Infect Everyone: Modelling Infection Spreading via Bootstrap Percolation

William Kellough

Introduction to Bootstrap Percolation

Bootstrap percolation is a discrete time process on a graph where every vertex is in one of two states: healthy or infected. If a vertex becomes infected during the process, it stays infected. Given a fixed threshold r, this process is called an *r-neighbourhood bootstrap process* if every healthy vertex becomes infected when it is adjacent to r infected vertices. An initial set of infected vertices that eventually infects the entire graph it is contained in is called a *percolating* set for the graph. The minimum number of vertices needed to obtain a percolating set for a graph G in an r-neighbourhood process is denoted $m(G, r)$.

Bootstrap percolation processes have applications in numerous scientific fields. For example, a bootstrap percolation process can be used to model diseases spreading through networks of people. Bootstrap percolation is also used in quantum physics for modelling the orientational ordering processes of specific molecules, in fluid mechanics it is used to model how a fluid can flow through an environment, and in computer science bootstrap percolation is used to prevent failures in large computer memory storage systems. These are just some of the many examples that demonstrate how varied the applications for bootstrap percolation processes are and that studying them is worthwhile.

Introduction to Kneser Graphs

Let $n \in \mathbb{N}$ and $k \le n$ be a positive integer. Consider the set $\{1, \ldots, n\}$. A *Kneser graph* on $\binom{n}{k}$ vertices is a graph where each vertex corresponds to a k-element subset of $\{1, \ldots, n\}$ and two vertices are adjacent if their corresponding k-element subsets are disjoint. Such a graph is denoted $K(n, k)$. An example of a Kneser graph that is well known among graph theory students is the Petersen graph. The Petersen graph is obtained by using 2-element subsets of the set $\{1, 2, 3, 4, 5\}$. This is demonstrated below.
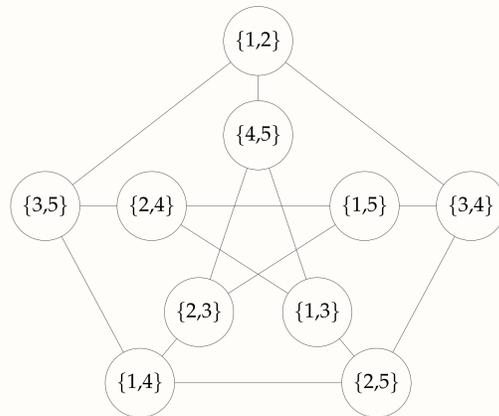

Using the aforementioned notation, the Petersen graph can be denoted as $K(5, 2)$. Using the colour red to represent infected vertices, the following is an example of a 2-neighbourhood bootstrap process on $K(5, 2)$ with a percolating set of size three.
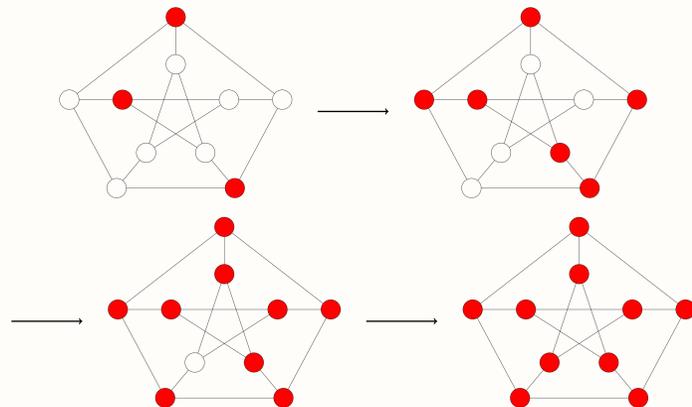
Bootstrap Process with Vaccinations

Consider a variation of the bootstrap process where the infection spreads the same way but at every discrete time step exactly one healthy vertex can be *vaccinated* and become immune to infection indefinitely. It is assumed that only healthy vertices are allowed to be vaccinated. Under this new process, one could ask how many vertices can be saved from infection. Using green to represent a vertex being vaccinated, here is an illustration of the same 2-neighbourhood bootstrap process on $K(5, 2)$ as before but with vaccinations.
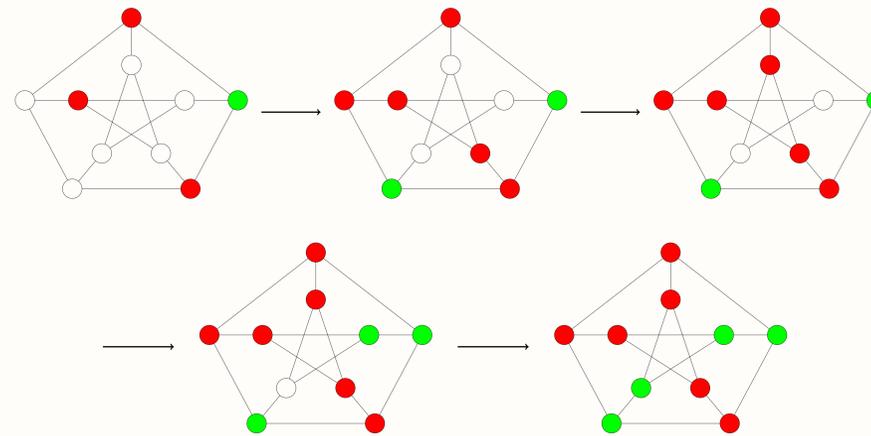
The largest number of vertices that are guaranteed to be saved through vaccinations in an r-neighbourhood bootstrap process for any initial minimum sized percolating set in a graph G is denoted $MSV(G, r)$.

Objectives

The main objective of this research is to determine the minimum number of initially infected vertices needed for the infection to spread through an the entire Kneser graph, $K(n, k)$, for arbitrarily large n. That is, for fixed k what is the value of $m(K(n, k), r)$ where $n \to \infty$? After finding the value of $m(K(n, k), r)$ as $n \to \infty$, the next objective is to find $MSV(K(n, k), r)$ as $n \to \infty$ for the same fixed k.

Results

Let r be the threshold for infection and let $k \in \mathbb{N}$.

Case 1: $r > \binom{n-k}{k}$**.** Note that the number of neighbours every vertex in $K(n, k)$ has is $\binom{n-k}{k}$ by the definition of a Kneser graph. Thus for the entire graph to get infected, every vertex needs to already be infected. So whenever $r > \binom{n-k}{k}$, $m(K(n, k), r) = \binom{n}{k}$ and $MSV(K(n, k), r)) = 0$.

Case 2: $r = \binom{n-k}{k}$**.** By using the Erdős–Ko–Rado Theorem [2], it can be shown that when $r = \binom{n-k}{k}$, $m(K(n, k), r)) = \binom{n-k}{k})) = \binom{n}{k} - \binom{n-1}{k-1}$ and $MSV(K(n, k), r) = 1$.

Case 3: $k \ge \frac{n}{2}$**.** When $k \ge \frac{n}{2}$, $m(K(n, k), r) = \binom{n}{k}$ and $MSV(K(n, k), r) = 0$ since the graph $K(n, k)$ has no edges. For the case where $k = \frac{n}{2}$ it can be shown that $m(K(n, k), r) = \frac{\binom{n}{k}}{2}$ and $MSV(K(n, k), r) = 1$ when $r = 1$. If $k = \frac{n}{2}$ and $r > 1$ then $m(K(n, k), r) = \binom{n}{k}$ and $MSV(K(n, k), r) = 0$

Case 4: $k = 1$**.** The graph $K(n, 1)$ is simply the graph on n vertices where every possible edge is in the graph. It follows that when $0 < r < n$, $m(K(n, 1), r) = r$ and $MSV(K(n, 1), r) = 1$.

Case 5: $2 \le r < \binom{n-k}{k}$ **and k is a fixed constant.** The two following theorems are the main results of this research and handle many of these cases.

Theorem. Let $r = \binom{\lfloor cn \rfloor - b}{k}$ where $c \in \mathbb{R}$ such that $0 < c < 1$. Fix $k \ge 2$. There exists an $N_0 \in \mathbb{N}$ such that for any $n > N_0$, $m(K(n, k), r) = r$.

Theorem. Let $r = \binom{\lfloor cn \rfloor - b}{k}$ where $c \in \mathbb{R}$ such that $0 < c < 1$. Fix $k \ge 2$. There exists an $N_0 \in \mathbb{N}$ such that for any $n > N_0$, $MSV(K(n, k), r) = 2$.

Diagram of Results

The following diagram illustrates all of the known values of $m(K(n, k), r)$ for fixed k as $n \to \infty$.
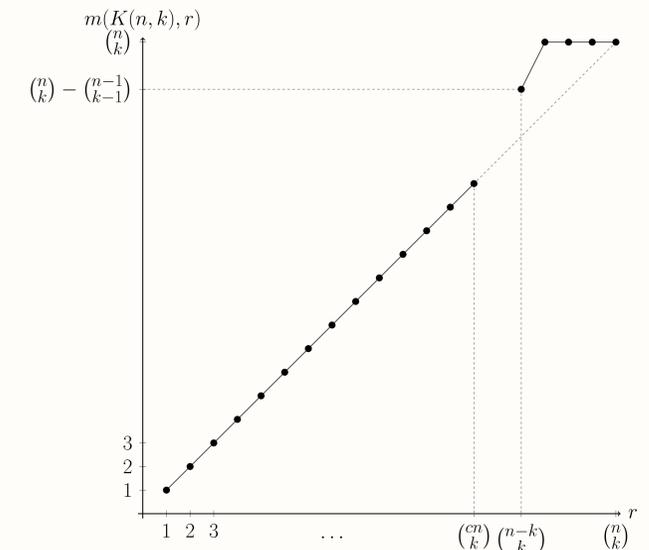
Open Problems

- For arbitrarily large n and fixed $k \in \mathbb{N}$, the values of $m(K(n, k), r)$ and $MSV(K(n, k), r)$ are currently unknown for $\binom{cn}{k} < r < \binom{n-k}{k}$ where $c < 1$. Will the relationship between r and $m(K(n, k), r)$ continue to be linear in this interval? Will the values for $MSV(K(n, k), r)$ continue to stay at 2 or will they converge to some other value?

- This research analyzed the values of $m(K(n, k), r)$ only for arbitrarily large n. Is it possible to identify the values of $m(K(n, k), r)$ and $MSV(K(n, k), r)$ for all $n \in \mathbb{N}$?

- The properties of Kneser graphs allowed for the values of $m(K(n, k), r)$ and $MSV(K(n, k), r)$ to be found. For example, one helpful property of Kneser graphs is that the diameter of the Kneser graph $K(n, k)$ is $\left\lceil \frac{k-1}{n-2k} \right\rceil + 1$ for any $k < \frac{n}{2}$ [3]. Are there any other families of graphs where the value of $m(G, r)$ can be found using the properties of the graphs in that family? As an example, one family of graphs that have been studied are square grid graphs [1].

- For finding the $MSV(K(n, k), r)$ values, only one vaccination was allowed for each time step. How many vertices can be saved if more than one vaccination was allowed in each time step? Given that $m(K(n, k), r)$ is known, what is the minimum number of vaccinations needed in each time step to stop the infection from spreading to unvaccinated, healthy vertices? These same questions could be asked about any graphs and not just Kneser graphs.

References

[1] J. Balogh, G. Pete, Random disease on the square grid, *Random Structures Algorithms* **13** (1998), 409–422.

[2] P. Erdős, C. Ko, and R. Rado, Intersection theorems for systems of finite sets, *Quart. J. Math. Oxford Ser. (2)* **12** (1961), 313–320.

[3] M. Valencia-Pabon, and J. Vera, On the diameter of Kneser graphs, *Discrete Math.* **305** (2005), 383–385.

Acknowledgement