 200 building locations to over 300.

Phase change diagrams (Figs. 4 and 6) show infestations generally following three phases: an initial phase of limited local infestation, followed by rapid growth and saturation. In our simulations the rapid growth phases begin at relatively low β values. This is consistent with real-life data and network theory.

It is well known that in the Erdős-Rényi random model a giant component begins to form at average degree of 1 (Barabási 2016). Although our city networks are not random, we see a similar phenomenon here. If a node has a degree of 200, a 0.5% spreading probability implies a migration from the average node to one other node. This level of migration is sufficient to begin the phase transition to rapid growth. In all cases, a β of 3.5% was sufficient to increase the infestation rate to above 10%. Infestations quickly reached over 90% of buildings with spreading parameter β values as low as 5% in the Providence $r = 550$ network.

Our simulations show that even small increases in r and β lead to greatly increased infestation rates within a connected component. Given that true infestation rates are not 100%, it is logical to seek to determine what is stopping rodents.

We do not know the true infestation rates of Tulsa or Providence, although common wisdom and pest control companies maintain that it is large. By contrast, our citywide simulations showed unexpectedly low infestation rates. For a sparse city like Tulsa the rates were approximately 3%. For a dense city like Providence, rates were between 25% and 54%. Interestingly, studies of dense cities using diverse methods such as trapping, analyzing government-reported rodent citing data, and monitoring rodent bites produced results similar to ours.

In our simulations, the spread of rodents was determined by two factors. First, cities consist of disjoint components. Infestations are contained within these connected components. This is the reason the simulation showed low infestation rates for Tulsa. It does not explain results for the Providence network, where over 60% of the city is in a single component. Within our simulations, the spreading parameter β determines the large component infestation rate. The true β parameter is unknown, so our simulations tested a range of values. Within this range, the true β for a city must result in a plausible infestation rate. Our simulations estimate the true β at between 0.5% and 8%, depending on the true r. Realistically, this makes sense. A β as high as 50% implies an unlikely scenario in which 150 rodents migrate from the *average* building in Providence (where $r = 550$). A true β of 1% implies that two or three rodents migrate from a building, which is still enough to result in a component infestation rate of over 40%.

Fig. 7 Providence Immunization. **a** Providence Largest Component. **b** First Detachment. Removal of only 10 nodes divides the largest component in two. Relevant removed nodes are shown in red, and the detached cluster is shown in orange. **c** Second Detachment. Removal of two additional nodes disconnects a third cluster, which is shown in gray
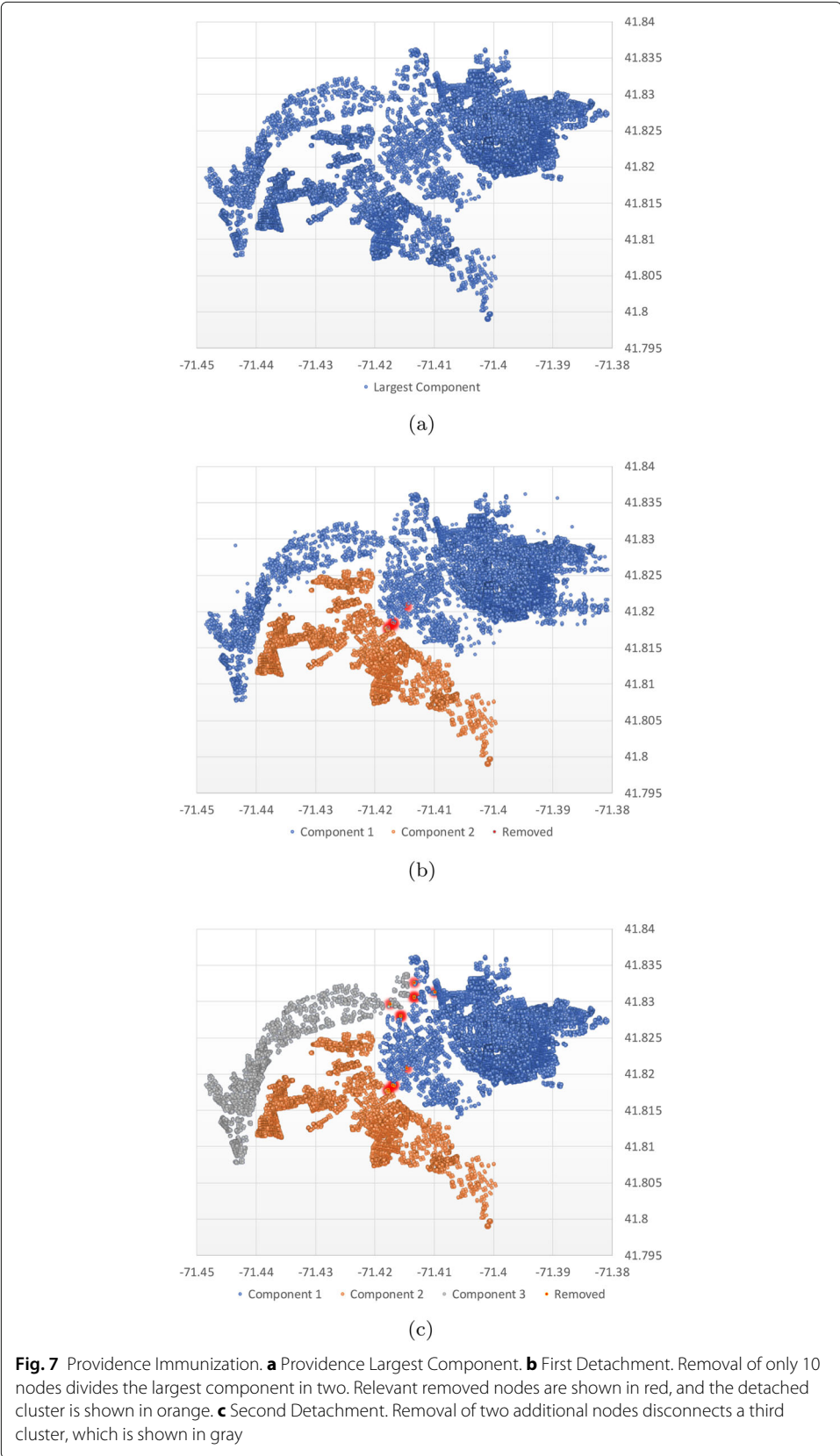
Immunization provides a way to prevent spreading, and the disconnection of the large component in Providence is an example of that. Breaking up the large component occurred when immunizing only 11 nodes with high betweenness centralities, like the United Church of Christ and an office building on State Street. Setting traps and exterminating pests can have a large impact on spreading, way beyond the immediate local effects. In the current example, immunizing 11 nodes would have reduced city-wide spreading by 39%.

Conclusion

In this paper, we gathered data using Open Street Map and Overpass Turbo to generate networks using coordinates of buildings in two cities, Providence and Tulsa. With the gathered coordinates, we converted the raw data into networks and simulated the spread of rodents throughout the networks. Our results show that even small increases in spreading parameter β and migration radius r result in large increases in infestation rates, meaning that dense cities, like Providence, are more prone to rodent infestations. Sparse cities, such as Tulsa, are less likely to allow the spread of rodents. Scientific studies estimate the migration radius r of a rodent to be 450 feet. Our simulations confirm the plausibility of that estimate. The real-life spreading parameter β was estimated to be between 0.5% and 8%. We then performed an immunization algorithm on the Providence network, which split the largest component and provided exact coordinate locations that would be useful for rodent traps and exterminations. Based on previous simulation results, we determined that the break up of the largest component in Providence would meaningfully reduce the infestation rates for the whole city.

References
Antulov-Fantulin N, Lancic A, Stefancic H, Sikic M (2012) Fastsir algorithm: A fast algorithm for simulation of epidemic spread in large networks by using SIR compartment model. CoRR abs/1202.1639. 1202.1639. Accessed date 18 Oct 2019

Aral S, Walker D (2011) Creating social contagion through viral product design: A randomized trial of peer influence in networks. Manag Sci 57(9):1623–1639

Avin C, Ercal G (2007) On the cover time and mixing time of random geometric graphs. Theor Comput Sci 380(1-2):2–22

Ayyad C, Mateu J, Tamayo-Uria I (2018) Non-linear spatial modeling of rat sightings in relation to urban multi-source foci. J Infect Public Health 11(5):667–676

Baker PJ, Ansell RJ, Dodds PA, Webber CE, Harris S (2003) Factors affecting the distribution of small mammals in an urban area. Mamm Rev 33(1):95–100

Balister P, Sarkar A, Bollobás B (2008) Percolation, connectivity, coverage and colouring of random geometric graphs. In: Handbook of Large-scale Random Networks. Springer, Berlin. pp 117–142

Barabási A-L (2016) Network Science. Cambridge university press, Cambridge

Battersby J (2004) Public health policy–can there be an economic imperative? an examination of one such issue. J Environ Health Res 3(1):19–28

Bearman PS, Moody J, Stovel K (2002) Chains of affection: The structure of adolescent romantic and sexual networks. American journal of sociology 110(1):44–91

Bennett J (2010) OpenStreetMap. Packt Publishing Ltd, Birmingham

Biondi A, Guedes RNC, Wan F-H, Desneux N (2018) Ecology, worldwide spread, and management of the invasive south american tomato pinworm, tuta absoluta: past, present, and future. Ann Rev Entomol 63:239–258

Cavia R, Cueto GR, Suárez OV (2009) Changes in rodent communities according to the landscape structure in an urban ecosystem. Landsc Urban Plan 90(1-2):11–19

Chen D, Lü L, Shang M-S, Zhang Y-C, Zhou T (2012) Identifying influential nodes in complex networks. Physica A Stat Mech Appl 391(4):1777–1787

Childs JE, McLafferty SL, Sadek R, Miller GL, Khan AS, DuPree ER, Advani R, Glass GE (1998) Epidemiology of rodent bites and prediction of rat infestation in new york city. Am J Epidemiol 148(1):78–87

Craft ME, Caillaud D (2011) Network models: an underutilized tool in wildlife epidemiology? Interdisc Perspect Infect Dis 2011:12

Dall J, Christensen M (2002) Random geometric graphs. Phys Rev E 66:016121. https://doi.org/10.1103/PhysRevE.66.016121

Dean KR, Krauer F, Walløe L, Lingjaerde OC, Bramanti B, Stenseth NC, Schmid BV (2018) Human ectoparasites and the spread of plague in europe during the second pandemic. Proc Natl Acad Sci 115(6):1304–1309

Díaz J, Mitsche D, Pérez-Giménez X (2009) Large connectivity for dynamic random geometric graphs. IEEE Trans Mob Comput 8(6):821–835. https://doi.org/10.1109/TMC.2009.42

Drummond D (2001) Rodents and biodeterioration. Int Biodeterior Biodegrad 48(1-4):105–111

Ebel H, Mielsch L-I, Bornholdt S (2002) Scale-free topology of e-mail networks. Phys Rev E 66(3):035103

Elsässer R, Gasieniec L, Sauerwald T (2008) On radio broadcasting in random geometric graphs. In: International Symposium on Distributed Computing. Springer, Berlin. pp 212–226

Fan H, Zipf A, Fu Q, Neis P (2014) Quality assessment for building footprints data on openstreetmap. Int J Geogr Inf Sci 28(4):700–719

Feng AYT, Himsworth CG (2014) The secret life of the city rat: a review of the ecology of urban norway and black rats (rattus norvegicus and rattus rattus). Urban Ecosyst 17(1):149–162. https://doi.org/10.1007/s11252-013-0305-4

Flaxman AD, Frieze AM, Vera J (2006) A geometric preferential attachment model of networks. Internet Math 3(2):187–205

Floyd C (2019) A social network and text analysis of charles darwin's correspondence, 1835-1842. PhD thesis

French NR, Hayden P, Tagami TY (1968) Dispersal in a Population of Desert Rodents. J Mammal 49(2):272–280. https://doi.org/10.2307/1377984. http://oup.prod.sis.lan/jmammal/article-pdf/49/2/272/2495232/49-2-272.pdf

van der Gaag MA, Vos F, Saatkamp HW, van Boven M, van Beek P, Huirne RB (2004) A state-transition simulation model for the spread of salmonella in the pork supply chain. Eur J Oper Res 156(3):782–798

Gilbert EN (1961) Random plane networks. J Soc Ind Appl Math 9(4):533–543

Gubler DJ, Reiter P, Ebi KL, Yap W, Nasci R, Patz JA (2001) Climate variability and change in the united states: potential impacts on vector-and rodent-borne diseases. Environ Health Perspect 109(suppl 2):223–233

Hagberg A, Swart P, S Chult D (2008) Exploring network structure, dynamics, and function using networkx. Technical report. Los Alamos National Lab.(LANL), Los Alamos

Haklay M (2010) How good is volunteered geographical information? a comparative study of openstreetmap and ordnance survey datasets. Environ Plan B Plan Des 37(4):682–703

Himsolt M (2000) Graphlet: design and implementation of a graph editor. Software: Practice and Experience 30(11):1303-1324

Holme P, Kim BJ, Yoon CN, Han SK (2002) Attack vulnerability of complex networks. Phys Rev E 65(5):056109

Jia X (2004) Wireless networks and random geometric graphs. In: 7th International Symposium on Parallel Architectures, Algorithms and Networks, 2004. Proceedings., Hong Kong. pp 575–579. https://doi.org/10.1109/ISPAN.2004.1300540

Kenniche H, Ravelomananana V (2010) Random Geometric Graphs as model of Wireless Sensor Networks. In: The 2nd International Conference on Computer and Automation Engineering (ICCAE), Singapore. pp 103–107. https://doi.org/10.1109/ICCAE.2010.5451758

Lopes PC, Block P, König B (2016) Infection-induced behavioural changes reduce connectivity and the potential for disease spread in wild mice contact networks. Sci Rep 6:31790

Lund M (2015) Commensal rodents, Rodent pests and their control. CABI, Oxfordshire

Masi E, Pino FA, Maria das Graças SS, Genehr L, Albuquerque JOM, Bancher AM, Alves JCM (2010) Socioeconomic and environmental risk factors for urban rodent infestation in sao paulo, brazil. J Pest Sci 83(3):231–241

Matta J, Ercal G, Sinha K (2019) Comparing the speed and accuracy of approaches to betweenness centrality approximation. Comput Soc Networks 6(1):2

McLaughlin A, Bader DA (2014) Scalable and High Performance Betweenness Centrality on the GPU. In: SC '14: Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis, New Orleans. pp 572–583. https://doi.org/10.1109/SC.2014.52

Meehan AP (1984) Rats and Mice. Their Biology and Control. Rentokil Ltd., East Grinstead

Mills JN, Childs JE (1998) Ecologic studies of rodent reservoirs: their relevance for human health. Emerg Infect Dis 4(4):529

Molefi RKK (2001) Of rats, fleas, and peoples: towards a history of bubonic plague in southern Africa, 1890-1950. Pula: Botswana Journal of African Studies 15(2):259–267

Newman MEJ (2002) Spread of epidemic disease on networks. Phys Rev E 66:016128. https://doi.org/10.1103/PhysRevE.66.016128

Onnela J-P, Saramäki J, Hyvönen J, Szabó G, Lazer D, Kaski K, Kertész J, Barabási A-L (2007) Structure and tie strengths in mobile communication networks. Proc Natl Acad Sci 104(18):7332–7336

Penrose M (2003) Random Geometric Graphs vol. 5. Oxford university press, Oxford

Preciado VM, Jadbabaie A (2009) Spectral analysis of virus spreading in random geometric networks. In: Proceedings of the 48h IEEE Conference on Decision and Control (CDC) held jointly with 2009 28th Chinese Control Conference, Shanghai. pp 4802–4807. https://doi.org/10.1109/CDC.2009.5400615

Quy RJ, Cowan DP, Haynes PJ, Sturdee AP, Chalmers RM, Bodley-Tickell AT, Bull SA (1999) The norway rat as a reservoir host of cryptosporidium parvum. J Wildl Dis 35(4):660–670

Raifer M (2018) Overpass Turbo. Overpass API. https://overpass-turbo.eu/. Accessed date 18 Oct 2019

Rocha LE, Liljeros F, Holme P (2011) Simulated epidemics in an empirical spatiotemporal network of 50,185 sexual contacts. PLoS Comput Biol 7(3):1001109

Ruan Z, Tang M, Liu Z (2012) Epidemic spreading with information-driven vaccination. Phys Rev E 86(3):036117

Schneeberger A, Mercer CH, Gregson SA, Ferguson NM, Nyamukapa CA, Anderson RM, Johnson AM, Garnett GP (2004) Scale-free networks and sexually transmitted diseases: a description of observed patterns of sexual contacts in britain and zimbabwe. Sex Transm Dis 31(6):380–387

Siksna A (1997) The effects of block size and form in North America and Australia. Urban Morphol 1:19–33

Sloot PM, Ivanov SV, Boukhanovsky AV, van de Vijver DA, Boucher CA (2008) Stochastic simulation of hiv population dynamics through complex network modelling. Int J Comput Math 85(8):1175–1187

Stangl P (2015) Block size-based measures of street connectivity: A critical assessment and new approach. Urban Des Intern 20. https://doi.org/10.1057/udi.2013.36

Ter Meulen J, Lukashevich I, Sidibe K, Inapogui A, Marx M, Dorlemann A, Yansane M, Koulemou K, Chang-Claude J, Schmitz H (1996) Hunting of peridomestic rodents and consumption of their meat as possible risk factors for rodent-to-human transmission of lassa virus in the republic of guinea. Am J Trop Med Hyg 55(6):661–666

Titball RW, Leary SE (1998) Plague. Br Med Bull 54(3):625–633

Traweger D, Travnitzky R, Moser C, Walzer C, Bernatzky G (2006) Habitat preferences and distribution of the brown rat (rattus norvegicus berk.) in the city of salzburg (austria): implications for an urban rat management. J Pest Sci 79(3):113–125

Tsai TF (1987) Hemorrhagic fever with renal syndrome: mode of transmission to humans. Lab Anim Sci 37(4):428–430

Weng L, Flammini A, Vespignani A, Menczer F (2012) Competition among memes in a world with limited attention. Sci Rep 2:335

Würbel H (2001) Ideal homes? housing effects on rodent brain and behaviour. Trends Neurosci 24(4):207–211

Zhang Y, Zhang H, Sun W, Pan C (2014) Connectivity analysis for vehicular ad hoc network based on the Exponential Random Geometric Graphs. In: 2014 IEEE Intelligent Vehicles Symposium Proceedings, Dearborn. pp 993–998. https://doi.org/10.1109/IVS.2014.6856464

Zhao L, Cui H, Qiu X, Wang X, Wang J (2013) Sir rumor spreading model in the new media age. Physica A Stat Mech Appl 392(4):995–1003

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.