1 Overview

In this course we will try to develop a mathematical understanding for the way in which information can be transmitted through social networks. For this purpose we will introduce mathematical models of social networks, communication, and human behavior, and then develop algorithms and incentive-based mechanisms that have provable guarantees. The premise of the course can summarized as follows.

*Human interactions follow mathematical patterns that can be leveraged by algorithms and mechanisms for acquiring, disseminating, and learning information.*

In the course we will broadly focus on three main topics: acquisition, dissemination, and learning of information. In the first part of information acquisition we will be interested in the question of how to design algorithms that find information using a social network. In the second part, we will be interested in how to design algorithms that can trigger large information cascades. In the last part, we will study mathematical models of learning that attempt to capture the way in which people update their beliefs based on their knowledge of beliefs of others in the network.

2 Information Acquisition Through Social Networks

In the first part of the course, we will be interested in understanding how one can acquire information using a social network: given a piece of information held by an individual, how can we design algorithms that find this information efficiently?

In 1967 while a graduate student at Harvard, Stanley Milgram performed an experiment in an attempt to verify the theory of “small world”. The theory of “small world” argues that every two people are separated by a short chain of acquaintances. To verify this theory, Milgram selected a set of people at random in the Midwest and asked them to forward an envelope to a specific person they did not know in Massachusetts. After some trial and error, Milgram launched an experiment where a good fraction of the envelopes reached their destination, in a mean number of steps of six, which coined the term “six-degrees of separation”.

What kind of graphs allow for these short paths to exist? What kind of graphs allow for people to be able to send information to strangers through acquaintances so efficiently? We will attempt to answer these questions in the next few lectures.
3 Graphs as Models of “Small World” Networks

Throughout this course, we will model social networks as graphs. A graph \( G \) consists of a set of nodes (or vertices) \( V \) and a set of edges \( E \). If \( u, v \in V \) and \( (u, v) \in E \), then they are connected by an edge. For an example of a graph, the Facebook network can be drawn as a graph where people are the nodes, and two people are connected by an edge if they are “friends”. We will sometime introduce weights on the graph, either on edges or nodes.

We will be interested in describing models that are somewhat descriptive of social networks. In particular, we would like to understand the topological properties that explain phenomena like the one observed in Milgram’s experiment.

As a first attempt for a social network model, we can consider the clique: a graph in which every node is connected to every other node in the graph. This model fails to capture our intuitive understanding of a social network, as the degree of every node is very large. We therefore wish to explore models in which every node has constant degree.

As a second attempt, we can consider trees with constant degree. This model however fails to capture another intuitive feature of social networks and it is that some of our friends are likely to be friends themselves. In their seminal paper, Watts and Strogatz [?] formalized this using the following definition:

**Definition 1 (Clustering coefficient).** Given a graph \( G = (V, E) \), the clustering coefficient of a vertex \( v \), denoted \( C(v) \) is the fraction, over all pairs of neighbors of \( v \), of those pairs who are neighbors of each other. Formally,

\[
C(v) = \frac{|\{(u, w) \in E : u, w \in N(v)\}|}{\binom{d(v)}{2}}.
\]

Intuitively, if you were to pick two of your friends uniformly at random, then the probability that they are also friends is your clustering coefficient. We will define the clustering coefficient of a graph to be the average clustering coefficient of its nodes.

For real-world social networks, we expect reasonably high clustering coefficients; ideally constant (that is, independent of the graph size). Note that trees and cycles have a clustering coefficient of zero, and so they also are not descriptive of the underlying structure of social networks.

4 Constructing a Small World model

We will now construct a model for small world networks which has two main building blocks: ring lattices and \( d \)-regular random graphs.

4.1 Ring lattices

A ring lattice is a graph which is obtained by taking a cycle and connecting each vertex to its neighbors two “hops” away, giving a 4-regular graph; We can generalize the definition to other even constants greater than 4 (connecting each vertex to its neighbors three “hops” away, giving a 6-regular graph, and so on).
In the problem set we will show that the clustering coefficient of a $d$-regular ring lattice is $\frac{3}{4} \left(\frac{d}{d-2}\right)$. For a 4-regular lattice it is easy to see that the clustering coefficient is $1/2$. The diameter of a $d$-regular lattice, however, is $\Theta \left(\frac{n}{d}\right)$. So the ring lattice has a high clustering coefficient, which nicely captures that aspect of social networks. For the purposes of explaining short distances the ring lattice fails since it has a large diameter.

4.2 Expansion and random graphs

We’d like to formalize a graph property that will guarantee small diameter (short distances).

**Definition 2.** The expansion of a graph $G = (V,E)$ is the minimum, over all cuts we can make (dividing the graph in two pieces), of the number of edges crossing the cut divided by the number of vertices in the smaller half of the cut. Formally, it is

$$\alpha = \min_{S \subseteq V, 1 \leq |S| \leq \frac{n}{2}} \frac{|e(S)|}{|S|}$$

where $e(S)$ is the set of edges leaving the set of nodes $S$.

**Theorem 3.** Suppose the graph $G$ is $d$-regular, for some $d \geq 3$ and has expansion $\alpha$. Then, the diameter of $G$ is $O \left(\frac{d}{\alpha} \log n\right)$.

**Proof.** We will show that any two vertices $s$ and $t$ are a distance at most $O \left(\frac{d}{\alpha} \log n\right)$ apart. Let $S_j$ be the set of vertices reachable from $s$ in at most $j$ steps. (We can think of $S_j$ as being formed by a breadth-first-search that starts at $s$.) Suppose that $|S_j| \leq \frac{n}{2}$.

Because $G$ is an $\alpha$-expander, there are at least $\alpha |S_j|$ edges leaving $S_j$. Consider all the vertices outside of $S_j$ that these edges touch (as they are the ones who will be added to $S_j$ to get $S_{j+1}$).
Each such vertex “uses up” at most $d$ of the edges (because $G$ is $d$-regular), so we will add at least $\frac{\alpha}{d}|S_j|$ new vertices to $S_j$ to get $S_{j+1}$. That is,

$$|S_{j+1}| \geq \left(1 + \frac{\alpha}{d}\right)|S_j|.$$ 

Because $S_0 = \{s\}$, this gives that

$$|S_j| \geq \left(1 + \frac{\alpha}{d}\right)^j.$$ 

Now pick $j = \frac{d}{\alpha} \log n$. Then we have

$$|S_{\frac{d}{\alpha} \log n}| \geq \left(1 + \frac{\alpha}{d}\right)^{\frac{d}{\alpha} \log n}.$$ 

Use the fact that $(1 + \frac{1}{k})^k \geq 2$ for $k \geq 1$ to get that $S_{\frac{d}{\alpha} \log n} \geq n$. \footnote{This assumes that $k = \frac{d}{\alpha}$ is bigger than 1. But if it is less, then we can fix the argument by taking $j = c\frac{d}{\alpha} \log n$ for some large $c$ in order to bring $(1 + \frac{\alpha}{d})^j$ back up to $\log n$. The constant $c$ only depends on $\frac{d}{\alpha}$, not $n$, so it gets eaten up by the big-O, so it doesn’t affect the conclusion.}

Therefore, the size of $S_j$ reaches at least $\frac{n}{2}$ before this point, i.e. before $j = \frac{d}{\alpha} \log n$. Now, by the exact same reasoning, if we start at $t$ and consider the sets $T_j$, we find that the size of $T_j$ reaches at least $\frac{n}{2}$ before $j = \frac{d}{\alpha} \log n$. But then $T_{\frac{d}{\alpha} \log n}$ and $S_{\frac{d}{\alpha} \log n}$ must have some vertex in common (since both are larger than $\frac{n}{2}$). That means there is a path from $s$ to $t$ of length at most $2\frac{d}{\alpha} \log n$, because we can go from $s$ to this common vertex in at most $\frac{d}{\alpha} \log n$ steps, and then similarly to get from this vertex to $t$.

So the above theorem implies that for $d \geq 3$, a $d$-regular graph with constant expansion $\alpha$ gives us the “small world” property that we want: It has small (logarithmic) diameter. So how do create $d$-regular graphs with good expansion? Using randomness, we can construct expanders quite easily: For $d \geq 3$, it turns out that $d$-random graphs have expansion that is a function of $d$ only, not of the number of vertices (we’ll skip the proof).\footnote{A $d$-regular random graph is a graph where all nodes have degree $d$ and every two nodes have the same likelihood to be connected. That is, fix some number of vertices $n$, consider all graphs on $n$ vertices that are $d$-regular, and pick one of these uniformly at random. According to the fact we stated, with very high probability, it will have constant expansion $\alpha$.} While $d$-regular random graphs have small diameter, they have low clustering coefficients (about $\frac{d}{n}$). So we have two types of graphs with different weaknesses.

**A model for small world networks.** To get the best of both worlds we can simply combine the ring-lattice with a $d$-regular random graph, for some constant $d \geq 3$. By that we mean that we’ll take a $d$-regular random graph and add the ring-lattice edges to the nodes of the random graph. The result will be a graph with a constant degree that has both short distances and high clustering coefficient.

**The Watts-Strogatz model.** The above construction is a simple variant of a model suggested by Watts and Strogatz \cite{watts1998collective}. In the Watts-Strogatz model we take the ring lattice and rewire every edge to a random node in the graph with some probability $p$. For $p = 0$ we have a ring-lattice,
and $p = 1$ a regular random graph. Watts and Strogatz showed that there is a region of $p$ s.t. the model has both short distances and a high clustering coefficient.

In summary:

<table>
<thead>
<tr>
<th>Model</th>
<th>Diameter</th>
<th>Clustering Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>ring lattice</td>
<td>bad (too big)</td>
<td>good (constant)</td>
</tr>
<tr>
<td>$d$-regular random</td>
<td>good (logarithmic)</td>
<td>bad (too small)</td>
</tr>
<tr>
<td>Watts-Strogatz</td>
<td>good</td>
<td>good</td>
</tr>
</tbody>
</table>