GitHub Repository

1 Basic Formulation

Idea: We attempt to estimate the reliability of both the source and children nodes and utilize the influence probabilities to select the set $\mathcal{S}$ for the Oracle.

1.1 Motivation:

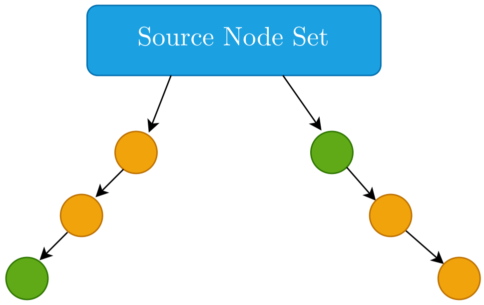


Figure 1: ICM Diffusion Model with high-reliability children nodes denoted in green

In this model, we identify two types of children nodes: reliable and semi-reliable. If the propagation follows the right path, we immediately remove the news from the network when an immediate reliable child flags it. However, there's a tradeoff between the cost to send to the Oracle and the misinformation spread if the left path is taken.

1.2 Notation

- X^t: Set of news x generated at time step t
- A^t: Set of active news at time step t $\{$news not given to Oracle until timestep $t\}$
- S^t: Set of news selected and given to the Oracle at timestep t
- $\pi^t(a)$: Set of affected nodes until time t by news a
- $\psi^t(a)$: Set of nodes/users who have flagged news a until time t
- $Y_u(x)$: label by user u assigned to x with true label $Y^*(x)$
- $\theta_{u,f} = P(Y_u(x) = f | Y^*(x) = f) \quad \theta_{u,\bar{f}} = P(Y_u(x) = \bar{f} | Y^*(x) = \bar{f})$
- $\beta^t(s)$: Reliability of source node s
- $\text{val}^t(a) : \mathbb{E}_{\mathcal{D}}\left(\left|\pi^\infty(a)\right| \,\middle|\, \pi^m(a), m \le t \right) - \left|\pi^t(a)\right|$ for an active news a

1.3 Objective Function

$$UTIL(T) = \mathbb{E}\left[\sum_{t \in [T]} \sum_{x \in S^t} \mathbb{I}\{x = \text{fake}\} \text{val}^t(x) \right]$$

To determine S^t, the algorithm will utilize the reliability Θ_u for every node u through which news a has propagated (using $\pi^t(a)$ and flags for a), along with the belief parameter $\beta^t(s)$ for the source node of a and the Diffusion Model $\mathcal{D}$

2 Our Problem Setting

There is no restriction on the number of news items given to the Oracle at any given timestep, but there is a cost associated with each verification. Here, we are introducing control over the number of news items that we can select at every timestep to be given to the Oracle.

2.1 Motivation

Note that we can have a situation where

$$\mathcal{P}^t(a = \text{fake}) \cdot \text{val}^t(a) \leq \mathcal{P}^{t+1}(a = \text{fake}) \cdot \text{val}^{t+1}(a), \quad a = \arg\max_x \left\{\mathcal{P}^t(x = \text{fake}) \cdot \text{val}^t(x)\right\}$$

holds and we may choose to select this news a at a later timestep.

2.2 Objective Function

$$\text{UTIL}(T) = \mathbb{E}\left[\sum_{t\in[T]} \left(\sum_{x\in S^t} \mathbb{I}\{x = \text{fake}\} \cdot \text{val}^t(x) - \gamma|S^t|\right)\right]$$

Here γ is the trade-off parameter

3 Single News Propagation

We consider a single news item propagating through the social network. Our objective is to determine if or when we need to send it to the Oracle, given our knowledge of the distribution model, user reliability, and source reliability. We dynamically update the probability $\mathcal{P}$ of the news being fake to make this decision

$$\mathcal{P}^t = \mathbb{P}(x = f|\psi(t), \theta_{u,f}, \theta_{u,\bar{f}}, \mathcal{S}(x)) = \frac{\prod_{u\in\psi(t)} \theta_{u,f} \prod_{u\in\pi(t)-\psi(t)} (1-\theta_{u,f})\beta}{\prod_{u\in\psi(t)} \theta_{u,f} \prod_{u\in\pi(t)-\psi(t)} (1-\theta_{u,f})\beta + \prod_{u\in\psi(t)} (1-\theta_{u,\bar{f}}) \prod_{u\in\pi(t)-\psi(t)} (\theta_{u,\bar{f}})(1-\beta)}$$

Consider a single news item x spreading on a social network, which hasn't been queried to the Oracle until time t. Now, consider the following scenarios at time t:

3.1 Idealistic Scenario

Let's consider a scenario in which an extremely reliable node ($\theta_{u,f} = 1$ and $\theta_{u,\bar{f}} = 1$) is a.s. to get infected at time $t+1$. We then have two possible optimal actions

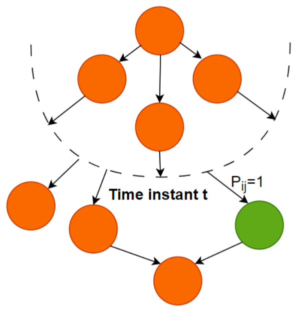


Figure 2: Idealistic Scenario at Time t

1. If news sent to the Oracle at time step t, then the utility function is given by

$$UTIL(T) = \mathcal{P}^t \text{val}(t) - \gamma$$

2. whereas the utility function if we decide to wait for another time step, the utility function will be given by

$$UTIL(T) = \mathcal{P}^t(\text{val}(t+1) - \gamma)$$

Therefore if $(1 - \mathcal{P}^t)\gamma - \mathcal{P}^t(\text{val}(t) - \text{val}(t+1)) > 0$ it is optimal not to query the at t

3.2 Non-idealistic Scenario

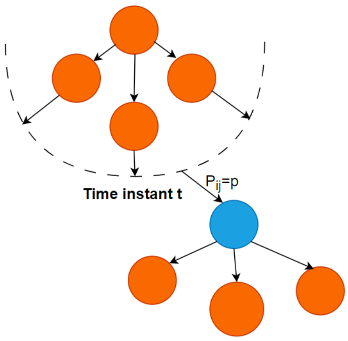


Figure 3: Non-Idealistic Scenario at Time t

$$\begin{aligned}\mathcal{P}_{\Theta_1} &= \frac{\Theta_f \prod\limits_{u \in \psi(t)} \theta_{u,f} \prod\limits_{u \in \pi(t)-\psi(t)} (1-\theta_{u,f})\beta}{\Theta_f \prod\limits_{u \in \psi(t)} \theta_{u,f} \prod\limits_{u \in \pi(t)-\psi(t)} (1-\theta_{u,f})\beta + (1-\Theta_{\bar{f}}) \prod\limits_{u \in \psi(t)} (1-\theta_{u,\bar{f}}) \prod\limits_{u \in \pi(t)-\psi(t)} (\theta_{u,\bar{f}})(1-\beta)} \\ &= \frac{1}{1+(\frac{1}{\mathcal{P}^t}-1)\frac{1-\Theta_{\bar{f}}}{\Theta_f}}\end{aligned}$$

$$\begin{aligned}\mathcal{P}_{\Theta_2} &= \frac{(1-\Theta_f) \prod\limits_{u \in \psi(t)} \theta_{u,f} \prod\limits_{u \in \pi(t)-\psi(t)} (1-\theta_{u,f})\beta}{(1-\Theta_f) \prod\limits_{u \in \psi(t)} \theta_{u,f} \prod\limits_{u \in \pi(t)-\psi(t)} (1-\theta_{u,f})\beta + \Theta_{\bar{f}} \prod\limits_{u \in \psi(t)} (1-\theta_{u,\bar{f}}) \prod\limits_{u \in \pi(t)-\psi(t)} (\theta_{u,\bar{f}})(1-\beta)} \\ &= \frac{1}{1+(\frac{1}{\mathcal{P}^t}-1)\frac{\Theta_{\bar{f}}}{1-\Theta_f}}\end{aligned}$$

Here $\mathcal{P}_{\Theta_1}$ and $\mathcal{P}_{\Theta_2}$ represent the probabilities of x being fake, given that the affected user u (Θ_f, $\Theta_{\bar{f}}$) flagged it and did not flag it, respectively.

As expected we observe that for if the user flags randomly ($\Theta_f = 0.5, \Theta_{\bar{f}} = 0.5$), rarely flags ($\Theta_f = 0, \Theta_{\bar{f}} = 1$) or always flags ($\Theta_f = 1, \Theta_{\bar{f}} = 0$) we don't gain advantage on waiting for another time step

Consider an algorithm $\mathcal{A}$ which send the news the to the Oracle with probability $U(\mathcal{P}^t)$ (U can be optimized wrt. to $\mathcal{P}^t$). If news not sent to Oracle at time t and $\mathcal{A}$ is to be applied at time $t+1$ for $p = 1$ then

$$\begin{aligned}UTIL(T) =& (\mathcal{P}^t\Theta_f + (1-\mathcal{P}^t)(1-\Theta_{\bar{f}})) \cdot U(\mathcal{P}_{\Theta_1}, t+1) \cdot (\mathcal{P}_{\Theta_1}(\text{val}(t+1)) - \gamma) \\ &+ (\mathcal{P}^t(1-\Theta_f) + (1-\mathcal{P}^t)\Theta_{\bar{f}}) \cdot U(\mathcal{P}_{\Theta_2}, t+1) \cdot (\mathcal{P}_{\Theta_2}(\text{val}(t+1)) - \gamma)\end{aligned}$$

4 Its All About Estimation: Approximating One-Step Look-Ahead

Let us consider a line-graph and let the current estimate of the news at time t being fake be $\mathcal{P}^t$. Now we have seen that the expected utility if we greedily act at $t+1$ is given by

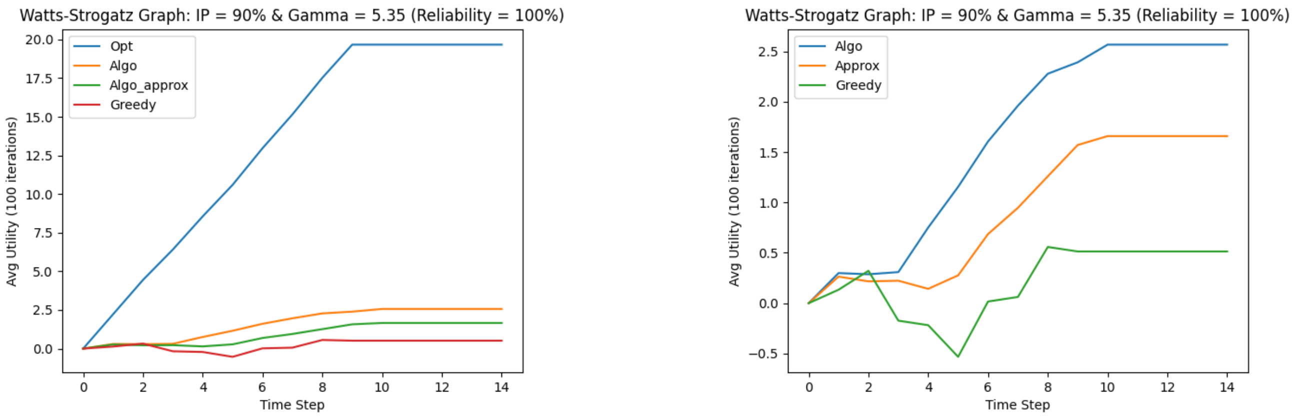
$$\begin{aligned}\mathbb{E}(UTIL(T+1)|\mathcal{H}^t) =& (\mathcal{P}^t\Theta_f + (1-\mathcal{P}^t)(1-\Theta_{\bar{f}})) \cdot q(\mathcal{P}_{\Theta_1}) \cdot (\mathcal{P}_{\Theta_1}(\text{val}(t+1)) - \gamma) \\ &+ (\mathcal{P}^t(1-\Theta_f) + (1-\mathcal{P}^t)\Theta_{\bar{f}}) \cdot q(\mathcal{P}_{\Theta_2}) \cdot (\mathcal{P}_{\Theta_2}(\text{val}(t+1)) - \gamma)\end{aligned}$$

Here $U(\mathcal{P}_{\Theta_1}, t)$ is a step-function evaluated on $\mathcal{P}_{\Theta_1}(\text{val}(t+1)) - \gamma$. This equation can also be interpreted in the following manner (both are inter convertible mathematically),

$$\begin{aligned}&\mathcal{P}^t \cdot \left(\mathbb{E}\left[U\Big(\text{Est}(t+1 \mid \text{Est}(t), Y^*(x) = f) \cdot (\text{val}(t+1) - \gamma)\Big) \cdot (\text{val}(t+1) - \gamma)\right]\right) \\ &\quad + (1-\mathcal{P}^t) \cdot \left(\mathbb{E}\left[U\Big(\text{Est}(t+1 \mid \text{Est}(t), Y^*(x) = \bar{f}) \cdot (\text{val}(t+1) - \gamma)\Big) \cdot (-\gamma)\right]\right)\end{aligned}$$

For graphs with high influence probabilities $\text{val}(t+1)$ is a nearly a constant constant and the independence assumption holds. The main approximation is taking the Expectation inside the non-linear step-function

$$\approx \mathcal{P}^t \cdot \Bigg(U\Big(\mathbb{E}\Big[\text{Est}(t+1 \mid \text{Est}(t), Y^*(x) = f)\Big] \cdot \mathbb{E}\Big[(\text{val}(t+1) - \gamma)\Big]\Big) \cdot \mathbb{E}\Big[(\text{val}(t+1) - \gamma)\Big]\Bigg)$$
$$+(1-\mathcal{P}^t) \cdot U\Big(\mathbb{E}\Big[\text{Est}(t+1 \mid \text{Est}(t), Y^*(x) = f)\Big] \cdot \mathbb{E}\Big[(\text{val}(t+1) - \gamma)\Big]\Big) \cdot (-\gamma)\Big]\Big)$$



(a) Gamma = 5 (b) Gamma = 5.35%

Figure 4

5 Approximating Flagging Activity

Note that if we have all n users with flagging accuracies θ, its accuracy given by the majority algorithm is given by (probability that at least half of the flags are correct)

$$\Theta = \sum_{r=\lfloor n/2 \rfloor + 1}^{n} {}^nC_r \theta^r (1-\theta)^{(n-r)} + \frac{{}^nC_{n/2}\theta^{n/2}(1-\theta)^{n/2}}{2}$$

Now, note that this approximation has to be altered for unsymmetrical θs, as the majority algorithm may perform worse than the accuracy of the best user! For example, consider the case where the accuracies are 0.9, 0.8, and 0.7. In this case, the maximum reliability of the set of users 2 and 3 is bounded by 0.88 $(1-(0.2)(0.3))$, which is less than 0.9. This implies that irrespective of what the other users flag, we are better off trusting user 1. Note here that the probability that at least one of the users flags it correctly upper bounds the reliability of the collection of users.

6 Regret Calculations

6.1 Logarithmic Regret Bound on Source Reliability Algorithms

Consider a scenario where we focus solely on assessing the reliability of news sources to make optimal decisions aimed at maximizing the expected utility function. In this scenario, decisions regarding querying the Oracle are made immediately upon news generation, and once made, these decisions are adhered to, regardless of any potential increase in the expected further spread in subsequent time steps. We will subsequently prove that an algorithm based on Thompson Sampling attains a logarithmic regret.

For simplicity we prove for a single source node, but the bound easily extends if there are multiple source nodes. We consider a scenario where val $\beta - \gamma' < 0$ and val $> \gamma'$, that is the optimal algorithm never queries the Oracle on any news. A similar regret bound is obtained even for the other case. For ease of further calculations let

$\gamma = \frac{\gamma'}{val}$ (Regret obtained just gets scaled). Here $\gamma < 1$, else it is obvious that it is optimal to never query the Oracle.

Notice here that this problem of ours can be converted into a 2-armed bandit, where the action here is to choose or not whether to query to the Oracle for the news. Also in this setting, the rewards for one of the arms is deterministic and 0 . Assume $L = \frac{4}{(\gamma-\beta)^2}\ln T$ sub-optimal pulls for the learning algorithm until time t_j and $k(t)$ be the total no. of sub-optimal pulls until time t, then the number of further subsequent sub-optimal pulls Y_j are given by

$$Y_j = \sum_{t=t_j}^{T} \mathbf{1}_{\{\theta_2(t)>\gamma, k(t)\geq L\}}$$
$$\mathbb{E}[Y_j] \leq \sum_{t=t_j}^{T} P(\bar{\theta}_2(t) > \frac{\beta+\gamma}{2}, k(t) \geq L) + \sum_{t=t_j}^{T} P(\theta_2(t) > \gamma, k(t) \geq L, \bar{\theta}_2(t) \leq \frac{\beta+\gamma}{2})$$

Here $\bar{\theta}_2(t)$ is the empirical mean of the source's unreliability ($\frac{\text{no. of fake news generated until } t}{\text{total no. of news generated until } t}$). Let us try bounding each of these terms. For the first term we have,

$$\begin{aligned} P(\bar{\theta}_2(t) > \frac{\beta+\gamma}{2}, k(t) \geq L) &\leq \sum_{i=L}^{T} P(k(t)=i, \bar{\theta}_2(t) > \frac{\beta+\gamma}{2}) \\ &\leq \sum_{i=L}^{T} P(k(t)=i, \bar{\theta}_2(t) - \beta > \frac{\gamma-\beta}{2}) \\ &\leq \sum_{i=L}^{T} e^{-2i(\frac{\gamma-\beta}{2})^2} \\ &\leq Te^{-2L(\frac{\gamma-\beta}{2})^2} \leq \frac{1}{T} \end{aligned}$$

Observe that $\theta_2(t)$ is a $Beta(\bar{\theta}_2(t)^{*}k(t)+1, k(t) - \bar{\theta}_2(t)^{*}k(t)+1)$. Similarly for the second term,

$$\begin{aligned} P(\theta_2(t) > \gamma, k(t) \geq L, \bar{\theta}_2(t) \leq \frac{\beta+\gamma}{2}) &\leq \sum_{i=L}^{T} P(\theta_2(t) > \gamma, k(t) = i, \bar{\theta}_2(t) \leq \frac{\beta+\gamma}{2}) \\ &\leq \sum_{i=L}^{T} P(\theta_2(t) > \gamma, k(t) = i, \bar{\theta}_2(t) \leq \frac{\beta+\gamma}{2}) \\ &\leq \sum_{i=L}^{T} P(Bin(i+1,\gamma) \leq i\ (\frac{\beta+\gamma}{2}), k(t)=i) \qquad \leq Te^{-2L(\frac{\gamma-\beta}{2})^2} \leq \frac{1}{T} \end{aligned}$$

While going from the 2^{nd} inequality to the 3^{rd} we have used

$$P(Beta(\alpha,\beta) > y) = P(Bin(\alpha+\beta-1, y) \leq \alpha - 1)$$

Hence our total number of sub-optimal pulls is bounded by $\frac{4}{(\gamma-\beta)^2}\ln T + 2$, also the sub optimality gap for each pull is given $\delta = \gamma - \beta$. Therefore the regret suffered is $\frac{4}{\gamma-\beta}\ln T + 2(\gamma-\beta)$

6.2 Bounding Regret for N-step Look-ahead

Consider a one-step look-ahead approach (generalizable to multi-step look-ahead!), which allows for a maximum of 2 steps to make the decision on whether or not to query the oracle. We consider a scenario where multiple news are disjointly spreading in the graphical model, as many intricacies will be involved in proving a bound for such a case (querying one of the news to get better accuracy for others).We will utilize the following result on RL using Posterior Sampling from Osband et al.

6.2.1 Defining the Parameters of the MDP

- **State Space:** The state of the process is defined by the set of infected nodes for each active news in the network and their flags set by each affected user and the source node for each active news. Therefore for each active news v, each user u can take 4 possible "values" $x_{uv} = \{-1, 0, 1, s\}$ where

- $x_{uv} = -1$ indicates user u is unaffected by active news v
- $x_{uv} = 0$ indicates user u is affected by active news v but hasn't flagged it
- $x_{uv} = 1$ indicates user u is affected by active news v and has flagged it
- $x_{uv} = s$ indicates user u is the source node for active news v.

If there are N active news in the network, a state is given by the matrix $[[x_{uv}]]_{M\times N}$.

- **Action Space:** If there are M active news in the network, the cardinality of the action space is 2^N since there are 2^N ways of selecting N items.
- **Transition Probabilities**: The transition probabilities to states depends on the influence probabilities from the current set of affected nodes, the flags set by users for the current set of active news (along with source reliability), the flagging accuracies for the nodes that can be affected in the next time step, and also another degree of randomness for the seeds of new active news. Rewards here are solely the function of current state and action and depend on the flags set by users (along with source reliability) and the expected future spread for the set of active news decided to be queried by the algorithm.

Notation: For every episode k of fixed horizon τ, we run the optimal policy μ_{M_k} for the MDP M_k, whose parameters (rewards and transition probabilities) are obtained from posterior sampling (Thompson sampling) based on the history H_{t_k}.

Algorithm 1: Posterior Sampling for Reinforcement Learning (PSRL)

Input: Prior distribution f, $t = 1$
for *episodes* $k = 1, 2, \ldots$ **do**
  Sample $M_k \sim f(\cdot|H_{t_k})$;
  Compute $\mu_k = \mu_{M_k}$;
  for *time steps* $j = 1, \ldots, \tau$ **do**
    Sample and apply $a_t = \mu_k(s_t, j)$;
    Observe r_t and s_{t+1};
    $t = t + 1$;
  end
end

We define the regret incurred by a reinforcement learning algorithm π up to time T to be

$$\mathrm{Regret}(T, \pi) := \frac{1}{T}\sum_{k=1}^{T/\tau_e} \Delta_k,$$

where Δ_k denotes regret over the kth episode, defined with respect to the MDP M^* by

$$\Delta_k = \sum_{s\in\mathcal{S}} \rho(s) \left(V^{M^*}_{\mu^*,1}(s) - V^{M^* \mu_k,1}(s) \right),$$

with $\mu^* = \mu^{M^*}$ and $\mu_k \sim \pi_k(H_{t_k})$ and ρ the initial state distribution

Theorem If f is the distribution of M^*, then,

$$E\left[\mathrm{Regret}(T, \pi^{PS}_{\tau})\right] = O\left(\tau S\sqrt{AT\log(SAT)}\right)$$

Now, let us consider an MDP whose states encompass all possible spreads through the network for Multiple news N spreading disjointly through the network and the flags set by the corresponding affected users. The action space consists of 2^N actions: querying each of the news M and not querying it. Here we are considering a horizon of 2 with a multiple news spreading disjointly through the network.

Claim: An optimal policy (which knows the user parameters and the network transition probabilities) will carry out a one-step look ahead approach at $t = 1$ and at $t = 2$ (final step of the episode) will act greedily for each of the news N. Therefore an algorithm implementing PSRL will suffer a $O(\sqrt{T\log(T)})$ regret.

Note: The theorem holds irrespective of the starting state distribution for the episodes it runs.

Let us consider a single news spreading in the network. Also since a maximal of 2 time steps are permitted to query the news, we can remove a news from the set of active news after 2 time steps of its generation. If we now consider the starting state of the subsequent episode to be the terminal state of the current episode, we are essentially running the algorithm on our setting, which is a single news spreading.

We can extend the theorem for spreads of news in the network which are disjoint upto 2 time steps. In this scenario, the corresponding learning algorithm (via posterior sampling) suffers a $O(\sqrt{T\log(T)})$ regret wrt. the optimal algorithm (that knows the user reliability as well as the social network) that takes a maximum of 2 steps to make its decision for any news. Here we did not allow the news to be queried 2 steps after generation. This is a restriction, but it can be resolved by using a multi-step look-ahead approach.

Though we could set the reward function in M^* with horizon 1 MDP such that it only gets a reward if the expected future look-ahead is less than the current expected value, but it wouldn't be possible to get a sampling of such a reward from the Oracle.

7 Simulation on Watts-Strogatz Graph

"It provides a way to generate small-world networks, which are characterized by both high clustering coefficient (like regular lattices) and short average path lengths (like random graphs)." **Random Rewiring:** "The model then introduces random rewiring of edges with a probability p. During rewiring, each edge in the lattice is considered for rewiring with probability p. If rewiring occurs, one end of the edge is detached from its current node and reattached to a different randomly chosen node, potentially resulting in long-range connections."

We have carried simulations on the above mentioned graph with 100% influence probabilities so as to consider the critical case for evaluating the performance of $\mathrm{Algo_approx}$.

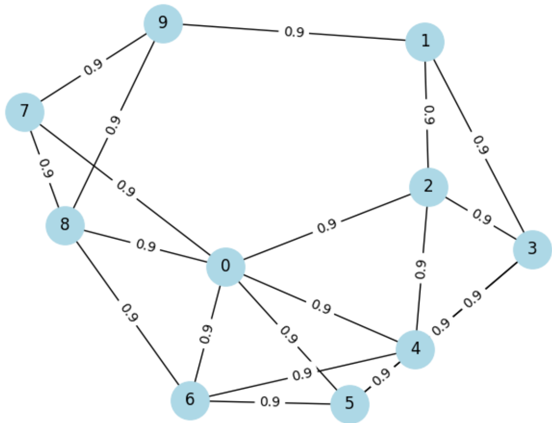
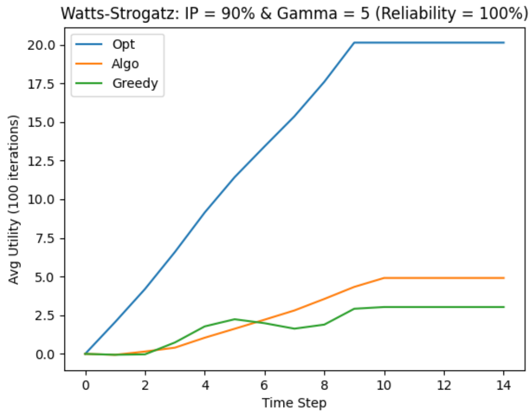
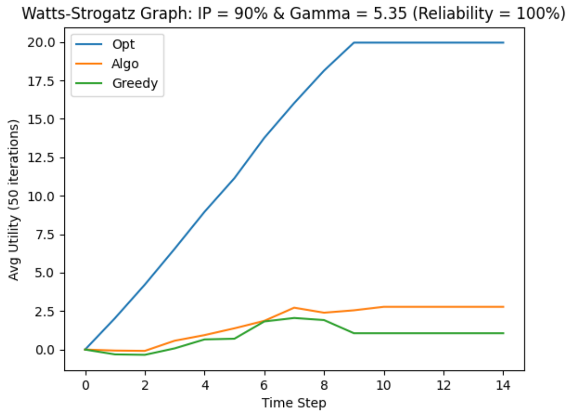


Figure 5: Watts-Strogatz graph with 10 nodes

For 10 nodes with influence probabilities 0.95 and a rewiring probability of 0.2 and 2 nearest-neighbors, we have the following results.



(a) Gamma = 5



(b) Gamma = 5.35

Figure 6

For $\gamma = 5.35$ the average utility is 1.064, while it is 2.78 for Algo (news generated for 9 timesteps). We observe that increase in γ has a higher effect on the Greedy algorithm, which is intuitive as the effective decrease for one-step look ahead is $\Delta\gamma \cdot \beta$ whereas for Greedy its simply $\Delta\gamma$.

Whenever the source nodes are $0, 4$, both the Greedy as well as Algo query the Oracle instantaneously, where as when source nodes are $1, 9, 7, 3, 5$ Algo waits for another time step before querying, whereas Greedy instantaneously queries. Surprisingly even if the source node is $2, 6, 8, 9$ Algo doesn't query since with probability atleast 0.35 the val(t + 1) is 6, which results in a Utility of atleast 0.2.

8 Simulation on Zachary's Karate Club Graph

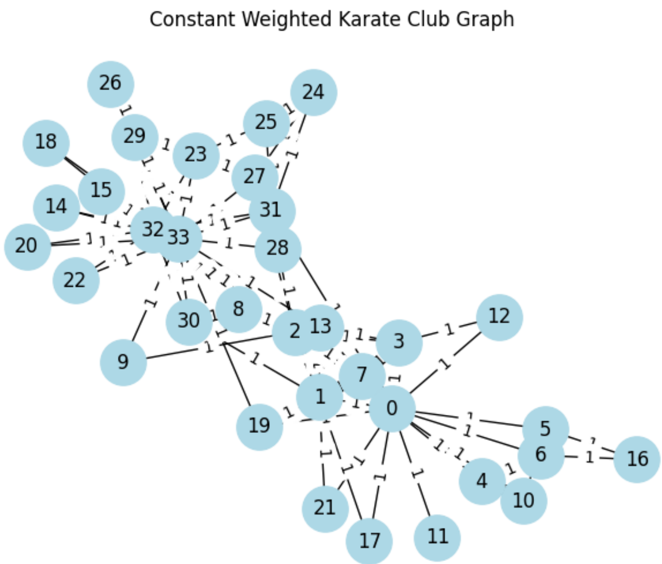


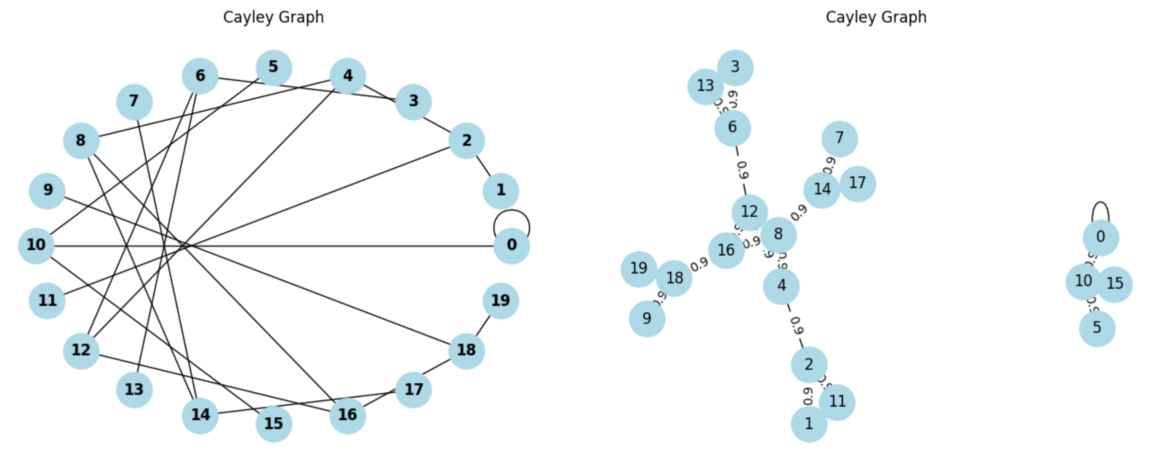
Figure 7: Zachary's Karate club graph

For Karate graphs we try to observe the effect of reducing influence probabilities on the performance gain on our prescribed Algorithm.

9 Simulation Results on Cayley Graphs

"Cayley graphs are excellent models for small-world networks, in the sense that with suitable choice of relevant parameters, they can be adapted to possess the distinguishing characteristics of such networks. By suitably choosing the parameters of the Cayley-graph models, they can be made to mimic many real networks of the types found in social, technological, and biological domains."

The simulation results below were obtained for a Cayley graph shown below with a diffusion probability of 0.7, $\beta = 0.6$, and $\gamma = 4$, where all the user reliabilities are 0.85. While carrying out simulations for gamma we realized that, for small values of gamma one-step look ahead approach is equally effective as the greedy approach, therefore only after a threshold does this approach become more effective.



(a) Cayley Graph for $Z20$ group with generator 2 (b) Alternate view of 8a

Figure 8: Cayley Graph for $Z20$ group with a single generator

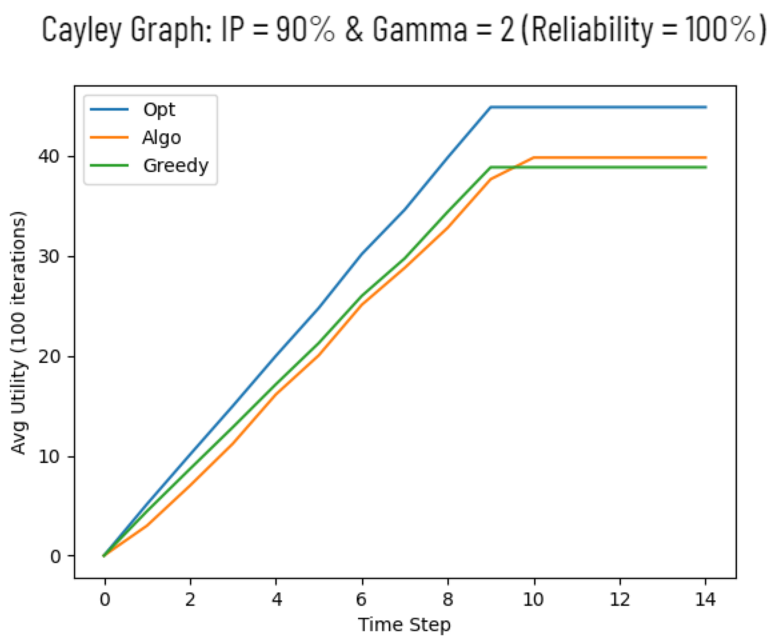


Figure 9: Performance Analysis of Algorithm for Realistic Social Networks

10 Simulation Results on Binary Trees

Note that as γ decreases, the one-step look ahead does not provide much advantage, and hence the greedy approach performs as well as our algorithm (verified by simulation for $\gamma = 0.5$). The simulation results below were obtained from a Binary Tree with a diffusion probability of 0.6, $\beta = 0.6$, and $\gamma = 2.5$.

1. High Reliability: Probability distribution of $[0.005, 0.005, 0.99]$ with values $[[0.65, 0.65], [0.1, 0.9], [0.9999, 0.9999]]$
2. Uniformly distributed: Probability distribution of $[0.33, 0.33, 0.34]$ with values $[[0.65, 0.65], [0.1, 0.9], [0.9999, 0.9999]]$
3. Low Reliability: Probability distribution of $[0.45, 0.45, 0.1]$ with values $[[0.65, 0.65], [0.1, 0.9], [0.9999, 0.9999]]$
4. Adversarial: Probability distribution of $[0.01, 0.99, 0]$ with values $[[0.65, 0.65], [0.1, 0.9], [0.9999, 0.9999]]$

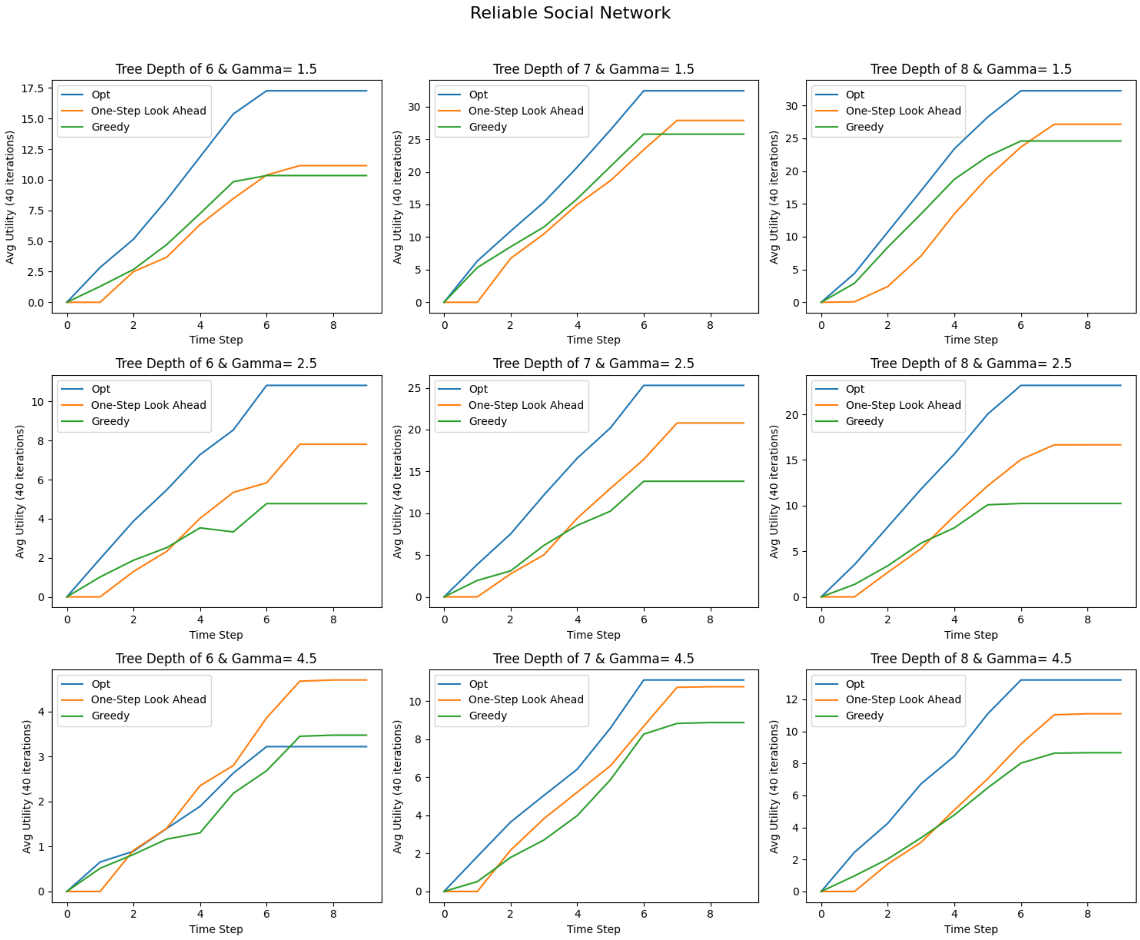


Figure 10: Performance Analysis of Algorithm with Depth of Tree and γ

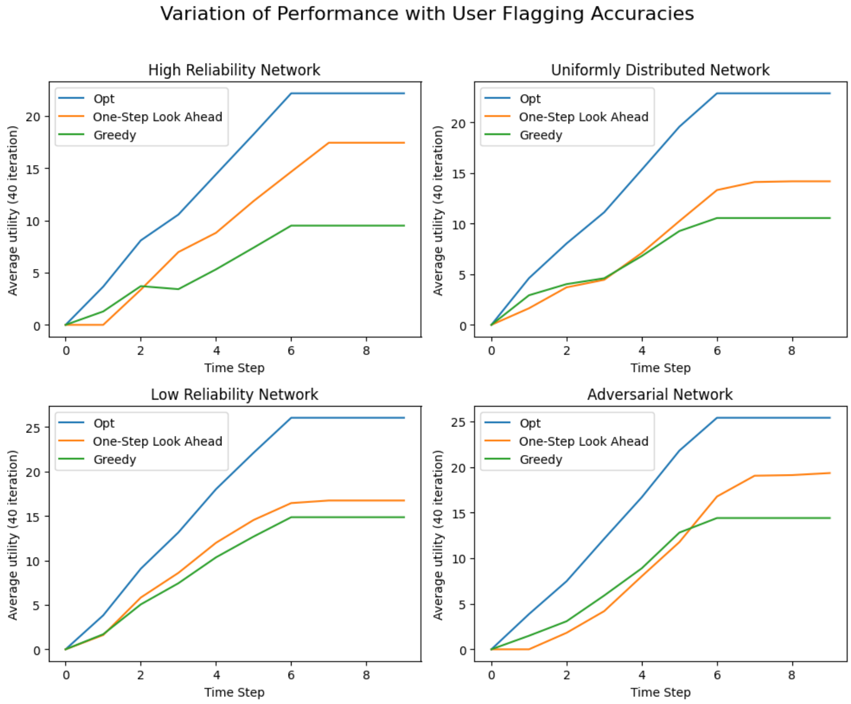


Figure 11: Performance Analysis of Algorithm with Variation in Social Network

11 Extension To Multi Oracle Model

Utility if sent to the Perfect Oracle $\mathcal{O}$,

$$UTIL(T) = \mathcal{P}^t \text{val}(t) - \gamma$$

Here the faulty oracle $\mathcal{O}'$ acts similar to a user with imperfect flagging accuracy. If we assume the Oracle to have parameters $(\Theta_f, \Theta_{\bar{f}})$, Utility after sending

$$UTIL(T) = (\mathcal{P}^t \Theta_f + (1 - \mathcal{P}^t)(1 - \Theta_{\bar{f}})) \cdot (\mathcal{P}_{\Theta_1} \text{val}(t) - (1 - \mathcal{P}_{\Theta_1})\bar{\gamma}) - \gamma'$$

where $\bar{\gamma}$ is the cost of removing a genuine news from the network. For the given parameters, we can find the range of probabilities (we get a messy quadratic in $\mathcal{P}^t$) for which querying to $\mathcal{O}'$ is optimal.

At every time-step, we evaluate these utilities and send the news to the corresponding Oracle if the expected utility is positive. Note that here we assume that we abide by the decision given by the Oracle.

An interesting problem statement could involve learning the faulty Oracle parameters using the user flags on news and queries directed to the ideal Oracle.

12 Regret Bounds Wrt Opt

Let us first consider a scenario where only a single news is propagating in the network from a given source node. The expected utility for Opt is given by

$$UTIL_{Opt} = \mathbb{E}\Big[\beta \cdot (\text{val}(0) - \gamma)\Big]^+$$

whereas if we consider a scenario where all users are reliable, then the utility of our algorithm is given by the expression

$$UTIL_{Algo} = \max\left\{\mathbb{E}\Big[\beta \cdot \text{val}(0) - \gamma\Big]^+, \mathbb{E}\Big[\beta \cdot (\text{val}(1) - \gamma)\Big]^+\right\}$$

$$R_T \leq \mathbb{E}\Big[\beta \cdot (\text{val}(0) - \text{val}(1))\Big]$$

13 Alternative Frameworks

- A case where there are adversarial users in the network, and the adversary has complete information regarding the message propagation, such as the flagging activity of users and the Diffusion model structure.
- A framework where there are two oracles: a noisy oracle (with lower verification cost) and an ideal oracle.
- Framework where we can remove a news from the network without verifying from the Oracle {risk of removing correct news}