Graph Bootstrap Percolation

A. Babb
S. Martinez
T. Wang

Graph Bootstrap Percolati

Case studies on selected graphs and algorit

Andy Babb, Stephanie Martinez, Tim
Advisor: Connor Anderson

Department of Mathematics, Northeastern Ur

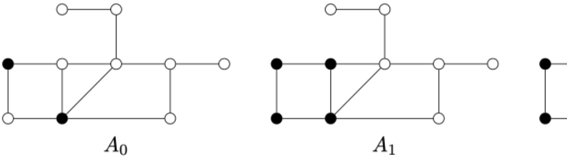
Summer REU Final Presentations – Jun

Introduction

Consider the following activation process:

- Let $G = (V, E)$ be a graph, whose vertice in an "active" or "inactive" state. Fix a tl and choose a seed set $\;_0 \subseteq V$.
- Denoting the adjacent vertices, or neighbo vertex v by $N(v)$, activations (also called spread through the system as follows:

$$\;_{i+1} = \;_i \cup \{v \in V \; : \; |N(v) \cap \;_i|$$



A simple example

Motivation

Bootstrap percolation can model spread in real

- Vertices and edges represent people and r
while the percolation process represents th
ideas, rumors, or trends.
- Which groups at minimum should be adve
maximal influence throughout a communit
- How can we reduce the spread of ideas or

Bootstrap percolation problems

In the language of bootstrap percolation:

- Minimum contagious sets
  - What is the smallest possible seed set $_0$
    $| \; _0|$) required to eventually infect an ent
- Spread minimization
  - If we must initially infect a fixed number
    choose a seed set $_0 \subseteq V$ of fixed cardin
    vertices should be infected to minimize s

Minimum contagious sets

Let $G(V, E)$ be a graph, and fix a threshold k
percolation process.

- A contagious set is a seed set ${}_0$ that eve
  the entire graph (i.e., $\langle\ {}_0\rangle = V$).
- A contagious set is called minimum if ther
  smaller contagious set, and its size is deno
- Goal: Find $m(G, k)$.
- One solution: First, compute all possible c
  by exhaustively finding all seed sets ${}_0$ for
  $\langle\ {}_0\rangle = V$. Then, $m(G, k)$ is equivalent to
  of the smallest ${}_0$ (there may be more tha

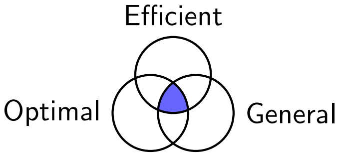
Spread minimization

Let $G(V, E)$ be a graph and fix a threshold k
percolation process. If we must initially infect
of vertices (i.e., choose a seed set $_0 \subseteq V$ of
which vertices should be infected to minimize

- For any seed set $_0 \subseteq V$, we denote the
  that eventually become infected by $\langle \ _0\rangle$.
- Goal: Find a seed set $_0$ of size n that mi
- One solution: Compute $\langle \ _0\rangle$ for all possib
  on the graph to find the smallest $|\langle \ _0\rangle|$ by

Challenges

- Spread minimization is believed to be NP-
this remains unproven)
- This implies there is unlikely to be an effic
general solution to the problem
- Hence, most approaches focus on specific
(sacrificing generality) or finding a subopt
- Past work on the problem focuses on man
supporting the suspected NP-hardness of



Some interesting cases

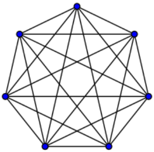
In order to better understand bootstrap percola
through percolation processes on specific kinds

- Complete graphs
- Petersen graph
- Bipartite graphs

Complete graphs

In K_n, the complete graph of n vertices, each v
to every other vertex. Given a seed set $\;_0$ and
whether any other vertices can be infected is p

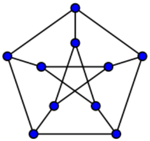
- If $|\;_0| < k$, no vertices have k or more inf
so no spread occurs. Hence $|\langle\;_0\rangle| = |\;_0|$.
- Otherwise, $|\;_0| \geq k$, and all vertices have
infected neighbors. Therefore, the entire g
infected in the next step, and $|\langle\;_0\rangle| = |V$



K_7, the complete graph of 7 verti

Petersen graph

The Petersen graph is a peculiar graph with 10
edges. It is helpful to divide its vertices into tw
the "inner" vertices and the "outer" vertices.



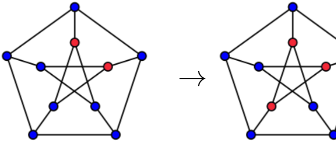
The Petersen graph

We examined outcomes for $|\langle \; {}_0\rangle|$ with thresho
di erent seed sets ${}_0$.

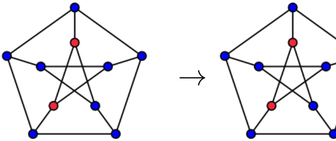
Petersen graph

Considering the "inner" vertices when $| \ _0| = 2$

- Selecting two nearby vertices results in at subsequent infection.



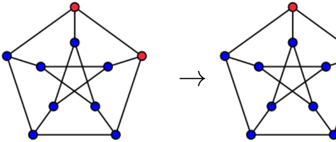
- Selecting not so nearby vertices as $\ _0$ resu infections, and $|\langle \ _0\rangle| = | \ _0|$.



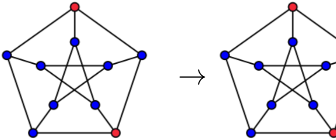
Petersen graph

Considering the "outer" vertices when $|\ _0| = 2$

- Selecting two nearby vertices results in no infections.



- Selecting not so nearby vertices as $_0$ resu one subsequent infection, and $|\langle\ _0\rangle| = |$



Petersen graph

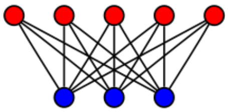
Why does this happen?

- There is an automorphism on the Peterse[n]
preserves its shape, and maps the inner ve[rtices]
outer vertices.
- Any choice of two nearby inner vertices m[aps]
vertices that are not so nearby.
- Therefore, the opposite rule applies after [a]
automorphism.

Bipartite graphs

A bipartite graph $K_{m\ n}$ is a graph whose vertex
split into two disjoint subsets and B, where
$|B| = n$.

- A bipartite graph is called complete if eve
connects to every vertex in B.
- Percolation on the complete bipartite grap
similarly to the complete graph K_n.
- We explored percolation on bipartite grap
the percolation process on complete bipar
di erent combinations of m and n.



The complete bipartite graph K_3

Complete bipartite graphs

For $m, n \geq k$, a large enough and well distribut
will infect the entire graph $K_{m\ n}$.

- If $|\ _0| < k$, every vertex in $K_{m\ n}$ has less t
neighbors and remains inactive.
- If $|\ _0| \geq k$, $\ _0$ can be entirely in subset
the entire graph.
  - If $\ _0$ is chosen to be in subset , subset
$|\ _0| \geq k$ infected neighbors and its vertic
activated during the first iteration of per
  - During the second iteration, all vertices i
$n \geq k$ infected neighbors and be activate

General bipartite graphs

For any general bipartite graph $K_{m\ n}$, $m \geq n$, if
we can choose a seed set such that no new ver
activated during the percolation process.

- For a graph where $m, n \geq \lceil \frac{\lfloor\ 0 \rfloor}{2} \rceil$, split the
in half. If the seed set is odd sized, let the
di er by 1 vertex.
- For a graph where $n < \lceil \frac{\lfloor\ 0 \rfloor}{2} \rceil$, activate all
as much of subset A as necessary.

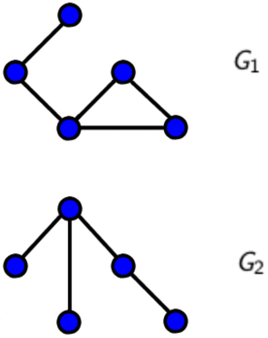
Erdős–Renyi random graphs

The Erdős–Rényi graph is a random graph repr
$G = G(n, p)$. n is the number of vertices and
associated probability.

- Edges of all possible combinations of verti
  G with probability p
- We examined only the cases in which this
  a connected graph – any isolated vertex m
  seed set to ever get activated
- Con: Does not accurately portray social c

Example of Erdős–Renyi graphs

- Two possible Erdős–Rényi graphs for $G(5$
- All possible edges among 5 vertices are se probability 0 3



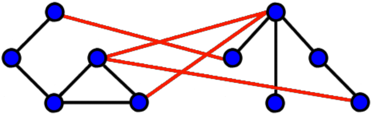
Stochastic block model

A stochastic block model (SBM) graph is the
Erdős–Rényi graphs G_1 and G_2. There is an ac
probability q and two di erent communities G_1

- We have a graph $G = (n_1, n_2, p_1, p_2, q)$
- n_i is the number of vertices per graph
- p_i is the probability per graph for intra-co
- q is the probability for inter-connected edg
  communities G_1 and G_2

Example stochastic block model grap

- We union the two Erdős–Rényi graphs fro
- Edges between vertices of di erent commu with probability q



Stochastic block model

In Bootstrap percolation on the stochastic bloc
(Torrisi, Giovanni Luca, Michele Garetto, and E
2022), asymptotic bounds are given for p in te

$$1/n_i \ll p_i \ll 1/({n_i}^{1/k})$$

Additionally, a critical number of seeds per co
given in terms of p, n, and k.

$$g_i = \ 1 \quad \frac{1}{k}) \left(\frac{(k \quad 1)!}{n_i p_i^k}\right)^{\frac{1}{k \quad 1}}$$

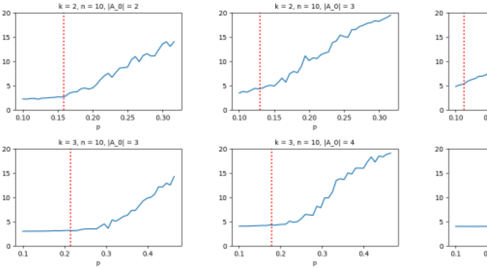
- When the the size of the seed set $| \quad _0| =$
"phase transition" in which the graph is r

A brute-force approach

We generated computations to observe the beh
percolation in the SBM model which includes r
cases. Generating these computations would al
see how choosing a random seed set performs
We set the following parameters,

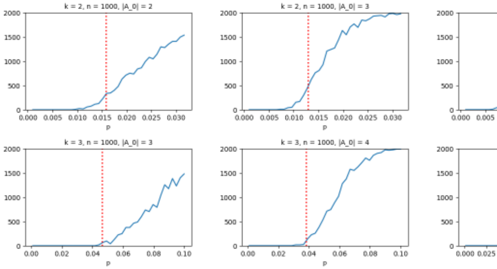
- $n_1 = n_2$ and $p_1 = p_2$
- Sizes for n of 10, 50, 100, 500, 1000 verti
- $2 \leq |\ _0| \leq 4$ and $2 \leq k \leq 4$
- k is bounded below by $|\ _0|$
- q is a constant multiple of p
- 40 p-values bounded between $1/n_i$ and $1/$

SBM computations, 10 vertices



Randomly chosen seed sets for n =

SBM computations, 1000 vertices



Randomly chosen seed sets for $n =$

SBM Results

The red dotted line indicates when the critical
g_i is reached and equal to $| \quad _0|$ for some value
be derived by solving for p in $g_i = \quad 1 \quad \frac{1}{k}) \left(\frac{(k}{r}\right.$
randomly chosen seed set, it can be seen that,

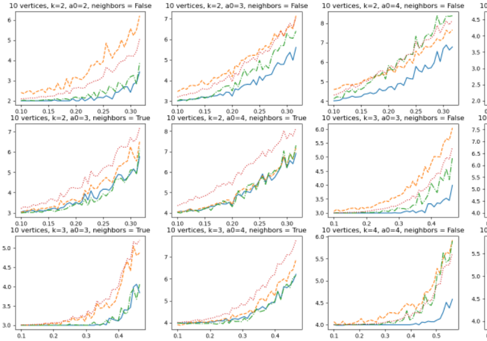
- When $| \quad _0| < g_i$, there is little to no sprea
- When $| \quad _0| > g_i$, we see the phase transiti $|\langle \quad _0\rangle|$ gets larger
- So far, it looks like percolation in the SBM performs accordingly to defined bounds!

An algorithmic approach

An attempt to choose $_0$ based on degree of v
di erent ways. We hoped to minimize the kno
edges between active and inactive vertices

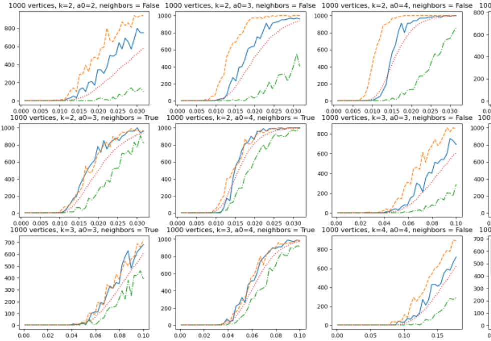
- There are two simple cases - picking the l
vertices, and picking the smallest degree v
- Can either also choose to infect the neigh
chosen vertices or not
- Looked into our own algorithm that choo
with "greatest degree less than remaining

Algorithms, 10 vertices



The algorithms on random graphs of 1

Algorithms, 1000 vertices



The algorithms on random graphs of 100

Results

Empirically, choosing the vertices with smallest
considering neighbors is the best algorithm for
activation of a random graph

- Considered the same p-values as in the SE
- 10 vertex case is unique, potentially due t
  - Higher variance
  - Both $|\ _0|$ and k being close to the numb

Next Steps

We plan to continue looking into the algorithm
better understand the probability behind why c
algorithms perform as they do, and how likely
perform within a certain threshold.

Questions

Thank you!
Any questions?