TemporalPrestige

Nicholas Winsley

2025-03-19

Contents

Introduction **1**

Data **4**

Methods **6**
Network Measures of Centrality 6
Time-ordered networks 10
Network-Based Disease Transmission Simulation 18
Simulation 21

Results **24**
Model Selection of Centrality and Prestige Measures 27
Variable Selection 33
Constant Transmission Probability 37

References **41**

Introduction

Epidemiology is a subject of much contemporary relevance with the recent COVID-19 pandemic highlighting the importance of effective methods for combating the spread of disease, particularly in the early stages of a pandemic. One effective method is contact tracing, which aims to identify close personal contacts of an infected case. By uncovering close contact events, contact tracing can be used to identify people at high risk of infection and foresee future growth or contraction of the epidemic. This can help inform interventions (e.g. quarantining, disease tests, vaccination). Traditionally, contact tracing has focused on

confirmed cases, which are reported to authorities. When a new case is reported, an official asks the patient to recall all recent contacts, and these recent contacts can then be followed up with or notified. Although this method has been shown to be effective (Fetzer and Graeber (2021)), it is labor-intensive and does not give a complete picture. Recently, digital contact tracing (DCT) has emerged as a cost-effective (albeit in some ways unreliable) alternative to conventional contact tracing methods. In some of its most recent forms, DCT uses portable bluetooth devices which detect close contacts between carriers of the device (e.g., Chambers et al. (2021)). In addition, DCT can provide indicators of proximity between the interacting individuals and duration of the contact event. With the emergence of DCT technology, early prophylactic identification and quarantining of high-risk individuals could be practicable in the future. This study investigates statistical methods for the identification of high-risk individuals, using DCT data.

Governments typically wish to intervene for individuals who are most instrumental in the spread of a disease or most susceptible to adverse effects of contracting the disease. Although both groups of individuals could be considered as "high risk", our focus will be on individuals who have a greater influence on the spread of a disease and use "high risk" to refer to this group. Here, risk is a combination of the likelihood of a particular individual becoming contagious due to prior contacts and the likelihood of the individual spreading the illness via future contacts. In practice, future contacts are not known. Consequently, this study will primarily focus on the likelihood of becoming infected in identifying high-risk individuals. (We note that the methods presented here are easily adapted for individuals who are more susceptible.) **Ryan note: This sounds more like a description of "high-risk" being those who are most susceptible than those who are more likely to spread the disease. Also, while I get your point on "future contacts", in our simulations we have information on how many individuals would be infected if a particular individuals was our seed, and we are implicitly assuming that the temporal network essentially maintains a similar structure in the future.**

The contacts that individuals have over time can be represented using a temporal network, or graph, where each contact event is recorded as the triple (i, j, t), where i and j denote the

interacting individuals and t denotes the time of the interaction. Note that we assume that (i, j, t) is semantically no different than (j, i, t) (i.e., contacts are symmetric). Corresponding to each contact event (i, j, t) is some associated measure of risk for that event. For the DCT data we consider, this measure of risk is simply a function of the proximity and duration of the interaction, but it could include other factors (e.g., disease status of or disease prevention measures taken by the individuals) if these were recorded in the contact tracing.

Statistical methods are often used (in combination with contact tracing) to identify high-risk individuals in a network. This study utilises two common methods for the analysis of contact tracing data: social network analysis and simulation. Social network analysis (SNA) is an interdisciplinary approach for the study of entities and their relationships with each other. SNA involves constructing a network to model a real-world situation of interest and calculating metrics or fitting models to understand key aspects of the network structure. When considering network metrics, these can be broadly categorized into two types: population-level (or group-level) measures and individual-level measures. This study is primarily concerned with individual-level measures, focusing specifically on centrality, a measure which is meant to capture the influence of a particular individual within a network. Prestige (sometimes called status or rank) is a similar individual-level measure specific to directed networks, which only considers incoming edges/contacts. (In the context of a directed network, centrality is based on outgoing edges/contacts.) Although SNA has been used in many fields since its creation in the 1930s, it's value in epidemiology only became apparent in 1985, when Klovdahl (1985) applied SNA to AIDS data.

Simulation is an effective method for ascertaining properties of a temporal network. Some classical models have been deterministic (e.g., the differential equation model of Kermack and McKendrick (1927) is a notable example). However, deterministic models typically rely on simplifying assumptions and thus do not capture the full granularity of the network. Compartmental models are a popular simulation approach in which the population is divided into groups, and individuals transition between groups over time. The susceptible-infected-recovered (SIR) model is a quintessential compartmental model in which all individuals are initially susceptible, and individuals may become infected due to contact with a contagious

individual. **Ryan note: Should provide a citation for SIR model.** Once infected, individuals transition to the recovered state at a predictable recovery rate, where they stay for the remainder of the simulation. Recovery rates are typically sampled from a probability distribution, which may be estimable by exogenous information (e.g. medical knowledge, recovery rates for similar diseases). Key metrics are averaged over many simulations to approximate a true underlying distribution. A plethora of summary metrics have been applied to simulated epidemics. Macdonald (1952) introduced the reproductive number, which is defined as the number of cases resulting from a single infection. Holme (2018) used the time for the disease to go extinct (i.e., no new cases can occur).

This study aims to address several key research questions:

1. Can we make simplifying assumptions to reduce computation time of simulations?
2. Which centrality measures are most effective when applied to epidemiology?
3. How can we extend these measures to cases where contact risk varies?

Simulation is important for answering questions 2. and 3. as it provides a sort of "ground-truth" against which centrality measures can be compared. **Ryan note: Not sure you really need this last sentence. Simulation is being used for all three questions.**

Data

In 2020, the New Zealand Ministry of Health (MoH) commissioned a pilot study of the CovidCard, a portable device which used bluetooth technology to record contacts between carriers of the device. Adults 19 years of age or older who live in Ngontotahā West and East were recruited to participate in a seven-day study. Additionally, people who live outside these boundaries but work within the Ngontotahā Village were also permitted to take part in the trial. Ngontotahā was chosen because it met several key criteria, namely compactness, geographical isolation, small population size and high sociodemographic diversity. In total, 1,191 people participated in the study. At the end of the trial period, a subset of 158 participants from the main trial were contacted by MoH case investigators to establish

contacts that they had over the trial period using a modified version of the MoH case investigation protocol. Work carried out by Admiraal et al. (2022) compared the CovidCard to conventional case investigation methods and found a greater rate of reciprocal interactions identified by the CovidCard. In short, the study concluded that the CovidCard is a highly effective supplementary contact tracing approach, and we use CovidCard data to compare existing network centrality metrics as well as develop novel alternatives.

The CovidCard is a bluetooth device developed for detecting close-contact events between carriers. Each card advertises it's presence and detects signals from other cards. Algorithms evaluate the radio signal strength indicator (RSSI) of close-contacts in real-time, and the signal strength is aggregated over 15 minute time intervals. The raw RSSI values are transformed into distance estimates by the path loss model, proposed by Seidel and Rappaport (1992), for which

$$RSSI \propto -20 \log_{10}(distance)$$

Noise in distance estimates is subsequently reduced by signal processing methods, most commonly Kalman Filters.

Each interval was classified as either < 1 meter, < 2 meter and < 4 meter proximity, and the total number of intervals belonging to each class was summed over a two-hour period. The cards can hold up to 128 contact events in short-term cache memory at any given time, of which some are recorded in long-term flash memory. An interaction was recorded in flash memory if it was longer than 2 minutes in duration, and the RSSI exceeded -62dBm (roughly corresponding to a distance of less than 4 meters). For more details on the CovidCard, see Admiraal et al. (2022).

The last day of the trial period saw an anomalously high number of close-contacts, most likely because participants congregated at a single location for card collection. **Ryan note: Actually, because all cards were collected at the same location, so cards were probably just thrown in a box together.** For this reason, contact events which occurred on the last day of the trial were omitted. Participants who could not be cross-verified by

case investigation were removed. If two cards registered the same contact event, and gave conflicting proximity values, one of the proximity values was arbitrarily removed. Contact dates were converted to numeric times by calculating the time elapsed (in hours) between the contact date and the start of the trial. Some cards were collected before the last day of the trial, resulting in an anomalous number of contacts during card collection. For this reason, all contact events which occurred on the day in which they were uploaded were removed. Data on the proximity classes was processed to form non-overlapping categories. For instance, the number of 15-intervals with a distance less than 4 metres, $n_{<4}$, was transformed by subtracting $n_{<2}$ and $n_{<1}$ to get $n_{\geq 3,<4}$; the total number of 15-minute intervals between 3 and 4 metres. By doing this, we get a categorical variable on which further statistical models are based.

Methods

Network Measures of Centrality

We first describe key measures for networks, in particular focusing on measures of centrality and prestige, or rank. Centrality and rank are meant to represent the influence of a given node in a graph, although the exact meaning of a "central" or "prestigious" node varies depending on the context.

Degree centrality

Degree centrality is a simple metric for static networks where we only consider the number of neighbors of a given node (i.e., other nodes for which a given node shares an edge/contact). Let A denote the adjacency matrix for a static network (i.e., $A_{ij} = 1$ if i and j are neighbours and 0 otherwise). Then degree centrality is defined as:

$$C_D(i) = \sum_{j=1}^{N} A_{ij} \tag{1}$$

Degree centrality is simple and easily calculated, however it does not incorporate knowledge

of the entire network structure. Improving this point is the motivation for our next centrality measure. **Ryan note: What is degree centrality meant to capture/reflect in terms of "importance" of a node?**

Closeness centrality

For undirected graphs, Bavelas (1950) proposed closeness centrality

$$C_c(i) = \frac{N-1}{\sum_{j \neq i} d(i,j)} \tag{2}$$

where N is the network size (i.e., number of nodes) and $d(u, v)$ is the distance of the shortest path between j and i. As an example, for the network shown in Figure 1, the closeness centrality of node E is $\frac{4}{1+1+2+3} = \frac{4}{7}$. Closeness centrality can be interpreted as the efficiency with which a node can access all other nodes in a network.

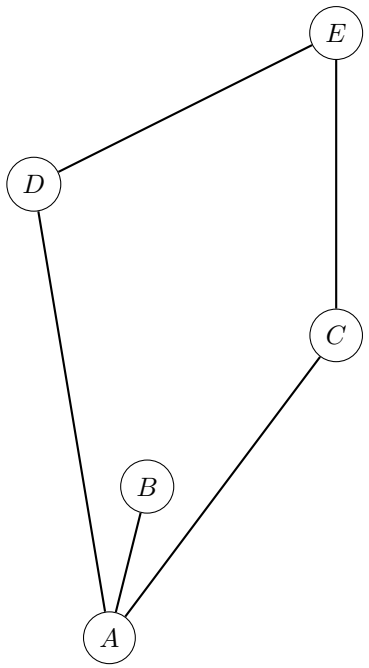


Figure 1: A simple undirected graph

Betweenness centrality

In some situations, the effect of removing a node on transmission through a network may be highly important (for instance, when quarantining individuals during a pandemic). This is the primary motivation behind betweenness centrality (Freeman 1977), which is defined as

$$C_B(i) = \frac{1}{(N-1)(N-2)} \sum_{j \neq k \neq i} \frac{\sigma_{j,k}(i)}{\sigma_{j,j}} \tag{3}$$

where $\sigma_{j,k}$ is the number of shortest paths (geodesics) between j and k, and $\sigma_{j,k}(i)$ is the number of such paths that pass through i. The denominator term $(N-1)(N-2)$ is a normalising constant, ensuring that betweenness centrality values range between 0 and 1.

Percolation centrality

In practice, additional information around the "percolation state" of nodes may be known. For instance, in epidemiology we may know that certain individuals are infected. To incorporate knowledge of percolation state into centrality metrics, Piraveenan, Prokopenko, and Hossain (2013) proposed percolation centrality:

$$C_P^t(i) = \frac{1}{(N-1)(N-2)} \sum_{j \neq k \neq i} \frac{\sigma_{j,k}(i)}{\sigma_{j,k}} \frac{x_j^t}{\left(\sum_{\ell=1}^{N} x_\ell^t\right) - x_i^t} \tag{4}$$

where x_ℓ^t is the percolation state of node ℓ at time t and $\sigma_{j,k}$ and $\sigma_{j,k}(i)$ are as defined for betweenness centrality. The percolation state ranges from 0 to 1 with 1 meaning that the individual is certainly infected and 0 meaning that the individual is certainly uninfected. A value between 0 and 1 could be used to represent a probability of infection or a proportion of a township which is infected.

Adjusted percolation centrality

We propose a variant of percolation centrality which we will call adjusted percolation centrality. Adjusted percolation centrality only considers paths which do not pass through any percolated nodes. By doing this, it ensures that ineffectual paths are not considered (i.e., node k can only

become percolated due to node j being percolated and not due to some already percolated intermediary node on the path). We define an unpercolated path as a path where no incident nodes are percolated except for the origin and terminal nodes, which may have any percolation state. Then adjusted percolation centrality is given by

$$C_{AP}^{t}(i) = \frac{1}{M_P} \sum_{s \neq v \neq r} \frac{\sigma_{j,k}^{P}(i)}{\sigma_{j,k}^{P}} \frac{x_j^t}{\left(\sum_{\ell=1}^{N} x_\ell^t\right) - x_i^t} \tag{5}$$

where M_P is the number of dyads (or pairs of nodes) (a, b) where there is an unpercolated path between a and b, and $\sigma_{j,k}^{P}$ is the number of shortest unpercolated paths between nodes j and k.

Katz centrality

Closeness and betweenness centrality focus on shortest paths (geodesics), but the path by which a disease is transmitted need not be the shortest path. Katz centrality (Katz 1953) is an alternative centrality measure which considers all paths between nodes and is given by

$$C_K(i) = \sum_{j=1}^{N} \left(\alpha A + \alpha^2 A^2 + \alpha^3 A^3 + \ldots\right)_{ji} \tag{6}$$

where A denotes the adjacency matrix and $0 \leq \alpha \leq 1$. A^n is simply the n-step adjacency matrix for the network, and α^n decreases exponentially with n, meaning that longer paths are downweighted. The infinite sum $(\alpha A + \alpha^2 A^2 + \ldots)$ converges when $\alpha \leq \frac{1}{\rho(A)}$, where $\rho(A)$ is the spectral radius of A (i.e., the largest absolute value of any of its eigenvalues). Under this condition, the sum converges to $(I - \alpha A)^{-1} - I$, where I denotes the $N \times N$ identity matrix. If we are only interested in paths of length n or less, we can adjust the formula accordingly to get $(I - \alpha^{n+1} A^{n+1})(I - \alpha A)^{-1}) - I$. This result is convenient for temporal networks, where a finite number of steps in each snapshot is often assumed.

Temporal Katz centrality

Temporal Katz centrality, proposed by Grindrod et al. (2011), is a relatively straightforward extension of @ref(eq:katz). Let A_t denote the adjacency matrix at time t. For a temporal network with snapshots at times $1, 2, 3, \ldots, T-1, T$, the temporal Katz centrality is given by the matrix product

$$C_{TK}(i) = (I - \alpha A_1)^{-1}(I - \alpha A_2)^{-1}...(I - \alpha A_{T-1})^{-1}(I - \alpha A_T)^{-1} \tag{7}$$

The centrality for a given node is likewise calculated by the row sum of this matrix product.

Time-ordered networks

Generalisation of conventional centrality measures to temporal networks requires a high-granularity representation of the network as a graph. Kempe, Kleinberg, and Kumar (2000) proposed a graph where the edge weights are contact times. However, this model fails to account for differential rates of transmission. Kim and Anderson (2012) proposed a more general solution using a time-ordered directed graph. Consider a network of N nodes for which M edges are observed over T time points. Without loss of generality, discretise the contact times to get a list, $t = (0, 1, 2, \ldots, T)$. We can construct a time-ordered graph where each node appears $T + 1$ times. Denote by i_t the node i at time t. In this graph, a directed edge from i_t to j_{t+1} only exists if $j = i$ or there is a contact between i and j at time t. We can construct this graph for any temporal network without loss of information. In practice, computational constraints may require aggregation of contact times and thus loss of information.

To illustrate this idea, consider a simple temporal network with five nodes (a, b, c, d, e) and contacts observed at two time points, as shown in Table @ref(tab:Table1). Figures 2 and 3 show snapshots of the network at these two time points.

Origin node	Terminal node	Time point
a	b	1

Origin node	Terminal node	Time point
b	d	2
a	c	2
d	e	1

Table (#tab:Table1): Temporal network of five nodes observed at two time points.

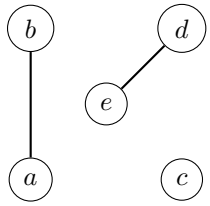


Figure 2: Snapshot of the network when $t = 1$

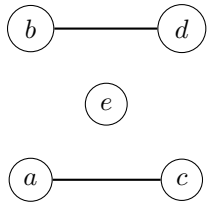


Figure 3: Snapshot of the network when $t = 2$

This can be represent using a time-ordered direct network, as shown in Figure 4.

The distance of the temporal shortest path length over time interval $[a, b]$, denoted by $d_{a,b}(i, j)$ is defined as the smallest $d = b - n$, where $a \leq n \leq b$ and there is a path from i_n to j_b. Thus, in Figure 1, the shortest path distance $d_{1,3}(a, b)$ is two, with the temporal shortest path being $a_1 \longrightarrow b_2 \longrightarrow b_3$. By representing a temporal network as a high-granularity digraph, we can generalise conventional measures of prestige and centrality to temporal networks.

Proximity prestige

This study is primarily concerned with the likelihood of infection, as estimating transmission requires future contact tracing data, which is usually not known. Likelihood of infection is roughly analogous to the idea of prestige, also known as rank. We will use the terms "prestige"

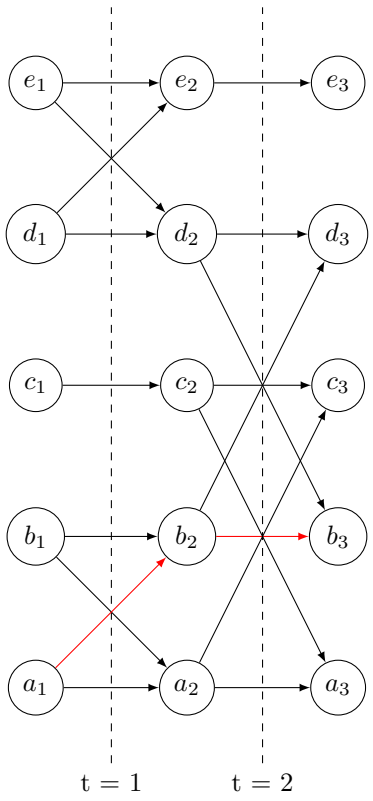


Figure 4: A simple temporal network represented as a digraph. The temporal shortest path from a to b is shown in red.

and "rank" interchangeably throughout this paper. In a directional network, a prestigious node is the object of many in-ties (i.e., has many incoming connections). In the context of a directed network, this is a distinct concept from centrality, which relates to out-ties (i.e., outgoing connections).

Many conventional measures of centrality are ill-defined in directional graphs, owing to the fact that directional graphs are not necessarily strongly-connected. Due to this limitation, we usually only consider nodes in the influence domain of a given node (i.e., the set of all nodes that can be reached from the node [when considering measures of centrality] or the set of all nodes that can reach the node [when considering measures of prestige]). Lin (1976) proposed proximity prestige, defined as

$$P_P(i) = \frac{I_i/(N-1)}{\sum d(j,i)/I_i} \tag{8}$$

where I_i is the size of the influence domain of node i, and N is the network size. Intuitively, this is the proportion of the network covered by the influence domain, divided by the average distance from other nodes to nodes i over the influence domain. When the influence domain is empty, the proximity prestige is defined to be 0.

Temporal proximity prestige

For temporal networks, we propose a modified version of proximity prestige given by

$$TPP_{[0,b]}(i) = \sum_{0 \le t < b} \frac{I_{t,b,i}/(N-1)}{\sum_{j \in I_{t,b,i}} d_{t,b}(j,i)/I_{t,b,i}} \tag{9}$$

where $I_{t,b,i}$ is the influence domain of i over the time interval [0, b]. We will call this the temporal proximity prestige up to time b. The temporal proximity prestige can be normalized by dividing by b.

Temporal closeness

Kim and Anderson (2012) proposed temporal closeness centrality, a similar metric which considers all time intervals $[t, b], t \in [0, b-1]$.

$$TC_{[a,b]}(i) = \sum_{a \leq t < b} \sum_{j} \frac{1}{d_{t,b}(i,j)} \tag{10}$$

When i is unreachable from j over $[t, b]$, $d_{t,b}(j,i) = \infty$. We cover cases where the denominator is infinite by assuming that $\frac{1}{\infty} = 0$. Note that, as we are considering a directed network, $d_{t,b}(j,i)$ is not equivalent to $d_{t,b}(i,j)$. To turn @ref(eq:tc) into a prestige measure, we simply reverse the direction of the paths to get

$$TP^{P}_{[a,b]}(i) = \sum_{a \leq t < b} \sum_{j} \frac{1}{d_{t,b}(j,i)} \tag{11}$$

We will call this the temporal prestige. The temporal prestige can be normalized by dividing by $(N-1)(b-a)$.

Multiplicative Temporal Closeness Rank

When each edge is associated with a probability of transmission w, as may be the case in epidemiological models, the probability of a path may be of greater interest than the temporal length. (A longer path with greater probability of transmissions at each step may be more effective at spreading a disease than a shorter path with lower probabilities.) In this case, we can generalise existing methods by considering a directed network where the edge weights are the natural logarithm of the probability. Figure 5 shows a graph of this kind. **Ryan note: Why do you have values of positive 2 and 3 for a couple of these log-transformed edge weights?**

The probability of transmission from an individual to themself is assumed to be 1, and hence the natural logarithm becomes 0. Consider a path P starting at j_a and ending at i_b. The probability of this path is equal to $\prod_{k=a}^{b} E_k$, where E is the list of transmission probabilities of path P. The probability of path P can be calculated by:

$$e^{\sum_{w \in E} log(w)}$$

It can be shown that for the representation in Figure 5, all highest-probability paths to i_b

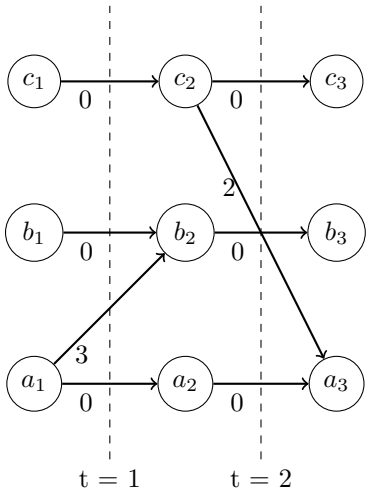


Figure 5: A simple temporal network represented as a digraph

can be calculated in $O\left(bN^2\right)$ time using a modified version of the reversed evolution network (REN) algorithm proposed by Hanke and Foraita (2017). Algorithm 1 shows the pseudocode for the REN algorithm.

Absorption Rank

Rocha and Masuda (2014) proposed TempoRank, an extension of the illustrious Google PageRank algorithm to temporal networks. TempoRank considers the stationary distribution of a random walk through the temporal network. However, TempoRank does not generalize well to epidemiological modelling, where infection may be a permanent state. Intuitively, some sort of aggregation over all previous contacts would be preferred. In this section, we propose an aggregate metric for temporal networks. Without loss of generality, we will consider a temporal network consisting of N nodes and contact times $t = 1, 2, 3, \ldots, T$. Let $p_{ijk}(t)$ denote the probability of transmission for the k^{th} contact between individual i and individual j at time t. Let $n_{ij}(t)$ denote the number of contacts between i and j at time t. Define the transition probability matrix $\mathbf{B}$ for each contact time t as

Result: Multiplicative Closeness Rank
Input : Data file with each row representing a contact and a probability of transmission
Output: Temporal Prestige for a given node

```
contacts <- List of contact times sorted in increasing order. Each contact time is a data
  structure with a map of nodes to out-neighbours.
while Data file has next line do
    Read line
    Add line to contacts using bisection search
end
tcp <- Temporal closeness prestige
reachable <- Set of reachable nodes
Add target node to reachable
sums <- Map of node id to log of the highest probability path (obtained by summing
  edge weights)
back <- Pointer to final contact time in contacts list
while back is not null do
    temp <- Map of updated path lengths for this iteration
    foreach node in reachable do
        out-neighbours <- back.out-neighbours[node]
        foreach neighbour in out-neighbours do
            Add neighbour to reachable set
            weight <- Edge weight of connection between node and neighbour
            temp[neighbour] = Max(temp[neighbour], sums[node] + weight)
        end
    end
    foreach node in reachable do
        sums[node] = Max(temp[node], sums[node])
        mcr += expsums[node]
    end
    Decrement back pointer
end
Return tcp
```

Algorithm 1: Modified version of the REN algorithm.

$$\mathbf{B}_{ij}(t) = \begin{cases} 1 & i = j, s_i(t) = 0 \\ 0 & i \neq j, s_{ij}(t) = 0 \\ \prod_{m \in N, m \neq i} \prod_{k=1}^{n_{im}(t)} (1 - p_{imk}(t)) & i = j, s_i(t) > 0 \\ (1 - \mathbf{B}_{ii}(t))(s_{ij}(t)/s_i(t)) & i \neq j, s_{ij}(t) > 0 \end{cases} \tag{12}$$

where $s_{ij}(t)$ and $s_i(t)$ are defined as:

$$s_{ij}(t) = 1 - \prod_{k=1}^{n_{ij}(t)} (1 - p_{ijk}(t)) \tag{13}$$

$$s_i(t) = \sum_{j \in N, j \neq i} s_{ij}(t) \tag{14}$$

Ryan note: You really need to give some insight into what $s_i(t)$ and $s_{ij}(t)$ represent.

Denote by $\mathbf{B}_i(t)$ the transition matrix obtained by taking $\mathbf{B}(t)$, and setting all entries in the i^{th} row to zero, except the diagonal entry (which is necessarily one). The walk $\mathbf{B}_i = (\mathbf{B}_i(1), \mathbf{B}_i(2), \cdots, \mathbf{B}_i(T))$ is an absorbing random walk, and the product $C_i^A(t) = \mathbf{B}_i(1)\mathbf{B}_i(2)\cdots\mathbf{B}_i(t)$ is the absorption rank for individual i at time t. The absorption rank of individual i is interpreted as the probability that a random walk through the temporal network passes through i.

If we assume a constant transmission probability p for all contact events, @ref(eq:absorb) reduces to:

$$\mathbf{B}_{ij}(t) = \begin{cases} 1 & i = j, s_i(t) = 0 \\ 0 & i \neq j, s_{ij}(t) = 0 \\ \prod_{m \in N, m \neq i} (1 - p)^{n_{ij}(t)} & i = j, s_i(t) > 0 \\ (1 - \mathbf{B}_{ii}(t))(s_{ij}(t)/s_i(t)) & i \neq j, s_{ij}(t) > 0 \end{cases} \tag{15}$$

where $s_{ij}(t)$ is defined as:

$$s_{ij}(t) = 1 - (1-p)^{n_{ij}(t)} \tag{16}$$

and $s_i(t)$ is similarly defined as:

$$s_i(t) = \sum_{j \in N, j \neq i} s_{ij}(t) \tag{17}$$

In practice, p may be estimable from domain-specific knowledge, but the value of p is generally not particularly important. **Ryan comment: Why not? Simply in terms of preserving ranks?** However, values close to 0 or 1 may cause the output values to cluster together, making differences difficult to detect. **Ryan note: You're going to need to clarify this last sentence. Differences in...?**

Network-Based Disease Transmission Simulation

Here, we describe a novel approach for network-based disease transmission modelling using static networks, which we treat as "ground truth" to validate adjusted percolation centrality. **Ryan note: I think it makes sense to have this with your simulations section and not have this described as being used simply "to validate adjusted percolation centrality". It's an approach for efficiently simulating a dynamic network. I think we might want to discuss what you mean by "static" network in this case, as I've most commonly seen "static" used to describe a network with unchanging/fixed nodes and edges. In this case, the set of nodes is the same, but the edges are technically changing with time.** A social network is modeled by a static graph, where each node represents an individual, and each edge represents an ongoing, time-independent relationship between two individuals. It is assumed that contacts between any two individuals i and j follow a Poisson process with a rate parameter of λ_{ij}, and each contact has a constant probability β of being infectious. Then the sojourn times (i.e., time between contacts) of contacts between i and j follow an exponential distribution with a rate of λ_{ij}, and the sojourn times of *infectious* contacts between any two nodes i and j follow independent exponential distributions with a fixed rate parameter of $\beta\lambda_{ij}$. Due to the memoryless property of the

exponential distribution, the time until the next infectious contact, conditional on the current time, follows the same distribution. By these assumptions, epidemics can be efficiently simulated in the most general case without prior knowledge of contacts. **Ryan note: What exactly do you mean by this last sentence? Surely it would be prudent to estimate λ_{ij} from temporal network data if available.**

Cconsider a simulation ending at time T, which starts with a set of infected nodes. We iterate through all neighbours of each infected node, and sample a time until the next infectious contact, adding it to a sorted list of infection times as we go. At each subsequent step, the smallest infection time is selected from the list. If this time is greater than T, the simulation stops. Otherwise, we sample an infection time for each neighbour of the infected node. If the sampled infection time for a given node is greater than its current infection time, it is ignored.

Figure @ref(fig:netsim) shows an example static graph representing a small social network, where the edge weights (shown next to edges) correspond to the rate parameter $\beta\lambda_{ij}$ for an exponential distribution. Consider a simulation on this graph which terminates at the time T = 100 and for which only node A is initially infected. Suppose we sample infectious sojourn times t = 60, 40, and 50 for neighbours C, B and D, respectively. In the next step, the infection time of C (t = 40) is removed from the list, and the time t = 70 is sampled for node E. Thus, the time to infection for E is t = 40 + 70 = 110. Then, the time to infection of node D (t = 50) is removed from the list, and each neighbour is considered in turn. Suppose we sample an infectious sojourn time of 70 for node E, hence the new infection time is t = 50 + 70 = 120 which is greater than 110, thus it is ignored. Likewise, suppose the time t = 40 is sampled for the infectious sojourn time from D to B. The new infection time for B is t = 50 + 40 = 90 which is greater than 60, thus it is ignored.

In general, contact rate parameters λ_{ij} may be estimable from contact tracing data, and the constant probability of infection may be estimable from incidence data. It should be clear that this method is not a replacement for simulations on contact tracing data, and should only be used as a crude alternative for extremely large datasets where it is plausible that the time between contact events follows an exponential distribution.

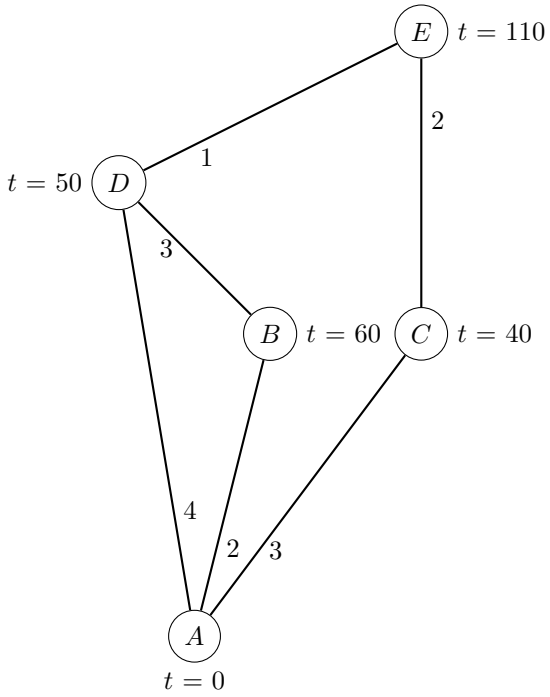


Figure 6: A simple social network represented as a static graph

Ryan note: Is this the only case where you've used this simulation approach? If so, then I think that the last sentence in the preceding paragraph is followed by something along the lines of, "Although we demonstrate how this simulation approach can be used, we turn our attention to approaches that are less restrictive in their assumptions. . . ."

The Erdos-Renyi model (Erdos, Rényi, et al. 1960) with $p = 0.4$, 0.45, and 0.5 (i.e., an edge is constructed between any pair of nodes with constant probability p) was used to generate 100 random networks of 40 nodes for each value of p. For each simulated network, we ensured that the network was connected so as to ensure that all measures of centrality could potentially be calculated. In each network, four randomly selected nodes were initially infected. **Ryan note: Further details required. What are you using for contact rates and probability of infection?** The importance of a given node was tested empirically by calculating the difference in average reproductive number between graphs which included and excluded the node in question. **Ryan note: You should be defining the reproductive number and giving a citation here or somewhere previously. (Personally, I think it makes sense to clearly define what measures you will consider for simulated networks**

and then explaining that the goal is to see which measures of centrality most strongly correlate with these measures.) A high difference in reproductive numbers indicates that the node is instrumental in the spread of the disease. Figure @ref(fig:perccent) shows the results for the three considered edge inclusion probabilities ($p = 0.4$, 0.45, and 0.5). Both variants of percolation centrality showed similar levels of correlations with the empirical measure of importance. **Ryan note: What is the takeaway message that we are meant to get from this? Why does this matter? Also, for the graph have your vertical axis be 'r' and your legend title be "Centrality measure".**

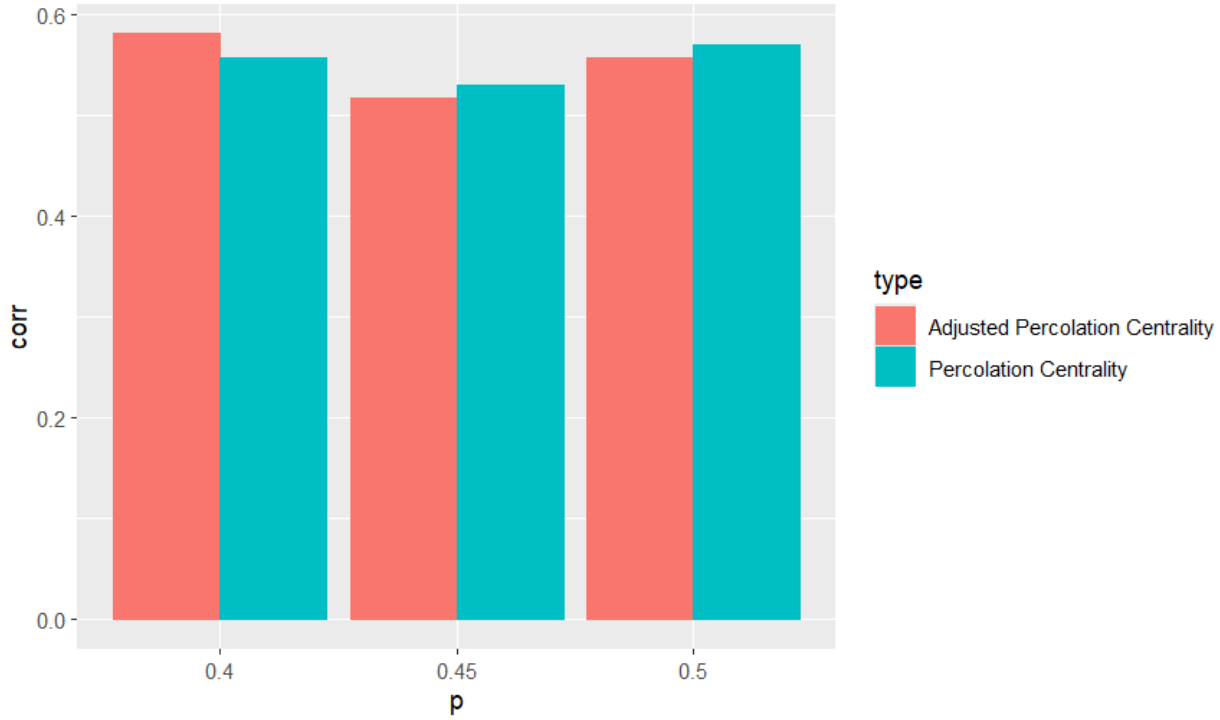


Figure 7: Correlation between the difference of reproductive numbers and the two variants of percolation centrality for different edge inclusion probabilities

Simulation

Stochastic simulations typically follow the standard Markovian framework in which we assume transmission depends only on the current state of the network. The contact times can be thought of as discrete snapshots of the network, and it is assumed that during each snapshot, only one "hop" can occur. In other words, if we have two contacts events (i, k, t) and (k, j, t)

at time t, node i cannot infect node j via node k at time t. The standard Markovian framework typically employs simplifying assumptions to ensure ease of implementation. The susceptible-infected-recovered (SIR) compartmental model is a common implementation which assumes that all individuals are susceptible at the beginning, and, once infected, they remain infectious for a recovery period. Once recovered, individuals cannot be infected again. Other common implementations of the compartmental model include the susceptible-infected (SI) model (i.e., no recovery) and the susceptible-infected-susceptible (SIS) model (i.e., recovery does not proffer immunity).

Algorithms

Here, we describe simulation algorithms for the SIR model. Simulations were carried out using the event-based algorithm first proposed by Kiss et al. (2017) and described in detail by Holme (2021). To understand this algorithm, we first describe a naive approach:

Naive Algorithm

1. Initialize all starting, or "seed", nodes as infectious.
2. Initialize all non-starting nodes as susceptible.
3. Traverse the contacts in increasing order of time.
4. Whenever a node becomes infected, change its state to infectious.
5. Stop the simulation when there are no more contacts or no more nodes can be infected.

Now consider two nodes, i and j, and suppose that i is infected at time $t = 0$. There could theoretically be thousands of contacts between i and j, however, only one of these contacts is infectious. Clearly we can avoid many unnecessary iterations if we find the first infectious contact in a single step. Now suppose that all contacts are assumed to have a constant transmission probability of β. Then the index of the first infectious contact between i and j follows a geometric distribution. The probability that the k^{th} contact is infectious is given by

$$\beta(1-\beta)^{k-1}$$

One can sample k by

$$\left\lceil \frac{\log(1-X)}{\log(1-\beta)} \right\rceil \tag{18}$$

where β is the fixed transmission probability and $X \sim \text{Uniform}(0,1)$.

The event-based algorithm relies on several user-defined data structures:

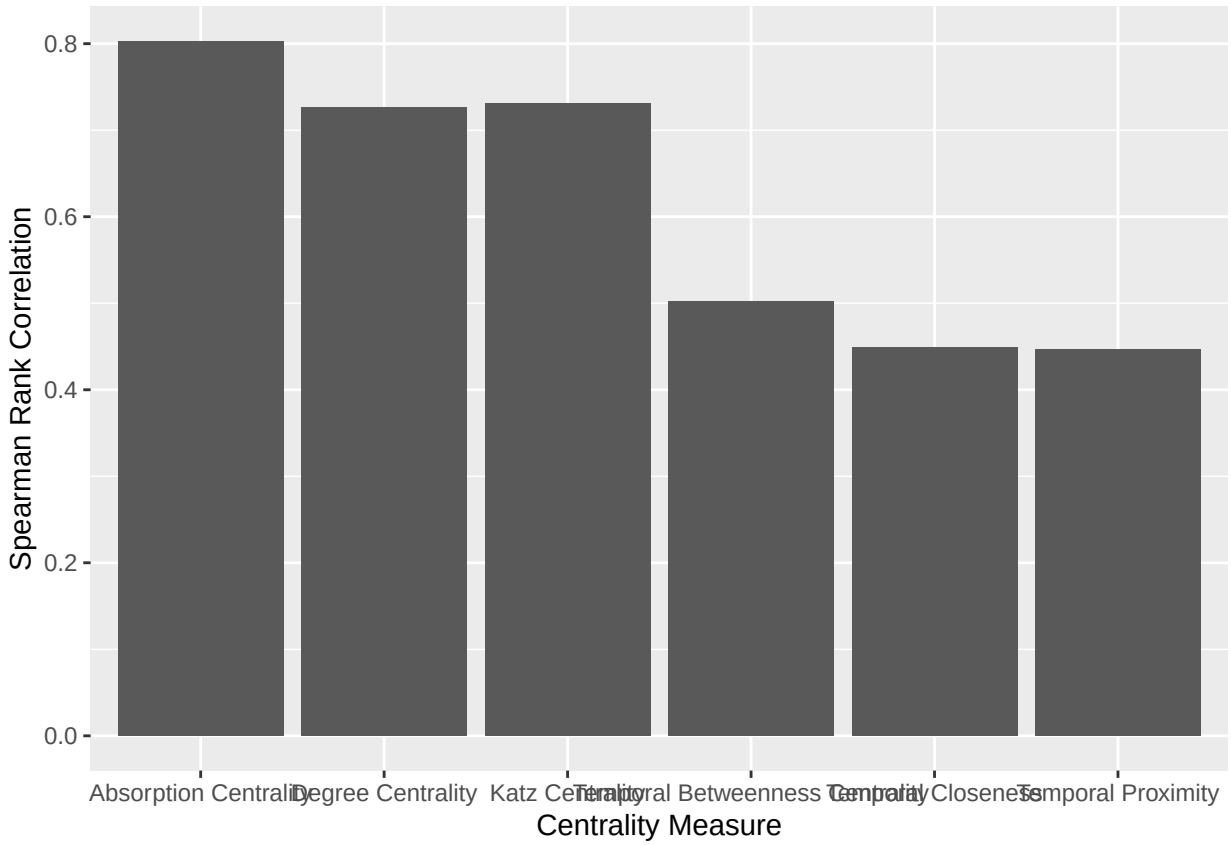
- **Node:** A data structure with a unique identifier (ID) and an infection time (**t_inf**) attribute. In addition, each node contains a list of neighbours and a list of contact times (sorted in increasing order) for each neighbour, as well as a list of associated transmission probabilities if necessary. Note that the contact times must be sorted in increasing order to enable a bisection search. (The significance of this will become clear later.)
- **Heap:** A min-heap data structure for storing node IDs of infected nodes, ordered by earliest infection time, with $\text{O}(\log(n))$ "up-heap" and "down-heap" operations, where n is the number of elements in the heap. A min-heap is conceptually a tree where the root node is the lowest order element, which in this case is the earliest unprocessed infection event. The exact implementation of the heap is unimportant, however the min-heap property must be restored, whenever a node is added or removed, via the *up-heap* and *down-heap* operations.

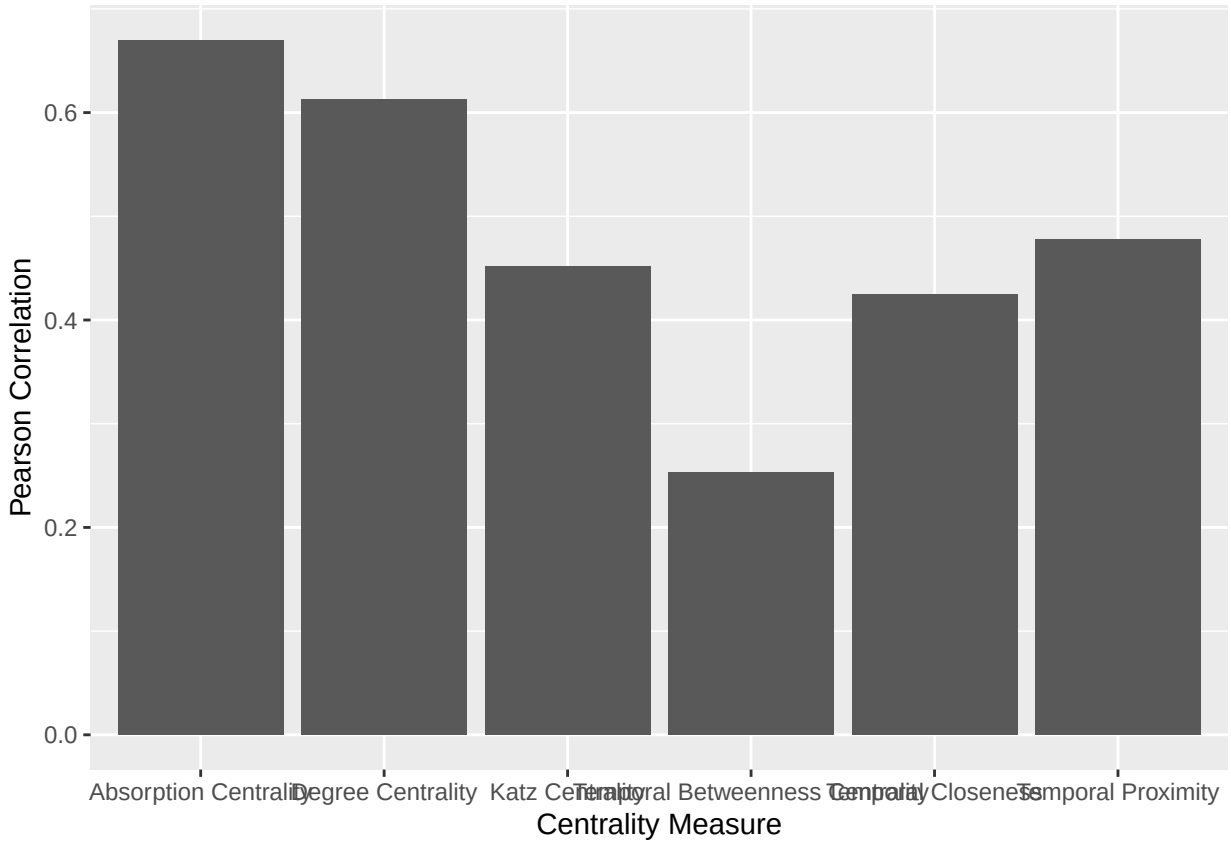
At the beginning of the simulation, **t_inf** is set to T (i.e., the end time of the simulation) for all nodes. All start nodes are added to the heap and their infection times (**t_inf**) are set to 0. At each step, the earliest infection event is removed from the heap and processed. This is repeated until the heap is empty. When an infected node is processed, its neighbours are considered in turn. Suppose node i is infected at time $\text{t\_inf}_i = 20$, and node j is a neighbour of node i. If $\text{t\_inf}_j < 20$ then node j is skipped. Otherwise, a bisection search is carried out to find the earliest contact time t such that $t \geq 20$. Suppose there are m possible infectious contacts between nodes i and j, and we sample an integer k by (18). If $k \geq m$ then we continue, otherwise we add node j to the heap and update t_inf_j.

Ryan note: How are you using the CovidCard data for these simulations? There seems to be a sizeable gap here in clearly explaining how simulations work for the specific data you are using. You should be explaining here that you are carrying out a total of 1,000 simulations where you treat each node as the seed node, leading to the 751,000 total simulations noted in the results section. How many time points/steps are considered, and how does that relate to the underlying data? How do you treat contact events (symmetric rather than directed)? How is the transmission probability estimated for a contact event? How are potential differences in the two sources of information for a contact event rectified? Any descriptive statistics to provide an indication of what the underlying network data are like (e.g., network density, degree distribution, etc., and how those change over time? Is the network connected? etc.

Results

Five centrality measures were calculated for the CovidCard dataset, namely temporal closeness (TP) and temporal proximity prestige (TPP), absorption rank, temporal Katz centrality and temporal degree centrality. **Ryan comment: Why these five? Explain.** For a fair comparison, all centrality measures were calculated by assuming constant transmission probabilities for every contact event. This was done because temporal Katz centrality does not generalise to varying transmission probabilities. These measures were correlated to the observed number of times each individual was infected over 751,000 simulations (shown in Figure 8). **Ryan note: It does not appear that you have a caption that shows for this figure. Additionally, you have six centrality measures here, not five. (You will want to rotate the centrality names so that they can be read. At the moment there is far too much overlap.) Finally, do we actually care about Pearson correlation?** The dataset contained 751 individuals in total, and 1000 simulations for each of the 751 possible starting node were carried out.


Spearman Rank Correlation
0.8
0.6
0.4
0.2
0.0
Absorption Centrality
Degree Centrality
Katz Centrality
Temporal Betweenness Centrality
Temporal Closeness
Temporal Proximity
Centrality Measure


Pearson Correlation
0.6
0.4
0.2
0.0
Absorption Centrality
Degree Centrality
Katz Centrality
Temporal Betweenness Centrality
Temporal Closeness
Temporal Proximity
Centrality Measure

The Spearman correlations are generally large, with absorption rank showing a larger correlation ($\rho \approx 0.8$) compared to other metrics. **Ryan note: Can you provide any insight into why this may be the case? Can you point to aspects of the network that would help to explain why certain measures performed better than others?** Conversely, the metrics generally showed lower Pearson correlations, especially Katz centrality. This suggests that the metrics should not be used for prediction, rather they should be used for ranking the relative importance of individuals within the same network. TP and TPP performed similarly with regards to ranking individuals, however TPP outperformed TP in terms of Pearson correlation.

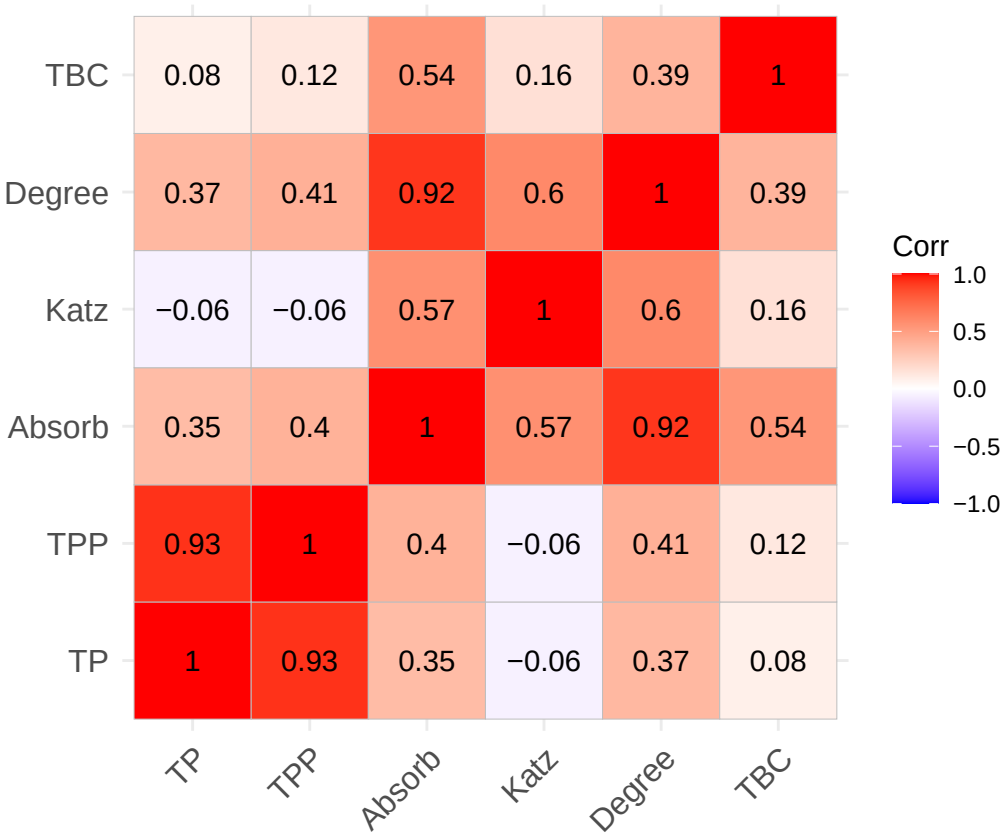


Figure 8: Correlations between the five centrality measures.

Figure 8 shows the Pearson correlations between the five metrics. **Ryan note: Again, we have six measures.** With the exception of proximity rank and closeness rank, **Ryan note: Aim for consistency in how you are referring to these measures, otherwise your reader may get confused.** the measures show low dependence with many weak

correlations. This suggests that the measures could be tapping into different effects, and our rankings could be improved by combining many metrics as predictors in a regression model. **Ryan note: Good. This helps properly motivate your next section.**

Model Selection of Centrality and Prestige Measures

We now turn our attention to examining whether a combination of centrality and prestige measures provides greater predictive ability of nodes' likelihood to become infected. Initially, we might consider a linear model, but, as seen in Figure 9, the relationship between infectivity and centrality and prestige measures is clearly non-linear with non-normal residuals.

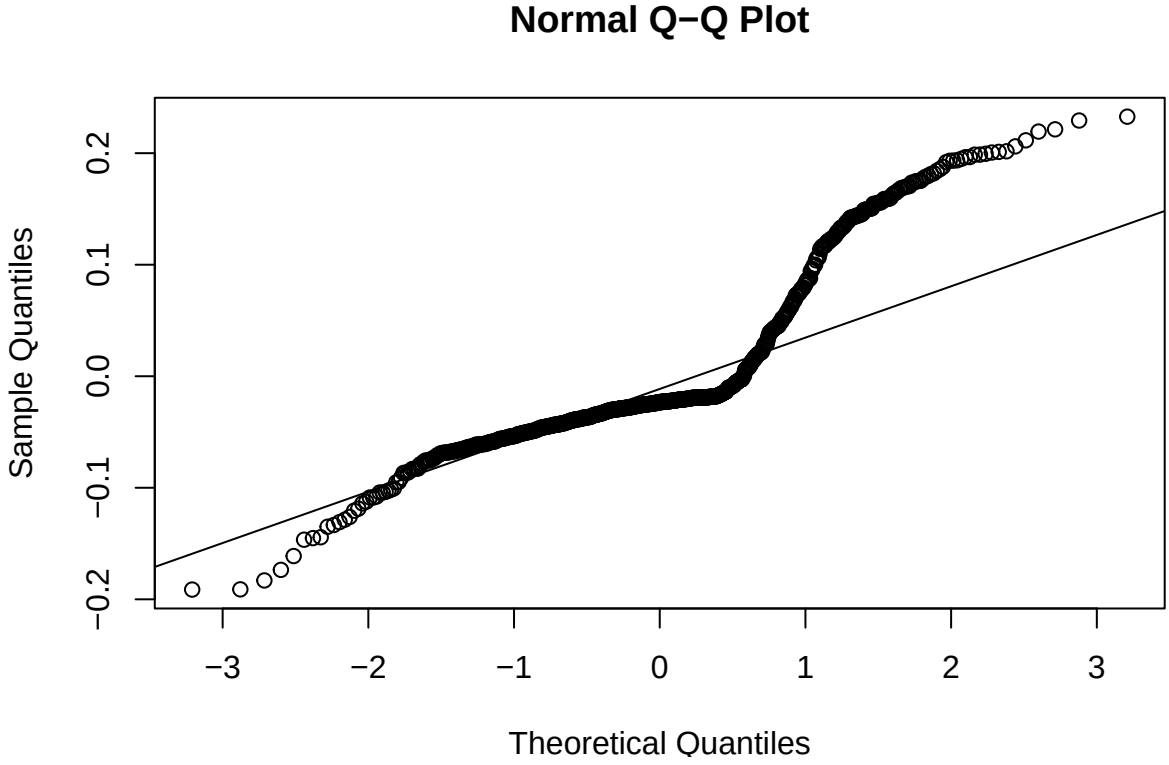


Figure 9: Standardized residuals of the model plotted against theoretical quantiles of the normal distribution.

Examining pairwise relationships (shown in Figure 10), the assumptions of linear regression clearly cannot be satisfied with common transformations, as the output variable ("likelihood", plotted in the bottom-right corner) is strongly bimodal. **Ryan note: This is where descriptive statistics might be helpful in elucidating the reason for this bimodality.** We turn our attention to modelling approaches that can be more effective when the distribution

of the response is non-standard. In particular, we consider gradient boosting, decision trees, and polynomial regression.

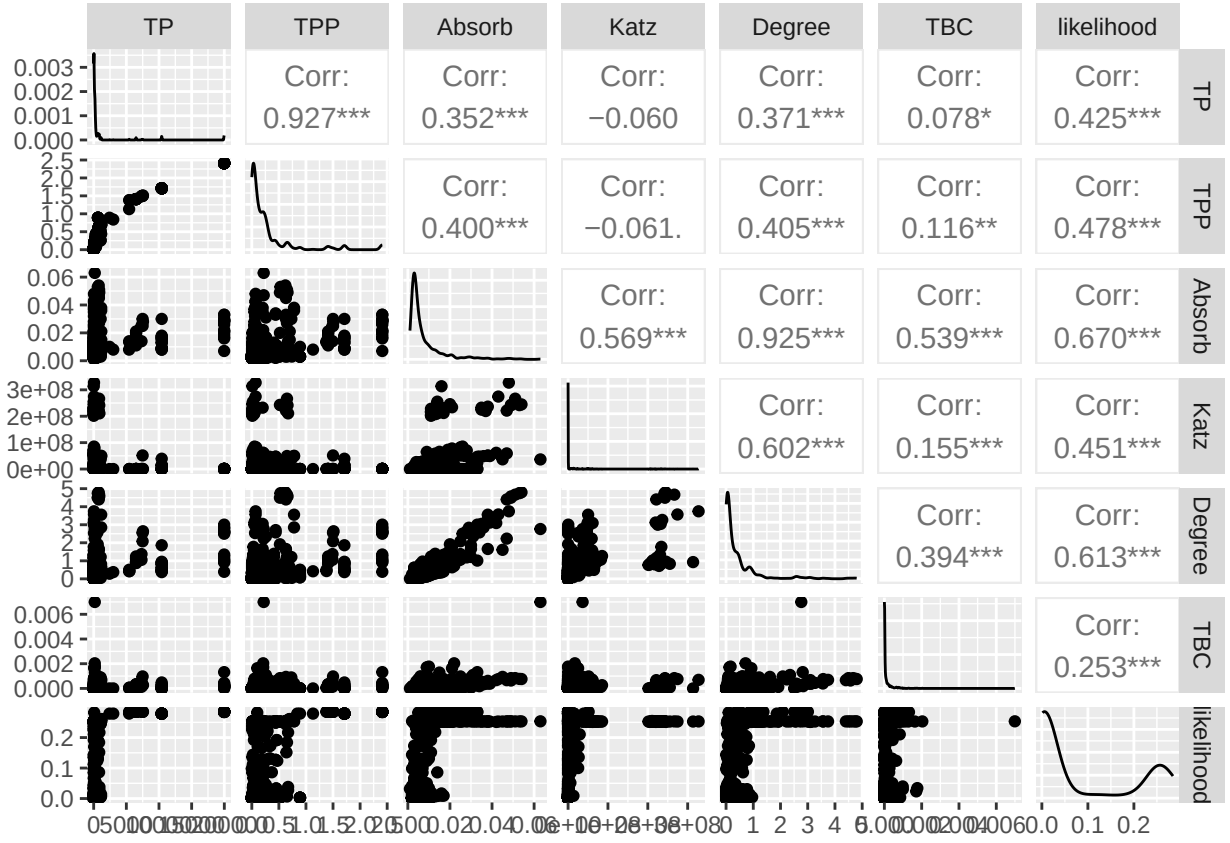


Figure 10: Pairwise relationships between the centrality metrics and the output variable (likelihood)

Of these approaches, polynomial regression is the most similar to linear regression, but it allows for a non-linear relationship between predictors and the response by introduction of polynomial expansions of predictors (e.g., $y = \beta_1 X + \beta_2 X^2 + \beta_3 X^3$). Decision trees split the data into progressively smaller subgroups using decision rules (e.g., $X > 40$), and each group (sometimes called a "leaf node") is assigned a predicted value. **Ryan note: How so?** The prediction for a given observation is then the predicted value of its respective leaf node, which is determined by the decision rules. At each branch of the tree, the optimal decision rule is found by minimising a splitting criteria (e.g., mean squared error). Gradient boosting takes a weighted average of predictions for many decision trees. This may result in better predictions, especially when the number of predictors is large or the data is very noisy.

The data were randomly split into training and test sets comprising 75% and 25% of the data, respectively. The three classes of models (polynomial regression, decision trees, gradient boosting) using all six centrality and prestige measures were trained on the training set and evaluated on the test set. Hyperparameters **Ryan note: What are the hyperparameters? For which model(s)?** were optimized using 10-fold cross validation. Figures 11 to 13 show the predicted values plotted against observed values for the three algorithms. **Ryan note: You only have a figure caption for Figure 11. Personally, I don't think any of these three figures are particularly important, and my suggestion is that they not be included.** Figures 14 and 15 show the Spearman and Pearson correlations **Ryan note: of...? be clear what the results represent.**. We note that these correlations are less than what was obtained by absorption rank in Figure 8. **Ryan note: What are these correlations based off of? Is this really a fair comparison if you are basing this strictly on the training and/or test data, not the full dataset like was done previously. Why a negative correlation with gradient boosting? (Also, be sure to have a consistent way of referring to these. Nowhere have you talked about XGBoost Regression, but that is how you label gradient boosting in your plots.)** In the previous case **Ryan note: What previous case?** gradient boosting and polynomial regression both performed poorly, whereas the decision tree performed similarly to absorption rank.

Ryan note: Not sure about the importance of reporting these results. Why not just report results from where you have performed five-fold cross-validation using all possible combinations of predictors, which includes this case?

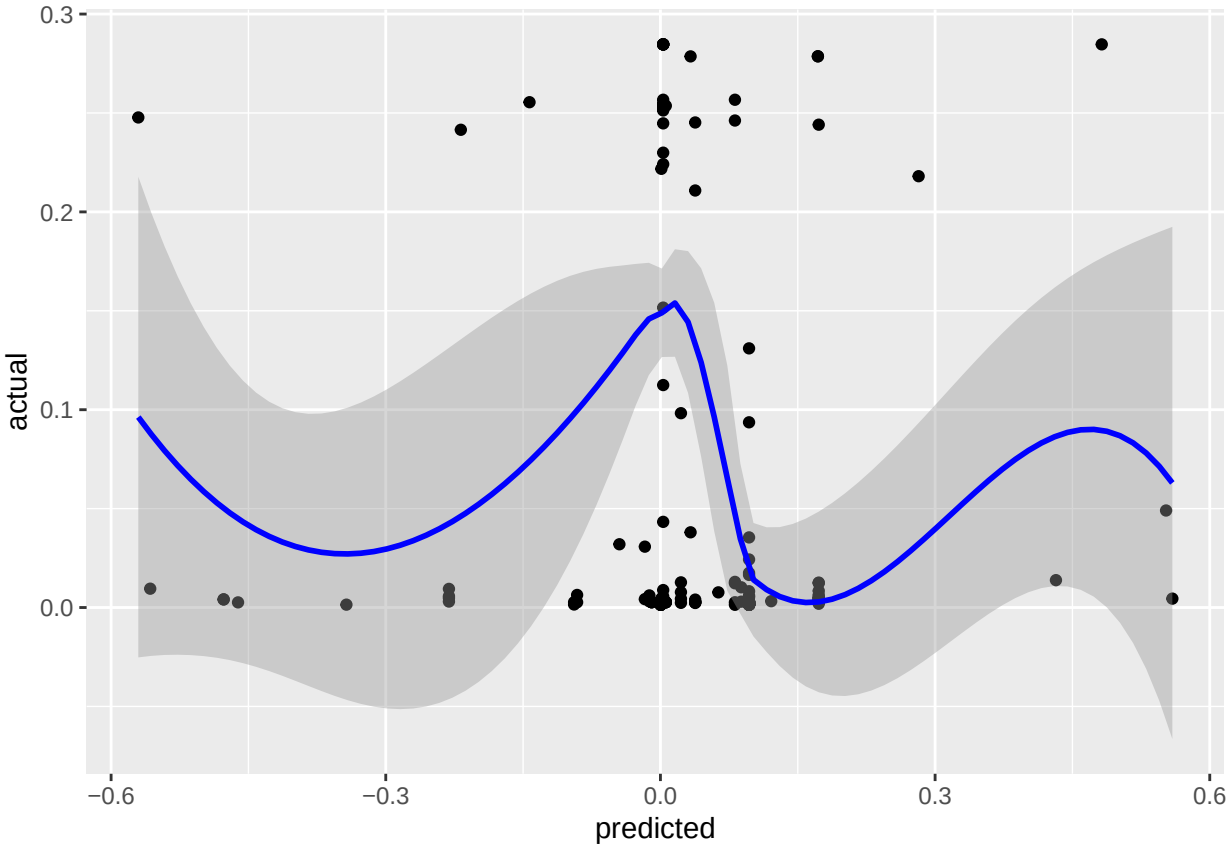
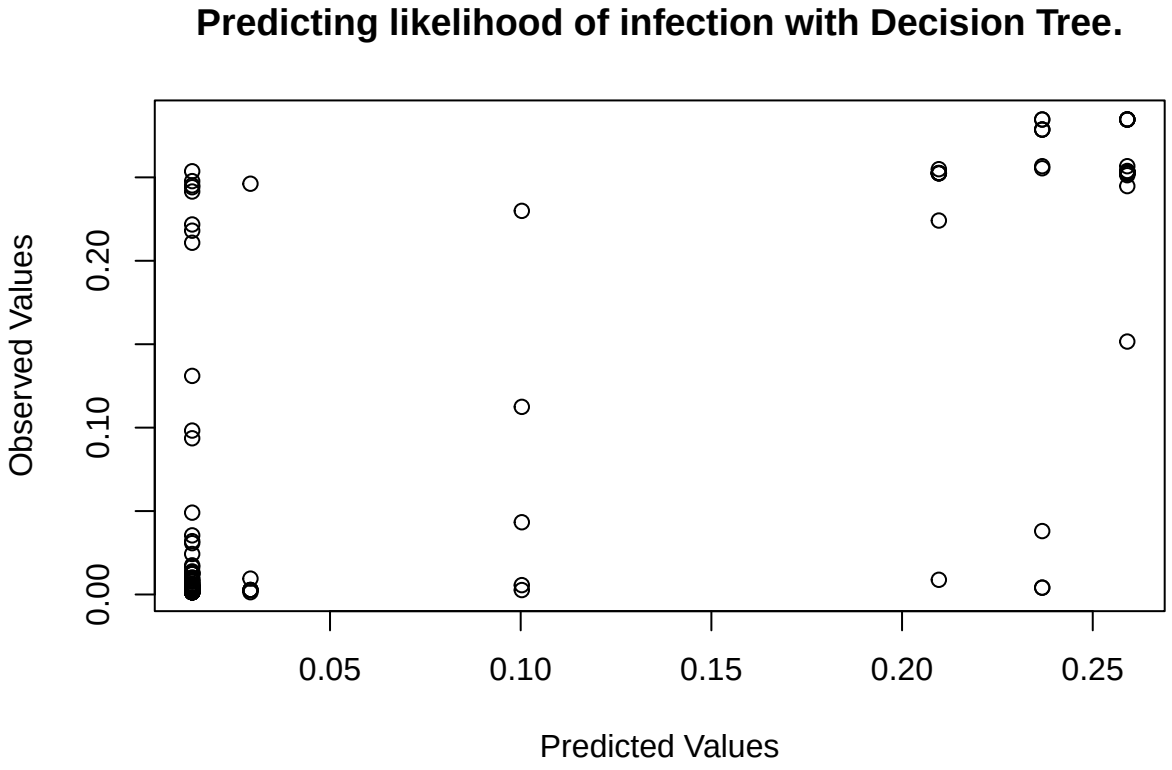
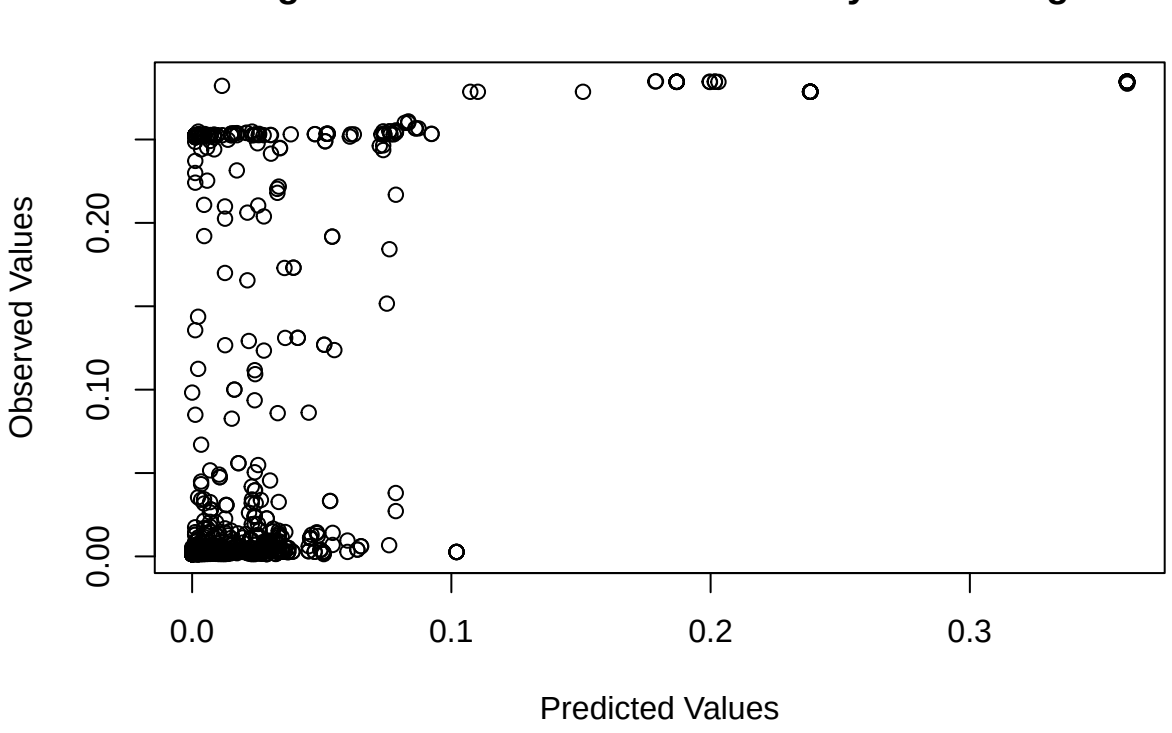
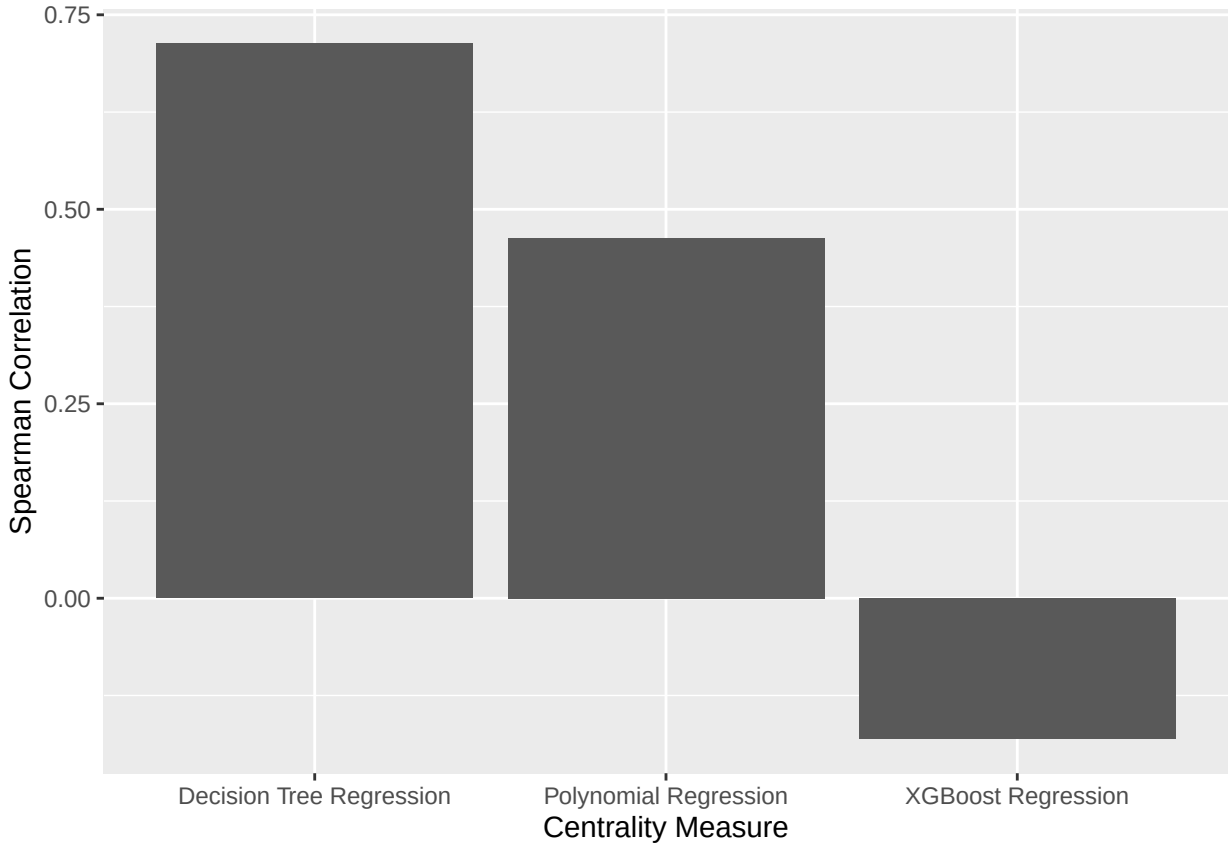
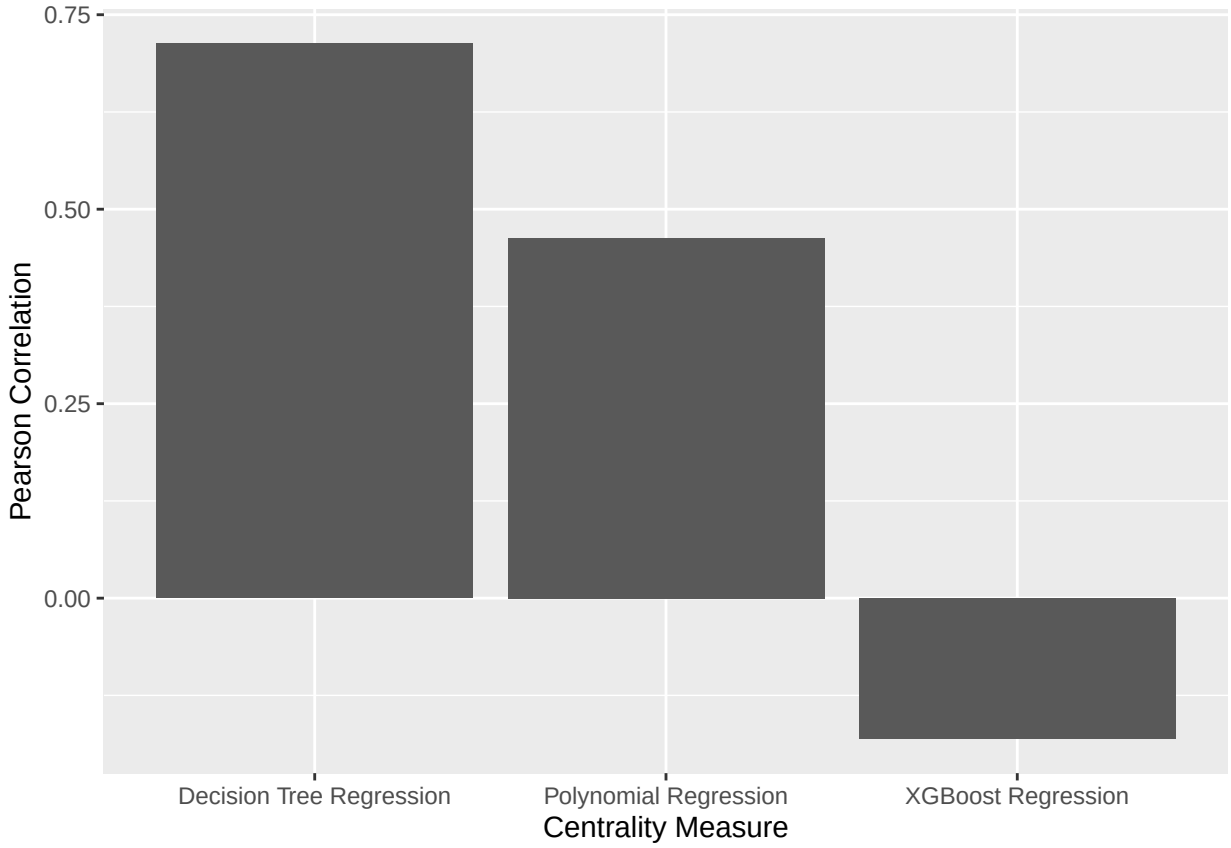


Figure 11: Predicted versus observed values after training Gradient Boosting model and applying to test set.


Predicting likelihood of infection with Decision Tree.
Observed Values
0.00
0.10
0.20
0.05
0.10
0.15
0.20
0.25
Predicted Values

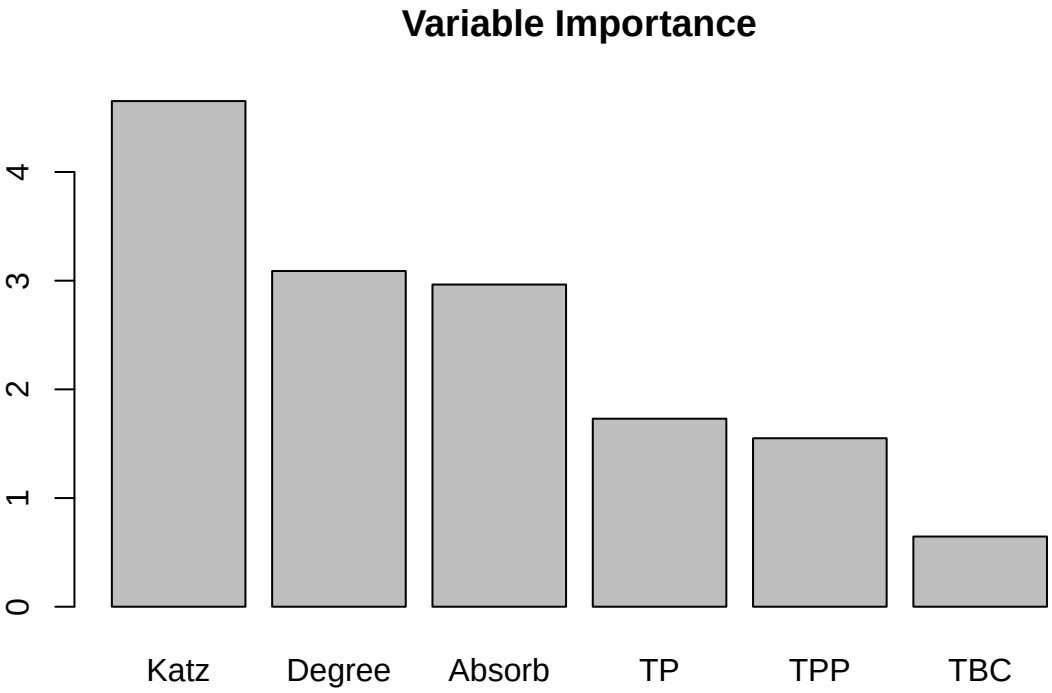

Predicting likelihood of infection with Polynomial Regression
Observed Values
0.00
0.10
0.20
0.0
0.1
0.2
0.3
Predicted Values


0.75
0.50
0.25
0.00
Spearman Correlation
Decision Tree Regression
Polynomial Regression
XGBoost Regression
Centrality Measure


0.75
0.50
0.25
0.00
Pearson Correlation
Decision Tree Regression
Polynomial Regression
XGBoost Regression
Centrality Measure

Variable Selection

If some predictors in a model are unimportant, we risk reducing model performance due to overfitting. In the previous models we considered, we included all centrality measures. Previously, we saw that absorption rank, temporal degree centrality, and temporal Katz centrality correlated more strongly with risk of infection than TP and TPP (Figure 8). Additionally, some of these measures of centrality and prestige were strongly correlated with each other (e.g., TP and TPP with a Pearson correlation of 0.93), suggesting strong multicollinearity of predictors and redundancy through the inclusion of all measures of centrality and prestige. Consequently, we consider variable selection methods to see if a subset of measures of centrality and prestige is more effective in predicting the risk of infection.



We test all possible combinations of two or more variables by 5-fold cross validation. **Ryan note: Did you do this for polynomial regression, decision trees, and gradient boosting or simply decision trees?** Each model was tested with four different values of the complexity parameter, a regularization criterion which limits the complexity of the tree in order to prevent over-fitting. **Ryan note: Only used for decision trees, correct? What values were used and why?** Models were evaluated using mean squared error (MSE), and the five best-performing models were tested against a holdout set. **Ryan note: No. The**

cross-validation results are used to select the "best" model, and the holdout set is used strictly to provide an unbiased estimate of test MSE. Note that, in the interest of model parsimony, it is common to consider all models with MSE lying within 1 standard error of the optimal model and choose the simplest model. (You will want to include the standard error of MSE in your Table 2.) Also, it is common to use repeated 5-fold cross-validation, so maybe repeating 10 or 20 times.

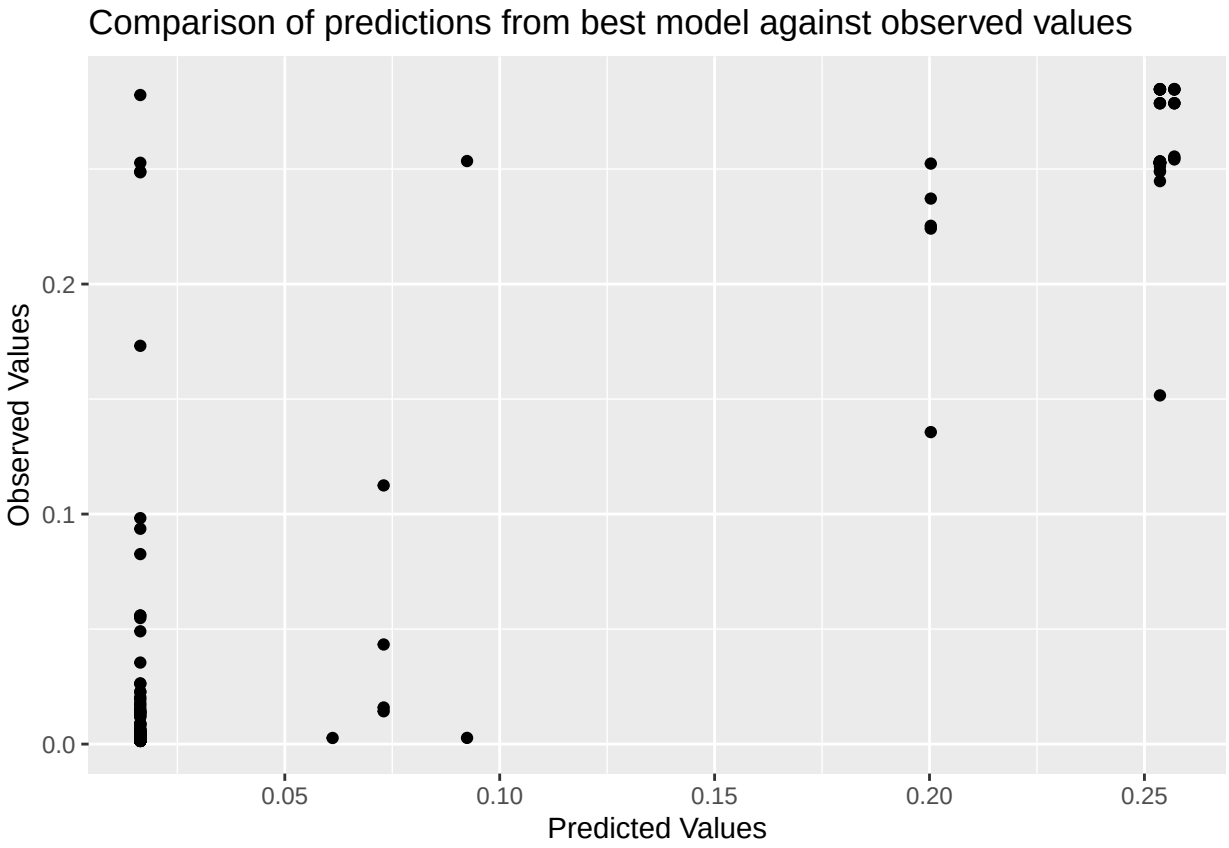
Table 2: Five best performing feature combinations and their MSE

MSE	TP	TPP	Absorb	Katz	Degree	TBC
0.002412	✗	✓	✓	✓	✓	✗
0.002497	✓	✓	✓	✓	✓	✗
0.002507	✓	✗	✓	✓	✗	✓
0.002552	✓	✗	✓	✗	✗	✗
0.002583	✗	✓	✓	✓	✓	✓

The best-performing model used absorption rank, temporal Katz centrality, and temporal degree centrality. **Ryan note: As noted in the previous paragraph, you will want to look at all models lying within 1 SE of this model. No issues with showing the top, say, five models in a table in line with what you have done.** As expected, absorption rank was included in all five models, however the inclusion of Temporal Proximity Prestige did not exclude Temporal Prestige. **Ryan note: This is really an inconsequential statement.*** Figure 12 compares the Root Mean Squared Error (RMSE) **Ryan note: Why switching from MSE to RMSE? Keep things consistent throughout.** of the final model to that of the individual predictors. **Ryan note: Using which class of model(s)?** Temporal Katz centrality and temporal prestige were omitted from Figure 12 due to extremely large RMSE values. **(Ryan note: Still important to report values for temporal Katz centrality and TP, even if omitted from the figure.)**

Ryan note: Remove figure titled "Comparison of predictions from best model

against observed values" on p. 34.



The decision tree model **Ryan note: Again, did you only consider decision trees? If so, then you're really talking only about a model including multiple measures of centrality and prestige.** has the lowest RMSE, followed closely by absorption rank and temporal betweenness centrality. **Ryan note: For these latter two models, is this based on a decision tree as well?** It should be clear that MSE rankings are only meaningful in the context of prediction, and these results should be interpreted accordingly. When ranking the relative importance of nodes in a network, other metrics should be given greater importance (e.g., Spearman rank correlation).

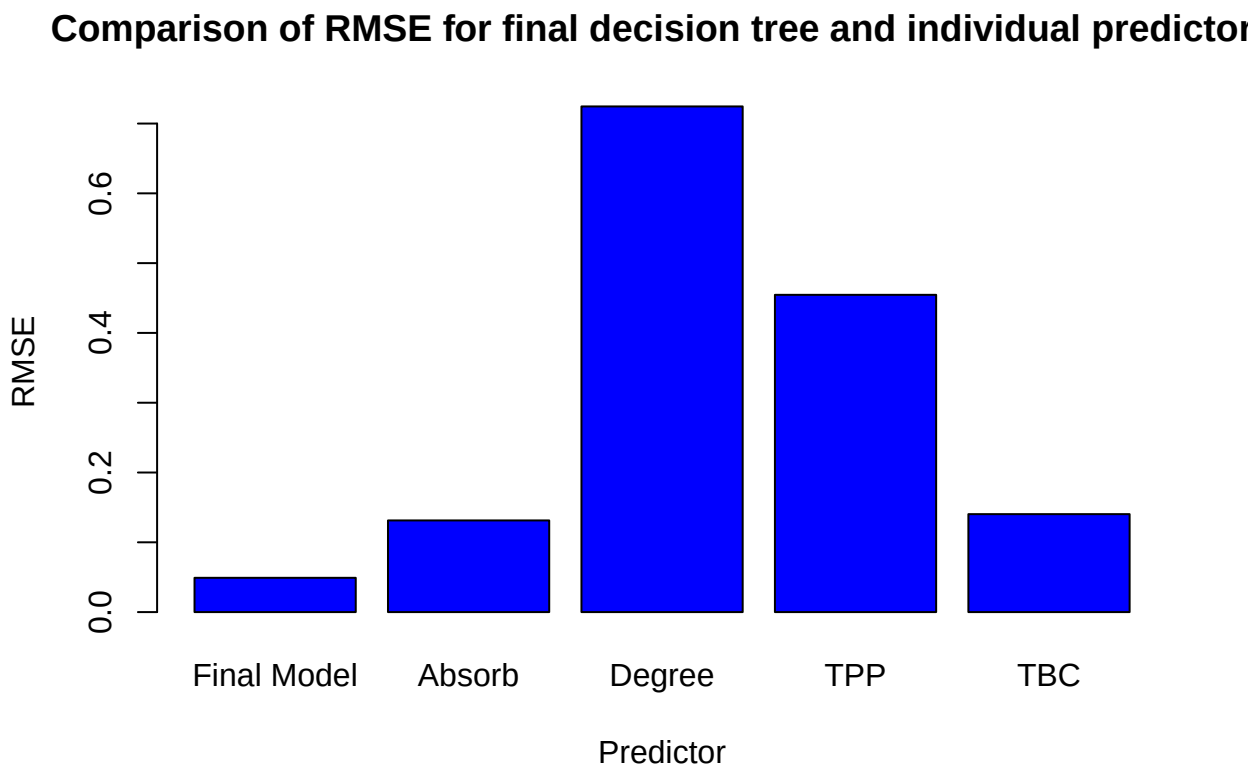
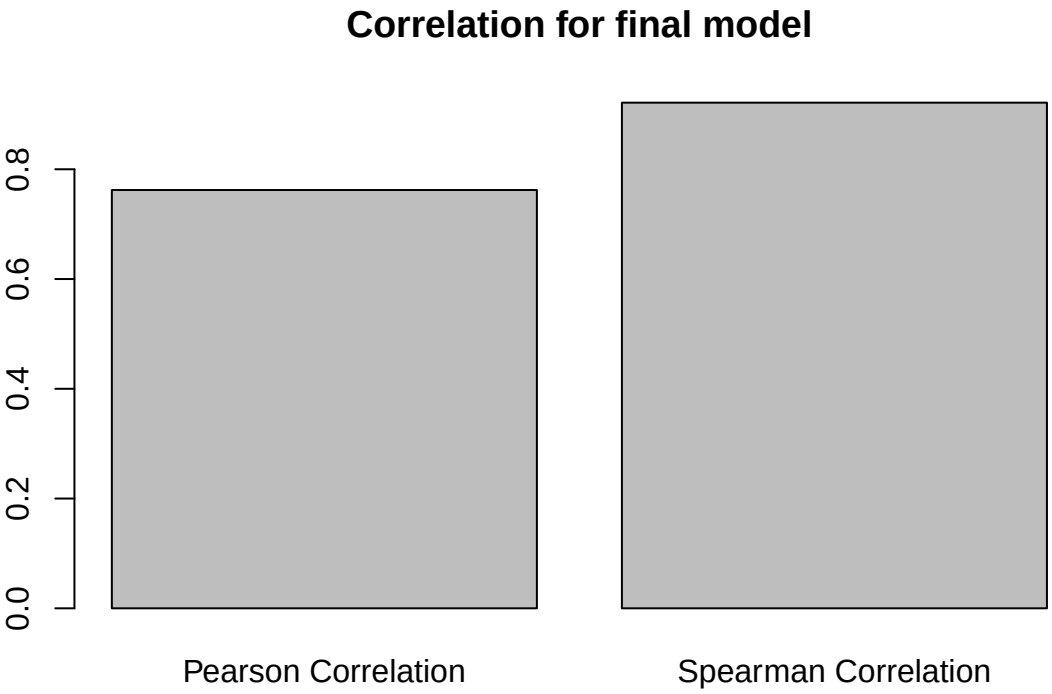


Figure 12: Comparison of RMSE for final decision tree and individual predictors



Ryan note: What does this figure mean? You need to explain.

Constant Transmission Probability

The constant transmission probability assumption is implemented where applicable because it greatly speeds up disease transmission simulation algorithms. Here, we present results applying a constant transmission probability. Probabilities were estimated by using proximity categories, which we will call classes, as predictors in a generalized linear model. The number of 15-minute intervals spent in each class is given over the two hour observation period. **Ryan note: Probably worth very briefly reminding the reader of the way that data were recorded before then describing the data used in estimating constant transmission probabilities.** Transmission probabilities were then calculated by a logistic regression model of the form:

$$log\left(\frac{\pi(x)}{(1-\pi(x))}\right) = \alpha + \beta_1 c_1 + \beta_2 c_2 + \beta_3 c_3$$

where $\pi(x)$ is the transmission probability of the contact and c_i is the number of 15-minute time intervals spent in class i during the two hour observation period. The classes are indexed in increasing order of proximity (i.e., class 3 indicates greater proximity than class 2), therefore the coefficients were chosen to satisfy the constraint $\beta_3 \geq \beta_2 \geq \beta_1$. One may treat the classes as levels of an ordinal variable and assume a proportional relationship, which reduces the model to:

$$log\left(\frac{\pi(x)}{(1-\pi(x))}\right) = \alpha + \beta_1(c_1 + 2c_2 + 3c_3)$$

However, this simplified model was rejected due to insufficient evidence to suggest a proportional relationship.

Ryan note: So what values were used for α and the βs. Why?

To test the constant transmission probability assumption, we compare the observed probability of infection **Ryan note: By "observed", do you actually mean "simulated"?** for 751 individuals for two different algorithms run on the same temporal network. In the first algorithm, the probability of transmission is calculated separately for every contact. In the

second algorithm, we assume that for two nodes i and j, the probability of transmission is a constant value p_{ij} **Ryan note: You used π to denote the probability in your formulation of the logistic regression model. Either change that to p or change p to π going forward.**, regardless of the duration and proximity of the contact event, and this probability is estimated by averaging the event-specific probabilities p_{ijt} (where t again denotes a time index) over all time points. As in the first algorithm, event-specific probabilities are calculated by the same logistic regression model.

In total, 1,000 simulations were run for each algorithm. For each run, the indicator variable, I_i, is 1 if node i was infected, and 0 otherwise. The observed probability. **Ryan note: Again, this is not an observed probability. You continue to use "observed" probability below, so you will need to change that as well.** of infection for node i is the average of I_i over all 1,000 runs.

Denote by $\hat{P^{\mathbf{1}}}$ the vector of observed probabilities for algorithm 1 (and likewise $\hat{P^{\mathbf{2}}}$ for algorithm 2). **Ryan note: But you have 1,000 simulations, so you need an index for that as well.** Thus, $\hat{P_i^{\mathbf{1}}} - \hat{P_i^{\mathbf{2}}}$ is the observed difference between the two algorithms for node i. If we are interested in whether the constant transmission probability leads to significantly different results in terms of our outcome of interest—infection—then we can simply use the Wald statistic

$$Z = \frac{\hat{P_i^{\mathbf{1}}} - \hat{P_i^{\mathbf{2}}}}{\hat{P}_i(1 - \hat{P}_i)(\frac{1}{1000} + \frac{1}{1000})}$$

where $\hat{P}_i$ is the pooled proportion under the null hypothesis ($P_i^{\mathbf{1}} = P_i^{\mathbf{2}}$). **Ryan note: You describe the Wald statistic, but you use ultimately use Fisher's exact test. I would describe here that the Wald test requires a sufficiently large sample size, which is a function of the sample size as well as the estimated proportion.** If we wished to test for equality of the vectors $P^{\mathbf{1}}$ and $P^{\mathbf{2}}$ using a series of Wald tests for each node i, we would need to adjust the individual test significance levels to preserve the familywise Type I error rate. We briefly discuss commonly employed adjustment methods and their use cases.

Bonferroni correction

The Bonferroni correction, proposed by Dunn (1961), is an adjustment which controls the familywise Type I error rate. For a given significance level α, the Bonferroni correction guarantees a family-wise Type I error rate α. It does this by setting the test-wise significance level to $\frac{\alpha}{N}$, where N is the number of tests. The Bonferroni correction is ideal for this experiment because it makes no assumptions about independence between the individual tests, but that comes with the drawback of being overly conservative, meaning that this comes at the tradeoff of statistical power.

Holm's method

Holm's method, proposed by Holm (1979), is a powerful alternative to the Bonferroni correction. Holm's method tests the hypotheses iteratively, updating the p-value at each step. First, we sort the list of p-values in increasing order, and we begin sequential significance tests from the lowest p-value. The algorithm starts with a significance level of $\frac{\alpha}{N}$. If the first result is non-significant, we test the second result with a significance level of $\frac{\alpha}{N-1}$. In general, the i^{th} test statistic is tested with a significance level of $\frac{\alpha}{N-i+1}$. Holm's method also guarantees a familywise Type I error rate of α, but it offers greater power than the Bonferroni correction. **Ryan note: What does Holm's method require of the tests to be valid?**

Hochberg procedure

The Hochberg procedure is similar to Holm's method, however it assumes non-negative correlation between tests. We begin by testing the largest p-value, adjusted for a single comparison. If the p-value is insignificant, we test the next largest p-value sequentially. In general, the i^{th} largest p-value is tested by adjusting for i comparisons. The Hochberg procedure we guarantee a greater power than Holm's method and the Bonferroni Correction.

Pairwise comparisons were carried out with Fisher's Exact test with simulated (rather than exact) p-values, due to computational constraints. **Ryan note: As mentioned previouly, it makes sense to talk about Fisher's exact test earlier with the Wald statistic.** The Z-test was rejected due to the presence of proportions close to 0, which violated the

Table 1: Number of rejections for each method

Method	Number of Rejections	Proportion of Rejections
Bonferroni	24	0.03196
Holm	24	0.03196
Hochberg	24	0.03196

$np > 5$ assumption. Multiple comparisons were controlled for by the three methods. The results are shown in Table ??. **Ryan note: Construct a standard Markdown table, otherwise your table numbering is going to be way off.**

The null hypothesis was rejected 24 times for all multiple comparison adjustments, which suggests that the constant transmission probability assumption is invalid. It is worth noting that the Bonferroni Correction is extremely conservative for 751 comparisons, therefore this result provides strong evidence of a difference between the algorithms. Alger (2020) promotes the use of multiple experiments to increase the reliability of a conclusion. Accordingly, we use a goodness of fit test to support the hypothesis that applying the constant transmission probability assumption produces different results.

Goodness-of-fit test

In this section, we test the constant transmission probability assumption by a likelihood ratio test. First, we sum I_i over all runs to get the total times infected, n_i. By doing this for every node, we obtain a contingency table of the form:

Node	1	2	3	4	...
Algorithm 1 (Constant Transmission)	n_{11}	n_{12}	n_{13}	n_{14}	...
Algorithm 2 (Varying Transmission)	n_{21}	n_{22}	n_{23}	n_{24}	...

Where n_{ij} is the number of times node j is infected for algorithm i. Denote by n_{i+} the i^{th} row sum, and similarly by n_{+j} the j^{th} column sum. Under the assumption of independence between the two variables, algorithm and node, the expected value of the cell count, E_{ij}, is given by:

$$E_{ij} = \frac{n_{i+}n_{+j}}{n}$$

where $n = \sum_i n_{i+}$. The likelihood ratio test statistic

$$G = 2 \sum_{ij} n_{ij} \log \left(\frac{n_{ij}}{E_{ij}} \right) \tag{19}$$

asymptotically converges to a χ^2 distribution with $(I-1)(J-1)$ degrees of freedom, where I and J are the number of rows and columns respectively in the contingency table. **Ryan note:

Table 3: Likelihood Ratio Test Results

statistic	p.value	parameter	method
2330.044	0	750	Log likelihood ratio (G-test) test of independence without correction

The likelihood ratio test (Table 2??) corroborates the conclusion that the two algorithms produce different outcomes, with $p - value < 1 \times 10^{-20}$. **Ryan note: Provide the p-value in scientific notation so that it is not rounded to 0.**

Ryan note: Discussion section? What have we learned? Limitations? Future research?

References

Admiraal, Ryan, Jules Millen, Ankit Patel, and Tim Chambers. 2022. "A Case Study of Bluetooth Technology as a Supplemental Tool in Contact Tracing." *Journal of Healthcare Informatics Research* 6 (2): 208–27.

Alger, Bradley E. 2020. "Scientific Hypothesis-Testing Strengthens Neuroscience Research." *Eneuro*. Society for Neuroscience.

Bavelas, Alex. 1950. "Communication Patterns in Task-Oriented Groups." *Journal of the Acoustical Society of America*.

Chambers, T., A. Anglemyer, R. Egan, S. Derrett, T. Maclennan K.and Emery, A. Patel,

and J. Millen. 2021. "Research Report—Te Whatu Trial of the Bluetooth-Enabled Contact Tracing Card. Wellington (NZL):" Wellington, NZ: New Zealand Ministry of Health. https://www.health.govt.nz/system/files/documents/pages/20201218-te-whatu-research-report.pdf.

Dunn, Olive Jean. 1961. "Multiple Comparisons Among Means." *Journal of the American Statistical Association* 56 (293): 52–64.

Erdos, Paul, Alfréd Rényi, et al. 1960. "On the Evolution of Random Graphs." *Publ. Math. Inst. Hung. Acad. Sci* 5 (1): 17–60.

Fetzer, Thiemo, and Thomas Graeber. 2021. "Measuring the Scientific Effectiveness of Contact Tracing: Evidence from a Natural Experiment." *Proceedings of the National Academy of Sciences* 118 (33): e2100814118.

Freeman, Linton C. 1977. "A Set of Measures of Centrality Based on Betweenness." *Sociometry*, 35–41.

Grindrod, Peter, Mark C Parsons, Desmond J Higham, and Ernesto Estrada. 2011. "Communicability Across Evolving Networks." *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics* 83 (4): 046120.

Hanke, Moritz, and Ronja Foraita. 2017. "Clone Temporal Centrality Measures for Incomplete Sequences of Graph Snapshots." *BMC Bioinformatics* 18: 1–18.

Holm, Sture. 1979. "A Simple Sequentially Rejective Multiple Test Procedure." *Scandinavian Journal of Statistics*, 65–70.

Holme, Petter. 2018. "Objective Measures for Sentinel Surveillance in Network Epidemiology." *Physical Review E* 98 (2): 022313.

———. 2021. "Fast and Principled Simulations of the SIR Model on Temporal Networks." *Plos One* 16 (2): e0246961.

Katz, Leo. 1953. "A New Status Index Derived from Sociometric Analysis." *Psychometrika* 18 (1): 39–43.

Kempe, David, Jon Kleinberg, and Amit Kumar. 2000. "Connectivity and Inference Problems for Temporal Networks." In *Proceedings of the Thirty-Second Annual ACM Symposium on Theory of Computing*, 504–13.

Kermack, William Ogilvy, and Anderson G McKendrick. 1927. "A Contribution to the

Mathematical Theory of Epidemics." *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* 115 (772): 700–721.

Kim, Hyoungshick, and Ross Anderson. 2012. "Temporal Node Centrality in Complex Networks." *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics* 85 (2): 026107.

Kiss, István Z, Joel C Miller, Péter L Simon, et al. 2017. "Mathematics of Epidemics on Networks." *Cham: Springer* 598 (2017): 31.

Klovdahl, Alden S. 1985. "Social Networks and the Spread of Infectious Diseases: The AIDS Example." *Social Science & Medicine* 21 (11): 1203–16.

Lin, Nan. 1976. "Foundations of Social Research." *(No Title)*.

Macdonald, George. 1952. "The Analysis of Equilibrium in Malaria." *Tropical Diseases Bulletin* 49 (9): 813–29.

Piraveenan, Mahendra, Mikhail Prokopenko, and Liaquat Hossain. 2013. "Percolation Centrality: Quantifying Graph-Theoretic Impact of Nodes During Percolation in Networks." *PloS One* 8 (1): e53095.

Rocha, Luis EC, and Naoki Masuda. 2014. "Random Walk Centrality for Temporal Networks." *New Journal of Physics* 16 (6): 063023.

Seidel, Scott Y, and Theodore S Rappaport. 1992. "914 MHz Path Loss Prediction Models for Indoor Wireless Communications in Multifloored Buildings." *IEEE Transactions on Antennas and Propagation* 40 (2): 207–17.