Park Slope or Bed-Stuy?: Gentrification on Various Network Structures

Erin Masatsugu and Emily Wang

March 2016

1 Project Description

Major metropolitan areas such as the San Francisco Bay Area and New York City have recently undergone the process of gentrification, in which certain previously "undesirable" neighborhoods of the city become increasingly popular and, more notably, extremely expensive. This process can be driven by any number of factors: development can revitalize a previously quiet region, or a major company can establish its headquarters in a neighborhood, resulting in an influx of employees seeking to live nearby. Once the process of gentrification has begun, the flow of people into the new neighborhood is often self-propagating: as an individual's social circle begins to move into this new neighborhood, it becomes an increasingly convenient and desirable place for this individual to live. Moreover, a large number of friends living in a new neighborhood may signal a neighborhood "on the rise," and local information may consequently signify a desirable global property such as monetary value.

We characterize this process (in its simplest formulation) as follows: consider a network $G = (V, E)$ in which nodes in V represent individuals and edges in E represent social ties. Let a node i's friends (individuals in the same neighborhood as i will be called "friends" in order to avoid confusion between the idea of neighbors sharing ties and nodes occupying the same neighborhood), be represented by the expression $N(i)$. At the beginning of the simulation, agents are randomly assigned to one of two neighborhoods, A and B. An individual must decide, at each time step, whether to live in neighborhood A or B, and this decision is based off of the neighborhoods in which its neighbors currently reside. More formally, for node i with some constants k and n,

$$Pr(i \in A) = \frac{(k + |j \in N(i) \cap j \in A|)^n}{(k + |j \in N(i)|)^n}$$

Note that this model resembles the model presented by Becker et al. for *I. humilis* path selection, with the proportion of neighbors occupying a certain neighborhood replacing pheromone concentration. This decision making process, where neighborhood selections are propagated across social ties, is a slight variation on the DeGroot model in opinion dynamics, where edges between nodes represent entries in the "trust matrix." However, in our project, individuals will hold a discrete neighborhood choice at each time step rather than a distribution over neighborhood choices, and this selection is made stochastically rather than deterministically. The DeGroot model is a model of social learning and opinion dynamics and falls under the broader subset of non-Bayesian models of opinion formation. Equilibrium behavior in the DeGroot model on a non-primitive graph involves either complete convergence (all agents holding the same "neighborhood choice" or opinion) or periodic fluctuation; when we examine equilibrium behavior in our sample networks, it may be interesting to test whether the ratio of residents living in neighborhood A versus neighborhood B can remain near-constant near the end of our simulations.

We wish to simulate this process on Erdos-Renyi networks, small world networks, and preferential attachment (scale free) networks. Does equilibrium behavior exist, and if so, how does it differ depending on initial conditions (the proportion of nodes assigned to each neighborhood and the centrality of nodes assigned to each neighborhood) and network structure? We will study how initial conditions, network structure, and the choice of constants k and n affect the eventual proportion of nodes living in each neighborhood. We will also analyze the centrality of the nodes ultimately residing in each neighborhood and whether neighborhood choice is adopted by clusters of nodes.

In addition to running simulations on our initial model, we will extend the model to include budgets and price. In this modified model, each individual i has some amount of money m_i. At each timestep t, a house in neighborhood A has some value $a(t)$ and a house in neighborhood B has some value $b(t)$. If an individual currently residing in neighborhood B wishes to switch to neighborhood A, this switch will result in an updated budget $m_i = m_i + b(t) - a(t)$ at timestep $t + 1$. Similarly, a switch from neighborhood A to neighborhood B will result in an updated budget $m_i = m_i + a(t) - b(t)$. Individuals cannot switch neighborhoods when they cannot afford to do so, i.e. the switch would result in a negative budget. The value functions $a(t)$ and $b(t)$ can be defined solely as a function of time (such that the price of neighborhood a may gradually increase and the price of neighborhood b may gradually decrease), or the value functions can also take into account some measure of demand, such as the number of individuals currently residing in the neighborhood. We will study how the introduction of prices and budgets affects equilibrium behavior (and in cases where equilibrium initially did not occur, if the introduction of prices forced equilibrium behavior).

Another important extension involves the incorporation of more details about the complex process of neighborhood choice and population flow. What happens when we allow the population to be heterogeneous, and the constants k and n that define the neighborhood selection rule differ between individuals? This variation can capture the difference between the "early adopters" eager to seek out new neighborhoods and more risk-averse individuals. What happens when, instead of agents observing only the discrete opinion of their friends (what neighborhood their friends live in), they are able to see the probability distribution with which each neighborhood was selected?

Finally, two further extensions—simulating behavior with multiple neighborhoods and adding edge weights to capture differing degrees of closeness in relationships—may also produce interesting results in simulation.

2 Background and Related Work

This project takes inspiration from the work done by Michael Kearns et al. in examining the role of network structures on graph-coloring, a coordination game whose goal is social-differentiation. Kearns inspected different types of small-world networks (simple cycles, 5-chord cycles and 20-chord cycles), a network with a few highly-connected "hubs" (leader cycles), and preferential attachment networks. The group found that the network strongly influenced the outcome of these coloring games and made some interesting observations. For example, they found that it is more difficult to color a preferential attachment graph than a small-world graph, and that within small-world graphs, it is more difficult to color graphs with larger network average distances (the average shortest distance over all pairs of vertices). In preferential attachment graphs, increasing the connectivity of the graph (allowing agents to see a global state of the network, for example) also made the coloring problem more difficult [3].

For this project, as mentioned, we focus on the role of network structures on the outcomes of gentrification. Gentrification is not a coordination game; there is no ideal outcome, except to reach a stable equilibrium in which agents are no longer switching between neighborhoods. Rather, gentrification is an example of *opinion dynamics*, a field of study concerned with a "fragmentation of the patterns of opinions" [1]. Every agent has a certain opinion on a matter – which political party you align with, what field of study you want to pursue, or, in our case, which neighborhood you want to live in. This opinion is influenced, to a certain degree, by those of your neighbors. Hegselmann and Krause have done much work in this field, have defined three main outcomes of opinion dynamics, three main ways in which the opinions of a population might reach equilibrium: consensus, polarization or fragmentation, where consensus and polarization are specific types of fragmentation in which the entire population adopts the same opinion or the population is split into exactly 2 opinions, respectively.

The most classic model of opinion dynamics is one in which an agent's opinion is the weighted average of the opinions of its neighbors, such that if $x(t)$ is an agent's opinion at time t, then $x(t+1) = Ax(t)$ where A is some fixed stochastic matrix. Note that a stochastic matrix is one in which each row sums to 1. This is formally known as the DeGroot Model of belief updating [2]. The intuition behind this model is that agents will imitate their neighbors. With this classic model, Hegselmann was able to analytically prove that if at some point in time two agents both put a positive weight on the third agent, the model will completely reach a consensus in which all agents share the same opinion. Likewise, if every agent disregards the opinion of all of its neighbors, the model will never reach consensus. They also show that if the subgroups of agents are all primitive (the population can be divided into groups where agents in that group only value the opinions of each other), then the model will reach a "stable opinion pattern" – while consensus might not be reached, all agents will eventually reach a stable opinion.

A more expressive model of opinion dynamics is one in which A is neither constant nor time-dependent. Rather, the structure changes as the opinions of an agent's neighbors change. For example, Hegselman introduced a "bounded confidence" model in which each agent will only take into account neighbor opinions that differ by at most a certain value ϵ. The most interesting conclusions drawn with this particular model were based on simulations; the simulation began with 625 randomly distributed opinions, and ran for 15 different update steps. With 38 opinions remained when $\epsilon = .01$ (normal fragmentation), 2 opinions remained when $\epsilon = .15$ (polarization) and 1 opinion remained when $\epsilon = .25$ (consensus). Note that this simulation was run on a completely connected network. With regards to this paper, the introduction of every agent having a budget and each neighborhood having a cost associated with moving there yields a non-linear model; because as more people move to a neighborhood, that neighborhood becomes more expensive, thus changing the likelihood with which you will move there.

It is also important to quickly explain the different networks that we will inspect: small-world networks, Erdos Renyi (random) networks, and preferential attachment networks. *Small-world networks* are networks in which most nodes aren't neighbors with eachother, but the networks are rich in short paths, and these short paths chain together to connect otherwise "distant" nodes. Social networks are a common example of small-world networks; the common act of "Facebook stalking" in which one finds himself or herself on a friend's friend's cousin's brother's friend's Facebook, demonstrates how people are effective at using short connections to reach socially distant others. *Erdos Renyi networks*, otherwise known as *random networks*, are networks that are constructed randomly. There are two main models for generating such graphs: randomly selecting from the set of all graphs with n vertices and m edges, or starting with n nodes and adding m edges, where each edge has an equal chance of being selected. *Preferential attachment networks* are networks in which the probability of an edge being drawn from node a is a function of the number of edges already adjacent to node b – that is, those edges with higher degrees are more likely to accumulate more adjacent edges. This model lends itself well to the development of "hubs," which are nodes with many adjacent edges that are can be very influential in opinion dynamics. For example, if the opinion is which political candidate you'll vote for, a celebrity or even a news station or newspaper might serve as a hub.

3 Project Timeline

Our proposed timeline follows the schedule below:

- During week 1, we want to write code to generate random networks with specified structures (small world, preferential attachment, Erdos Renyi) and specified size. Moreover, we want to write the code that will simulate agent behavior according to the ant-bridge rule. At this point, we will more precisely decide the various details of the simulation process. In particular, we hope to pinpoint how many simulations to run for each condition, conditions for when to consider a simulation “complete,” how large the randomized networks should be, the density of each network, and the parameters for the initial random conditions. Note that we may alter these details as we begin the project.

- During week 2, we will run simulations with random initial conditions. If initial conditions seem to dramatically affect outcomes, then we will experiment further with initial conditions in order to get a sense of what initial conditions produce what types of outcomes. If simulation outcomes do not appear to vary significantly based on initial conditions, then we will extend the simulation to include budgets and changing neighborhood prices.
- Hopefully, if all goes smoothly, week 3 will involve simulating heterogeneous agents with different behavior (some more likely to move into new neighborhoods, some less likely).
- If all goes well up to this point, then we will conclude the project by exploring further extensions such as extending the problem to multiple neighborhoods and adding edges with differing weights to the network.
- Throughout the fourth and fifth weeks, we will also produce the final writeup.

4 Further Work

We would first be interested in completing the possible extensions that we listed in the project timeline that were not explored due to time constraints.

If we complete all of the elements listed in the proposed timeline, we would also be interested in seeing how the results extend to core-periphery network structures. Core-periphery networks, which consist of a densely connected “core” and some number of “periphery” nodes, have been demonstrated to mirror the structure of real-life social networks when partitioned by socioeconomic status, with middle-class and upper-class individuals occupying the core, and lower-class individuals existing in the less connected periphery. In particular, we would like to study how our proposed simulation of gentrification with budgets and moving prices functions in a core-periphery network, and whether eventual outcomes mimic the real-life socioeconomic segregation of gentrified neighborhoods.

However, because models of core-periphery networks are less well-defined than small world, preferential attachment, and Erdos-Renyi networks, this extension would first require the development of a suitable core-periphery network and likely lies outside of the scope of this final project.

Another limitation of our current formulation of the problem is the fact that our simulations run on a fixed network. Because gentrification can also be caused (and exacerbated) by an influx of new entrants to a region, social networks would likely grow overtime. Extending the proposed model to account for changing networks would be another interesting direction to explore in future iterations of the project.

References

[1] Hegselmann, Rainer and Ulrich Krause, ”Opinion Dynamics and Bounded Confidence Models, Analysis, and Simulation,” *Journal of Artificial Societies and Social Simulation*. vol.5, no.3, 1993.

[2] Acemoglu, Daron and Asuman Ozdaglar, ”Opinion Dynamics and Learning in Social Networks,” *Dynamic Games and Applications*, vol.1, no.1, p.3-49, 2011.

[3] Kearns, Michael, Siddharth Suri and Nick Montfort, ”An Experimental Study of the Coloring Problem on Human Subject Networks,” *Science*, vol.313, p.824-27, 2006.

[4] Easley, David and Jon Kleinberg, *Networks, Crowds, and Markets: Reasoning About a Highly Connected World*. Cambridge University Press, 2010.