A Bayesian model of information cascades

Sriashalya Srivathsan[1], Stephen Cranefield[1], and Jeremy Pitt[2]

[1] University of Otago, Dunedin, New Zealand
ashal.srivathsan@postgrad.otago.ac.nz,
stephen.cranefield@otago.ac.nz
[2] Imperial College, London, UK
j.pitt@imperial.ac.uk

Abstract. An information cascade is a circumstance where agents make decisions in a sequential fashion by following other agents. Bikhchandani et al. predict that once a cascade starts it continues, even if it is wrong, until agents receive an external input such as public information. In an information cascade, even if an agent has its own personal choice, it is always overridden by observation of previous agents' actions. This could mean agents end up in a situation where they may act without valuing their own information. As information cascades can have serious social consequences, it is important to have a good understanding of what causes them. We present a detailed Bayesian model of the information gained by agents when observing the choices of other agents and their own private information. Compared to prior work, we remove the high impact of the first observed agent's action by incorporating a prior probability distribution over the information of unobserved agents and investigate an alternative model of choice to that considered in prior work: weighted random choice. Our results show that, in contrast to Bikhchandani's results, cascades will not necessarily occur and adding prior agents' information will delay the effects of cascades.

Keywords: Information cascade · Coordination · Probabilistic graphical model · Bayesian inference

1 Introduction

Propagation of opinions in society has a significant impact. In daily life it is clear that people are affected by others' views. For example, in electoral and financial campaigns, the spreading of news, opinion and rumours can have an enormous effect on the behaviour of the crowd. When people look at others' actions, or listen to others they update their assessment of the value of those actions and imitate accordingly.

Information cascades are a social phenomenon in which all individuals from some point in a sequence onwards make the same decision. It occurs when other people's prior choices can strongly impact the choices of those who follow and it is the result of solely following others while discounting their own opinion. Bikhchandani et al. [3] say that, "An informational cascade occurs when it is optimal for an individual, having observed the actions of those ahead of him, to follow the behavior of the preceding

individual without regard to his own information". This phenomenon can be observed, for example, when people choose a restaurant or school [2].

People are more attracted when they see that someone already adopted a trend or chosen a specific fashion. New communication technology such as social media has become a habit, helping to share new trends, fashion and ideas [6]. It passes very quickly across web technologies when someone shares a particular fashion he or she likes, and this is subsequently shared by friends [17,12]. Resharing posts develops an information cascade in social media and leads others to choose a product, movie, or a specific fashion. Furthermore, recent studies [8,15] have shown that e-marketing is driven by information cascades. Customer ratings and reviews make a huge impact in online shopping. Thus, often people become more aware of and start to follow other's opinions when they need to shop online, rather than following their own preferences. It is evident that when shopping in Amazon or Ebay, people rely heavily on others' choices since the quality of the goods is still an issue [21].

In addition, an information cascade can aid society to unite its members for environmental collective action such as minimising emission of carbon and global warming as well as social and political collective action. Lohmann [16] notes that cascades are observed in society when individuals join and protest against a regime to address their political concerns.

It is obvious that people often avoid their own preference or personal signals and choose to pursue others' choices as soon as a cascade begins. However, even though they have strong qualities, people's own preferences are concealed once the cascade starts. For instance, the algorithm for discovering the real rating of an Amazon product was proposed by Wang et al. [21] as people demand plausible and accurate online reviews. Wang et al. attempted as much as possible to eliminate the herding effect occurring through online purchasing using their algorithm. But the actual rating was very difficult to obtain.

Bikhchandani et al. [3] have presented a probabilistic model to explain how an information cascade comes about, in the context of uptake of a fashion or fad. Their model predicts that cascades are inevitable, once they start. It is assumed that there is a true value of the fashion (0 or 1), which may be perceived correctly (with some probability $p >= 0.5$) or incorrectly. A cascade can be either a *correct* cascade, where eventually all individuals make the choice that aligns with the true value, or an *incorrect* one, where eventually all choices go against the fashion's true value. The model uses a deterministic model of choice, which means that the agents choose the action (to adopt or reject the fashion) with the greatest evidence for its correctness, based on observations of other agents and their own private information, or make a 50/50 random choice if the evidence for "adopt" and "reject" is equal.

Bikhchandani et al. presented a high level analysis of their model, which lacks mathematical details. Our aim is to provide a detailed Bayesian model that accounts for the uncertainty about the private information of other agents. We also wish to investigate the impact of agents making choices non-deterministically, via random weighted choice. Indeed, this probabilistic choice model is claimed to be psychologically more realistic [18,20].

In the model of Bikhchandani et al., the first agent in the chain of observed agents is assumed to be the first to make a decision to adopt or reject the fashion. This makes its choice highly influential, as its action is known to correspond directly to its private signal. Moreover, this is unrealistic in many settings. An agent may know that the fashion has been around for a while, so it could estimate a number of prior agents whose actions were unobserved. Our model therefore introduces uses a prior probability distribution over the count of positive perceptions of the fashion by these unobserved agents. Even if this is a completely uninformative prior distribution, it weakens the dominance of the first agent.

2   Prior Work

Besides Bikhchandani et al., many other researchers have highlighted the importance of information cascades. Benerji et al. [2] introduced a model to investigate herd behaviour to understand how people adopt others' actions while ignoring their own information. They showed that people observe others' actions and tend to act in the same way because they believe the previous people have better information than them. Easley et al. [9] presented a theoretical and experimental study of herding behaviour. They presented a Bayesian model for sequential decision-making where people consider the counts of the previous actions and choose the highest one.

Vany et al. [7] developed an agent-based model to compare the theoretical aspects of Bikichandani's concept. Specially, they analysed how people are attracted to watch a specific movie as they have multiple choices. The key aspect of this model is how an agent will be highly influenced by the nearest neighbour's action as well as the popularity of different movies. In the same fashion, Lee et al. [11] examined how online reviews of a movie are socially influenced [19] by two different groups such as a general crowd and friends. They concluded that the reviews of the popular movies show a cascading behaviour. However, people tend to follow friend ratings regardless of popularity, since these have more impact than the general crowd rating. Similarly Liu et al. [14,13] analysed how information cascades occur in e-book marketing. Particularly, for both paid and free e-books they experimented to find the effect of information cascades, and found that an information cascade has a strong impact when selecting paid e-books in comparison with free books.

Anderson et al. [1] designed a laboratory experiment to show how an information cascade occurs. It became apparent that people at some stage discard their own knowledge and continue to follow others. Huber et al. [10] used functional magnetic resonance imaging of experimental participants' brains to research how a cascade can be stimulated by individual preference. They concluded from their findings that over-weighting personal information can trigger and stop the information cascade.

Watts [23,22] developed a model to show how agents' actions interact with neighbours' actions by setting a simple threshold rule. Cheng et al. [6] addressed the prediction of a cascade when sharing a picture post in Facebook. They found that the size of an information cascade can be predicted by its temporal (observed time) and structural (caption, language and content) features. Lu et al. [17] proposed a system

to analyse the collective behaviour of information cascades by analysing a huge social media data set.

3   The information cascade model of Bikhchandani et al.

Bikhchandani et al. [3] considered a sequence of individuals who will choose between accepting (A) or rejecting (R) a fashion or fad. The model assumes that there is some true value $V \in \{0, 1\}$ representing the benefit of following the fashion/fad. Every individual perceives a private signal $X_i \in \{H, L\}$, representing a "high" or "low" perception of V. As shown in Table 1 a correct perception of V has probability p for $V = 1$ and $1-p$ for $V = 0$, and vice versa for incorrect perceptions.

In the "basic model" considered by Bikhchandani et al., each agent considers the actions of earlier agents in the sequence, and takes those actions of adopt or reject as a proxy for what each of those agents perceived as a true value, which is unknown by observing agents. The agent can count all of the previous accepts or rejects, adds a count 0 or 1 (respectively) given his own information (L or H), and then chooses his action based on the action with the greatest count. Here we notice that the actions of agents earlier in the chain are repeatedly used as evidence by later agents, and their impact may therefore be exaggerated.

The "general model" of Bikhchandani et al. states that an agent will make its decision based on the expected value of V given the agent's private signal and observations of other agents' actions. Bikhchandani et al. gave the following definition: $V_{n+1}(x; A_n) \equiv E[V|X_{n+1} = x, X_i \in J_i(A_{i-1}, a_i)$, for all $i \leq n]$. Here x is the signal perceived by agent $n + 1$, A_{i-1} is the sequence of prior actions that agent i has observed and $J_i(A_{i-1}, a_i)$ is a set of signals that led individual i to choose action a_i. Individual $n + 1$ adopts if $V_{n+1}(x; A_n) \geq C$, where C is the cost of adoption[3]. However, no computational account is given of how this expected value is determined by the agent.

Bikhchandani et al. found that "An informational cascade occurs if an individual's action does not depend on his private signal". Hence, he ignores his private signal and will adopt based on the prior agents' actions alone. This is true for all subsequent agents too, so they follow their predecessors and create a cascade. If $V = 1$, a cascade can be either a correct cascade, where all adopt, or an incorrect cascade, where all reject, and vice versa for $V = 0$.

Bikhchandani et al. showed that once the cascade starts it will last forever even if it is incorrect. They also discussed the fragility of cascades. For instance, cascades can be broken if public information is revealed.

Our aim is to create a Bayesian model of an information cascade that captures the uncertainty about prior agents' private information. We avoid the strong influence of the first agent's action on later agent's action through the ability to model a probability distribution over the information of unobserved prior agents. The model of Bikhchandani et al. uses a deterministic model of choice in which an agent ranks options and

---

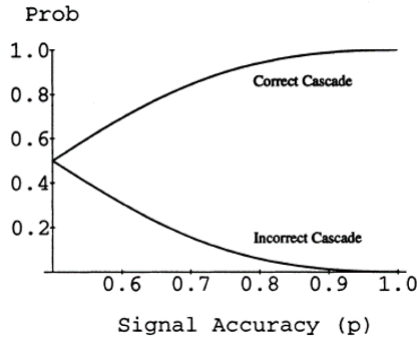[3] This cost isn't used in our model.

Fig. 1: Graph of probability of correct and incorrect cascade as a function of p in the simple model of Bikhchandani et al. [3]

Table 1: Signal probabilities [3]

Gain of adopting	$P(X_i = H \mid V)$	$P(X_i = L \mid V)$
$V = 1$	p	$1 - p$
$V = 0$	$1 - p$	p

chooses the highest ranked one. This deterministic technique[4] is commonly used in economic models [5]. Inspired by the work of Luce, [20,18] we consider the effects on cascades if agents use weighted non-deterministic choice when choosing to adopt or reject. This means that choice is probabilistic and an agent may choose randomly from a set of weighted choices, in proportion to their weights. This has been claimed to be more psychologically plausible than deterministic choice [20].

4   Our Model and Approach

In this section we present our Bayesian model of information cascades. We define the following variables:

- $V \in \{0, 1\}$, is the true value of the fad/fashion.
- $X_i \in \{H, L\}$, is the private signal of the agent: high or low. This represents agent i's possibly incorrect perception of V.
- $A_i \in \{A, R\}$, is the action of the agent: adopt or reject. C_i is a count of how many times H appears in $\{X_1, \cdots, X_n\}$. This will be probabilistically inferred by each agent since it cannot be directly perceived.
- k is a number of prior agents who made choices that were not observed by any of the agents. C_0 is the count of H signals observed by these prior agents. As this is unknown, our model uses an estimated probability distribution for C_0.

---

[4] Luce refers to this as *algebraic choice* due to the common use of algebraic rather than probabilistic models in economics [18].

Fig. 2: Probabilistic graphical model of information cascade from agent 4's viewpoint

- p is the signal accuracy, which is assumed to be the same for all agents.

The dependencies between these variables are shown using a probabilistic graphical model [4] in Figure 2, which is expressed from the viewpoint of agent 4. k and p are constants, but the other nodes are random variables. Shaded nodes represent observed variables, and X_i is known by agent i. A_4 is shown as a rectangle, as this is a decision node. Agent 4 will choose A if $P(V = 1 \mid X_4, A_1, A_2, A_3) > 0.5$, R if this probability is less than 0.5, and otherwise will make a 50/50 choice between A and R. An equivalent choice procedure involving C_i is given at the end of this section.

The information cascade model is used as follows. V and C_i are conditionally independent given C_{i-1} and X_i, therefore:

$$P(V, C_i \mid C_{i-1}, X_i) = p(V \mid C_{i-1}, X_i)\, p(C_i \mid C_{i-1}, X_i) \tag{1}$$

Each agent i maintains a joint probability distribution over V and C that is conditional on the actions observed so far. This is $p(V, C_0) = p(V)\, p(C_0 \mid V)$ for the first agent. We use a uniform prior over V and a binomial distribution for C_0 given V:

$$P(C_0 = c \mid V = 0) = (1 - p)^c p^{k-c} \text{ and } P(C_0 = c \mid V = 1) = p^c (1 - p)^{k-c} \tag{2}$$

For $i > 1$, agent i will observe the actions of all prior agents $j < i$ and compute the joint conditional distribution $p(V, C_j \mid C_0, A_1, \cdots, A_j)$. Agents act and are observed sequentially, so it will already have computed $p(V, C_{j-1} \mid C_0, A_1, \cdots, A_{j-1})$.[5] Once it observes A_j, it first uses Bayes' Theorem to compute $P(X_j \mid A_1, \cdots, A_j)$:

$$P(X_j \mid A_1, \cdots, A_j) \propto P(X_j \mid A_1, \cdots, A_{j-1})\, P(A_j \mid X_j, A_1, \cdots, A_{j-1}) \tag{3}$$

---

[5] C_0 appears as a condition whenever a sequence of action variables does. Henceforth, we omit it for brevity.

Table 2: $P(A_j \mid X_j, C_{j-1})$ for deterministic choice

X_j	Condition on C_{j-1}	$P(A_j\!=\!A)$	$P(A_j\!=\!R)$
	$P(C_{j-1}+1 > \frac{k+j}{2}) > P(C_{j-1}+1 < \frac{k+j}{2})$	1	0
H	$P(C_{j-1}+1 > \frac{k+j}{2}) < P(C_{j-1}+1 < \frac{k+j}{2})$	0	1
	$P(C_{j-1}+1 > \frac{k+j}{2}) = P(C_{j-1}+1 < \frac{k+j}{2})$	0.5	0.5
	$P(C_{j-1} > \frac{k+j}{2}) > P(C_{j-1} < \frac{k+j}{2})$	1	0
L	$P(C_{j-1} > \frac{k+j}{2}) < P(C_{j-1} < \frac{k+j}{2})$	0	1
	$P(C_{j-1} > \frac{k+j}{2}) = P(C_{j-1} < \frac{k+j}{2})$	0.5	0.5

Table 3: $P(A_j \mid X_j, C_{j-1})$ for non-deterministic choice

X_j	$P(A_j\!=\!A)$	$P(A_j\!=\!R)$
H	$P(C_{j-1}+1 > \frac{k+j}{2}) + \frac{1}{2}P(C_{j-1}+1 = \frac{k+j}{2})$	$P(C_{j-1}+1 < \frac{k+j}{2}) + \frac{1}{2}P(C_{j-1}+1 = \frac{k+j}{2})$
L	$P(C_{j-1} > \frac{k+j}{2}) + \frac{1}{2}P(C_{j-1} = \frac{k+j}{2})$	$P(C_{j-1} < \frac{k+j}{2}) + \frac{1}{2}P(C_{j-1} = \frac{k+j}{2})$

where:

$$
\begin{aligned}
P(X_j & \mid A_1, \cdots, A_{j-1}) \\
&= \sum_{v \in \{0,1\}} P(X_j \mid V = v, A_1, \cdots, A_{j-1}) \, P(V = v \mid A_1, \cdots, A_{j-1}) \\
&= \sum_{v \in \{0,1\}} P(X_j \mid V = v) \, P(V = v \mid A_1, \cdots, A_{j-1})
\end{aligned}
\tag{4}
$$

$$
P(A_j \mid X_j, A_1, \cdots, A_{j-1}) = P(A_j \mid X_j, C_{j-1})
\tag{5}
$$

$P(A_j \mid X_j, C_{j-1})$ is calculated as shown in Tables 2 and 3 for the deterministic and non-deterministic models of choice, respectively.

Agent i can compute $p(V, C_j \mid A_1, \cdots, A_j)$ as follows:

$$
\begin{aligned}
P(V = v, C_j = c \mid A_1, \cdots, A_j) = & \\
P(V = v, C_{j-1} = c \mid A_1, \cdots, A_{j-1}) \, & P(X_j = L \mid A_1, \cdots, A_j) + \\
P(V = v, C_{j-1} = c - 1 \mid A_1, \cdots, A_{j-1}) \, & P(X_j = H \mid A_1, \cdots, A_j)
\end{aligned}
\tag{6}
$$

When it is agent i's turn to act, it will know $P(V = v, C_{i-1} = c \mid A_1, \cdots, A_{i-1})$ and will have its own signal X_i to calculate $P(V = v, C_i \mid C_{i-1}, X_i)$. Since V and C_i are conditionally independent given C_{i-1} and X_i, we have:

$$
P(V = v, C_i \mid C_{i-1}, X_i) = P(V = v \mid C_{i-1}, X_i) P(C_i \mid C_{i-1}, X_i)
\tag{7}
$$

where

$$
P(V = v \mid C_{i-1}, X_i) \propto P(V = v \mid C_{i-1}) P(X_i \mid V = v, C_{i-1})
\tag{8}
$$

$$P(X_i = x \mid V = v, C_{i-1}) = P(X_i = x \mid V = v) = \begin{cases} p & \text{if } (v,x) \in \{(1,H),(0,L)\} \\ 1-p & \text{otherwise} \end{cases} \tag{9}$$

and

$$P(C_i = c \mid C_{i-1}, X_i = x) = \begin{cases} P(C_i = c-1) & \text{if } x = H \\ P(C_i = c) & \text{if } x = L \end{cases} \tag{10}$$

Finally, as agent i now has a probability distribution over C_i, it can choose between the options A and R by comparing $p_1 = P(C_i > \frac{k+i}{2} \mid \cdots)$, $p_2 = P(C_i < \frac{k+i}{2} \mid \cdots)$ and $p_3 = P(C_i = \frac{k+i}{2} \mid \cdots)$. For deterministic choice, if $p_1 > p_2$ it chooses A, and if $p_2 > p_1$ it chooses R. Otherwise it tosses an evenly weighted coin to choose between the two options. For non-deterministic choice, a random weighted choice is made with weights $p_1 + \frac{p_3}{2}$ for A and $p_2 + \frac{p_3}{2}$ for R.

5   Experiment and Results

5.1   Deterministic Model
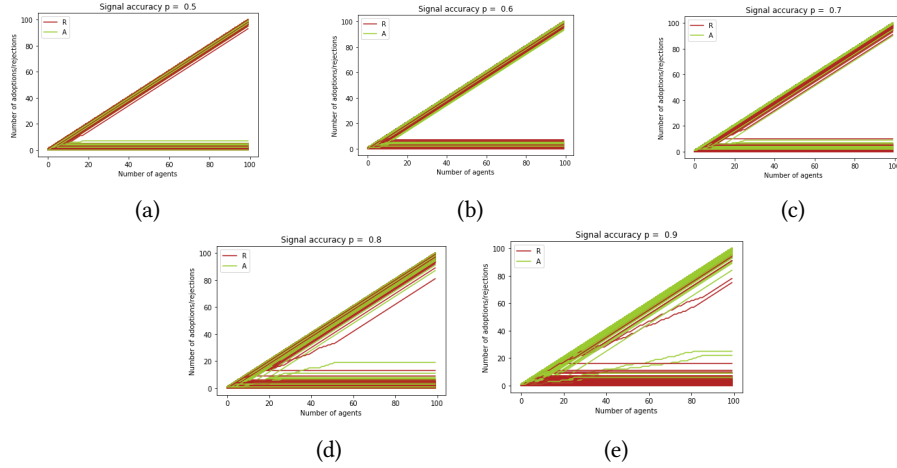
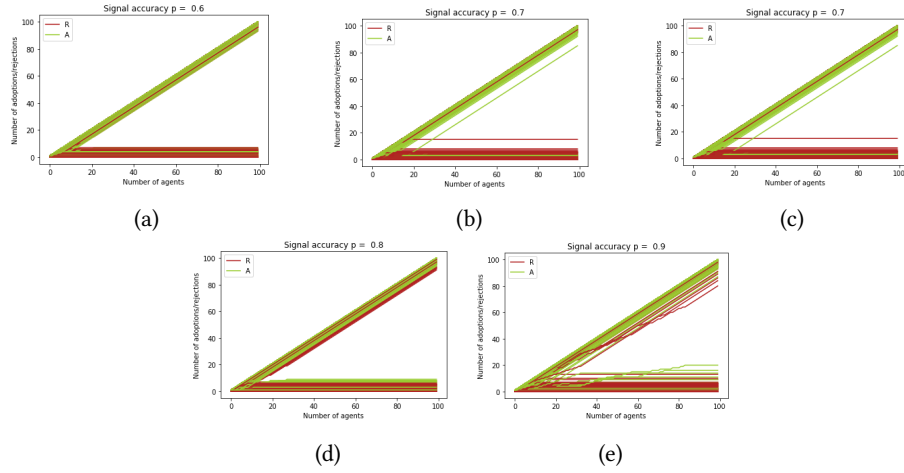Fig. 3: Cumulative frequencies of adopts and rejects for 100 agents over 1000 runs with $V = 1$ for each p with 1 prior agent for deterministic choice (best viewed in colour)

As a first step we implemented[6] a deterministic choice model [3] using our Bayesian approach. Our model uses the probability distribution $P(C_i \mid C_{i-1}, X_i)$ to choose actions.

---

[6] The implementation of our model in Python can be found at https://github.com/ashalya86/Information-cascade-models.

Fig. 4: Cumulative frequencies of adopts and rejects for 100 agents over 1000 runs with $V = 1$ for each p with 20 prior agents for deterministic choice (best viewed in colour)
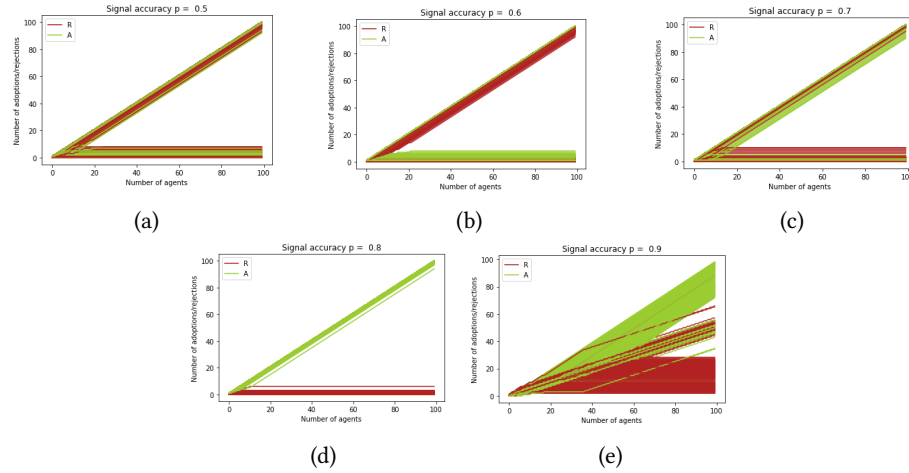
Fig. 5: Cumulative frequencies of adopts and rejects for 100 agents over 1000 runs with $V = 1$ for each p with 40 prior agents for deterministic choice (best viewed in colour)

We ran the simulation 1000 times for 100 agents for different values of $v \in \{0, 1\}$, $p \in \{0.5, 0.6, 0.7, 0.8, 0.9\}$ and $k \in \{1, 20, 40\}$. In a single graph we plot two lines, red and green, for each run. The green line indicates the cumulative frequency of acceptances (A) and the red line indicates the same for rejections (R). There are separate graphs for each value of $p \in \{0.5, 0.6, 0.7, 0.8, 0.9\}$. Figures 3, 4 and 5 show the outcome of cascades which start with 1, 20 and 40 prior agents, respectively, for $V = 1$.

The results for $V = 0$ are similar to those for $V = 1$, except that the colours are swapped as R is dominant when $V = 0$. It is evident from the graphs that there is always a high chance of cascades where everyone adopts when $V = 1$ and everyone rejects when $V = 0$. These graphs are similar to the results obtained from Bikhchandani's original model [3].

5.2 Non-deterministic Model

We wish to investigate how the deterministic model of choice in the model of Bikhchandani et al. impacts their results, given that non-deterministic choice has been described as more psychologically plausible [20]. We modified our implementation to pick an action using random weighted choice. As a result, even if the most likely optimal choice is to adopt, the agent could still choose the less likely one and reject.
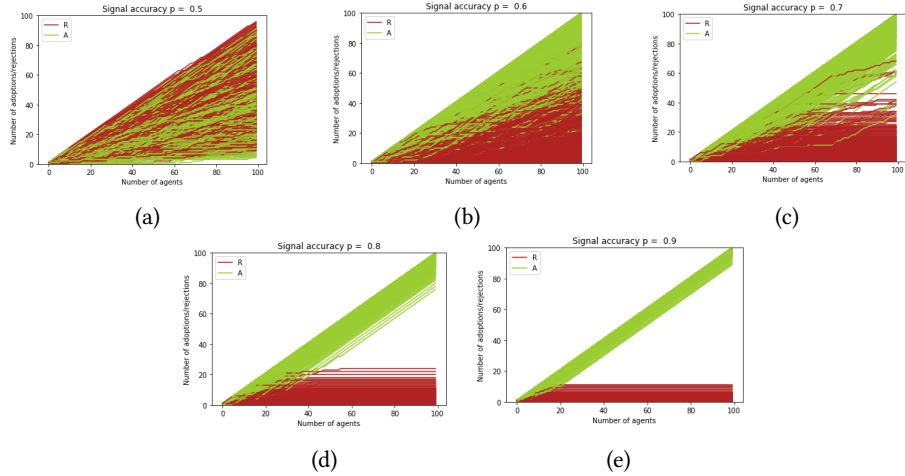

Fig. 6: Cumulative frequencies of adopts and rejects for 100 agents over 1000 runs with $V = 1$ for each p with 1 prior agent for non deterministic choice (best viewed in colour)

As for the deterministic choice model, in a single graph we plot cumulative frequency of acceptances (A) and rejections (R) for 1000 runs of 100 agents for each p. The plots of cascades which begin with 1, 20 and 40 prior agents for $V = 1$, are seen in Figures 6, 7, 8. The results for $V = 0$ are similar to $V = 1$, except that the colours are swapped as R is dominant when $V = 0$.

Simulation results suggest that for high accuracy perception of the true value of a choice ($p \in \{0.8, 0.9\}$) (Figures 6d, 6e, 7d and 7e), cascades still occur, but for lower accuracy perception, cascades are replaced by a bias towards one choice that is greater than would be expected if only the odds of correct vs. incorrect perception were considered (Figures 6a, 6b, 7a and 7b). Although the ratio of choosing As over Rs starts
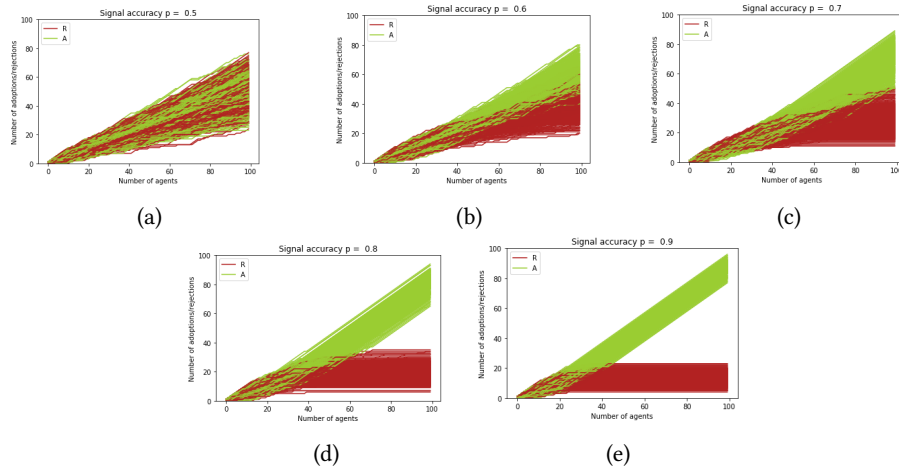
Fig. 7: Cumulative frequencies of adopts and rejects for 100 agents over 1000 runs with $V = 1$ for each p with 20 prior agents for non deterministic choice (best viewed in colour)

with 0.6 and 0.4 for $p = 0.6$, one choice becomes increasingly dominant and it still shows the possibility of no cascade occurring. (Figures 6b and 7b).

According to Bikhchandani's model, the first agent's private information can be uniquely determined from its action, which makes its choice highly influential. Therefore, our model starts by assuming there are some unobserved agents present and uses a prior distribution over C and V.

For instance, suppose there are assumed to be 20 unobserved prior agents. Then the model creates a uniform prior for V and a binomial distribution over the count of prior H signals given V, in terms of p, for 20 prior agents. We then do the Bayesian inference. We obtained a significant change in cascades while plotting the cascade with different numbers of prior agents. We notice that the prior agents delay the occurrence of cascades. While the proportions of two choices have wider deviation for 1 prior agent (Figures 6b, 6c, 6d and 6e), it gradually reduces for 20 prior agents (Figures 7b, 7c, 7d and 7e) and 30 (Figures 8b, 8c, 8d and 8e)

6   Conclusion

An information cascade happens when people observe the actions of their predecessors and try to follow these observations regardless of their own private information. We presented a full Bayesian account for the model of Bikhchandani et al. [3]. We maintain a probability distribution over V and the (unknown) counts of High signals received by the agents. Rather than always choosing the most likely option, agents make a weighted choice between Adopt and Reject. We do not assume that the first agent that was observed was the first to consider the fad. Instead, we incorporate prior
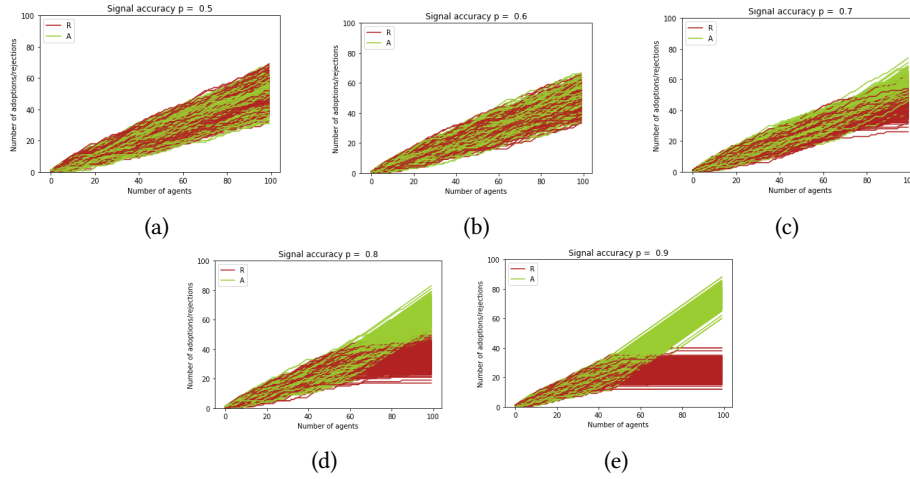
Fig. 8: Cumulative frequencies of adopts and rejects for 100 agents over 1000 runs with $V = 1$ for each p with 40 prior agents for non deterministic choice (best viewed in colour)

knowledge of unobserved agents. This means that first (observed) agent's choice is less dominant than in the earlier model. Our findings show that prior agents delay the occurrence of cascades. Furthermore, in contrast to the predictions of Bikhchandani et al., our results show that cascades will not necessarily occur. The graphs obtained show that, for lower accuracy perception, cascades occur with much less probability. However, as p gets high, there is a high chance of cascades.

While people may not be good at Bayesian reasoning, as used in our model, we believe that appropriate software using our model could support users to assess the evidence from others' choices, including the possible presence of unobserved agents. This could help to reduce the likelihood of cascades.

References

1. Anderson, L.R., Holt, C.A.: Information cascades in the laboratory. American Economic Review **87**, 847–862 (1996)
2. Banerjee, A.V.: A simple model of herd behavior. The Quarterly Journal of Economics **107**(3), 797–817 (1992). https://doi.org/10.2307/2118364
3. Bikhchandani, S., Hirshleifer, D., Welch, I.: A theory of fads, fashion, custom, and cultural change as informational cascades. Journal of Political Economy **100**(5), 992–1026 (1992)
4. Bishop, C.M.: Pattern recognition and machine learning. Springer-Verlag (2006)
5. Blavatskyy, P.R.: Dual choice axiom and probabilistic choice. Journal of Risk and Uncertainty **61**(1), 25–41 (2020). https://doi.org/10.1007/s11166-020-09332-7
6. Cheng, J., Adamic, L., Dow, P.A., Kleinberg, J.M., Leskovec, J.: Can cascades be predicted? In: Proceedings of the 23rd International Conference on the World Wide Web - WWW '14. pp. 925–936. ACM Press (2014). https://doi.org/10.1145/2566486.2567997

7. De Vany, A., Lee, C.: Information cascades in multi-agent models. Tech. rep., University of California, Irvine, Department of Economics (1999), http://pages.stern.nyu.edu/~wgreene/entertainmentandmedia/Cascades.pdf

8. Duan, Gu, Whinston: Informational cascades and software adoption on the Internet: An empirical investigation. MIS Quarterly **33**(1), 23 (2009). https://doi.org/10.2307/20650277

9. Easley, D., Kleinberg, J.: Networks, Crowds, and Markets: Reasoning about a highly connected world. Cambridge University Press, UK (2010)

10. Huber, R.E., Klucharev, V., Rieskamp, J.: Neural correlates of informational cascades: brain mechanisms of social influence on belief updating. Social Cognitive and Affective Neuroscience **10**(4), 589–597 (2015). https://doi.org/10.1093/scan/nsu090

11. Lee, Y.J., Hosanagar, K., Tan, Y.: Do I follow my friends or the crowd? information cascades in online movie ratings. Management Science **61**(9), 2241–2258 (2015). https://doi.org/10.1287/mnsc.2014.2082

12. Liu, C., Wang, W., Jiao, P., Chen, X., Sun, Y.: Cascade modeling with multihead self-attention. In: 2020 International Joint Conference on Neural Networks (IJCNN). pp. 1–8. IEEE (2020). https://doi.org/10.1109/IJCNN48605.2020.9207418

13. Liu, Q., Zhang, L.: Information cascades in online reading: an empirical investigation of panel data. Library Hi Tech **32**(4), 687–705 (2014). https://doi.org/10.1108/LHT-06-2014-0052

14. Liu, Q., Zhang, X., Li, Y.: The influence of information cascades on online reading behaviors of free and paid e-books. Library & Information Science Research **42**(1) (2020). https://doi.org/10.1016/j.lisr.2019.101001

15. Liu, Q., Zhang, X., Zhang, L., Zhao, Y.: The interaction effects of information cascades, word of mouth and recommendation systems on online reading behavior: an empirical investigation. Electronic Commerce Research **19**(3), 521–547 (2019). https://doi.org/10.1007/s10660-018-9312-0

16. Lohmann, S.: Collective action cascades: An informational rationale for the power in numbers. Journal of Economic Surveys **14**(5), 654–684 (2002)

17. Lu, Y., Yu, L., Zhang, T., Zang, C., Cui, P., Song, C., Zhu, W.: Exploring the collective human behavior in cascading systems: a comprehensive framework. Knowledge and Information Systems **62**(12), 4599–4623 (2020). https://doi.org/10.1007/s10115-020-01506-8

18. Luce, R.D., Suppes, P.: Preference, utility, and subjective probability. In: Luce, Duncan, R., Bush, R.R., Galanter, E. (eds.) Handbook of Mathematical Psychology, vol. 3, p. 162. John Wiley & Sons (1965)

19. Peng, S., Yang, A., Cao, L., Yu, S., Xie, D.: Social influence modeling using information theory in mobile social networks. Information Sciences **379**, 146–159 (2017). https://doi.org/10.1016/j.ins.2016.08.023

20. Pleskac, T.J.: Decision and choice: Luce's choice axiom. In: International Encyclopedia of the Social & Behavioral Sciences, pp. 895–900. Elsevier (2015). https://doi.org/10.1016/B978-0-08-097086-8.43031-X

21. Wang, T., Wang, D.: Why Amazon's ratings might mislead you: The story of herding effects. Big Data **2**(4), 196–204 (2014). https://doi.org/10.1089/big.2014.0063

22. Watts, D.J.: A simple model of global cascades on random networks. In: Proceedings of the National Academy of Science. pp. 4766–5771 (2022)

23. Watts, D.J., Dodds, P.S.: Influentials, networks, and public opinion formation. Journal of Consumer Research **34**(4), 441–458 (2007). https://doi.org/10.1086/518527