Social Network Analysis

Social Network Influence

Course Outline

- Graph Theory and Social Networks
- Visualizing Social Networks
- Game Theory
- Information Networks and the World Wide Web
- **Network Dynamics**
- Applications of SNA in various domains

Influence in a Social Network

- Information Cascade
- Epidemics
- Application – Viral Marketing
- Application – Churn Prediction

Following the Crowd

- When people are connected by a network, it becomes possible for them to influence each other's behavior and decisions
  - Restaurants to visit
  - Products to buy
- Information cascade has the potential to occur when people make decisions sequentially, with later people watching the actions of earlier people, and from these actions inferring something about what the earlier people know
  - Not mindless imitation. Rather, it is the result of drawing rational inferences from limited information
- Information cascades may be at least part of the explanation for many types of imitation in social settings.
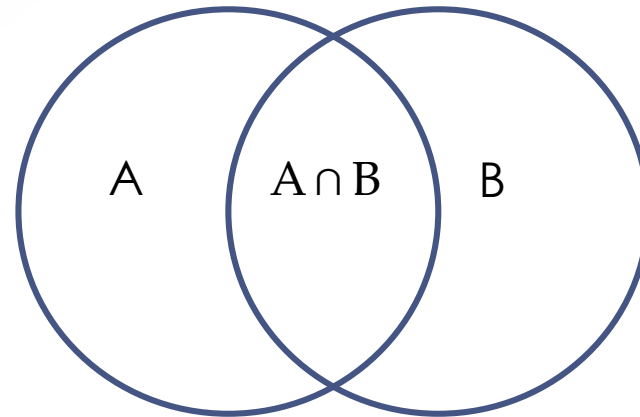
Simple Herding Experiment

- Experiment
  - There are 3 marbles of red or blue color
  - 2 possible scenarios: Majority-blue or Majority-red
  - There is a decision to be made by each student
  - Students make the decision sequentially, and each student can observe the choices made by those who acted earlier.
  - Each student has some private information that helps guide their decision (they can pick up a marble)

- Student Decision
  - First Student: Choose based on marble
  - Second Student: Choose based on marble
  - Third Student: Go by majority: If first 2 choose blue, will choose blue even if own marble red
  - Fourth Student: Follows majority

Bayes Rule

A   A ∩ B   B

- Pr(A | B) = Pr(A ∩ B)/Pr(B)
- Pr(B | A) = Pr(B ∩ A)/Pr(A) = Pr(A ∩ B)/Pr(A)
- Pr(A | B).Pr(B) = Pr(A ∩ B) = Pr(B | A).Pr(A)
- Pr(A | B)= Pr(A).Pr(B | A)/Pr(B)
- Pr(A): Prior probability of A
- Pr(A | B): Posterior probability of A given B

Bayes Rule in Herding Experiment

- $\Pr[\text{majority-blue}] = \Pr[\text{majority-red}] = \frac{1}{2}$
- $\Pr[\text{blue} \mid \text{majority-blue}] = \Pr[\text{red} \mid \text{majority-red}] = \frac{2}{3}$
- 1st student
  - $\Pr[\text{blue}] = \Pr[\text{majority-blue}] \cdot \Pr[\text{blue} \mid \text{majority-blue}] +$
    $\Pr[\text{majority-red}] \cdot \Pr[\text{blue} \mid \text{majority-red}]$
    $= \frac{1}{2} \cdot \frac{2}{3} + \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{2}$
  - $\Pr[\text{majority-blue} \mid \text{blue}]$
    $= \Pr[\text{majority-blue}] \cdot \Pr[\text{blue} \mid \text{majority-blue}] / \Pr[\text{blue}]$
    $= \frac{1}{2} \cdot \frac{2}{3} / \frac{1}{2} = \frac{2}{3}$

Bayes Rule in Herding Experiment

- 3rd student

$$\Pr\left[blue,\ blue,\ red\ |\ majority\text{-}blue\right] = \frac{2}{3} \cdot \frac{2}{3} \cdot \frac{1}{3} = \frac{4}{27}$$

$$
\begin{aligned}
\Pr\left[blue,\ blue,\ red\right] &= \Pr\left[majority\text{-}blue\right] \cdot \Pr\left[blue,\ blue,\ red\ |\ majority\text{-}blue\right] + \\
&\quad \Pr\left[majority\text{-}red\right] \cdot \Pr\left[blue,\ blue,\ red\ |\ majority\text{-}red\right] \\
&= \frac{1}{2} \cdot \frac{2}{3} \cdot \frac{2}{3} \cdot \frac{1}{3} + \frac{1}{2} \cdot \frac{1}{3} \cdot \frac{1}{3} \cdot \frac{2}{3} = \frac{6}{54} = \frac{1}{9}.
\end{aligned}
$$

$$\Pr\left[majority\text{-}blue\ |\ blue,\ blue,\ red\right] = \frac{\Pr\left[majority\text{-}blue\right] \cdot \Pr\left[blue,\ blue,\ red\ |\ majority\text{-}blue\right]}{\Pr\left[blue,\ blue,\ red\right]}$$

$$\Pr\left[majority\text{-}blue\ |\ blue,\ blue,\ red\right] = \frac{\frac{4}{27} \cdot \frac{1}{2}}{\frac{1}{9}} = \frac{2}{3}$$

A Simple General Cascade Model

- ❖ Consider a group of people (numbered 1,2,3, . . .) who will sequentially make decisions
  - ▪ Decision as a choice between accepting or rejecting some option
- ❖ States of the world
  - ▪ At the start we assume that by some initial random event the world is randomly placed into one of two possible states:
    - • A state in which the option is actually a good idea (G)
    - • A state in which the option is actually a bad idea (B)
  - ▪ $Pr(G) = p$, $Pr(B)=1-p$
- ❖ Payoff
  - ▪ Each individual receives a payoff based on her decision to accept or reject the option.
  - ▪ If the individual chooses to reject the option, she receives a payoff of 0.
  - ▪ If the option is a good idea, then the payoff obtained from accepting it is a positive number $v_g>0$
  - ▪ If the option is a bad idea, then the payoff is a negative number $v_b< 0$.
  - ▪ The expected payoff from accepting in the absence of other information is equal to 0
  - ▪ $v_g(p)+v_b(1-p) = 0$
- ❖ Signals
  - ▪ Before any decisions are made, each individual gets a private signal that provides information about whether accepting is a good idea or a bad idea
  - ▪ A high signal (H) suggesting that accepting is a good idea
  - ▪ A low signal (L) suggesting that accepting is a bad idea.
  - ▪ $Pr(H|G) = q > 0.5$
  - ▪ $Pr(L|G) = 1-q < 0.5$
  - ▪ $Pr(L|B) = q > 0.5$
  - ▪ $Pr(H|B) = 1-q < 0.5$

A Simple General Cascade Model

❖ **Individual decision**

$$
\begin{aligned}
\Pr[G \mid H] &= \frac{\Pr[G] \cdot \Pr[H \mid G]}{\Pr[H]} \\
&= \frac{\Pr[G] \cdot \Pr[H \mid G]}{\Pr[G] \cdot \Pr[H \mid G] + \Pr[B] \cdot \Pr[H \mid B]} \\
&= \frac{pq}{pq + (1-p)(1-q)} \\
&> p,
\end{aligned}
$$

since $\quad pq + (1-p)(1-q) < pq + (1-p)q = q$

❖ **Multiple signals**

- A sequence S of independently generated signals consisting of a high signals and b low signals
- Pr[G|S] > Pr[G] when a>b
- Pr[G|S] < Pr[G] when a<b
- Pr[G|S] = Pr[G] when a=b
- With a sequence of signals, individuals can decide according to a majority-vote over the signals they receive

A Simple General Cascade Model

$$\begin{aligned} \Pr[S] &= \Pr[G] \cdot \Pr[S \mid G] + \Pr[B] \cdot \Pr[S \mid B] \\ &= pq^a(1-q)^b + (1-p)(1-q)^a q^b. \end{aligned}$$

$$\Pr[G \mid S] = \frac{\Pr[G] \cdot \Pr[S \mid G]}{\Pr[S]}$$

$$\Pr[G \mid S] = \frac{pq^a(1-q)^b}{pq^a(1-q)^b + (1-p)(1-q)^a q^b}$$

- If we replace 2nd term of denominator by $(1-p)q^a(1-q)^b$, denominator becomes $q^a(1-q)^b$.
- Then $\Pr(G \mid S) = p$
- If $a > b$ then this replacement makes the denominator larger, since $q > 1/2$ and we now have more factors of q and fewer factors of $1-q$ Since the denominator gets larger, the overall expression gets smaller as it is converted to a value of p and therefore $\Pr[G \mid S] > p = \Pr[G]$
- If $a < b$ $\Pr[G \mid S] < p = \Pr[G]$
- If $a = b$ $\Pr[G \mid S] = p = \Pr[G]$

Sequential Decision Making and Cascades
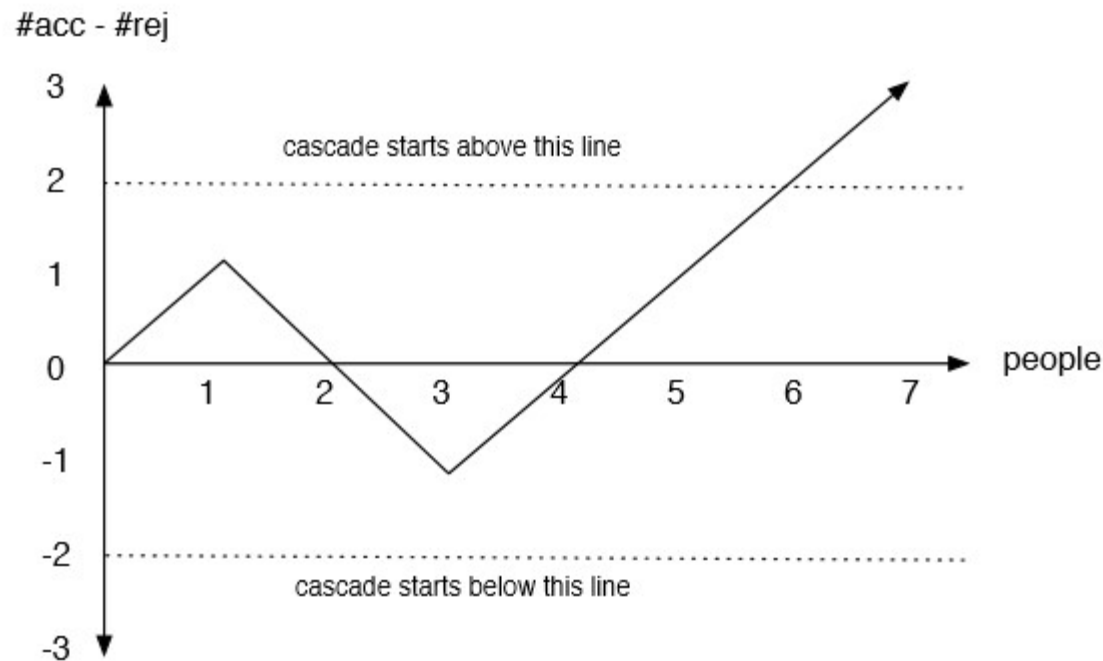
- As long as the number of acceptances differs from the number of rejections by at most one, each person in sequence is simply following their own private signal in deciding what to do.

- Once the number of acceptances differs from the number of rejections by two or more, a cascade takes over, and everyone simply follows the majority decision forever.

- A cascade begins when the difference between number of acceptances and rejections reaches 2

Lessons from Cascades

- **Cascades can be wrong:** If, for example, accepting the option is in fact a bad idea but the first two people happen to get high signals, a cascade of acceptances will start immediately, even though it is the wrong choice for the population.

- **Cascades can be based on very little information**: Since people ignore their private information once a cascade starts, only the pre-cascade information influences the behavior of the population. This means that if a cascade starts relatively quickly in a large population, most of the private information that is collectively available to the population (in the form of private signals to individuals) is not being used.

- **Cascades are fragile**: The previous point, that cascades can be based on relatively little information, makes them easy to start; but it can also make them easy to stop. One manifestation of this is that people who receive slightly superior information can overturn even long-lived cascades.

Exercise

1. In the Herding experiment what is the probability of a wrong information cascade forming after the first two students?

2. You have to choose between 2 mobiles A and B. Due to reputation of the brands the probability of A being better is 2/3. Before buying you read reviews comparing the two. The probability of a review recommending the right product is 3/4.
   a. You choose A after reading a review rating A higher. What is the probability that the you have chosen the right mobile?
   b. You choose B after 2 contradictory reviews. What is the probability that the you have chosen the right mobile?

Exercise - Answers

1. In the Herding experiment what is the probability of a wrong information cascade forming after the first two students?

$$½. 1/3.1/3 + ½. 1/3.1/3 = 1/9$$

2. You have to choose between 2 mobiles A and B. Due to reputation of the brands the probability A is better is 2/3. You read a review comparing the two. The probability of a review recommending the right product is 3/4.

   a. You choose A after reading a review rating A higher. What is the probability that the you have chosen the right mobile?

$$P(G|H) = P(G).P(H|G)/P(H) = 2/3.3/4/(2/3.3/4 + 1/3.1/4) = ½/(1/2 + 1/12)$$

   b. You choose A after 2 contradictory reviews. What is the probability that the you have chosen the right mobile?

$$P(G|S) = P(G).P(S|G)/P(S) = 2/3.3/4.1/4/(2/3.3/4.1/4 + 1/3.1/4.3/4) =2/3$$

References

1. Networks, Crowds, and Markets: Reasoning about a Highly Connected World *Reading Chapter 16*
   https://www.cs.cornell.edu/home/kleinber/networks-book/