Information Cascades in Social Networks via Dynamic System Analyses

Shao-Lun Huang and Kwang-Cheng Chen
Graduate Institute of Communication Engineering, National Taiwan University, Taipei, Taiwan
Email:{huangntu, ckc}@ntu.edu.tw

***Abstract*—Systematically analyzing the dynamic behaviors of social networks is one of the central topic in understanding the structure of large networks. In particular, the information cascade [1] introduced by Banerjee provides great insights in characterizing the opinion exchanging between network agents. Traditionally studies of information cascades focus on the Bayesian models, which are often difficult to model real world situations. In this paper, we attempt to study the information cascades from a non-Bayesian point of view. In particular, we consider a sequential decision model but with an arbitrary decision rule. We show that the fraction of agents in a network making any specific decision will converge. Thus, the agents in the network reach a sort of consensus with high probability, which allows us to predict the herd behaviors. In addition, we also apply our non-Bayesian model to different network structures, such as ER model and network with communities, in which the affect of information cascades are quantified. Finally, we simulate the decision process for multiple communities, which justifies our proposed model to comprehend real world complex user behaviors and dynamics.**

I. Introduction

Learning and predicting how the decentralized opinions of a large number of individuals are aggregated or cascaded is an important topic in studying the opinion dynamics in a social network. In a social network, it is usually assumed that the agents can acquire knowledges from environmental observables that are generated from the underlying state of world, and then perform social activities or decisions to affect the behaviors of the network, according to the available knowledges as well as agents' decision logics. For instance, in elections, the voters can observe the political platform announced by the candidates, the related contents from the news media, and the political tendency of their neighborhoods. Then, the electors decide who to vote according to their personal experiences and logics. This class of observation and decision processes are particularly suitable for modeling real world phenomenons. However, in practice, the observation and decision processes are dynamically changed through the underlying environmental signals, and the agents are mutually affected with each other by their activities. This causes extraordinary difficulties to thoroughly understand the dynamic structure of the herd behaviors in a social network, and hence prevent the prediction of future network behaviors from possible.

In this paper, our primary purpose is to develop a mathematical social learning model that not only allows analytical analyzations of the dynamic structures in social networks, but also capturing the critical features of the decision logics of human beings. Among literatures on social learning, the pioneer efforts are probably the Bayesian decision model proposed by Banerjee [1] and Bikhchandani, Hirshleifer, and Welch [2]. In Banerjee's work, he considered the model where agents are allowed to observe the decision made by all previous agents and then make their decisions sequentially. In addition, the agents follow the Bayesian strategy as their decision logic. With this simple social network model, Banerjee observes that, while every agent makes their decision rationally (Bayesian), the effect of observing other agents' actions may cause irrational decision results. This effect was termed as the *information cascade* [3]. After [1], the researches that aim to profound the insights of information cascades are primarily focusing on two directions. In [4]-[8], the authors extended the ordinary Bayesian decision model in [1] to more general payoff and cost functions. In addition, the authors in [9]-[11] considers the topology and connecting structure of the network in implementing the sequential decision, in which the distributed opinion aggregations under Bayesian decision rules are investigated.

On the other hand, comparing to the studies of the Bayesian models, the information cascades under general non-Bayesian decision rules have a relatively limited understanding [12]. In a sense, the Bayesian social learning provides a heuristic framework for analyzing the dynamic structure of social networks, which can be served as the natural benchmark for understanding the epidemic of herd behavior. However, the Bayesian model is often difficult to model realistic human decision strategies, since it has been shown in many literatures that the decision logic of human beings tends to follow rather irrational and stochastic mechanisms, e.g., see [13]-[18]. In addition, it was suggested in [19] that when social pressures are great, the decision of herds may act irrationally, namely, make the suboptimal decisions given the available information. These evidences demonstrate that Bayesian models is insufficient to model the decision patterns of human beings. Therefore, while Bayesian social learning often provides clean analytic results, these results may not be suitable for real world phenomenon of dynamics.

Facing these difficulties, in this paper, we propose a new class of non-Bayesian model that addresses the nature of human perceptions. Specifically, we consider the sequential decision process for n agents, where each agent i can first observe the decisions of previous agents and then make a binary decision $x_i \in \{0, 1\}$. Instead of employing the conventional Bayesian decision policy, we introduce here a two-step decision process to model the nature of human decisions. First, we model the observation mechanism of agents as a pre-decision filter $R_i = \frac{1}{i-1} \sum_{j=1}^{i-1} \mathcal{S}_{i,j}(x_j)$, which takes the observables $x_1, \dots, x_{i-1}$ as the input. Here, $\mathcal{S}_{i,j}$'s are some stochastic functions that map $\{0, 1\}$ to $\mathbb{R}$.

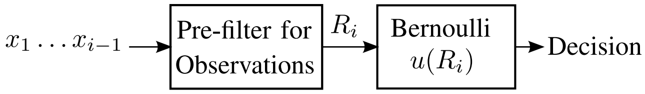


Fig. 1. The human decision model.

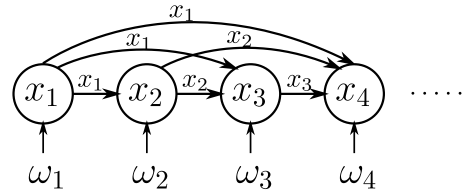


Fig. 2. The sequential decision model, where the links denote the observability of the decisions from previous agents.

In other words, while the agent i is allowed to observe $x_1, \dots, x_{i-1}$, we assume that most important information that is aware by agent i through the human perception system is the filtered observation R_i. This pre-decision filter model is motivated by the psychological fact that when receiving high dimensional signals, people tend to have the selective attention and consider the signals as low dimensional objects that may be thought of as projecting original high dimensional signals to low dimension spaces. After receiving the filter output R_i, the agents then make a randomized binary decision according to this output by their personal decision logics. Rigorously, we can describe this stochastic decision process as a Bernoulli random process with the parameter $u(R_i)$, where u is an utility function that models the decision patterns of agents. Figure 1 illustrates this decision model. In this paper, our goal is to study the asymptotic properties of our social dynamic model. Specifically, we are interested in understanding that for a large amount of agents, what is the fraction of agents, who will make a specific decision under the sequential decision process. This asymptotic behavior can be interpreted as the decentralized information aggregation from a large population that considers the dynamic social interactions among agents.

In order to study the structure of the asymptotic decision fraction, in this paper we investigate two particularly important classes of filters. First, we consider the linear filters with deterministic and identical coefficients, i.e., $\mathcal{S}_{i,j}(x_j) = x_j$, for all i, j. In this case, the filtered output R_i corresponds to the fraction of the agents among the first $i-1$ agents, who make the decision "1". In other words, we are assuming that the most significant information being aware by the agents regarding to previous decisions is the fraction of decisions made by previous agents. This characterizes a key feature of human decisions about social activities in current human societies. The most illustrative example is the opinion polls in elections, where the electors can only access the support rate of each candidate, but has no clue of whether some group of people support certain candidate or not. Second, we consider the stochastic filters, in which for all i and j, the $\mathcal{S}_{i,j}(0)$ and $\mathcal{S}_{i,j}(1)$ are i.i.d. according to some probability distributions $f_{\mathcal{S}_0}$ and $f_{\mathcal{S}_1}$. This i.i.d. stochastic filter, on the other hand, deals with a broad range of network structures, such as the network connectivity, degree distributions, and agents with different communities. For example, the famous Erdös-Rényi (ER) network model, where each agent is assumed to observe each of the previous agent with probability q, can be modeled in our setup by appropriately designing the distributions $f_{\mathcal{S}_0}$ and $f_{\mathcal{S}_1}$. The detail of this design will be presented in section III-B.

The study of our social dynamic model in general provides the fundamental characterizations and quantifications for the dynamics of information aggregations and cascades. In particular, we show in section III that for a large population, the opinions of of the agents will reach a consensus such that the fraction of the agents making the same decision will approach to a convergent set, which depends on the utility functions of the agents. This illustrates how the decisions of previous agents can affect the entire populations. In addition, our result provides a quantitive characterization of information cascades in herd decision for arbitrary decision strategies, and hence sheds the light on understanding the dynamic structures of the information cascades, and is potentially helpful in designing social interactive systems.

The rest of this paper is organized as follows. In section II, we introduce the setup of our social dynamic model and the problems formulation. Then, we consider in section III how the decisions of the agents in the social network can converge to a consensus, in which we show that the fraction of agents making the same decision will approach to a convergent set. We also study how the opinion can be exchanged between different communities via our model in section III-C. Finally, the simulations of our model is provides in section IV, which suggests how to apply our model to practical social network problems.

II. System Model and Problem Formulation

In this paper, we investigate the scenario that there are n agents in a social network, and the agents aim to sequentially make a binary decision x_i from $\{0, 1\}$. In order to make such decisions, the agents are allowed to observe two sorts of informations. First, it is assumed that each agent i observes an outcome ω_i from the environment or the ambient space, where the outcomes ω_i's are i.i.d. drawn from a probability density function $f_\Omega(\omega)$ defined on an outcome space Ω. Second, we assume that each agent i can also observe the decisions made by previous agents[1] $1, 2, \dots, i-1$. Here, we denote the observable previous decisions as an $i-1$ dimensional binary vector $\underline{x}^{(i)} = \{x_1, \dots, x_{i-1}\} \in \{0,1\}^{i-1}$. Figure 2 demonstrates the observation structure of this model. Then, each agent randomly makes a binary decision according to a utility function u_{θ_i} that maps the observations of agent i to a probability distribution on $\{0, 1\}$. Here, $\theta_i \in \Theta$ is a parameter that characterizes the behavior pattern of the agent i. Note that different decision logics of the agents are characterized by different utility functions that are indexed by θ_i. This parameter also allows us to model different communities of agents. In particular, we assume that agent i behaves as pattern θ_i with probability $f_\Theta(\theta_i)$, where f_Θ is a probability density function on Θ, and θ_i's are also mutually independent for different i.

In order to model the nature of human decision and also simplify the problem, we would like to assume that the observable vector $\underline{x}^{(i)}$ is first passed through a pre-decision filter with the filter output $R_i = \frac{1}{i-1}\sum_{j=1}^{i-1} \mathcal{S}_{i,j}(x_j)$. Then,

[1]For the first agent, we assume the decision is made as "1" with probability p, for some $p \in [0, 1]$.

after observing R_i and ω_i, the randomized decision rule for agent i is given as

$$x_i = \begin{cases} 1, & \text{with probability } u_{\theta_i}(R_i, \omega_i) \\ 0, & \text{with probability } 1 - u_{\theta_i}(R_i, \omega_i), \end{cases} \tag{1}$$

where $u_{\theta_i} : \mathbb{R} \times \Omega \mapsto [0, 1]$ is the utility function that characterizes the decision pattern of the agent i w.r.t. the observations. Note that here, we did not assume the utility function to take any particular form, but can be arbitrary utility functions. This allows us to deal with many real world problems, where the utility functions can be trained correspondingly. In addition, the randomized decision rule (1) adopted in our setup considers both the irrational and stochastic nature of human decisions, as well as other random factors and events from the environment that are usually very difficult to be precisely modeled. Therefore, our system setup practically models a broad range of decision processes in social networks.

Now, we would like to simplify the decision strategy (1) for operational convenience. First, note that the outcome ω_i and behavior pattern θ_i are independently drawn from the probability distributions f_Ω and f_Θ, so the probability of making the decision "1" for each agent i given the observation R_i can be represented as a simplified utility function $u : [0, 1] \mapsto [0, 1]$ that is specified by

$$u(R_i) = \int_{\theta \in \Theta} \int_{\omega \in \Omega} u_\theta(R_i, \omega) f_\Omega(\omega) f_\Theta(\theta) d\omega d\theta,$$

For convenience, we assume u to be a continuous function. In addition, the randomized decision rule of each agent i can be reduced to:

$$x_i = \begin{cases} 1, & \text{with probability } u(R_i) \\ 0, & \text{with probability } 1 - u(R_i). \end{cases} \tag{2}$$

The simplified decision rule (2) in effect averages out different decision patterns w.r.t. θ_i. Thus, we can virtually think of each agent operating with the same decision logic, but still have to keep in mind that physically the group of agents are composed of different types.

The main focus of this paper is to investigate the decentralized information aggregation and information cascade for the proposed social dynamic model (2). Specifically, we are interested in studying the fraction of agents in a large number of populations, who will make a specific decision. For this purpose, we demonstrate the information cascade and aggregation structure of our model for two particularly important classes of filters: (i) the deterministic filter with all $\mathcal{S}_{i,j}(x_j) = x_j$, and (ii) the stochastic filter, where each $\mathcal{S}_{i,j}(k)$ is i.i.d. according some distribution $f_{\mathcal{S}_k}$, for $k = 0, 1$. For the first case, the filtered output $R_i = \frac{1}{i-1} \sum_{j=1}^{i-1} x_j$ is the fraction of agents making the decision "1". Thus, the utility function $u(r)$ is a function that maps the closed interval $[0, 1]$ to itself. On the other hand, for the stochastic filter, the linear output can be an arbitrary real number, and hence the utility function is $u : \mathbb{R} \mapsto [0, 1]$. In section III, we will show the information cascade structure for these two scenarios.

III. The Information Cascades for Arbitrary Utility Functions

In this section, we consider the social dynamic model with the decision rule (2) with the deterministic filter, as well as the stochastic filter. In particular, we would like to address the question that for a large number of agents, after sequentially making their decisions, what will be the fraction of agents making a specific decision.

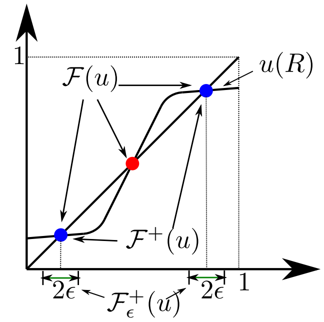


Fig. 3. The geometrical illustration of $\mathcal{F}(u)$ and $\mathcal{F}^+(u)$.

A. The Deterministic Filter

First, let us consider the deterministic filter with the filtered output $R_i = \frac{1}{i-1} \sum_{j=1}^{i-1} x_j$. We want to investigate the asymptotic behavior of the fraction of agents making decision "1" in our model. Before proceeding to more details, we would like to first introduce some definitions.

Definition 1. The fixed-point set $\mathcal{F}(u) \subseteq [0, 1]$ for an utility function u is defined as the set of points $\mathcal{F}(u) = \{r \in [0, 1] : u(r) = r\}$.

Since $u(r)$ is a smooth function, by the intermediate value theorem, the equation $u(r) = r$ has at least one solution. Thus, $\mathcal{F}(u)$ can never be an empty set. In particular, we are more interested in a specific subset in $\mathcal{F}(u)$.

Definition 2. The positive fixed-point set $\mathcal{F}^+(u) \subseteq \mathcal{F}(u)$ is defined as the set of points $r \in \mathcal{F}(u)$, such that the first order derivative $u'(r) \leq 1$. In addition, for an $\epsilon > 0$, we define the *ϵ-neighborhood* of a positive fixed-point set $\mathcal{F}^+(u)$, denoted as $\mathcal{F}^+_\epsilon(u)$, as the set of points $r \in [0, 1]$, such that for some $r' \in \mathcal{F}^+(u)$, it holds for $|r - r'| < \epsilon$.

Figure 3 illustrates the geometrical meaning of $\mathcal{F}(u)$, $\mathcal{F}^+(u)$, and $\mathcal{F}^+_\epsilon(u)$. Now, we are ready to state the main theorem of this section. The following theorem claims that the fraction of agents making the decision "1" converges to the points in the negative fixed-point set $\mathcal{F}^+(u)$, when the amount of agents approaches infinity.

Theorem 1. *Suppose that there are n agents and each agent i sequentially makes a binary decision X_i following the decision rule* (2)*, and we denote the fraction of the agents making decision "1" as a random variable R_n. Then, for any fixed t and $\epsilon > 0$, the random variable R_n converges to $\mathcal{F}^+_\epsilon(u)$ with probability 1, when $n \to \infty$. i.e.,*

$$\lim_{n \to \infty} \mathbb{P}\left[R_n \in \mathcal{F}^+_\epsilon(u)\right] = 1, \quad \forall \epsilon > 0. \tag{3}$$

Due to the space limitations, we will omit the technical proof, but present the intuition and idea behind this Theorem as follows. Suppose that after agent i making the decision, there are $i \cdot r_i$ agents making decision "1". Then, the agent $i + 1$ will make the decision "1" with probability $u(r_i)$. Now, if $u(r_i)$ is greater than r_i, then the ratio r_{i+1} will be

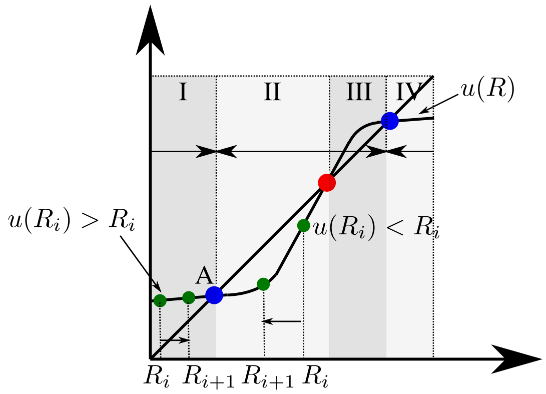


Fig. 4. The convergence of herd opinions.

more likely greater than r_i, since the probability of making decision "1" for agent $i+1$ is greater than r_i. Therefore, agent $i+1$'s decision will draw the ratio r_{i+1} away from r_i. Similarly, if $u(r_i)$ is less than r_i, then r_{i+1} will have the tendency to be less than r_i. With this intuition, we can see that if at some point the ratio r of agents making decision "1" is greater than $u(r)$, then as shown in Figure 4, the decisions of later agents will draw the ratio toward the fixed point r^* that has the slope $u'(r^*)$ no greater than 1. On the other hand, if $u(r) > r$, the later decisions will pull the ratio toward that fixed point. Moreover, this also provides the intuition of why the fixed points with slopes greater than 1 can not be convergent points: even at some point, the empirical ratio is close to a fixed point with slope greater than 1, the laters decisions of agents will at end draw the ratio away from that fixed point.

Remark 1. Theorem 1 demonstrates the potential of predicting herd behavior, and more importantly, adopting the herd behavior to match certain desired pattern in a social network. For instance, if it is desirable to have fewer decision "1" from the agents, i.e., to have the convergence appear at point A illustrated in Figure 4, then one shall adopt some agents to make decision "0" in the beginning stage, such that the ratio of agents making decision "1" is located in the region I or II. Then, the information cascades start to build up,, which leads the convergence of the ratio to happen at the desired place. This observation can be particularly useful in setting up the strategies to win elections, or constituting the political policy for governments.

B. The Stochastic Filter

Now, let us consider the case, where the $\mathcal{S}_{i,j}(0)$ and $\mathcal{S}_{i,j}(1)$ are i.i.d. random variables, for all i, j, and distributed with some probability distributions $f_{\mathcal{S}_0}$ and $f_{\mathcal{S}_1}$. This scenario is motivated from that in a practical social network, the agents are usually only interacting with a small fraction of other agents in the network. Therefore, there exists a network topology that indicates the connection between agents. In particular, the most widely adopted stochastic network model in network science is the Erdös-Rényi (ER) network model, which assumes that the connection between any pair of agents exists with the probability $q > 0$, and is independent to other pair of agents. In order to adopt the ER network model to our social dynamic model, we assume that each agent can now observe each previous agent's decision only independently with probability q. Then, the filtered output with the ER network topology can be written as

$$R_i = \frac{1}{L} \sum_{j=1}^{L} \mathcal{S}_{i,i_j}(X_{i_j}) \tag{4}$$

where $i_1, \ldots, i_L$ are the agents that the user i can observe. Here, L is a random variable distributed as the binomial distribution with the parameter q.

Now, we are interested in the asympototic properties of the information cascades with the network topology. For this purpose, let us first consider the distribution of R_i, when i grows large. Note that L is a binomial random variable, hence for large i, L is concentrated to the mean $i \cdot q$. Mathematically, for any $\varepsilon > 0$, there exists a $\gamma_1(\varepsilon) > 0$ independent to i, such that for large enough i,

$$\mathbb{P}\left[|L - i \cdot q| \geq \varepsilon\right] \leq \exp\left(-\gamma_1(\varepsilon) \cdot i\right). \tag{5}$$

Now, suppose that within the first $i-1$ agents, the fraction of the agents making the decision "1" is ρ_i. Then, the sum (4) can be written as

$$R_i = \frac{1}{L} \sum_{j=1}^{L} \mathcal{S}_j,$$

where $\mathcal{S}_j$ is the random variable that can be written as

$$\mathcal{S}_j = \rho_i \cdot \mathcal{S}_1 + (1 - \rho_i) \cdot \mathcal{S}_0,$$

and $\mathcal{S}_1$, $\mathcal{S}_0$ are the random variables with the probability distributions $f_{\mathcal{S}_1}$ and $f_{\mathcal{S}_0}$. Therefore, applying the law of large number to (5), we have for a fixed L and any $\varepsilon > 0$, there exists a $\gamma_2(\varepsilon) > 0$, such that for large enough i,

$$\mathbb{P}\left[|R_i - \mu(\rho_i)| \geq \varepsilon\right] \leq \exp\left(-\gamma_2(\varepsilon) \cdot L\right), \tag{6}$$

where

$$\mu(\rho_i) = \rho_i \cdot \mathbb{E}[\mathcal{S}_1] + (1 - \rho_i) \cdot \mathbb{E}[\mathcal{S}_0] \tag{7}$$

is the mean of $\mathcal{S}_j$. Therefore, from (5) and (6), we can see that for any $\varepsilon > 0$ there exists a $\gamma(\varepsilon) > 0$, such that

$$\mathbb{P}\left[|R_i - \mu(\rho_i)| \leq \varepsilon\right] \geq 1 - \exp\left(-\gamma(\varepsilon) \cdot i\right). \tag{8}$$

In other words, for large i, R_i is concentrated to $\mu(\rho)$ given by (7). Thus, the decision rule (2) can now be adopted with the ER network model as

$$x_i = \begin{cases} 1, & \text{with probability } u^*(\rho_i) \\ 0, & \text{with probability } 1 - u^*(\rho_i) \end{cases} \tag{9}$$

where $u^*(\rho_i) = u(\mu(\rho_i))$ is the adopted utility function with the network topology. From (9), we can come up with a similar asymptotic convergence result in the social network with the ER network model.

Theorem 2. *Suppose that there are n agents and each agent i sequentially makes a binary decision X_i following the decision rule* (2)*. In addition, we assume that every agent can observe each of the previous agent independently with probability q, and denote the fraction of the agents making decision "1" as a random variable R_n, then,*

$$\lim_{n \to \infty} \mathbb{P}\left[R_n \in \mathcal{F}_\epsilon^+(u^*)\right] = 1, \quad \forall \epsilon > 0, \tag{10}$$

where $u^*(\rho_i) = u(\mu(\rho_i))$ is the adopted utility function.

Remark 2. It turns out that the convergence of R_n only depends on the adopted utility function u^*, but independent to the connecting probability q in the ER model. This result can be interpreted as follows. Note that in the ER model, for any q that is fixed w.r.t. to the size of the network nodes n, the probability of existing a giant component approaches to 1 as n grows. Therefore, the network can be considered as almost connected, when q is some fixed quantity. Therefore, our result says that as far as the network is connected, the opinion of each agent can be propagated and shared by other agents, and hence the consensus can be reached according to their behavior pattern u^*. On the other hand, the connecting probability q can indeed affect the convergence rate in our model. This can be seen from (5) that for small q, each agent can typically observe fewer previous agents, and hence the right hand side of (6) will decade slower. This result also makes sense because when q is small, the degrees of the network nodes becomes small (while still connected). Thus, the information will be more difficult to be shared among agents, and the information cascading speed will be slow.

C. Information Cascading Model With Communities

Now, let us consider that there are M communities in the social network, and each agent belongs to one community. We assume that each agent belongs to the community m with probability p_m, which is independent to other agents. In addition, to model the network topological structure, we also assume that an agent i in the community m_1 can observe the decision of an agent j with probability $q_{m_1 m_2}$, if the agent j belongs to the community m_2. Our goal is again to characterize the fraction of agents making the decision "1", for a large number of agents with the community and network topology. For the sake convenience, we consider in this section the deterministic filter with binary decision space as in section III-A, where the results can be easily carried to the stochastic filter case. Moreover, we assume that an agent i in the community m make the decision "1" with probability $u^{(m)}(R_i)$, where R_i is the fraction of the agents among the observable agents of the agent i, who make the decision "1", and $u^{(m)}$ is the utility function as in (2). This specifies the behavior patterns of different communities.

Now, let us consider the first n agents, who are in the community m and have made their decisions. We write the fraction of agents, who make the decision "1" among these n agents, as $R_n(m)$. Then, following similar large deviation arguments as in section III-B, we can find that as n approaches to infinity, $R_n(1), \ldots, R_n(M)$ convergence with probability 1 to the region $\mathcal{R} = \{R(m)\}_{m=1}^{M}$, which satisfies that for all $m \leq M$,

(i) $u^{(m)}\left(\sum_{m'=1}^{M} Q_{m'}^{(m)} R(m')\right) = R(m)$,

(ii) $u^{'(m)}\left(\sum_{m'=1}^{M} Q_{m'}^{(m)} R(m')\right) \leq \left(Q_m^{(m)}\right)^{-1}$,

where

$$Q_{m'}^{(m)} = \frac{p_{m'} \cdot q_{m,m'}}{\sum_{m'=1}^{M} p_{m'} \cdot q_{m,m'}}. \quad (11)$$

In fact, $Q_{m'}^{(m)}$ can be interpreted as the fraction of agents among all the observed agents of an agent i in community m, that is in community m'. Thus, the above constraint (i) can be understood as the convergent points similar to Theorem 1, but adjusted according to the community and network structure. Moreover, from (11), we can see that $Q_{m'}^{(m)}$ is proportional to $q_{m,m'}$. This tells that when the agents in a community m' have a great visibility by the agents in another community m, the behavior pattern of the agents in m will be significantly affected by the agents in m'. In addition, this effect can be quantified by the equation (i). In a nutshell, the region $\mathcal{R}$ quantifies how the decision patterns of the agents in different communities can affect with each other, and hence allows the prediction of the herd behaviors in different communities.

IV. Experimental Simulations

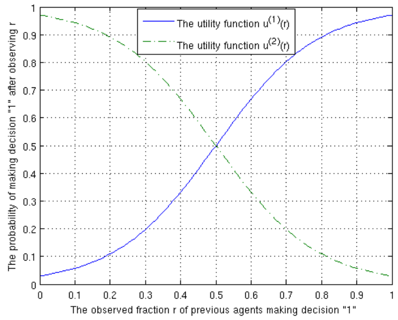
In this section, we consider agents belonging to two communities 1 and 2, such that the agents in community 1 can observe the agents in the same community with probability $q_{11} = 1$, and the agents in community B with probability $q_{12} = 0.5$, and similarly $q_{22} = 1$ and $q_{21} = 0.5$. Then, the normalized observing fractions of agents $Q_1^{(1)} = 2/3$, $Q_1^{(2)} = 1/3$, $Q_2^{(1)} = 1/3$, and $Q_2^{(2)} = 2/3$. In addition, we employ the sigmoid functions as the decision pattern of the agents in community 1, since the sigmoid function appropriately model the humans' decision logics in many real world problems as suggested in [20]. Specifically, we assume the utility function $u^{(1)}$ of the decision pattern of the agents in the first community as

$$u^{(1)}(r) = \frac{1}{1 + \exp(-7(r - 0.5))}.$$

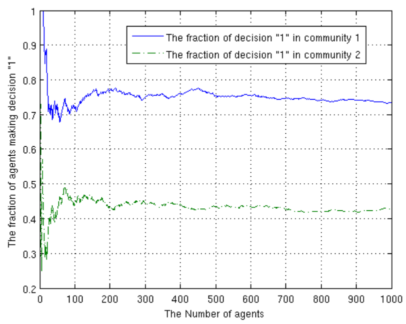
Moreover, we take $u^{(2)}(r) = u^{(1)}(1 - r)$ to model the scenario, where the agents in community 2 tends to have counter opinions from community 1. The utility functions of communities 1 and 2 are illustrated as in Figure 5(a), where the utility functions show a completely different nature in making decisions.

We run the sequential decision process for agents in two communities and consider the fraction of agents making the decision "1" among the first k agents, for $k = 1, \ldots, 1000$, in these two communities. The simulation result is shown in Figure 5(b). From Figure 5(b), we can see that the fraction agents making decision "1" for the community 1 converges to about 0.728 and the fraction converges to about 0.437 for the community 2. This is a point in the convergent region $\mathcal{R}$ that satisfies (i) and (ii) in section III-C with the set of $Q_m^{(m')}$, $u^{(1)}$ and $u^{(2)}$ specified in this section. Moreover, we can also see that the convergence of the fractions to the convergent point is quite fast (close to the convergent point at around 50 agents). Therefore, while we assume a large number of agents to derive the results in previous sections, the simulation shows that our results can also apply to a median size social network. Thus, the convergent region we derived in section III-C is in fact quite accurate and also robust in predicting herd behaviors in practical social networks.

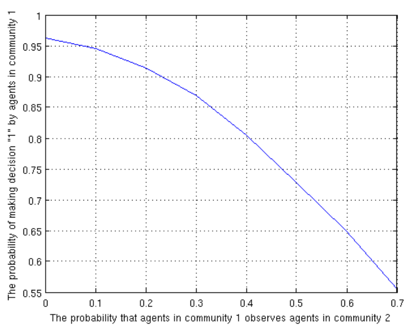
Finally, Figure 5(c) plots the fraction of the agents in community 1 making the decision "1" w.r.t. different q_{12}, i.e., different probability of observing the agents in community 2 (we set $q_{12} = q_{21}$, and $q_{11} = q_{22} = 1$). This figure shows that without observing agents in community 2 ($q_{12} = 0$), the



(a)



(b)



(c)

Fig. 5. (a) The utility functions for 2 communities. (b) The fraction of decision "1" w.r.t. the number of agents. (c) The fraction of agents making decision "1" w.r.t. different q_{12}.

agents in community 1 making decision "1" with the fraction 0.962, which is the solution of $r = u^{(1)}(r)$. However, as q_{12} grows, the opinions of the agents in community 1 are more and more affected by the counter opinions from community 2, and the fraction of agents making decision "1" goes down. This result is intuitive, while our results in this paper provide a quantitive way to characterize how much the opinions in one community can affect the other.

V. Conclusions

In this paper, we proposed a novel model to study social dynamic problems. We consider the sequential decision scenario that happens in many real world situation such as the committee meeting. We show that in a large population, the fraction of agent making a specific decision will converge to a set, which depends on the decision utility function of the agents. In addition, we also present the extension of our model to widely adopted network models, such as the ER network model and decision with communities. Our results illustrates that how the decision made by previous agents can affect the decisions of the entire population, and how the opinions of one community can affect the other community. This provides a quantitive view for information cascades, which can be potentially useful in understanding the dynamics of social networks.

Acknowledgment

This research was supported in part by the US Air Force under the grant AOARD-14-4053 and by the Ministry of Science and Technology under the grant MOST-103-2221-E-002-022-MY3.

References

[1] Banerjee A., "A Simple Model of Herd Behavior," *The Quarterly Journal of Economics*, vol. 107, pp. 797-817, 1992.

[2] Bikhchandani S., Hirshleifer D., and Welch I., "A Theory of Fads, Fashion, Custom, and Cultural Change as Information Cascades," *The Journal of Political Economy*, vol. 100, pp. 992-1026, 1992.

[3] Easley D. and Kleinberg J., "Networks, Crowds and Markets: Reasoning about a Highly Connected World," Cambridge University Press, 2010.

[4] Banerjee A. and Fudenberg D., "Word-of-mouth Learning," *Games and Economic Behavior*, vol. 46, pp. 1-22, 2004.

[5] Bala V. and Goyal S., "Learning from Neighbours," *The Review of Economic Studies*, vol. 65, no. 3, pp. 595-621, 1998.

[6] Bala V. and Goyal S.,"Conformism and Diversity under Social Learning," *Economic Theory*, vol. 17, pp. 101-120, 2001.

[7] Bikhchandani S., Hirshleifer D., and Welch I., "Learning from the Behavior of Others: Conformity, Fats, and Informational Cascades," *The Journal of Economic Perspectives*, vol. 12, pp. 151-170, 1998.

[8] Smith L. and Sorensen P., "Pathological Outcomes of Observational Learning," *Econometrica*, vol. 68, no. 2, pp. 371-398, 2000.

[9] Acemoglu D., Dahleh M., Lobel I., and Ozdaglar A., "Bayesian Learning in Social Networks," *Review of Economic Studies*, vol. 78, pp. 1201-1236, 2011

[10] Gale D. and Kariv S., "Bayesian Learning in Social Networks," *Games and Economic Behavior*, vol. 45, no. 2, pp. 329-346, 2003.

[11] Gill D., Sgroi D., De C. and Grubb M., "Sequential Decisions with Tests," *Games and Economic Behavior*, vol. 63, no. 2, pp. 663-678, 2008.

[12] Jadbabaie A., Molavi P., Sandroni A., Tahbaz-Salehi A., "Non-Bayesian social learning," *Games and Economic Behavior*, vol. 76, no. 1, pp. 210-225, 2012

[13] Yair Pinto, Andries R. van der Leij, Ilja G. Sligte, Victor A. F. Lamme and H. Steven Scholte, "Bottom-up and top-down attention are independent," *Journal Of Vision*, Vol. 13, No. 3, Dec. 2013

[14] Donald Eric Broadbent, "Perception and communication," Pergaman Press, New York, 1958

[15] J. A. Gray and A. A. I. Wedderburn, "Grouping strategies with simultaneous stimuli," *Quarterly Journal of Experimental Psychology*, 1960

[16] A. Tversky and D. Kahneman, "Judgement under uncertainty: Heuristics and biases," *Science*, Vol. 185, pp. 1124-1131, 1974

[17] Ward Edwards, "Conservatism in human information processing," *Formal representations of human judgment*, Wiley, New York, 1968

[18] Mark A. Gluck and Gordon H. Bower, "From Conditioning to Category Learning: An Adaptive Network Model," *Journal of Experimental Psychology: General*, Vol. 117, No. 3, pp. 227-247, 1988

[19] Schiller, R.J., "Conversation, Information and Herd Behavior," *Rhetoric and Economic Behavior* , 85 (3), pp. 181-185, 1995.

[20] K.C. Chen, M. Chiang, H.V. Poor, "From Technological Networks to Social Networks," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 9, pp. 548-572, September Supplement 2013.