Impact of Social Connectivity on Herding Behavior

Deepanshu Vasal

University of Texas, Austin

dvasal@umich.edu

Motivating Example [Shirky, 2008]

- Small protests began in Leipzig, Germany in 1989 with few activists challenging the German Democratic Republic
- Slowly the numbers started rising, and by September 1989, they become too big to be quashed by the government
- The number of protestors grew to 100,000 by October 1989 and to 400,000 by first Monday of November 1989

Motivating Example [Shirky, 2008]

- Small protests began in Leipzig, Germany in 1989 with few activists challenging the German Democratic Republic
- Slowly the numbers started rising, and by September 1989, they become too big to be quashed by the government
- The number of protestors grew to 100,000 by October 1989 and to 400,000 by first Monday of November 1989
- Two days later, Berlin wall was dismantled

Source:bloomberg.com

Motivating Example [Beal et al., 1957]

Source:cshl.edu

- During great depression and till mid 1930s, it was observed that adoption rate of drought-resistant hybrid corn was low
- This was despite the fact that the seeds from newer technology faired significantly better
- After conducting surveys, it was found that farmers valued opinions of their neighbors more than the sales person's word and thus the group as a whole did not adopt the new technology despite it being better.

Background

- **Classical Bayesian learning:** *single decision maker*; makes noisy observations of the state of the system; eventually learns the true state.

[1] [Bikhchandani et al., 1992, Smith and Sørensen, 2000, Acemoglu et al., 2011, Smith and Sørensen, 2013, Le et al., 2014]

Background

- **Classical Bayesian learning:** *single decision maker*; makes noisy observations of the state of the system; eventually learns the true state.
- More interesting scenario is **learning over a social network:** *multiple decision makers*; act *strategically* based on their own private information and actions of previous users;[1].

[1] [Bikhchandani et al., 1992, Smith and Sørensen, 2000, Acemoglu et al., 2011, Smith and Sørensen, 2013, Le et al., 2014]

Background

- **Classical Bayesian learning:** *single decision maker*; makes noisy observations of the state of the system; eventually learns the true state.
- More interesting scenario is **learning over a social network:** *multiple decision makers*; act *strategically* based on their own private information and actions of previous users;[1].
- Occurrences of **informational cascades**: users discard their private information and follow majority action of users before them

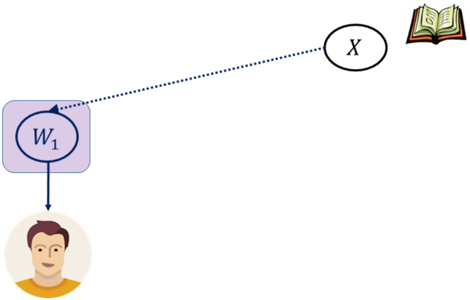
[1][Bikhchandani et al., 1992, Smith and Sørensen, 2000, Acemoglu et al., 2011, Smith and Sørensen, 2013, Le et al., 2014]

Background

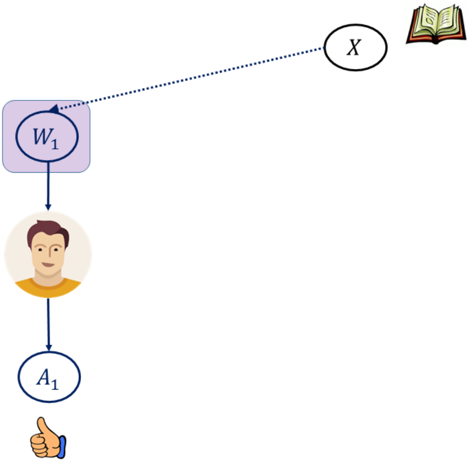
- **Classical Bayesian learning:** *single decision maker*; makes noisy observations of the state of the system; eventually learns the true state.
- More interesting scenario is **learning over a social network:** *multiple decision makers*; act *strategically* based on their own private information and actions of previous users;[1].
- Occurrences of **informational cascades**: users discard their private information and follow majority action of users before them
- An informational cascade could be good or bad for the team/society

[1] [Bikhchandani et al., 1992, Smith and Sørensen, 2000, Acemoglu et al., 2011, Smith and Sørensen, 2013, Le et al., 2014]

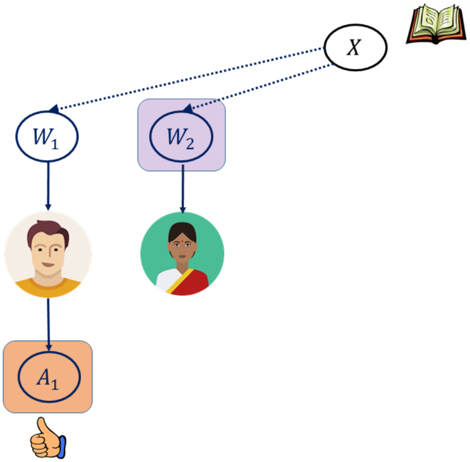
[Bikhchandani et al., 1992]'s sequential model



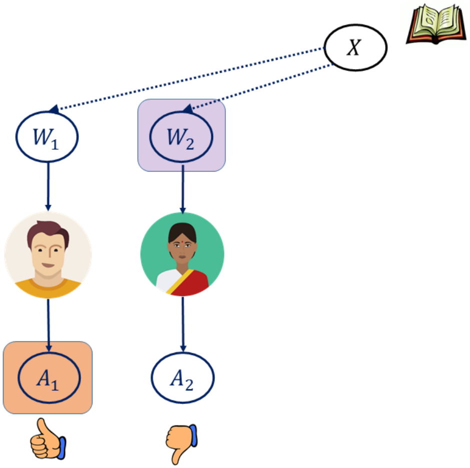
[Bikhchandani et al., 1992]'s sequential model



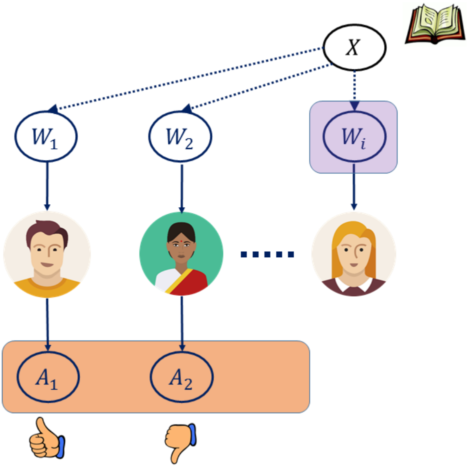
[Bikhchandani et al., 1992]'s sequential model



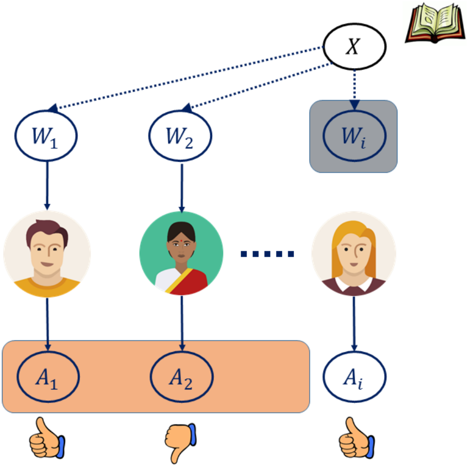
[Bikhchandani et al., 1992]'s sequential model



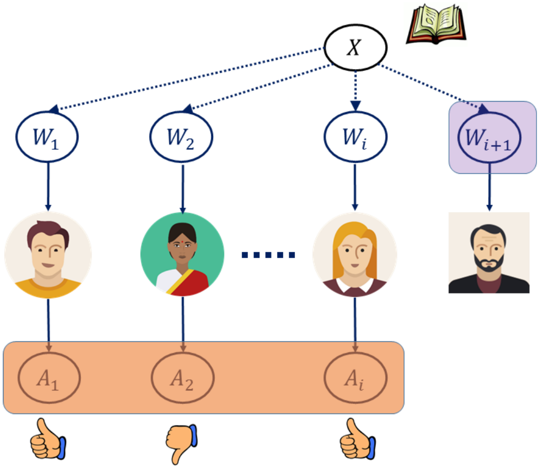
[Bikhchandani et al., 1992]'s sequential model



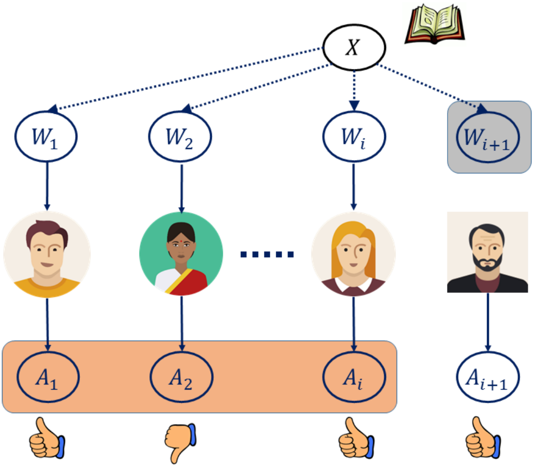
[Bikhchandani et al., 1992]'s sequential model



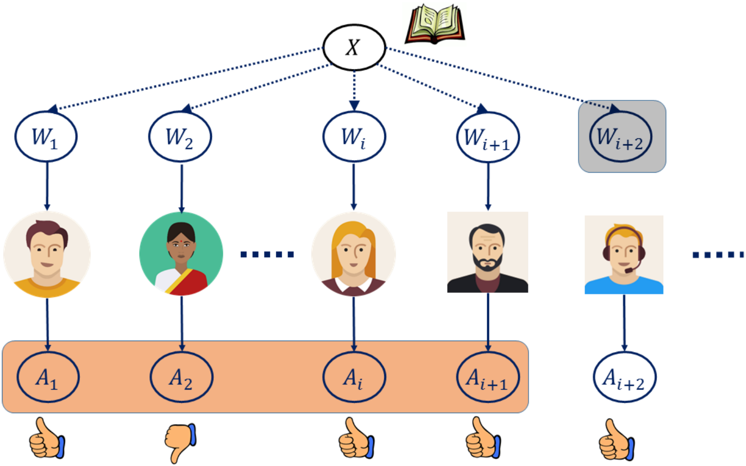
[Bikhchandani et al., 1992]'s sequential model



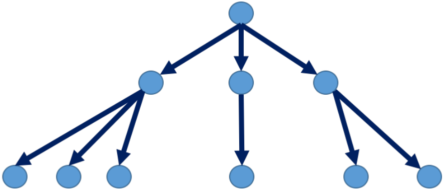
[Bikhchandani et al., 1992]'s sequential model



[Bikhchandani et al., 1992]'s sequential model



Informational cascades in random trees



- Goal: Understand cascading behavior in large random networks
- Locally such networks behave like trees
- Model: *Myopic* players appear on a random tree, observe a private signal and actions of their ancestors
- Using the theory of Multi-type Galton Watson branching process, characterize herding behavior

Model

- Let there be a product whose value is $V \in \{0, 1\}$ such that $P(V = 0) = 1/2$

Model

- Let there be a product whose value is $V \in \{0, 1\}$ such that $P(V = 0) = 1/2$
- Infinite myopic selfish players appear sequentially on nodes of a tree with random degree with given probability generating function (PGF) ϕ_D.

Model

- Let there be a product whose value is $V \in \{0,1\}$ such that $P(V=0)=1/2$
- Infinite myopic selfish players appear sequentially on nodes of a tree with random degree with given probability generating function (PGF) ϕ_D.
- Notation: We denote k^{th} player at stage t of the tree by p_t^k

Model

- Let there be a product whose value is $V \in \{0, 1\}$ such that $P(V = 0) = 1/2$
- Infinite myopic selfish players appear sequentially on nodes of a tree with random degree with given probability generating function (PGF) ϕ_D.
- Notation: We denote k^{th} player at stage t of the tree by p_t^k
- Player p_t^k privately observes $x_t^k \sim Q(\cdot|V = v)$

Model

- Let there be a product whose value is $V \in \{0,1\}$ such that $P(V=0) = 1/2$
- Infinite myopic selfish players appear sequentially on nodes of a tree with random degree with given probability generating function (PGF) ϕ_D.
- Notation: We denote k^{th} player at stage t of the tree by p_t^k
- Player p_t^k privately observes $x_t^k \sim Q(\cdot|V=v)$
- Equivalently she observes $q_t^k := P(V=1|x_t^k)$ where q_t^k has CDF $F^v, v = 1,2$, $supp(F^v) = [\underline{b}, \overline{b}] \subseteq [0,1]$, and $0 \le \underline{b} < 1/2 < \overline{b} \le 1$.

Model

- Let there be a product whose value is $V \in \{0, 1\}$ such that $P(V = 0) = 1/2$
- Infinite myopic selfish players appear sequentially on nodes of a tree with random degree with given probability generating function (PGF) ϕ_D.
- Notation: We denote k^{th} player at stage t of the tree by p_t^k
- Player p_t^k privately observes $x_t^k \sim Q(\cdot|V = v)$
- Equivalently she observes $q_t^k := P(V = 1|x_t^k)$ where q_t^k has CDF $F^v, v = 1, 2$, $supp(F^v) = [\underline{b}, \bar{b}] \subseteq [0, 1]$, and $0 \leq \underline{b} < 1/2 < \bar{b} \leq 1$.
- She observes actions of her ancestors, denoted by $a^{\mathcal{P}(p_t^k)}$, and her own private observation, x_t^k or q_t^k

Model

- Let there be a product whose value is $V \in \{0,1\}$ such that $P(V=0)=1/2$
- Infinite myopic selfish players appear sequentially on nodes of a tree with random degree with given probability generating function (PGF) ϕ_D.
- Notation: We denote k^{th} player at stage t of the tree by p_t^k
- Player p_t^k privately observes $x_t^k \sim Q(\cdot|V=v)$
- Equivalently she observes $q_t^k := P(V=1|x_t^k)$ where q_t^k has CDF $F^v, v=1,2$, $supp(F^v)=[\underline{b},\bar{b}] \subseteq [0,1]$, and $0 \le \underline{b} < 1/2 < \bar{b} \le 1$.
- She observes actions of her ancestors, denoted by $a^{\mathcal{P}(p_t^k)}$, and her own private observation, x_t^k or q_t^k
- She takes an action $a_t^k \in \{0,1\}$ to either buy the product ($a_t^k = 1$) or not buy the product ($a_t^k = 0$) and then she leaves the system

Model

- Let there be a product whose value is $V \in \{0,1\}$ such that $P(V=0)=1/2$
- Infinite myopic selfish players appear sequentially on nodes of a tree with random degree with given probability generating function (PGF) ϕ_D.
- Notation: We denote k^{th} player at stage t of the tree by p_t^k
- Player p_t^k privately observes $x_t^k \sim Q(\cdot|V=v)$
- Equivalently she observes $q_t^k := P(V=1|x_t^k)$ where q_t^k has CDF $F^v, v=1,2$, $supp(F^v)=[\underline{b},\bar{b}] \subseteq [0,1]$, and $0 \le \underline{b} < 1/2 < \bar{b} \le 1$.
- She observes actions of her ancestors, denoted by $a^{\mathcal{P}(p_t^k)}$, and her own private observation, x_t^k or q_t^k
- She takes an action $a_t^k \in \{0,1\}$ to either buy the product ($a_t^k=1$) or not buy the product ($a_t^k=0$) and then she leaves the system
- She gets a reward $R_t^k(a_t^k, v) = \begin{cases} 1 \text{ if } a_t^k=1, v=1, \\ -1 \text{ if } a_t^k=1, v=0, \\ 0 \text{ if } a_t^k=0 \end{cases}$

Perfect Bayesian Equilibrium

- Decision strategy:

$$\mathbb{E}\{R_t^k | a^{\mathcal{P}(p_t^k)}, x_t^k, a_t^k\} = \begin{cases} 2P(V = 1 | a^{\mathcal{P}(p_t^k)}, x_t^k) - 1 \text{ if } a_t^k = 1 \\ 0 \text{ if } a_t^k = 0. \end{cases} \quad (1)$$

[2] where ties are broken in favor of user's private information.

[3] [Smith and Sörensen, 2000, Acemoglu et al., 2011]

Perfect Bayesian Equilibrium

- Decision strategy:

$$\mathbb{E}\{R_t^k | a^{\mathcal{P}(p_t^k)}, x_t^k, a_t^k\} = \begin{cases} 2P(V=1|a^{\mathcal{P}(p_t^k)}, x_t^k) - 1 \text{ if } a_t^k = 1 \\ 0 \text{ if } a_t^k = 0. \end{cases} \quad (1)$$

- User p_t^k takes action $a_t^k = 1$ if $P(V=1|a^{\mathcal{P}(p_t^k)}, x_t^k) > 1/2$. [2]

[2] where ties are broken in favor of user's private information.

[3] [Smith and Sörensen, 2000, Acemoglu et al., 2011]

Perfect Bayesian Equilibrium

- Decision strategy:

$$\mathbb{E}\{R_t^k | a^{\mathcal{P}(p_t^k)}, x_t^k, a_t^k\} = \begin{cases} 2P(V = 1 | a^{\mathcal{P}(p_t^k)}, x_t^k) - 1 \text{ if } a_t^k = 1 \\ 0 \text{ if } a_t^k = 0. \end{cases} \quad (1)$$

- User p_t^k takes action $a_t^k = 1$ if $P(V = 1 | a^{\mathcal{P}(p_t^k)}, x_t^k) > 1/2$. [2]
- Let common belief $\pi_t^k := P(V = 1 | a^{\mathcal{P}(p_t^k)})$ and private belief $q_t^k := P(V = 1 | x_t^k)$. Then [3]

$$P(V = 1 | a^{\mathcal{P}(p_t^k)}, x_t^k) > 1/2 \iff \pi_t^k + q_t^k > 1 \quad (2)$$

[2] where ties are broken in favor of user's private information.
[3] [Smith and Sörensen, 2000, Acemoglu et al., 2011]

Perfect Bayesian Equilibrium

- Decision strategy:

$$\mathbb{E}\{R_t^k | a^{\mathcal{P}(p_t^k)}, x_t^k, a_t^k\} = \begin{cases} 2P(V = 1 | a^{\mathcal{P}(p_t^k)}, x_t^k) - 1 \text{ if } a_t^k = 1 \\ 0 \text{ if } a_t^k = 0. \end{cases} \quad (1)$$

- User p_t^k takes action $a_t^k = 1$ if $P(V = 1 | a^{\mathcal{P}(p_t^k)}, x_t^k) > 1/2$. [2]
- Let common belief $\pi_t^k := P(V = 1 | a^{\mathcal{P}(p_t^k)})$ and private belief $q_t^k := P(V = 1 | x_t^k)$. Then [3]

$$P(V = 1 | a^{\mathcal{P}(p_t^k)}, x_t^k) > 1/2 \iff \pi_t^k + q_t^k > 1 \quad (2)$$

- Then the equilibrium policy (PBE), g, is

$$a_t^k = \begin{cases} 1 \text{ if } q_t^k > 1 - \pi_t^k \text{ or } q_t^k = 1 - \pi_t^k \text{ and } q_t^k \geq 1/2 \\ 0 \text{ if } q_t^k < 1 - \pi_t^k \text{ or } q_t^k = 1 - \pi_t^k \text{ and } q_t^k < 1/2. \end{cases} \quad (3)$$

[2] where ties are broken in favor of user's private information.
[3] [Smith and Sörensen, 2000, Acemoglu et al., 2011]

Common belief update

Lemma

Under the equilibrium policy g, common belief π_t^k is updated as

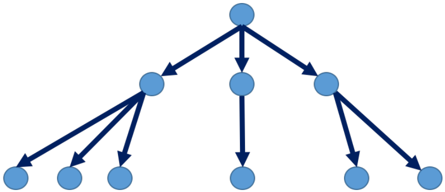
$$\pi'_{t+1} = \begin{cases} \psi_0(\pi_t^k) & \text{if } a_t^k = 0 \\ \psi_1(\pi_t^k) & \text{if } a_t^k = 1. \end{cases} \tag{4}$$

Proof: Bayes' Rule ■

Note: For $(1-\pi_t^k)$ outside the support of F^v, then the update of π_t^k for both actions leads to same belief i.e.

$$\pi'_{t+1} = \psi_a(\pi_t^k) = \pi_t^k \qquad \text{if } (1-\pi_t^k) \notin [\underline{b}, \bar{b}] \tag{5}$$

Preliminaries: Single type Galton-Watson Process



- A Galton-Watson process is defined as a branching process which starts with one node and each node of the tree independently gives birth to D children, where D is a discrete random variable with known probability generating function (PGF) ϕ_D i.e. $\phi_D(s) = \mathbb{E}[s^D]$.
- **Proposition** The probability of extinction is the smallest nonnegative root t of the fixed-point (FP) equation

$$\phi_D(t) = t \tag{6}$$

- If $\mathbb{E}[D] > 1$, then the FP equation has a unique root less than 1. If $\mathbb{E}[D] \leq 1$, then the only root is 1.

Preliminaries: Multi-Type Galton-Watson Process

- A multi-type Galton-Watson process is defined as a branching process where each node of the tree could be of multiple type (finite, countable or uncountable) and gives birth to children of any of the types with a given PGF.
- Suppose every node has a type $x \in [0,1]$ and it gives birth to n children of types $(y_1, \dots, y_n)$ with probability $P_1^{(n)}(dy^n|x)$, where $y_i \in [0,1]$.
- Define probability generating functional

$$G_1(\xi|x) := \sum_{n=0}^{\infty} \int_{X^n} \xi(y_1)\dots\xi(y_n) P_1^{(n)}(dy^n|x) \tag{7}$$

- Then the asymptotic extinction probability, $q(x)$, is given by the minimal non-negative solution of the functional equation [Moyal, 1962]

$$\forall x, \quad \xi(x) = G_1(\xi|x) \tag{8}$$

Information cascades as extinction probability in a branching process

- Let π_t^k be *type* of player p_t^k
- A player of type m, where $1-m \in [\underline{b}, \overline{b}]$, has D children of type $\psi_0(m)$ with probability $F^1(1-m)$, or D children of type $\psi_1(m)$ with probability $(1-F^1(1-m))$
- For $1-m < \underline{b}$, player p_t^k has D children of type of type $\psi_1(m)$ with probability 1, and for $1-m > \overline{b}$, player p_t^k has D children of type of type $\psi_0(m)$ with probability 1
- We define types $1-m \notin [\underline{b}, \overline{b}]$ as "extinct" from the point of view of the branching process

Lemma

Informational cascades are equivalent to extinction probability in the branching process defined above.

Probability of falling into cascade

- Using multi-type Galton-Watson branching process theory the probability of falling into a cascade is given by $q(1/2)$, where $q(\cdot)$ is given by the minimal non-negative solution of the functional equation

$$\xi(x) = G_1(\xi|x) \tag{9}$$

where

$$G_1(\xi|x) = \sum_{n=0}^{\infty} \int_{X^n} \xi(y_1)\ldots\xi(y_n) P_1^{(n)}(dy^n|x) \tag{10a}$$

$$= F^1(1-x)\phi_D(\xi(\psi_0(x))) + (1 - F^1(1-x))\phi_D(\xi(\psi_1(x))). \tag{10b}$$

Special case:BSC with erroneous actions

- Users read actions of their ancestors erroneously through a binary symmetric channel with error probability $\epsilon \in [0, 0.5]$
- Each user observes o_t^k instead of a_t^k where $P(O_t^k \neq A_t^k) = \epsilon$.
- It is shown in [Le et al., 2014] that $z_n := (\#1's - \#'0s)$ (instead of π_t!) is a sufficient statistics for common observation history
- It is also shown that a positive (or negative) cascade occurs when z_n becomes greater (smaller) than k_0 $(-k_0)$, $k_0 := \lfloor \log_{\frac{1-a}{a}} \alpha \rfloor + 1$

Corollary

Let $q := [-q_k, \ldots, q_k]$ where q_i represent the probabilities of falling into a cascade starting from type i. Using (9) the minimal solution of the following fixed-point represents this probability, for $i = \{-k+1, \ldots, k-1\}$,

$$q_i = a\phi_D(q_{i-1}) + \bar{a}\phi_D(q_{i+1}) \qquad (11)$$

and $q_{-k} = q_k = 1$, $a = \epsilon(1-p) + (1-\epsilon)p$, $\alpha = \frac{1-p}{p}$

Probability of falling into cascades for BSC with erroneous actions: d-regular tree

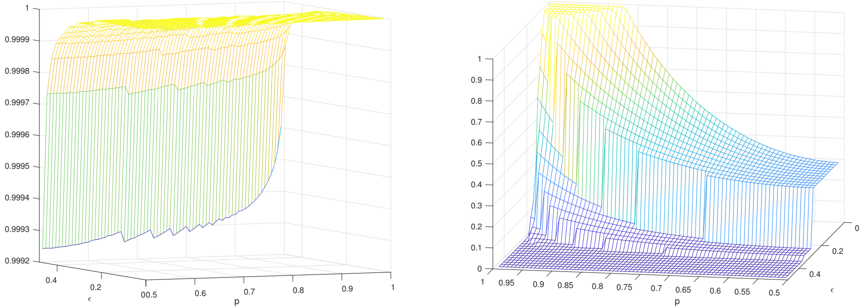


Figure : Probability of tree falling into cascade for d-regular tree with d=1 and d=2 where $0.5 < p < 1$ and $0 < \epsilon < 0.5$. Notice for smaller values of the channel and observation noise, the tree cascades with a higher probability.

Probability of falling into cascades for BSC with erroneous actions: Poisson tree

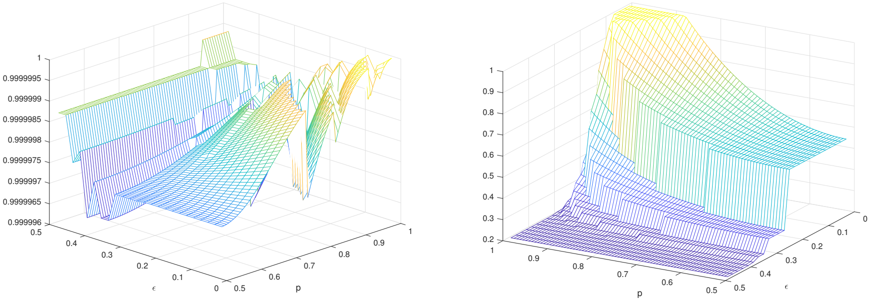


Figure : Probability of tree falling into cascade for tree with degree distribution Poisson (0.2) and Poisson(2)

Special case: BSC with perfectly observed actions

- Actions are perfectly observed
- There exist 5 possible states $\{-2, -1, 0, 1, 2\}$, where ± 2 are absorbing states
- The probability of cascades is given by q_0, where $(q_{-2}, q_{-1}, q_0, q_1, q_2)$ is the minimal solution of the fixed-point equation

$$
\begin{aligned}
q_{-2} &= q_2 = 1 && (12)\\
q_{-1} &= p + \bar{p}\phi_D(q_0) && (13)\\
q_0 &= p\phi_D(q_{-1}) + \bar{p}\phi_D(q_1) && (14)\\
q_1 &= p\phi_D(q_0) + \bar{p}. && (15)
\end{aligned}
$$

- the probability of occurrence of an information cascades is given by the smallest non-negative solution of the fixed-point equation

$$y = p\phi_D(p + \bar{p}\phi_D(y)) + \bar{p}\phi_D(p\phi_D(y) + \bar{p}) \tag{16}$$

Special case: BSC with perfectly observed actions

Corollary

The tree cascades with probability 1 if and only if $\mathbb{E}[D] \leq \frac{1}{\sqrt{2p(1-p)}}$.

Special Case: D=1 a.s. The above condition is satisfied and thus the tree cascades with probability 1. This is an alternate proof of occurrence of informational cascades of [Bikhchandani et al., 1992] for the BSC channel.

Special Case: D=2 a.s. The tree cascades with probability 1 if $\frac{2+\sqrt{3}}{4} < p \leq 1$. For $\frac{1}{2} < p < \frac{2+\sqrt{3}}{4}$, the probability of occurrence of informational cascades is the smallest fixed-point of the following equation,

$$y = p(p + \bar{p}y^2)^2 + \bar{p}(py^2 + \bar{p})^2. \tag{17}$$

Concluding Remarks

- We study occurrence of information cascades on random trees which serve as approximation of large random graphs such as Erdös Rényi graph.
- Using multi-type Galton-Watson branching process, we characterize the probability of tree falling into a cascade
- Our model is a special case of [Acemoglu et al., 2011]. They provide sufficient conditions for "asymptotic learning" whereas we study probability of falling into a cascade.
- Our analysis confirms the observation of [Le et al., 2014] that there is no monotonicity of probability of cascades in channel noise.
- Our results indicate that groups that are less tightly knit, (i.e. have smaller $\mathbb{E}[D]$) (and as a result have lesser diversity of thought) tend to herd more than the groups that have more social connections.

Thank you

References I

Acemoglu, D., Dahleh, M. A., Lobel, I., and Ozdaglar, A. (2011).
Bayesian learning in social networks.
The Review of Economic Studies, 78(4):1201–1236.

Beal, G. M., Bohlen, J. M., et al. (1957).
The diffusion process.
Agricultural Experiment Station, Iowa State College.

Bikhchandani, S., Hirshleifer, D., and Welch, I. (1992).
A theory of fads, fashion, custom, and cultural change as informational cascades.
Journal of Political Economy, 100(5):pp. 992–1026.

Le, T. N., Subramanian, V., and Berry, R. (2014).
The impact of observation and action errors on informational cascades.
In *Decision and Control (CDC), 2014 IEEE 53rd Annual Conference on*, pages 1917–1922.

Moyal, J. (1962).
Multiplicative population chains.
In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, volume 266, pages 518–526. The Royal Society.

References II

Shirky, C. (2008).
Here Comes Everybody: The Power of Organizing Without Organizations.
New York: Penguin Press.

Smith, L. and Sørensen, P. (2000).
Pathological outcomes of observational learning.
Econometrica, 68(2):371–398.

Smith, L. and Sörensen, P. (2000).
Pathological outcomes of observational learning.
Econometrica, 68(2):371–398.

Smith, L. and Sørensen, P. N. (2013).
Rational social learning by random sampling.
Available at SSRN 1138095.