2019-10-28

Degree Centrality and the Probability of an Infectious Disease Outbreak in Towns within a Region

Elizabeth Hunter
Technological University Dublin, elizabeth.hunter@tudublin.ie

Brian Mac Namee
University College Dublin, brian.macnamee@ucd.ie

John D. Kelleher
Technological University Dublin, john.d.kelleher@tudublin.ie

Follow this and additional works at: https://arrow.tudublin.ie/scschcomcon

DEGREE CENTRALITY AND THE PROBABILITY OF AN INFECTIOUS DISEASE OUTBREAK IN TOWNS WITHIN A REGION

Elizabeth Hunter and John Kelleher
School of Computer Science
Technological University Dublin
Kevin Street
Dublin 8
Ireland
email: elizabeth.hunter@mydit.ie

Brian Mac Namee
School of Computer Science
University College Dublin
Belfield
Dublin 4
Ireland

KEYWORDS

Agent-based model, Health Sciences, Discrete Simulation, Compartmental, Centrality

ABSTRACT

Agent-based models can be used to help study the spread of infectious diseases within a population. As no individual town is in isolation, commuting patterns into and out of a town or city are a vital part of understanding the course of an outbreak within a town. Thus the centrality of a town in a network of towns, such as a county or an entire country, should be an important influence on an outbreak. We propose looking at the probability that an outbreak enters a given town in a region and comparing that probability to the centrality of the town. Our results show that as expected there is a relationship between centrality and outbreaks. Specifically, we found that the degree of centrality of a town affected the likelihood of an outbreak within the network spreading to the town. We also found that for towns where an outbreak begins the degree of centrality of the town affects how the outbreak spreads in the network.

INTRODUCTION

The spread of infectious diseases is often thought of as a potential threat to global security. While some believe the biggest threat is of a yet unknown disease there are many well known diseases that are spreading throughout the world today. Diseases previously thought near eradication, such as measles and mumps, have recently re-emerged as a threat due to a number of factors such as lack of vaccinations and global travel. The 2018-2019 influenza season was one of the longest seasons in recent years and the 2019-2020 season is predicted to be worse than normal; at the same time recurring outbreaks of diseases, such as Ebola in Africa, have the risk of spreading abroad; and the lack of sanitary conditions in certain parts of Los Angeles is leading to fears of the spread of diseases such as typhus. One of the best ways to stop an outbreak and prevent a public health emergency is through preparedness. In many cases it may be impossible to test how an intervention will work in a real outbreak. This is where modelling becomes important. Modelling creates a simplified system to study a more complicated real world system and allows us to test intervention strategies and preventative measures.

Within epidemiology research a number of modelling methods are commonly used, including equation based models, such as compartmental SIR models, and computer simulation models, such as agent-based models. In this paper we adopt agent-based models because they capture interactions between factors and emerging patterns that can affect the spread of an infectious disease. There are many factors that can influence the course of an outbreak. Knowing how factors such as population size or vaccination rates lead to different outbreaks between towns can help to lessen an outbreak when it starts or help to focus preventative strategies in the most vulnerable areas. Hunter et al. (2018) showed that other factors can influence the course of an outbreak and two towns that appear similar on paper might have very different outbreaks. While Hunter et al. (2018) study towns in isolation we propose looking at towns within a larger network or region, such as a county. Although understanding how towns differ when their population is closed is important in understanding which towns might be more susceptible to an outbreak if it enters the town, it is also necessary to study how commuting patterns can have an effect on the spread of the outbreak. Knowing how different towns within a network might be more or less susceptible to an outbreak could be essential in helping to stop the spread of a disease when it takes off. Resources might be better funnelled towards towns that have a higher probability of an outbreak in the county then those with a lower chance of an outbreak. We propose that it not just factors within an individual town that might make it more susceptible to an outbreak, such as population density, but the centrality of the town within its network, specifically the transportation network and the number of commuters travelling between each town.

To study this we take an agent-based model for the spread of diseases that is made up of a network of towns. For each town we look at the degree centrality of the town and how that affects the spread of the disease to that town. We first will introduce the model that is used for the study and the different components of the model, then we discuss the experiment: how centrality is calculated and how we account for the other differences in towns. Finally we discuss the results and the conclusions of the study.

MODEL

The model used in the study is a version of the Hunter et al. (2018) model scaled up to model a region made up of many towns instead of a single town. While a town model tells us interesting ideas about the susceptibility of a given town, it is limited in its capacity to capture the town as part of a larger system. In a real world scenario, everyone in the town would not work within the town, instead some would commute out to other towns and others would commute in. This could change the course of the outbreak. To scale up the model we made a number of changes and assumptions that we will outline in this section. We break the description of the model up into the four main components of an agent-based models in human infectious disease epidemiology outlined in Hunter et al. (2017).

Society

Agents are added into the town based on the population data from the Irish Central Statistics Office (CSO) (CSO 2014). The CSO data is at the small area level. Small areas are geographic census areas that contain between 50 to 200 dwellings. They are the smallest area over which the Irish census data is aggregated. For each small area we create a population that reflects the population statistics of that small area including age, sex, household size and economic status. Irish vaccination data is used to determine the percentage of each age group that have received vaccinations for the infectious disease being modelled. For example, if 90% of 1 year olds in Ireland had been given the MMR vaccination in 2011 and we are running a model for 2012, we give each agent in the model with an age of 2 a 90% chance of having been vaccinated. If an agent is vaccinated they are given a 97% chance of being immune to the disease. This takes into account vaccination failure and is based on the vaccine effectiveness rate for measles (Nelson and Williams 2007). Half of the agents with age less than 1 are given immunity to a disease to mimic passive immunity infants receive from their mothers (Nicoara et al. 1999). For any agents that have an age corresponding to a vaccination year not in our data we give a 99% chance of being immune. Prior to vaccination campaigns the majority of the population would have

Figure 1: An image of the environmental setup of the model for Leitrim County Ireland
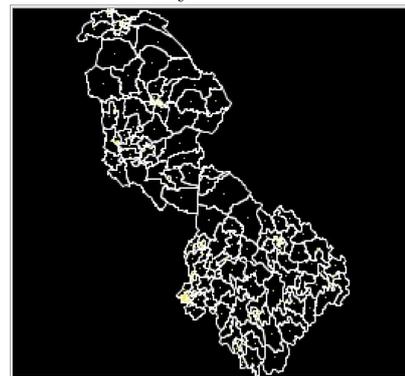
either had or been exposed to childhood diseases, such as measles, leaving them immune in later life. Agents are also given a set of social networks, a family network which connects them to all members of their household, and a school or work network that connects agents to other agents in their school or workplace.

Environment

The counties are created in Netlogo (Wilensky 1999). The Netlogo world is a two dimensional grid where the squares that make up the grid are referred to as patches. Patches in our model represent the small areas within a county. Figure 1 shows the environmental setup of the model for Leitrim County Ireland. All agents that are in a small area at a given time are physically in the same location. However, agents will keep track of their location within that small area. There are four possibilities for agents locations within a small area: home, work, school, or the community. Who an agent comes into contact with depends on that location. For example, if an agent is at home, they know they are at home and will only come into contact with other members of their household who are also at home. Agents in the community within a small area will come into contact with other agents in that small area in the community but will not come into contact with all agents in the community. If two agents from the same household are in the community there is a larger probability of them coming into contact then two agents who are in the same workplace or school who in turn have a larger probability of coming into contact than two agents who have no other connection. All agents within a small area patch have access to information about the patch they are in, including the number of primary and secondary schools in the small area. As well as the real world distances between the center of that small area and all other small areas in the model.

Transportation

Transportation in the town model is a simple model where agents move in a straight line between their current location and their desired destination. When not moving to home, school, or work, agents will pick a random location in the town to move to. It is a naive transportation model but for a small town it is a appropriate simplification. However, we do not feel that this simplification scales to a larger model. Moving randomly throughout a county is not realistic as moving randomly within one town as the distances between locations are much larger. Thus we use a gravity model to determine agent movements. Gravity models are a type of transportation model that is similar in formula to Newton's gravitation model. A traditional gravity model gives the interactions between two location pairs and determines those interactions based on the characteristics of a location and the distance between locations (Rodrigue et al. 2006). In our model, agents will move between home and school or work at certain predetermined times and will return home at predetermined times. On weekends, summers for students and after school or work hours agents will move through the community. An agents movements when they are deemed to be within the community are determined by our gravity model. The probability of an agent moving to another small area is proportional to the population density of the small area, an area that has a lot of other agents is more attractive, and inversely proportional to the distance to the small area from the agents current location, areas that are farther away are less attractive. We feel that this transportation model provides a more accurate simplification of movement within a larger area than that in the original town model.

Commuting patterns within the model are determined using CSO Place of Work, School or College - Census of Anonymity Records (POWSCAR) data (CSO 2017). This dataset provides information on the commuting patterns of people in Ireland and gives the number of people that commute from one electoral division to another. Electoral divisions are the census geographic area one step above the small areas.

Disease

The disease part of the model is the same as that in Hunter et al. (2018) and is based off of an SEIR compartment model that tracks agents' movements between susceptible, exposed, infected and recovered categories. Disease transmissions occurs as follows: if a susceptible agent comes into contact with an infectious agent, the susceptible agent has a percentage chance of becoming exposed. That percentage is used to determine if an agent will move from the susceptible state to the exposed state. This decision is made by drawing a random number between 0 and 1 from a uniform distribution

Figure 2: An image representing the disease model. Agents move from susceptible to infected to exposed than recovered
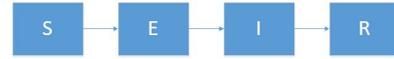
and comparing the random number with the percentage chance of the agent becoming exposed. If the random number drawn is less than the percentage chance of becoming exposed, the agent will move to the exposed state. If the number is greater than the percentage, the agent will remain susceptible. The agent will be in the exposed state for a predetermined period of time and once that period of time is over will move to the infected state. The agent will be in the infected state for a predetermined period of time and then will recover. The agent will stay in the recovered state for the rest of the model runtime and can no longer be infected. A representation of the disease model can be found in Figure 2.

We do not consider agent death in the model, however, as it is a short time frame, with only one outbreak, we do not feel that agents deaths or births would have a significant impact on the model. In addition, once an agent recovers they are removed from the susceptible, exposed, and infected cycle of the model and thus do not have an impact on the disease dynamics any more.

In this work we study a measles-like disease and the disease dynamics reflect that. On average an individual will stay exposed to measles before becoming infectious for 10 days (Nelson and Williams 2007). The time an agent remains exposed is determined for each agent from a normal distribution with a mean of 10 and a standard deviation of 0.5. Once infectious an individual will remain infectious for an average of 8 days (Nelson and Williams 2007). The time an agent remains infectious in the model reflects this and is determined for each agent from a normal distribution with a mean of 8 and a standard deviation of 0.5. We determine the infection rate, the percentage chance that a susceptible agent will be infected after contact with an infectious agent using the basic reproductive number R_0 for measles (12-18) (Nelson and Williams 2007). The basic reproductive number is defined as the number of individuals infected by one infectious individual in a completely susceptible population. It is the standard measure of transmission of a disease. The parameter can be broken down into three components, number of contacts per unit time (c), the transmission probability per contact (p), and the duration of the infectiousness (d) (Thomas and Weber 2001). As we can determine the number of contacts per unit time from our model and we know the duration of the infectiousness, we can determine the transmission probability per contact.

Schedule

The schedule for our model remains the same as the Hunter et al. (2018) model. The model is run on discrete time steps and runs from initialization to when no more agents are exposed or infected. Each time step represents two hours in a day, with 12 steps representing one day. During weekdays agents will travel to work or school at designated times and remain there until the time assigned for them to leave. Agents can then move throughout the county until they must return home at a certain time. On weekends, and summers for students, all agents can move throughout the county

EXPERIMENT

The goal of the study is to look how the centrality of a town within a region has an influence on an outbreak within the town. To do this we run our model 300 times for the county of Leitrim, Ireland. For each run the initially infected agent is selected at random from the susceptible agents in the county. We then look at how the outbreak spreads through the county and in particular how many outbreaks occur in each town. We look at a set of 16 towns within Leitrim that are made up of more than one small area and are located entirely in Leitrim. The towns are: Ballinamore, Carrigallen, Cloon, Dromahair, Drumod, Drumkeeran, Drumshanbo, Drumsna, Fenagh, Keshcarrigan, Kinlough, Leitrim, Lurganboy, Manorhamilton, Mohill and Tullaghan. An outbreak is defined using the World Health Organization's definition of a measles outbreak, which is two or more linked cases of measles. Thus for each town in the model we determine that there is an outbreak in a given run in that town if two or more agents from the town are infected. We also run a second experiment where we select the town that the outbreak starts in and look at where it spreads from there. Four different towns are chosen to start the infection in, two with high centrality, one with moderate centrality and one with very low centrality.

Centrality

To determine the centrality of the towns we use a weighted degree centrality. There are a number of different ways to determine the centrality of a point in a network. Some of the most common being betweenness, the number of paths that pass through each point, closeness, the proximity of a point to other points in the network, and degree, the number of links between each point in the network. As we want to look at centrality created by commuting between towns degree centrality was chosen, with commutes between different towns considered a link. Degree centrality can be directed based on if the links are coming into the town or out of the town, for the purpose of this study we do not consider the direction of the link. However, a link between two towns in a

county where one agent is commuting between the two towns should not have as much of contribution to the centrality of a town as a link between two towns where twenty agents are commuting between the two towns. To account for this we use a weighted degree centrality from Opsahl et al. (2010) that is calculated using a product of the number of links and the average weight of the links adjusted by a tuning parameter. Equation 1 shows the formula for weighted degree centrality with C_i being the centrality of town i, k_i the number of links into the town, s_i the number of agents commuting into or out of the town and α is the tuning parameter. The tuning parameter is used to determine the strength of the weight and the importance of individual link strength: when the tuning parameter is less than one the centrality measure favours more links into the town. If the total number of commuters is fixed a town with more links will have a higher centrality than a town with fewer links. When the tuning parameter is greater than one the centrality measure favours fewer links into the town. If the number of commuters is fixed a town with fewer links will have a higher centrality compared to a town with more links (Opsahl et al. 2010). For the purpose of this study we use an α less than one and set it at 0.5.
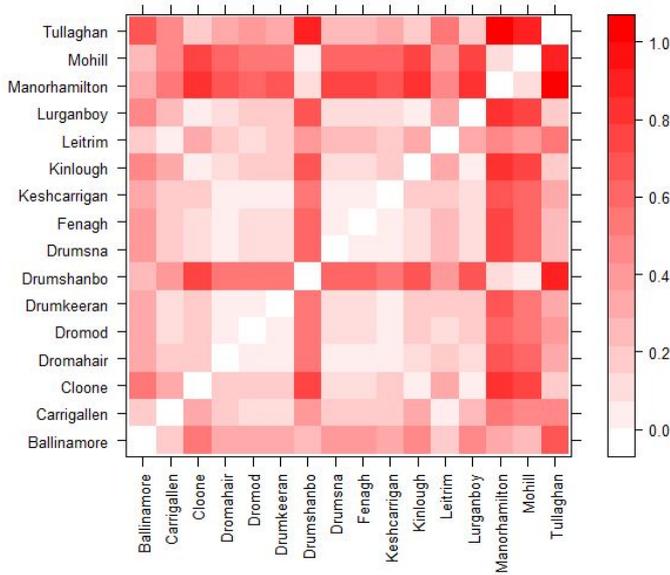
$$C_i = k_i * (\frac{s_i}{k_i})^\alpha \qquad (1)$$

The degree centrality calculated for each town is shown in Table 1.

Table 1: Centrality by town

Town	Centrality
Ballinamore	160.9
Carrigallen	118.9
Cloone	51.4
Dromahair	86.6
Drumod	96.5
Drumkeeran	90.7
Drumshanbo	208.7
Drumsna	75.1
Fenagh	74.5
Keshcarrigan	87.8
Kinlough	55.8
Leitrim	127.3
Lurganboy	58.0
Manorhamilton	232.8
Mohill	212.2
Tullaghan	17.4

We are interested in examining how the centrality of a town affects the outbreaks within a region. To explore this question we wish to control for other factors (town characteristics) other than centrality which may affect an outbreak. Consequently, a key step in our experiment is to identify pairs of towns that differ in terms of centrality but which are similar with respect to other characteristics. Once we have identified these town pairs we can examine how centrality affects a outbreak by

Figure 3: Distance Matrix showing the normalized difference in centrality between two towns

similar towns. Figure 4 presents a distance matrix which visualizes which towns are similar based on these categories. The lighter the square the more similar the towns are and the darker the more dissimilar.

Figure 4: Distance Matrix showing the normalized euclidean distance between towns


comparing how an outbreak spreads through the region when it starts in different towns. The first step in identifying these town pairs is to identify towns that differ in terms of centrality. To visualise which towns have the biggest difference in centrality we created a distance matrix shown in Figure 3 where the distance is the normalized difference between the centrality for each pair of towns. The more red the square, the greater the difference in centrality and the more white the square the less the difference in centrality between the towns.

Similar Towns

While comparing the results between two towns with different centrality might give an idea of how centrality influences an outbreak within a town, there are many other factors involved. Two towns with markedly different centrality might also be different in size or population density and this could be what is affecting the different results. To control for this we use euclidean distance to find towns that are similar in other characteristics that we believe might influence an outbreak. Each town is represented by a vector of quantitative characteristics: population size, town area (km^2), population density, number of small areas that make up the town, the number of secondary schools, the number of primary schools, the percent of susceptible agents in the town and the percent of agents who are students in the town. All categories except for the number of secondary schools and number of primary schools are standardized. The euclidean distance is then calculated between each of the 16 towns so that we can compare results between
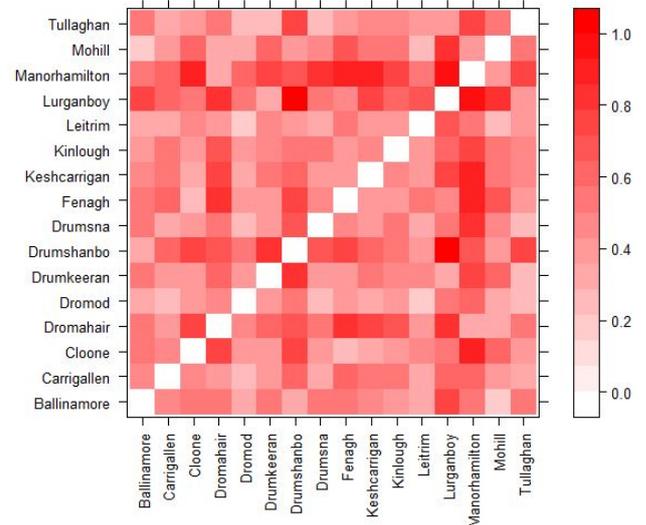
To be able to easily focus on the towns that have similar characteristics but different centralities we created a distance matrix that is the difference between the euclidean distance and the difference in centrality, Figure 5 presents this distance matrix. The more blue a square is in the matrix the smaller the euclidean distance and the larger the difference in centrality while the more red a square is the larger the euclidean distance but the more similar the centrality.

RESULTS

After running the model 300 times with a random start location for the outbreak, for each of the sixteen towns we calculate the percent of runs that results in at least one resident becoming infected and the percent of runs that lead to an outbreak (two or more agents infected who are residents of the town). The results along with the centrality for each town can be found in Table 2.

To look at the basic relationship between the degree centrality of each town and the outbreaks that occur in the towns we find the Pearson correlation. The correlation between the centrality of the town and at least one agent from the town getting sick is 0.66 and the correlation between the centrality of the town and an outbreak occurring in the town is 0.65. Although these are not strong correlations, they can be considered to show a moderate relationship (Ratner 2009). This is not unexpected as there are other factors besides centrality that

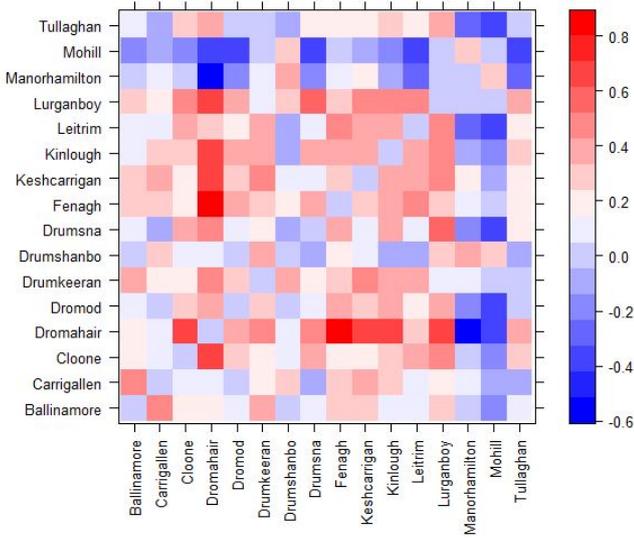Figure 5: Distance Matrix showing the difference between euclidean distance and centrality between towns

To further look into how the centrality of a town affects the outbreaks within a county we looked at repeated starts of the outbreak in the same town and looked at where the outbreak spreads and compare this to outbreaks that started in a similar town with different centrality. Examining Figure 3 Manorhamilton and Dromahair are similar towns with a large difference in centrality (the square at the intersection of these towns is deep blue). Similarly, Mohill and Tullaghan are similar towns with a large difference in centrality. The centrality of the four towns can be seen in Table 3. Table 4 lists the percent of runs where an infection that started in one of these 4 towns (Tullaghan, Drumahaire, Mohill, and Manorhamilton) resulted in an outbreak occuring in one of the 15 other towns in the Leitrim model. For example, Table 4 shows that 19% of the time that an infection started in Tullaghan and outbreak occurred in Ballinamore.

Table 3: Correlations between outbreaks started in each town and the centrality of the towns

Town	Centrality
Dromahair	86.6
Manorhamilton	232.8
Mohill	212.2
Tullaghan	17.4

Table 2: Results by town

Town	Centrality	One Sick	Outbreak
Ballinamore	160.9	38.0	24.0
		(32.5, 43.5)	(19.2, 28.8)
Carrigallen	118.9	27.3	17.3
		(22.3, 32.4)	(13.0, 21.6)
Cloone	51.4	20.3	7.7
		(15.8, 24.9)	(4.7, 10.7)
Dromahair	86.6	32.3	18.3
		(27.0, 37.6)	(13.9, 22.7)
Drumod	96.5	31.3	18.3
		(26.1, 36.6)	(11.3,19.4)
Drumkeeran	90.7	19.3	7.3
		(15.9,23.8)	(4.4,10.3)
Drumshanbo	208.7	34.0	18.6
		(28.6, 39.4)	(14.3, 23.1)
Drumsna	75.1	27.7	14.3
		(22.6, 32.7)	(10.4, 18.3)
Fenagh	74.5	26.3	11.7
		(21.3, 31.3)	(8.0, 15.3)
Keshcarrigan	87.8	22.7	9.3
		(17.9, 27.4)	(6.0, 12.6)
Kinlough	55.8	35.3	21.0
		(29.9, 40.7)	(16.4, 25.6)
Leitrim	127.3	32.3	18.0
		(27.0, 37.6)	(13.7, 22.3)
Lurganboy	58.0	23.7	6.7
		(18.9, 28.5)	(3.8, 9.5)
Manorhamilton	232.8	39.7	22.7
		(34.1, 45.2)	(17.9, 27.4)
Mohill	212.2	39.0	24.3
		(33.5, 44.5)	(19.5, 29.2)
Tullaghan	17.4	30.0	15.3
		(24.8, 35.3)	(11.3, 19.4)

Table 4: Results by town when the outbreak starts in one of four towns

Town	Tullaghan	Dromahair	Mohill	Manorhamilton
Ballinamore	19.0	20.3	24	15.7
	(14.6,23.4)	(15.8, 24.9)	(19.2,28.8)	(13.7,22.3)
Carrigallen	11.0	11.3	16.3	7.7
	(7.5,14.5)	(7.7,14.9)	(12.2,20.5)	(4.7,10.7)
Cloone	8.7	5.3	15.3	5.3
	(5.5, 11.9)	(2.8,7.9)	(11.3,19.4)	(2.8,7.9)
Dromahair	31.3	-	19.3	31
	(26.1, 36.6)	-	(14.9,23.8)	(25.8,36.2)
Drumod	13.0	15.3	22	7.7
	(9.2, 16.8)	(11.3,19.4)	(17.3,26.7)	(4.7,10.7)
Drumkeeran	10.7	14.7	7.3	15
	(7.2,14.2)	(10.7,18.7)	(4.4,10.3)	(11.0, 19.0)
Drumshanbo	15.3	22.7	18	18.7
	(11.3,19.4)	(17.9,27.4)	(13.7,22.3)	(14.3,23.1)
Drumsna	9.7	13.7	20.3	8
	(6.3,13.0)	(9.8,17.6)	(15.8,24.9)	(4.9,11.1)
Fenagh	8.0	10.3	10.3	7.3
	(4.9,11.1)	(6.9,13.8)	(6.9,13.8)	(4.4,10.3)
Keshcarrigan	8.0	10.7	6.3	6
	(4.9,11.1)	(7.2,14.2)	(3.6,9.1)	(3.3,8.7)
Kinlough	31.3	24.3	17.7	30.7
	(25.8,36.2)	(19.5,29.2)	(13.4,22.0)	(25.4,35.9)
Leitrim	14.3	19.3	18.7	16
	(10.4,18.3)	(14.8,23.8)	(14.3,23.1)	(11.9,20.1)
Lurganboy	24.3	14.3	7.3	27
	(19.2,28.8)	(10.4,18.3)	(4.4,10.3)	(22.0,32.0)
Manorhamilton	40.7	34.3	18	-
	35.1,46.2)	(29.0,39.7)	(13.7,22.3)	-
Mohill	25.0	19.3	-	16.7
	(19.8,29.5)	(14.9,23.8)	-	(12.4,20.9)
Tullaghan	-	18	16	26
	-	(13.7,22.3)	(11.9,20.1)	(21.0,30.9)

To get an idea of how the results vary and the outbreak spreads through the network when it starts in specific towns, we compare the outbreaks that occur in the other

fifteen other towns where the outbreak did not start with the centralities of those fifteen other towns. Similar to what was done for the initial analysis, we calculate the correlations between the outbreaks for each town in Table 4 and the centralities of these towns from Table 1. The correlations can be found in Table 5 separated by the town the outbreak starts in. The idea being that if the correlations are different we can learn something about how an outbreak spreads through a network based off of the centrality of the the town where it began.

Table 5: Correlations between outbreaks started in each town and the centrality of the towns

Start of Outbreak	Correlation
Dromahair	0.61
Manorhamilton	0.13
Mohill	0.34
Tullaghan	0.38

From Table 5 we can see that when the outbreak starts in a town with lower centrality, Dromahair and Tullaghan, there is a moderate correlation between the percent chance of an outbreak in another town and the centrality of that other town. However, the correlations are lower when the outbreak starts in a town with high centrality, Mohill and Manorhamilton. One way to interpret this is that when an outbreak starts in a town with high degree centrality, the centrality of other towns in the network do not have an effect on if the outbreak spreads to that town but when an outbreak starts in a town with low degree centrality the centrality of the other towns have a much higher influence on if the outbreak will spread there. The difference is more pronounced when comparing Dromahair and Manorhamilton versus Tullaghan and Mohill. This could be due to Tullaghan's very low degree centrality. As these are combinations of towns we have identified in the early analysis as having similar characteristics but different centrality we can conclude that most of the difference is due to the centrality and not outside factors.

CONCLUSION

The more that we are able to learn about how outbreaks move throughout a network of connected towns the more prepared we can be for a real world outbreak and not just a simulation. Understanding how well connected a town is in the network of a county or even the entire country might be able to inform the intervention strategies being used. If we know its more likely for an outbreak to spread to the towns with a higher degree centrality interventions could be focused in those areas. Finding out how susceptible a town is to an outbreak could save time and effort when trying to stop an outbreak that has already started or might save costs in focusing vaccination campaigns or other prevention strategies on those towns. Understanding how towns fit into a network of other towns and how that network influences the outbreak in an individual town is important in understanding how infectious diseases spread. Agent-based modelling allows us to understand these differences in towns without having to wait until a real outbreak occurs. In our analysis we were able to find a moderately strong positive linear relationship between the degree centrality of a town and the likelihood that an outbreak would spread to the town. In addition, the analysis shows that where an outbreak starts has an influence on what other towns will be affected. However, further work should be done looking into how other measures of centrality might influence the spread of the outbreak and might interact with degree centrality. For example, does the distance to the town where the outbreak starts affect the results and does closeness centrality interact with degree centrality to influence outbreaks spreading through a network. This could lead to an even greater understanding of how an outbreak will spread through a network and could prove invaluable.

References

CSO, 2014. *Census 2011 Boundary Files.* URL http://www.cso.ie/en/census/census2011boundaryfiles/. Date accessed 26.05.2016.

CSO, 2017. *Census 2016 Place of Work, School or College - Census of Anonymised Records (POWSCAR). Central Statistics Office.* URL https://www.cso.ie/en/census/census2016reports/powscar/.

Hunter E.; Mac Namee B.; and Kelleher J., 2018. *An open-data-driven agent-based model to simulate infectious disease outbreaks. PLOS ONE*, 13, no. 12, 1–35. 10.1371/journal.pone.0208775. URL https://doi.org/10.1371/journal.pone.0208775.

Hunter E.; Mac Namee B.; and Kelleher J.D., 2017. *A Taxonomy for Agent-Based Models in Human Infectious Disease Epidemiology. Journal of Artificial Societies and Social Simulation*, 20, no. 3, 2. ISSN 1460-7425. 10.18564/jasss.3414. URL http://jasss.soc.surrey.ac.uk/20/3/2.html.

Nelson K.E. and Williams C.M., 2007. *Infectious Disease Epidemiology Theory and Practice.* Jones and Bartlett Publishers.

Nicoara C.; Zach K.; Trachsel D.; Germann D.; and Matter L., 1999. *Decay of passively Acquired Maternal Antibodies against Measles, Mumps and Rubella Viruses. Clinical and Vaccine Immunology.* .

Opsahl T.; Agneessens F.; and Skvoretz J., 2010. *Node Centrality in Weighted Networks: Generalizing Degree and Shortest Paths. Social Networks*, 245-251, no. 3, 1–35. URL https://doi.org/10.1016/j.socnet.2010.03.006.

Ratner B., 2009. *The correlation coefficient: Its values range between +1/1, or do they? Journal of Targeting, Measurement and Analysis for Marketing.* .

Rodrigue J.P.; Comtois C.; and Slack B., 2006. *The Geography of Transport Systems.* Routledge, Taylor and Francis Group, London. ISBN 9780415354400.

Thomas J.C. and Weber D.J., 2001. *Epidemiologic Methods for the Study of Infectious Diseases*, Oxford University Press, chap. Concepts of Transmission and Dynamics. 61–62.

Wilensky U., 1999. *Netlogo. Center for Connected Learning and Computer-Based Modeling, Northwestern University Evanston, IL.* URL `https://ccl.northwestern.edu/netlogo/`.