Towards Combating Rumors in Social Networks: Models & Metrics

Rudra M. Tripathy

Department of Computer Science & Engineering,

I I T Delhi, New Delhi

tripathy@cse.iitd.ac.in

Amitabha Bagchi

Department of Computer Science & Engineering,

I I T Delhi, New Delhi

bagchi@cse.iitd.ac.in

Sameep Mehta

IBM Research-India, New Delhi

sameepmehta@in.ibm.com

Abstract

Rumor is a potentially harmful social phenomenon that has been observed in all human societies in all times. Social networking sites provide a platform for the rapid interchange of information and hence, for the rapid dissemination of unsubstantiated claims that are potentially harmful. In this paper, we study different methods for combating rumors in social networks actuated by the realization that authoritarian methods for fighting rumor have largely failed. Our major insight is that in situations where populations do not answer to the same authority, it is the trust that individuals place in their friends that must be leveraged to fight rumor. In other words, rumor is best combated by something which acts like itself, a message which spreads from one individual to another. We call such messages *anti-rumors*. We study three natural anti-rumor processes to counter the rumor and present mean field equations that characterize the system. Several metrics are proposed to capture the properties of rumor and anti-rumor processes. The metrics are geared to capture temporal evolution as well as global properties of the processes. We evaluate our methods by simulating rumor and anti-rumor processes on a large data set of around 10^5 nodes derived from the social networking site Twitter and on a synthetic network of the same size generated according to the Barabási-Albert model.

Keywords: Rumor, Anti-rumor, Diffusion, Social Networks.

1 Introduction

Sociologists use the term rumor to refer to an unverified account or explanation of events circulating from person to person and pertaining to an object, event, or issue of public concern [23]. Rumor is a potentially harmful social phenomenon that has been observed in human societies in all times. In Book IV of the *Aeneid*, the Latin poet Virgil refers to “Rumor, the swiftest of all evils,” going on further to give an intuitive characterization of its spread: “Speed lends her strength, and she finds vigor as she goes.” [32]. Since Virgil wrote his classic, enabling technologies for the spread of rumor have multiplied. Specifically online social networks (Twitter, Orkut, Facebook etc.) provide a platform for the rapid interchange of information and hence, for the rapid dissemination of unsubstantiated claims that are potentially harmful. In this paper we study ways of combating rumors in social networks actuated by the realization that authoritarian methods for fighting rumor have largely failed. Our major insight is that in situations where populations do not answer to the same authority, it is the trust that individuals place in their friends that must be leveraged to fight rumor. In other words, rumor is best combated by something which act like itself, a message which spreads from one individual to another. We call such messages *anti-rumors*. Using mathematical models proposed in the literature for the spread of rumor, we study different anti-rumors processes and present metrics for evaluating the efficacy of these methods in fighting the spread of rumor.

Social networking sites have certain intrinsic properties that make them an ideal medium for the spread of rumor. They have huge and distributed user-bases, clusters of users sharing the same interests, developing trust in each other, and seeking access to the same resources. Moreover, the platform openness makes it easy to deploy malicious applications. During times of crisis the use of social networking to spread information can have the harmful effect of allowing rumors to proliferate even faster and wider than they did earlier. In an era where email spam is easily detected by most users (would you send your bank account number to the widow of an assassinated African general?), mainly because it comes from people we do not know. However, it is possible for malicious users to leverage the trust we repose in our "friends" or "connections" on social networking sites in order to spread harmful content. In fact, as of now, all popular social networking sites have experienced some level of malicious use (see e.g. [22, 26]).

1.1 A note on assumptions

Our model makes one radical simplifying assumption that the anti-rumor is a message which is completely convincing. There are several gaps in this assumption. The basic question that arises is: How does a person distinguish an anti-rumor from a rumor? In some cases this may be possible: for example if the anti-rumor carries a proof of its correctness with it, or if it comes from an authenticated source. But in general this may not be possible. We do not account for this lacuna in our work. Instead we present our studies as a foundational step. It may be possible in future to develop more sophisticated models that take into account the difficulty in distinguishing between rumor and anti-rumor, but we believe that the basic underlying dynamics will be not too different from the ones we analyze here. We discuss the possibilities of pursuing such a direction in Section 5.

Further, we assume that a person who has once heard the anti-rumor will never again believe the rumor and will in fact want to spread the anti-rumor. This is based on a somewhat rose-tinted view of human nature. While it is true that there are people who make independent efforts to quell rumors because they realize the danger they may cause (see e.g. [20]), to believe that every person will want to actively spread an anti-rumor is unduly optimistic. In fact the truth might be far from this assumption: It is even possible that malicious users might intercept the anti-rumor and, adapting the rumor so that it becomes harder to refute and start spreading the adapted rumor. It is no doubt possible to model a variable response to anti-rumor spread and that might be a way to refine our model, but in this paper we only lay out the basic program, and do not study the model to that level of sophistication.

1.2 Main contributions

The main contributions are setting out three models that describe natural scenarios for the spread of anti-rumor to combat the spread of rumor. We describe these briefly now and in more detail in future sections.

Traditionally governments have tried to combat rumor in a centralized and sometimes authoritarian fashion. The mode has been of an authority broadcasting the anti-rumor message. This is expensive and not always effective (e.g. the USA remains plagued by rumors regarding childhood vaccinations, despite concerted advocacy through the media that it is false [17]) especially in culture or in situations where people mistrust the authority or its motives. Tripathy et al. [30] studied decentralized and semi-centralized techniques to combat rumor. The decentralized method called *Delayed start model*, models a situation where a local authority might discover a rumor n days after it starts and decide to spread an anti-rumor. This is a purely reactive methodology where a single agent, perhaps a local authority, takes cognizance of a rumor and decides to act against it. The semi-centralized method called *Beacon model*, models a situation where a set of vigilant agents, *beacons*, are on the lookout for the spread of rumors. Once a beacon receives a rumor, it immediately starts spreading anti-rumors to combat the rumor. The Beacon model is somewhat more proactive than the Delayed start model. In the Beacon model we assume that there is a loose confederation of agents, perhaps a set of local authorities, that realize the danger of rumor spread and have planted listening posts in the social network. These listening posts are like sleeper agents, they come alive only when they detect a rumor. In this article, we study the Delayed start model and the Beacon model more rigorously using a number of new metrics (explained in Section 2). We also study a fully decentralized model of an enlightened citizenry vigilant against rumors, called the *Neighborhood model*, in which any user with some probability can detect the rumor on receiving it and decide to warn his or her contacts about the spread of the rumor. The difference between the Neighborhood model and the previously studied models is that each agent of the social network can independently decide to start combating the rumor.

The models we study are Markovian in nature and so their evolution can be characterized by simple differential equations. This approach, known as the mean field approach, follows from an influential work of Kurtz [15] that shows the connection between differential equations and limit laws for Markov process. In Section 3 we present mean field characterizations of our three models and validate the quality of the approximation they provide. This allows us to study the anti rumor models at scales that would be prohibitively difficult to simulate.

1.3 Related Work

Rumor spread bears strong similarities to the virus spread since the spreading mechanisms are similar e.g. the Susceptible Infected Susceptible (SIS) model and the Susceptible Infected Recovered (SIR) model [7, 28, 36], the Voter model [5, 16], the Independent Cascade (IC) model [13] etc. We wish to clarify an important difference between our model and virus spread models. Although both, rumor & virus, spread through social (or physical) contacts containing them requires fundamentally different strategies. Rumor can be combated by spreading messages (anti-rumor) on the same network, whereas combating viruses requires vaccination of each individual who is to be protected. Hence, unlike the virus spread case, the same social network used by a rumor can be leveraged to fight against it.

The problem we study is also very similar to the well studied problem of competitive viral marketing [3, 4, 29]. In our setting, there is a competition between two processes, the rumor process and the anti-rumor process. Similarly, in competitive viral marketing there is competition between two products. But the key difference between these two domains is encapsulated in our assumption thate rumor will eventually be removed by the anti-rumor because there is some essential distinction between falsehood and truth. Although this is definitely not true for all rumors and all situations we believe that it covers a large and significant part of the space. In the case of competitive viral marketing there is no authority who can unequivocally decide which product is superior. Hence, in our models anti-rumor eventually cleans the network from rumor, whereas in competitive viral marketing situation a stable state is reached when all parties settle on a market share.

The problem of characterizing the rumor process in a social network has been studied (see e.g. [36]) using an epidemic like model called SIR model, however the rumor-control problem has not received wide attention. Habiba et al. [11] studied the problem of identifying good blockers to minimize the rumor spread in the network. This paper also uses the Independent Cascade model for diffusion. Ur et al. [31] show how attack strength can be increased in a social network by controlling the hubs of the network. Webb et al. [33] proposed a technique to find spammer profiles by installing social Honeypots. They found that spam profiles follow a distinct temporal pattern and their geographical locations overlap with the target location. Kimura et al. [14] studied the problem of minimizing the contamination spread in network by blocking the links of the network. In this paper we study different ways of combating rumors in social networks.

We would like to observe that while sharing several aspects with the papers described above, our work's first main original contribution is the idea that rumor can be combatted with rumor-like anti-rumor messages. Our second main original contribution is the formulation of the three anti-rumor spread models described above and a set of metrics for studying these models which will be discussed in detail in the next section.

2 Models and Metrics

Basic Notation We assume that the online social network is modeled as a directed graph $\mathcal{G} = \{V, E\}$. For each node V_i, the immediate neighbors are represented by set N_i and n represents the total number of nodes in the network. A variable s_i, the status of the node, is maintained for each node V_i. A node i with $s_i = 0$ is a susceptible nodes. This nodes is yet to believe the rumor or anti-rumor. If $s_i = 1$, then node i is an infected node. Such nodes, after believing the rumor, will spread the rumor in the network. When $s_i = 2$ we say i is a cured node whereas if $s_i = 3$ we call i a vaccinated node. Cured nodes are those nodes which were once infected and now believe the anti-rumor. Vaccinated nodes are those susceptible nodes who now believe the anti-rumor but were never infected with the rumor. These two types of nodes, cured and vaccinated, will spread the anti-rumor message in the network. $s_i^{(t)}$ denotes the status of node at time t.

2.1 Rumor spread model

To model rumor spread process, we use the *Independent Cascade (IC) Model* with a little variation. This is a well studied model for diffusion in social networks [13, 14, 24]. In the IC model, each node gets a single chance to infect its neighbors, whereas in our setting each node gets multiple chances to infect its neighbors. Therefore, the rumor can spread very fast. We call this the *Multi Try Independent Cascade Model* (MTICM). In MTICM, at time t, each infected node V_i tries to infect each of its uninfected neighbors $w \in N_i$ and succeeds with probability p_1. The rationale for choosing this model of for rumor spread is that this model is guaranteed to spread the rumor throughout the network. And in fact this model succeeds in spreading the rumor through the network in a relatively short time (see e.g. the work by Chierchetti et. al. [6]). This makes this model a pessimistic estimate of the power of rumor, and hence a more difficult adversary to deal with than other models, some of which cannot even guarantee network-wide spread of the rumor (see e.g. Sudbury's work on the SIR model on complete graphs [27].) The hope is that if if we can combat rumor which is spreading by this model then the rumor spread using a less powerful model can also be easily combated. On the optimistic side, we also use the MTICM to model the spread of anti-rumor. This assumes that an anti-rumor spreader, who is essentially an altruistic person, will be persistent in his or her actions. We postpone for the future a more sophisticated study that models closely the variations in altruistic behavior of real world actors.

Through the course of our experiments we have taken p_1 to be 0.01. If it succeeds then w will become infected at time $t + 1$. The process starts with 10 random infected nodes. These choices of parameters are admittedly arbitrary as is the assumption that these parameters are uniform across nodes and across time. Deriving meaningful values of such parameters is a research problem in itself, one that we do not intend to focus on in this paper. Our goal is to demonstrate the fundamental characteristics of the process we study with parameter values that satisfy our general intuition.

2.2 Anti-rumor spread models

In this subsection, we present the anti-rumor spread models. In our previous study [30] we presented preliminary analysis of the first two models *Delayed start model* and *Beacon model*. For making this article self-contained, we briefly explain these models. In this work, we present a new model and also evaluate the older models on larger dataset and the new metrics proposed in next subsection.

Delayed Start Model: Here we model the situation that an authority with limited jurisdiction detects the spread of rumor and then combats it by starting an independent cascade from a randomly selected infected node. We contend that there will always be a time lag between the start of rumor and its detection (and hence, the start of the anti-rumor). We parametrize our model with this delay, represented by d. Fig. 1(b) pictorially depicts the Delayed start model. The checkered node is the previously infected node from where the anti rumor starts. The information goes to all neighbors: infected (node A) as well as uninfected (Node B). The process starts from a single random infected node after a delay time d.

Beacon Model: Between the time an authority detects the spread of rumor and decides to combat it, the rumor continues apace. In order to proactively combat rumors, authorities may embed agents in the network that are capable of detecting the spread of rumor and are authorized to start anti-rumors as soon as they detect rumor. We refer these agents as *beacons*. The beacons spread the anti-rumor according to the Multi Try Independent Cascade Model (MTICM). Fig. 1(c) shows a Beacon model. Please note that in the current state of the network, beacon B1 will be inactive since it has not yet received the rumor. In Delayed start model, the starting time of anti-rumor process is fixed but here it depends upon the time when the beacon receives the rumor.

Neighborhood Model: In the previous models the anti-rumor originates from the nodes selected by some authority either before or after the rumor starts. In the current model any node V_i may decide, on receiving the rumor from a neighbor V_j, to refute it. This model is similar to the Beacon model. The difference lies in choosing the set of initial beacons. In the Beacon model, the initial set of beacons are chosen by some authority whereas in the Neighborhood model, the beacons are self created with some probability during rumor spreading process. The Beacon model with b number of beacons (out of total n nodes) is comparable in an expected sense to a Neighborhood model where a node refutes the rumor with probability $\frac{b}{n}$.

2.3 Metrics

In this section, we propose various metrics to evaluate and compare the efficacy of these models. We divide these metrics into two categories: *Time varying metrics* and *Lifetime metrics*.

2.3.1 Time varying metrics:

These metrics capture the temporal evolution of the system. We consider three time varying metrics:

Number of Infected ($I(t)$): This metric captures the number of infected nodes at time t. Mathematically, we can define this as: $I(t) = \big|\big\{V_i \mid V_i \in V(G)$ and $s_i^t = 1\big\}\big|$. It provides us with a handle on the rumor growth process. T_i denotes the time period for which a node V_i remains infected. T_i is defined as: max {t :$s_i^{(t)} = 1$ } - min {t : $s_i^{(t)} = 1$ }. Finally, $\Delta(t)$ is calculated as $I(t+1)$ - $I(t)$.

Number of Cured ($C(t)$): This metric captures the number of cured nodes at time t, i.e., the number of infected nodes who have accepted the anti-rumor and now recognize the rumor as false. Formally, $C(t) = \big|\big\{V_i \mid V_i \in V(G)$ and $s_i^t = 2\big\}\big|$.

Number of Vaccinated ($V(t)$): This metric captures the number of nodes who learn the truth about the rumor before the rumor reaches them and therefore, they are *vaccinated* against future encounters with the rumor. It can be defined as: $V(t) = \big|\big\{V_i \mid V_i \in V(G)$ and $s_i^t = 3\big\}\big|$.

Finally, the number of nodes which are not covered in the above metrics are known as **susceptible nodes** ($S(t)$). Formally, they are defined as $S(t) = \big|\big\{V_i \mid V_i \in V(G)$ and $s_i^t = 0\big\}\big|$. Anti-rumor process performs two operations, it makes the infected nodes cured and makes the susceptible nodes vaccinated. Similar metrics are used in several papers related to rumors, viruses and epidemics spread [18, 19, 34, 35, 36].

2.3.2 Lifetime metrics

These metrics capture the global properties and are calculated at the end of the process. We consider five lifetime metrics:

Duration of Infection ($D(G)$): This metric denotes the total time required for the anti-rumor process to kill the rumor process completely. $D(G) = \max\{t \mid \exists i : s_i^{(t)} = 1\}$. This metric is a well studied metric in epidemic and virus literature [10, 36].

Outbreak size ($R(G)$): Outbreak size captures the total number of nodes which are infected at some point of time. $R(G) = \left|\{V_i \mid \exists t \leq D(G) : s_i^{(t)} = 1\}\right|$ Grabowski et al. [10] also used this metric to measure the strength of epidemic spread where in SIS model.

Maximum infected time ($M(G)$): The *Maximum infected time* measures the life time of the rumor in the network, i.e., it measures the maximum duration for which any node continues to believe the rumor. It can be defined as $M(G) = \max\{T_i \mid \forall V_i \in V(G)\}$

Average infected time($A(G)$): The average infected time captures the average time for which the users continue to believe the rumor, i.e., $A(G) = \frac{1}{|V[G]|} \sum_{V_i \in V[G]} T_i$.

Point of Decline ($P(G)$): There are two independent cascade processes growing simultaneously: rumor process and anti-rumor process. If $\Delta(t) \geq 0$, then rumor process is growing and if $\Delta(t) < 0$, then rumor process is declining. Please recall $\Delta(t) = I(t+1)$ - $I(t)$. Point of decline measure the time point at which the number of users believing the rumor start to decline and therefore, marking the time when anti-rumor process gets stronger than rumor process. It can be defined as $P(G) = \min\{t \mid \Delta(t) < 0\}$. Note that the uniqueness of the point of decline is not *a priori* obvious. However, our studies have shown that once the number of rumor infected individuals starts decreasing, it does not increase again. This metric is also used in epidemic literature [10].

3 Mean Field Characterization

Clearly, our models of rumor and anti-rumor are Markovian processes. The influential work of Kurtz [15] demonstrated that the limit laws of such processes conform to the solutions of ordinary differential equations. This gave rise to the "mean field" way of analyzing Markov processes. The key insight can be loosely stated as follows: if we assume that neighborhood of a node behave like an "average" neighborhood, the evolution of this "mean field" process closely approximates the evolution of the Markov process.

Mean field theory has been a central tool for theoretical physics. Several foundation papers on social networks use this method e.g. Newman et al. [21] studied the small world model using mean-field, Barabási et al. [2] described mean-field rate equations for the scale-free model. In the study of sociological phenomena that involves diffusion on complex networks, systems like epidemic spread (e.g. Funk et al. [9]) and competitive viral marketing [29] have also been analyzed using mean-field theory. Of direct relevance to our current paper, is the work of Nekovee et al. [19] where mean-field theory was used to study the spread of rumor in a complex networks using the MakiThompson (MK) model. Sathe [25] also used mean-field methods to study the spreading behavior of rumor in LiveJournal.

Notation In order to describe the mean field equations for our models, we take into account the heterogeneity induced by nodes having varying numbers of neighbors by partitioning the sets of nodes with different states according to their indegree. The number of susceptible nodes of in-degree k at time t is denoted by $S_k(t)$ where k takes all values up to the maximum degree of the network. Similarly $I_k(t), C_k(t)$ and $V_k(t)$ denote the infected, cured and vaccinated nodes, respectively, at time t that have indegree k. The indegree distribution of the graph is denoted by $\{F(k)\}_{k\geq 0}$ i.e. the probability of a node having degree k is $F(k)$. Recall that p_1 is probability of a node accepting the rumor and p_2 is the corresponding quantity for the anti-rumor message.

The probability of a node being infected at time t is the fraction of the nodes of the network that are infected at time t. We denote this quantity by $\theta_1(t)$ where

$$\theta_1(t) = \frac{\sum_k kF(k)I_k(t)}{n \cdot \sum_k kF(k)}$$

Similarly $\theta_2(t)$ is the probability that any given link points to an anti-rumor spreading node (cured or vaccinated)

$$\theta_2(t) = \frac{\sum_k kF(k)(C_k(t) + V_k(t))}{n \cdot \sum_k kF(k)}$$

Delayed Start Model Noting that a susceptible node is no longer susceptible once it accepts either the rumor or anti-rumor we have We solved the rate equations (**??**)-(**??**) numerically and validated the evolution by comparing the solution with the results of simulations conducted on a Barabási-Albert graph with 10^6 nodes. The results, for a delay of 160, are presented in Fig. 2(a). We note that the mean field evolution is a very close approximation to the simulated evolution.

Beacon Model For this model we also have to keep track of the unactivated beacons at time time. We denote by $B_k(t)$ the number of unactivated beacons of indegree k at time t. As in the Delayed start model we have

$$\begin{aligned} \Delta(S_k(t)) &= -\{1-(1-p_1)^{k\theta_1(t)}(1-p_2)^{k\theta_2(t)}\}S_k(t) \qquad (1) \\ \Delta(I_k(t)) &= \{1-(1-p_1)^{k\theta_1(t)}\}S_k(t) - \{1-(1-p_2)^{k\theta_2(t)}\}I_k(t) \qquad (2) \end{aligned}$$

The difference arises in $\Delta(C_k(t))$ and $\Delta(V_k(t))$. A beacon, once it accepts the rumor begins to spread the anti-rumor, i.e., it begins to behave like a cured node. Hence

$$\begin{aligned} \Delta(C_k(t)) &= \{1-(1-p_2)^{k\theta_2(t)}\}I_k(t) + \{1-(1-p_1)^{k\theta_1(t)}\}B_k(t) \qquad (3) \\ \Delta(V_k(t)) &= \{1-(1-p_2)^{k\theta_2(t)}\}S_k(t) + \{1-(1-p_2)^{k\theta_2(t)}\}B_k(t) \qquad (4) \end{aligned}$$

Also, an unactivated beacon gets activated once it accepts the rumor or, by default, once it accepts the anti rumor, i.e.

$$\Delta(B_k(t)) = -\{1-(1-p_1)^{k\theta_1(t)}(1-p_2)^{k\theta_2(t)}\}B_k(t) \qquad (5)$$

Solving these equations numerically for the case where there are 500 beacons in a 90,000 node graph mentioned above and plotting against the simulated value (see Fig. 2(b)) we find that the mean field results are a close approximation to the simulated values.

Neighborhood Model In the neighborhood model we add another probability, the probability of refutation. We denote it by p_3, the probability that a node recognizes the rumor and decides to refute it. In this model the rate equations for susceptible nodes and vaccinated nodes are as before.

$$\Delta(S_k(t)) = -\{1-(1-p_1)^{k\theta_1(t)}(1-p_2)^{k\theta_2(t)}\}S_k(t) \tag{6}$$

$$\Delta(V_k(t)) = \{1-(1-p_2)^{k\theta_2(t)}\}S_k(t) \tag{7}$$

The difference occurs in the infected and cured nodes. When a susceptible node accepts the rumor, it enters the infected set with probability $1-p_3$ else with probability p_3 it enters the cured set and becomes a spreader of anti rumor. Hence

$$\Delta(I_k(t)) = (1-p_3)\{1-(1-p_1)^{k\theta_1(t)}\}S_k(t) - \{1-(1-p_2)^{k\theta_2(t)}\}I_k(t) \tag{8}$$

$$\Delta(C_k(t)) = \{1-(1-p_2)^{k\theta_2(t)}\}I_k(t) + p_3\{1-(1-p_1)^{k\theta_1(t)}\}S_k(t) \tag{9}$$

In Fig. 2 (c) we show the comparison between the simulated evolution and the mean field solution of the Neighborhood model for refutation probability $p_3 = 500/90000$. The approximation is found to be close to the simulation.

Discussion The key idea of studying the mean field approach is to study the evolution of the rumor versus anti rumor process at scales which are too large to simulate conveniently. Similar approach also used by Nekovee et al. [19] to understand the dynamics of rumor spreading in scale free network. The quality of the approximation grows with the increase in scale. In the case of the Delayed start model, the quality of approximation deteriorates as the delay time increases. This is because the variability of the number of nodes increases with increase in delay. Fig. 3(a) shows the correlation between growth of infected nodes in simulation and in mean field solutions. We can see that correlation coefficient is deteriorated as the delay time increases. Similarly in Fig. 3(b) we can see that the mean-squared error increases as the delay time increases. The quality of the approximation is poorer for the beacon model and the neighborhood model because apart from the size of the network, both these models have another scale parameter: the number of beacons and the probability of refutation, respectively. As these parameters grow the approximation improves. Fig. 4 shows the quality of approximation for the growth of infected node in the Beacon model. In Fig. 4(a) we find that as the number of beacons increase the correlation between the simulated results and mean field results increase. Fig. 4(b) shows decrease in the mean squared error between simulation and mean field solutions as the number of beacons increase. The mean field equations for the Neighborhood model behave similar to that of the Beacon model. In the next section we present results from simulation on the Synthetic graph described above and a network derived from Twitter. We supplement these results by deriving trends from the mean field results for various metrics on Synthetic data.

4 Experiments and Analysis

In this section we discuss our simulation results. In all our results, the values used for the analysis are averaged over 50 runs. We also present the results for mean field models and showcase that in most cases our simulation model and analytical model are extremely close. In some cases, we observe a deviation in actual values, however, the observed growth rate and trend are still consistent.

4.1 Data sets

Twitter Data We used Twitter's API to crawl the Twitter network. We started from the seed user *DLF IPL*, the official Twitter account of The Indian Premier League (IPL), an Indian cricket league sponsored by Delhi Leasing and Finance (DLF). Considering each user as a node and a "follower" relationship as a directed link, we created a directed network containing 100,500 nodes and 2,465,836 edges. Since in this work we intend to find the impact of the proposed anti-rumor models for combating rumor, therefore, we are interested in the case where a rumor can possibly spread through entire network. Therefore, we have taken the maximal strongly connected component (with 89,999 nodes) and pruned the remaining 10,501 nodes (and corresponding edges) from the original data set. The reason this pruning is done is to ensure that each node can potentially send data to every node under consideration. If this were not done then there would be nodes which can either only receive messages from some part of the network or can only send messages to some part of the network. We pruned away these nodes to get rid of degenerate behavior. After pruning, the

Twitter graph contained 89,999 nodes and 2,262,104 edges with average degree 24 and diameter 17.

Synthetic Data We use the Barábasi-Albert (BA) model [1] to generate the synthetic data. The BA model is a random graph model which generates a scale-free graph (means the graph with power-law degree distribution) by incorporating growth and preferential attachment. The key idea behind preferential attachment is that the more connected a node is, the more likely it is to receive new links. In BA model, starting with an initial graph of a few nodes, a new node is added at each time step and is connected to other nodes based on their degrees i.e., higher degree nodes have a higher probability of connecting from new node.

A scale-free graph of size 90,000 (same size as that of the Twitter data) and having power-law coefficient of 2.54 is generated using BA model . We use BA model because, it is a widely accepted model for social network analysis due to its scale free properties which is also there in our Twitter graph and it is easy to derive the metrics for this model without simulation (using mean-field equation).

We have used NetworkX [12] for creating networks for both the Synthetic as well as the Twitter data. NetworkX is a Python package for creation, manipulation, and study of the structure, dynamics, and functions of complex networks.

4.2 Delayed Start Model

To get a preliminary idea about the efficacy of the Delayed start model, we looked at the behavior of the rumor process for the Twitter graph as well as the Synthetic graph. The results for the Delayed start model with a delay time of 40 for both the Twitter graph and the Synthetic graph are shown in Fig. 5. Growth of the infected nodes are similar for both the graph.

The key thing to note in Fig. 5 is the presence of a single point of decline for both networks. The growth of number of infected nodes, $I(t)$, is slow to start with but as the time goes the $I(t)$ increase very fast. Initially there are few nodes who believe the rumor but as the number of nodes who believe the rumor grows, the probability of accepting a rumor for a susceptible node increases which makes the rate of growth rise. This is because with the increase of infected nodes, the number infected neighbors of the susceptible nodes also increase. So, the probability of accepting a rumor increases. After certain time point the number of infected nodes start to decline, that point is called the point of decline $P(G)$. Beyond this point, not only does the rumor-affected population not grow it declines very rapidly, because the anti-rumor process is growing inside and outside the rumor process using the neighbors links. Once the anti-rumor process out-perform rumor process, in the succeeding steps the number of susceptible neighbors of the rumor accepted nodes decreases, which makes the growth reduce further. So after the point of decline, it requires very little time to completely remove the rumor. In combating rumor strategies, the point of decline is an important parameter. A detailed study about this parameter is presented in the succeeding sections.

The sharp behavior noted above is further reinforced by studying the ratio of the number of anti-rumor accepted nodes to the number of rumor accepted nodes, i.e., $C(t)+V(t)$ to $I(t)$ (Fig. 6(a) and Fig. 6(b)). It is clear that the ratio lies between zero and infinity, zero until anti-rumor process starts and infinity when all rumor infected nodes are cured. After the anti-rumor process starts, we see a sharp growth in this ratio, because the anti-rumor process starts killing the rumor process which breaks the growth of rumor process. However, in the Twitter we can observe that for higher values of delay time, after initial spurt the growth of the ratio slowsdown. The reasons for this behavior may be stated as, when the delay time increases the number of infected nodes increases and in the Twitter graph some infected nodes remain infected for longer period of time, because a lot of nodes in the Twitter graph have in-degree 1. These nodes get less chance to be cured than a higher in-degree node. Finally, Fig. 6(c) shows the results obtained for the metric by solving mean field equations. The analytical solution in this case matches not only in trend with the simulation results but the actual values are very close. We would like to point that the trends in the Twitter graph very closely match with the trends shown in the mean field solutions.

Next, we study how average infected time, $A(G)$, varies with the delay time. The results are shown in Fig. 7(a). The trend is very intuitive, as the time to start anti-rumor process increases, $A(G)$ increases sub-linearly but for the higher value of delay the average infected time grows much faster. This is because, the rumor accepted nodes remain infected for a longer period and also get more chances to infect other nodes, hence the delay time is an important parameter for controlling rumor. There is hardly any difference between the Twitter data and the Synthetic data. The $A(G)$ values of the Twitter graph are slightly more than that of the Synthetic graph, which indicates that in the Twitter graph some nodes remain infected for a longer period of time because of the lower in-degrees of these nodes. We remind the readers that both our graphs, the Twitter graph and the Synthetic graph, follow power-law degree distributions with power-law coefficients 2.34 and 2.54 respectively. Being a real graph, the Twitter graph contains a lot of variation in the nodes degree. The maximum infected time, $M(G)$, (Fig. 7(b)) grows very slowly. This implies that the anti-rumor is able to arrest the rumor effectively in the sense that no single node's duration of believing the rumor is disproportionately large because of the increase in delay, even though the entire populations average belief time does get affected significantly by delay in starting the anti rumor. Fig. 7(d) present the change in point of decline, $P(G)$, with increase in delay. It is evident that for smaller values of delay, $P(G)$ increases almost linearly, i.e, the time taken to tackle the rumor depends linearly on the delay. However, for larger values of delay, $P(G)$ increases but the increase is very slow because for higher delay the rumor spreads through the entire network and then stops because there are no nodes left to infect, so when the anti-rumor process starts it is the only process which is running. Therefore tracking this metric for higher values of delay does not make sense. Fig. 7(c) shows the trend in out-break size ($R(G)$). We can see a sharp behavior for smaller values of delay. i.e., a small increase in the delay time will cause a exponential rise in the infection, which shows the importance of delay time in the combating process. Finally, Fig. 7(e) and Fig. 7(f) shows the results for $P(G)$ and $R(G)$ for mean field model. The mean field model conforms to the simulation observations for the Synthetic as well as the Twitter graph. Therefore, the lifetime metrics can be studied for large scale graph with only mean field equations.

We know that out-break size, $R(G)$, represents the total number of infected nodes in the combating process. However, to study temporal behavior the time varying metric, $\frac{I(t)}{R(G)}$ is better suited. This ratio captures the growth of the rumor relative to its overall reach. The values of 1 at time t implies that, the rumor has attained the maximum strength at t. In Fig. 8, we have shown the results for this ratio. In Fig. 8(a), we can see that even if both rumor and anti-rumor process start at same time, at least 70% of the total out-break size, $R(G)$ is found to be infected at a particular period of time. In other words, this implies that even after the rumor begins declining, it is still able to infect almost one-third as many nodes as it did when it was on the rise. If the ratio touches 1 then the anti-rumor process is insignificant, since all nodes are already infected. These curves also show the single point of decline. Comparing the results of the Twitter data and the Synthetic data we observe that the maximum value of the ratio increases with increase in delay time in both case but the maximum values occur earlier in the Twitter graph compared to the Synthetic graph which again supports the faster spread in the Twitter graph. Fig. 8(c) presents the results of the mean field model. The overall trend for different delay factor matches with the Twitter as well as the Synthetic graph trends. However, in this result we can see that the maximum value of the ratio is less compared to the Twitter graph as well as to the Synthetic graph. The reason may be, the mean field is an approximation and the approximation is closer to the original solution for larger values of the scale parameter i.e. the number of nodes.

4.3 Beacon Model

We start with studying the growth of infected nodes $I(t)$. The results for synthetic and real graph are shown in Fig. 9. In this case we have used only 10 beacons. As in the Delayed start model, a single point of decline is also seen in the Beacon model. In Fig. 9, we can see that the growth of $I(t)$ is initially slow to start with but after certain point it grows exponentially. Because as the number of infected nodes increase, the probability of accepting a rumor also increase. That means, if the beacons are able to detect the rumor early then we can control the rumor easily. We also observe that after the point of decline there is sharp decline in the growth of infected nodes. That is, once the anti-rumor process gets hold of rumor process, it decimates it quickly. The growth of the rumor process for the Twitter graph and the Synthetic graph are almost same, but in the Twitter graph growth rate is slightly higher than the Synthetic graph.

Similar to the Delayed start model we have looked at the values of $S(t)$, $C(t)$ and $V(t)$ together with $I(t)$. The resultant graphs are shown in Figure-10.

Initially, the ratio $\frac{C(t)+V(t)}{I(t)}$ is zero because there is only one process (rumor) until the beacon(s) get activated, after that the ratio starts to increase. When the ratio approaches 1, then there is true competition, since the number of rumor and anti-rumor spreaders are same. Hence, from that point the increase in ratio slowsdown. The bend in the upper part of the curve in Fig. 10(a) suggests that there are a few nodes in the Twitter network which remain infected for a long period of time. The variation of the ratio in the Twitter graph Fig. 10(a) is almost similar to that of the Synthetic graph Fig. 10(b), but growth of the ratio is slightly higher in the Twitter graph. The increase in the ratio may be due to decrease in the number of infected nodes (oppositely increase in the number of cured node). Fig. 10(c) shows the results derived using the mean field equations. It is interesting to note that, the mean field solutions not only match the simulation results for the Synthetic graph but are quite close to the results for the Twitter graph. The small deviation from the mean field solution is due to noisy real world data, but looking at the overall trends the mean field equations provide good approximations.

Lifetime metrics for the Beacon model are shown in Fig. 11. We start with average infected time $A(G)$ shown in Fig. 11(a). The $A(G)$ value decreases as the number of beacons increase, but the decrease is slower for a higher number of beacons. This observation can be explained as follows: the anti-rumor process begins by the beacon nodes, when a beacon is activated the number of nodes who accept the anti-rumor grows centered around that beacon node, which results in a component formed by that beacon node. As the number of beacons increase these components start overlapping which make the effectiveness of some of the beacons to reduce. The Synthetic graph also shows a similar trend as the Twitter graph, but the $A(G)$ values are slightly higher in the Synthetic graph because of relatively slower growth process in the Synthetic graph which is already seen in Fig. 9. The maximum infected times, $M(G)$, for different number of beacons are shown in Fig. 11(b). The maximum time of infection displays a gentler trend. As the number of beacons increases the $M(G)$ values decreases but slowly, which means even though we start the anti-rumor process early by planting more number of beacons, there are some nodes which remain infected for a longer period of time. Both the Twitter graph and the Synthetic graph follow similar trends. Almost similar results were observed for the other two metrics point of decline $R(G)$ and out-break size $P(G)$ (Fig. 11(c) and Fig. 11(d)). Both $P(G)$ and $R(G)$ values decrease with increase in number of beacons. Overall trends between the Twitter graph and the Synthetic graph are similar but slightly higher $P(G)$ values and lower $R(G)$ values are observed for the Synthetic graph. This further added to the slower growth process in the Synthetic graph. The results for the $P(G)$ and $R(G)$ values of the mean field equations are shown in Fig. 11(e) and Fig. 11(f). The overall trend of mean field model matches with the Synthetic as well as the Twitter data. The results show that, after a sharp decrease in the values as the number of beacons increase the curve is almost flat, which suggest that there is some upper bound on the use of number of beacons in the combating process after that increasing the number of beacons does not contributes a significant change in the results.

Next we study the ratio $\frac{I(t)}{R(G)}$, i.e., the fraction of the outbreak size achieved at time t. We see that although the beacons sit on the boundary of the rumor's spread (by their very definition), they do not effectively encircle the rumor: even after the fraction starts decline, a large number of nodes do get infected. In fact in the Twitter graph, Fig. 12 (a), around 40% of the total number of infected nodes infect after the decline starts (for the cases of 50-beacons and 100 beacons). This metric is insignificant for the cases where the ratio is close to 1. The results for the Synthetic network Fig. 12(b) are almost similar to that of the Twitter network. The mean field solutions for higher number of beacons Fig. 12(c) are not only matches with simulation using the Synthetic network but also it is very close to the Twitter network as well.

4.4 Beacon Vs Delayed Start Model

Next we compare the performance of the Beacon model with the Delayed start model. The basic difference in the two models is how the anti-rumor process begins. In the Delayed start model we fix a time when the anti-rumor process begins whereas in the Beacon model, we fix the locations of the beacons but cannot know for sure when they are going to be activated. We have found that average time for the beacon (in case of 1-Beacon model) to activate is close to 45 for the Twitter graph and close to 60 for the Synthetic graphs. Therefore, it is meaningful to compare the 1-Beacon model and the Delayed start model with delay time 45 for the Twitter graph and with delay time 60 for the Synthetic graph. We observe that the out-break size, $R(G)$, value for the 1-Beacon model

(Fig. 11(c)) is around 68,000 whereas for the Delayed start model (having delay time 45) (Fig. 7(c)) the value is much higher than 68,000 (in fact it is close to 82,000). Similarly, the average infected time, $A(G)$, in case of the Beacon model is also less than the Delayed start model. These observations can be explained as follows: a beacon is activated on the way of the rumor growth process, i.e., the anti-rumor process starts at the edge of rumor growth process and hence it will limits the growth of rumor process. In case of the Delayed start model, the anti-rumor process begins at any infected node after a particular delay time. The neighbors of that node may be infected but not effective, because probably they could not able to further spread the rumor as their neighbor are already infected. Therefore from these experimental observations we may conclude that the Beacon model is able to combat the rumor more effectively. Similar results are obtained for the Synthetic data (comparing the 1-Beacon model and the Delayed start model with delay time 60). Examining the plots for the mean-field solution Fig. 11(e) and Fig. 11(f) for the Beacon model and the Delayed start model Fig. 7(e) and Fig. 7(f) we observed that in the Delayed start model as the delay time increases, the point of decline, $P(G)$, values increase linearly but not so sharp behavior seen in the Beacon model. Similarly, the sharp behavior in out-break size, $R(G)$, for the Delayed start model is not seen in the Beacon model. From these observations, we can conclude that the Beacon model is more effective than the Delayed start model under similar settings.

4.5 Neighborhood Model

In the Neighborhood model, a user may detect a message as rumor with some probability while receiving it and decide to warn its neighbors about the spread of rumor. If we compare this to the Beacon model it is as if a node decides to be a beacon rather being decided a priori. We can compare the Beacon model with b beacons with the Neighborhood model with probability of detection of rumor as $\frac{b}{n}$ where n is the size of the graph.

First, we study the growth of infected nodes for both the Twitter and the Synthetic graph. The graphs are shown in Fig. 13. In this case we consider the probability of detecting the rumor is $\frac{10}{90000}$, which is equivalent to the Beacon model with 10 beacons. The overall growth process is similar to that of the other two models, i.e., there is single point of decline and there is also a sharp behavior after the point of decline. The rate of growth for the Twitter graph is faster than the Synthetic graph. Comparing this result with that of the Beacon model Fig. 9, we find that rate of rumor growth in both the models is almost similar and in fact slightly lower for the Neighborhood model. This is quite interesting because, in the Beacon model we required some authorities to select the beacons nodes but for the Neighborhood model there are no authorities involved. The slight improvement in the Neighborhood model is observed because higher degree nodes are more likely to become beacons and in social network the higher degree nodes are the most influential node in term of information spread [8]. On the other hand in the Beacon model there is no control over the selection of the beacon nodes. Therefore, from these experimental observation we may conclude that the Neighborhood model is more efficient in arresting rumor.

We study other time varying metrics as shown in Fig. 14. The behavior of the ratio $\frac{C(t)+V(t)}{I(t)}$ is almost similar to the Beacon model. Initially, the ratio is zero because there is only one process (rumor) until any node(s) refute the rumor, after that the ratio starts to increase. Comparing these results with the results obtained by the Beacon model, Fig. 11, we have found that the time required to activate a beacon is slightly greater than the time required by the first node to refute the rumor. The growth rate of the ratio is much faster initially but slowdown later in the Twitter graph, Fig. 14(a), which is not observed with the Synthetic graph, Fig. 14(b). This may be due to the noise in real data. A lot of nodes in the Twitter graph are having in-degree 1. Therefore, these nodes take more time to be cured. The mean field solution, Fig. 14(c), is similar to the solution obtained by simulation.

Next, we study the behavior of the life time metrics. The results are shown in Fig. 15. The average infected time $A(G)$, Fig. 15(a), and the maximum infected time $M(G)$, Fig. 15(b), decrease very slowly with the increase in refutation probability. By comparing the results of the Twitter graph with the Synthetic graph we observe that, the $A(G)$ values of the Twitter graph are less than that of the Synthetic graph whereas the $M(G)$ values show the completely opposite behavior. This may be explained as: in the Twitter graph both rumor and anti-rumor spread very fast. Therefore, even when a larger number of nodes are infected, the infection lasts for a very short duration of time compared to the

Synthetic graph. However, there are some nodes in the Twitter network which are very loosely connected to other part of the network. When such nodes get infected, this take a long time to be cured which increases the maximum infected time. Comparing these results with the results obtained by the Beacon model Fig. 11(a) and Fig. 11(b), we find that both the models give similar results. Similar observation can be made for the point of decline $P(G)$ metric in Fig. 15(d). The $P(G)$ values decrease slowly and the value is greater for the Synthetic graph than for the Twitter graph. The outbreak size, $R(G)$ in Fig. 15(c) of both the graphs are quite close and decrease with the increase in the refutation probability. The mean field solutions Fig. 15(e) and Fig. 15(f) match the simulation quite closely.

The last property we study is the ratio between number of infected, $I(t)$ and out-break size, $R(G)$. The results for both the Twitter graph and the Synthetic graph are shown in Fig. 16. As we can see, the behavior of the ratio $\frac{I(t)}{R(G)}$ is almost similar to that of the Beacon model (Fig. 12) and the Delayed start model (Fig. 8). However, by looking at these results closely we can see slightly lower values for the Neighborhood model. The reason for this is already discussed previously, i.e., the higher degree nodes are more likely to become beacon and also in the Neighborhood model the anti-rumor process is growing inside the region of the rumor process. The solutions of mean field are shown in Fig. 16 (c) which resembles very closely with the Synthetic as well as the Twitter graph for higher values of refute probability.

4.6 Observations

In this article, we study three natural anti-rumor processes to counter the rumor process. Simulating in a real graph (derived from Twitter) as well as synthetic graph (generated using Barabási-Albert model), we study the temporal evolution as well as global properties of these anti-rumor processes. We have also presented mean field equations that characterize the system. In all the three models, we observe a sharp growth in the rumor process after a slow start. However, once the growth of rumor starts decline, within a very short period of time the rumor is completely removed from the network. This observation suggest that, once we detect the rumor (no matter in which ways) due to fast growth power of social networks we can able to conquer the rumor. However, the life time metrics behave differently for different models. In the Delayed start model, we find the point of decline $P(G)$ grows linearly and the out-break size $R(G)$ grows exponentially with delay time, but the other two models show different behaviors. In the Beacon model we observe that, the $P(G)$ and $R(G)$ values decrease very slowly as the number of beacons increase. Because, when a beacon is activated the number of nodes who accept the anti-rumor grows centered around that beacon node, which results in a component formed by that beacon node. As the number of beacons increase these components start overlapping which make the effectiveness of some of the beacons to reduce. Comparing the results of 1-Beacon model with the Delayed start model with delay 45 (average time for the beacon to active) we find that the average infected time $A(G)$, the out-break size $R(G)$ and the point of decline $p(G)$ values are less for the

1-Beacon model compared to the values for the Delayed start model with delay 45. The reasons for this can explained as: in the Delayed start model, the anti-rumor process starts from an infected node and that node may have already been infected for a long period of time. Therefore, the node lies in the region of infected nodes. Therefore, by starting the anti-rumor process from this node may decrease the number of infected nodes but not able to contain the rumor process by vaccinating susceptible node. In the Beacon model, the beacon nodes activate themselves while receiving the rumor, i.e., the anti-rumor process actually starts at the edge of rumor process. Therefore, the beacon nodes can effectively contain rumor by vaccinating susceptible nodes as well as curing infected nodes. In the Neighborhood model similar results are obtained, i.e. the $P(G)$ and $R(G)$ values decrease very slowly as the number of refutation probability increases. The Beacon model and the Neighborhood model show similar results and, even sometimes the Neighborhood model performs better than the Beacon model. The slight improvement may be because, in the Neighborhood model, a higher degree node has greater probability to become a beacon. However, even though the results for both the models are same, we strongly believe that in large social networking sites the Neighborhood model is most natural and effective way to combat rumor because, in case of the Beacon model some authorities are required to select the beacon nodes whereas, the beacons are self created in the Neighborhood model.

5 Conclusions

The main contribution of this paper is to study various anti-rumor strategies in online social networks. The guiding insight of our work is that since social networks span multiple nations with no governing authority, the only way that rumors can be quelled is by using the power of social relationships. In other words we study processes which are by nature decentralized and work in same fashion as rumor. We have studied a reactive situation where there is a time lag in the detection of rumor and a local authority attempts to stop the rumor by starting an anti-rumor once the rumor is detected. We found that the time lag is an important parameter. We also studied a proactive situation where beacons embedded in the network detect and fight rumor. Further, we studied the situation where individual citizens attempt to fight the rumor, no role of any authority is assumed and found that this way of fighting rumor is the most effective among other models. We believe our work is a first step in the direction of studying the efficacy of natural processes that can be employed to fight the spread of rumors in social networks, and may, by virtue of their distributed and organic nature succeed where authoritarian strategies have clearly failed.

References

[1] A. L. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286:509, 1999.

[2] A. L. Barabási, R. Albert, and H. Jeong. Mean-field theory for scale-free random

networks. *Physica A*, 272(1-2):173–187, 1999.

[3] S. Bharathi, D. Kempe, and M. Salek. Competitive influence maximization in social networks. In *WINE'07: Proceedings of the 3rd international conference on Internet and network economics*, pages 306–311, Berlin, Heidelberg, 2007. Springer-Verlag.

[4] T. Carnes, C. Nagarajan, S. M. Wild, and A. van Zuylen. Maximizing influence in a competitive social network: a follower's perspective. In *ICEC '07: Proceedings of the 9th international conference on Electronic commerce*, pages 351–360, New York, NY, USA, 2007. ACM.

[5] C. Castellano, D. Vilone, and A. Vespignani. Incomplete ordering of the voter model on small-world networks. *Europhys. Lett*, 63(1):153, 2003.

[6] F. Chierchetti, S. Lattanzi, and A. Panconesi. Almost tight bounds for rumour spreading with conductance. In *STOC '10: Proceedings of the 42nd ACM symposium on Theory of computing*, pages 399–408, 2010.

[7] D. J. Daley and D. G. Kendall. Epidemics and rumours. *Nature*, 204(4963):1118, December 1964.

[8] E. Even-Dar and A. Shapira. A note on maximizing the spread of influence in social networks. In *WINE'07: Proceedings of the 3rd international conference on Internet and network economics*, pages 281–286. Springer-Verlag, 2007.

[9] S. Funk, E. Gilad, C. Watkins, and V. A. A. Jansen. The spread of awareness and its impact on epidemic outbreaks. *Nat. Acad. Sci. Proc*, 106(16):6872–6877, April 2009.

[10] A. Grabowski and M. Rosinska. The sis model for assessment of epidemic control in a social network. *Acta Phys. Pol. B*, 37:1521–1536, May 2006.

[11] H. Habiba, Y. Yu, T. Y. Berger-Wolf, and J. Saia. Finding spread blockers in dynamic networks. In *SNAKDD'08: Proceedings of the 2nd international conference on Advances in social network mining and analysis*, pages 55–76. Springer-Verlag, 2010.

[12] A. Hagberg, D. Schult, and P. Swart. Networkx. High productivity software for complex networks., 2010. https://networkx.lanl.gov/.

[13] D. Kempe, J. Kleinberg, and E. Tardos. Maximizing the spread of influence through a social network. In *KDD '03: Proceedings of the 9th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 137–146. ACM, 2003.

[14] M. Kimura, K. Saito, and H. Motoda. Blocking links to minimize contamination spread in a social network. *ACM Transactions on Knowledge Discovery from Data*, 3(2):1–23, 2009.

[15] T. G. Kurtz. Solutions of ordinary differential equations as limits of pure Markov processes. *J. Appl. Probab*, 7:49–58, 1970.

[16] C. M. Mizell and L. M. Sander. A generalized voter model on complex networks. *J. Stat. Phys*, 136(1):59–71, 2009.

[17] C. Mooney. Why does the vaccine/autism controversy live on? Discover Magazine. Science, Technology and The Future, 6th May 2009.

[18] Y. Moreno, M. Nekovee, and A. F. Pacheco. Dynamics of rumor spreading in complex networks. *Phys. Rev. E*, 69(6), June 2004.

[19] M. Nekovee, Y. Moreno, G. Bianconi, and M. Marsili. Theory of rumour spreading in complex social networks. *Physica A*, 374(1):457–470, January 2007.

[20] T. N. Network. As phones jam, twitter, facebook save the day. Times of India, 14th July 2011.

[21] M. Newman, C. Moore, and D. Watts. Mean-field solution of the small-world network model. *Phys. Rev. E*, 84(14), April 2000.

[22] H. Nolte. Facebook and tinyurl.com - avoid scammers. http://www.examiner.com/, 25th May 2009.

[23] W. A. Peterson and N. Gist. Rumor and public opinion. *Am. J. Sociol*, 57(2):159–167, 1951.

[24] K. Saito, R. Nakano, and M. Kimura. Prediction of information diffusion probabilities for independent cascade model. In *KES '08: Proceedings of the 12th international conference on Knowledge-Based Intelligent Information and Engineering Systems, Part III*, pages 67–75, Berlin, Heidelberg, 2008. Springer-Verlag.

[25] S. Sathe. Rumor spreading in LiveJournal. Mini Project Report for Doctoral Course *Dynamical Networks*, School of Computer and Communication Sciences, EPFL, 2008. Downloaded from http://lanoswww.epfl.ch/studinfo/courses/Dynamical_Networks/.

[26] B. Stone and N. Cohen. Social networks spread Iranian defiance online., 15th June 2009. New York Times.

[27] A. Sudbury. The proportion of the population never hearing a rumour. *J. Appl. Probab*, 22(2):443–446, June 1985.

[28] R. Thompson, R. C. Estrada, D. Daugherty, and A. Clintron-Arias. A deterministic approach to the spread of rumors. Technical Report BU-1642-M, Mathematical and Theoretical Biology Institute, Los Alamos National Laboratatory, Centre for Non-Linear Studies, 2003.

[29] M. Tomochi, H. Murata, and M. Kono. A consumer-based model of competitive diffusion: the multiplicative effects of global and local network externalities. *J. Evol. Econ*, 15(3):273–295, 08 2005.

[30] R. M. Tripathy, A. Bagchi, and S. Mehta. A study of rumor control strategies on social networks. In *CIKM '10: Proceedings of the 19th ACM international conference on Information and knowledge management*, pages 1817–1820. ACM, 2010.

[31] B. E. Ur and V. Ganapathy. Evaluating attack amplification in online social networks. In *W2SP'09: 2009 Web 2.0 Security and Privacy Workshop*, Oakland, California, May 2009.

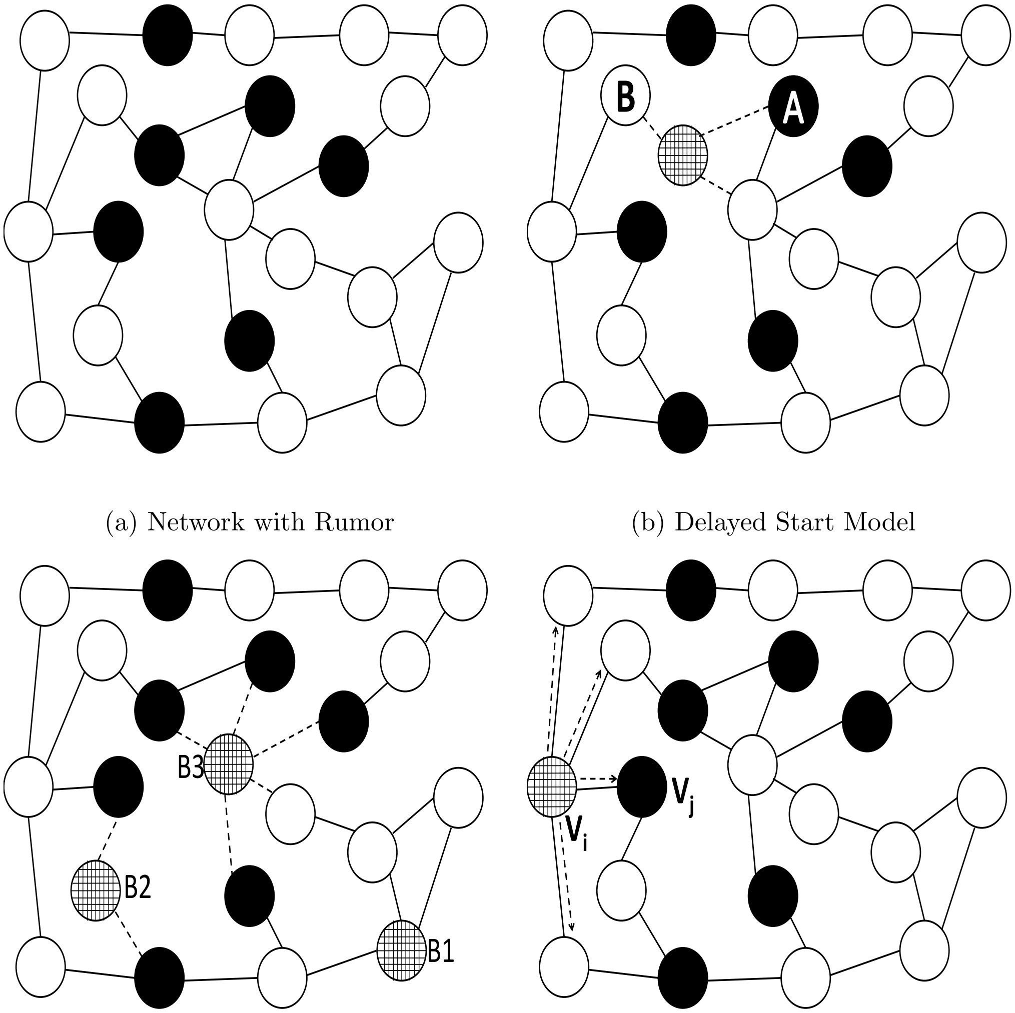
[32] Virgil. *Eclogues, Georgics, Aeneid 1-6.* Harvard University Press, 1916. Translated by H. R. Fairclough.

[33] S. Webb, J. Caverlee, and C. Pu. Social honeypots: Making friends with a spammer near you. In *CEAS '08: Proceedings of the 5th Conference on Email and Anti-Spam*, Mountain View, CA, August 2008.

[34] M. M. Williamson and J. Leveille. An epidemiological model of virus spread and cleanup. In *VB '2003: Proceedings of the 13th Virus Bulletin International Conference*, Toronto, Canada, 2003.

[35] D. H. Zanette. Critical behavior of propagation on small-world networks. *Phys. Rev. E*, 64(5), October 2001.

[36] D. H. Zanette. Dynamics of rumor propagation on small-world networks. *Phys. Rev. E*, 65(4), March 2002.



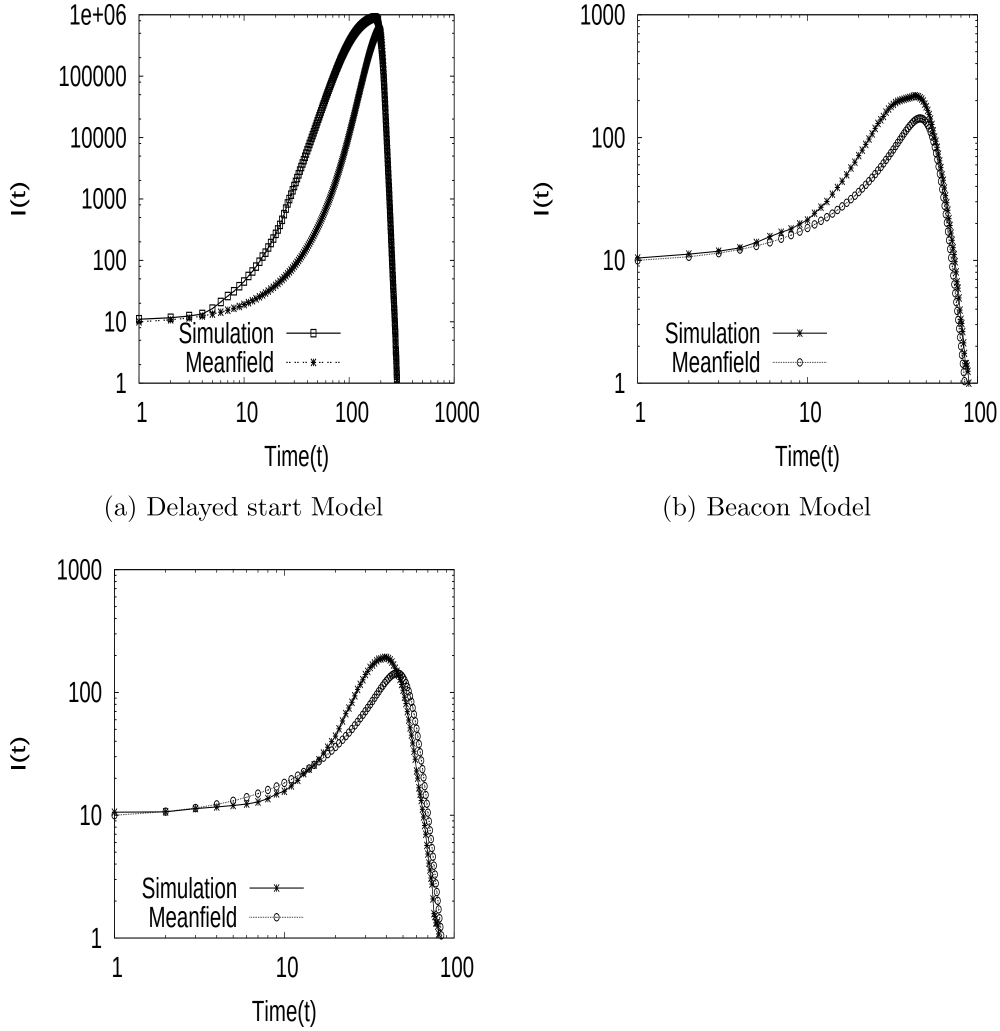
(a) Network with Rumor

(b) Delayed Start Model

(c) Beacon Model

(d) Neighborhood Model

Figure 1: Various Rumor and Anti-rumor models

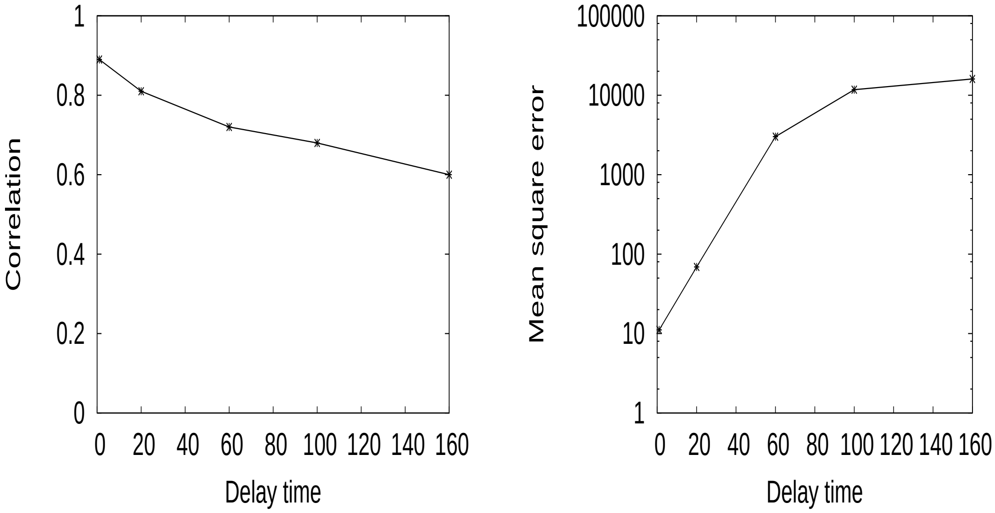


(a) Delayed start Model

(b) Beacon Model

(C) Neighborhood Model

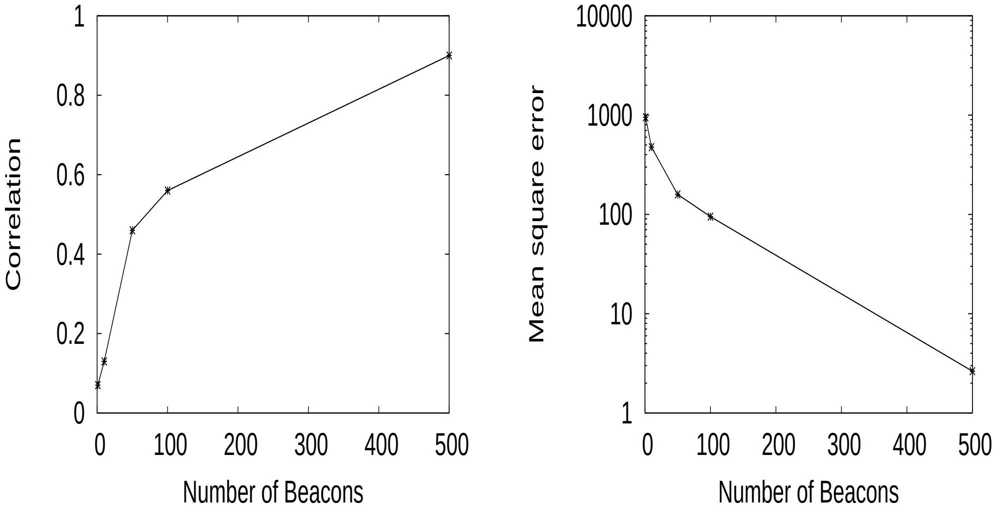
Figure 2: Simulation vs Mean-field evolution



(a) Correlation

(b) Mean squared error

Figure 3: Quality of Approximation: Delayed start model



(a) Correlation

(b) Mean squared error

Figure 4: Quality of Approximation: Beacon model

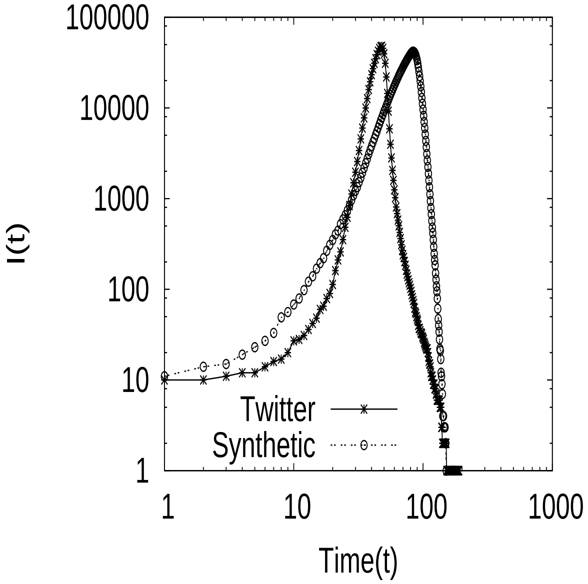
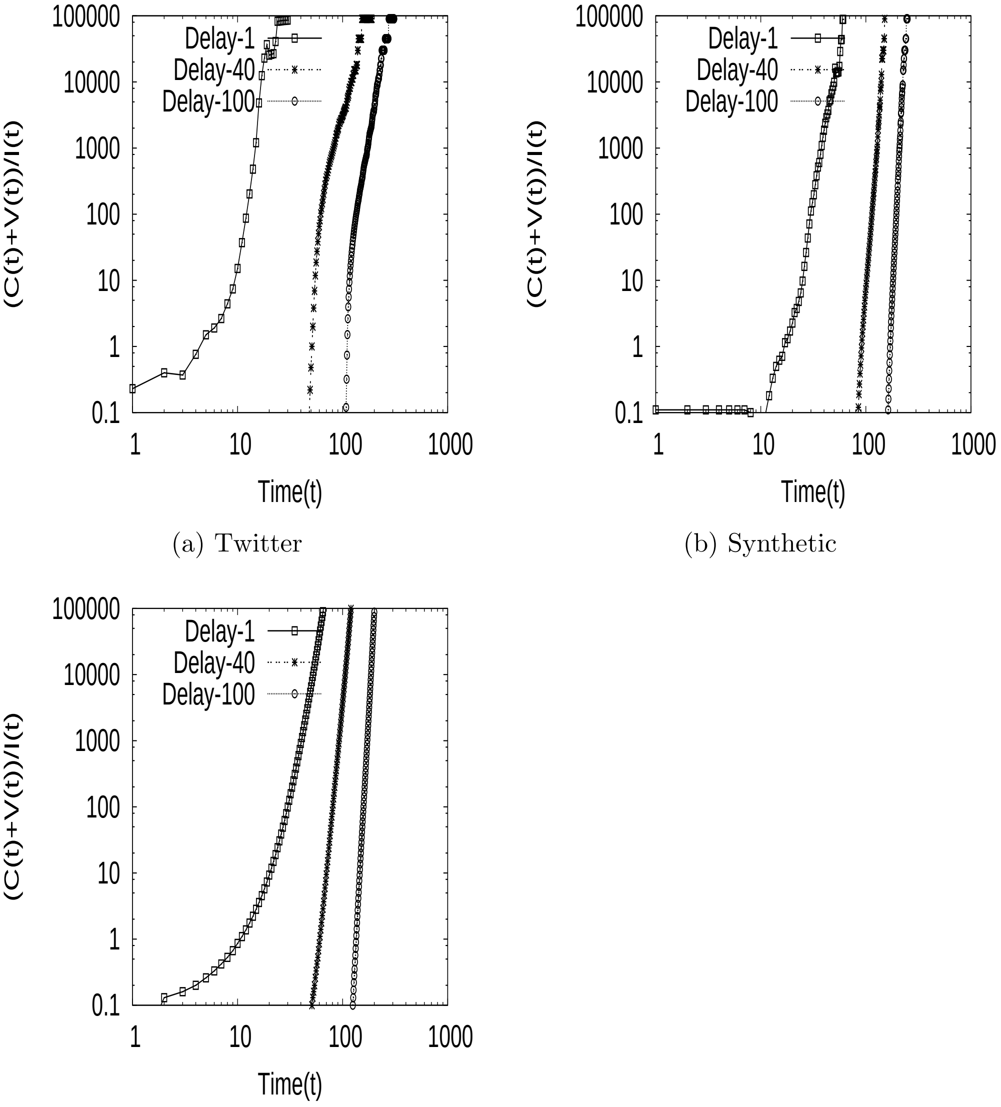


Figure 5: Rumor growth: Delayed Start Model

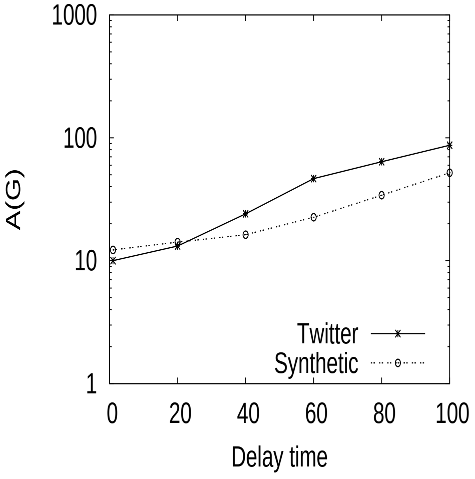


(a) Twitter

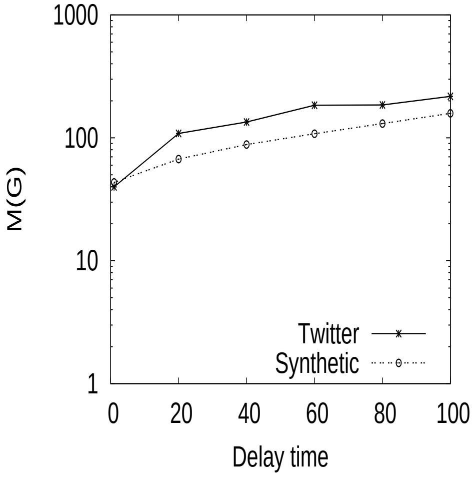
(b) Synthetic

(c) Mean field

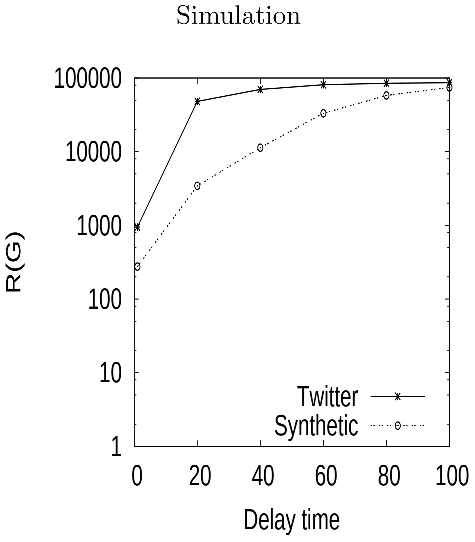
Figure 6: Behavior of time varying metrics: Delayed Start Model



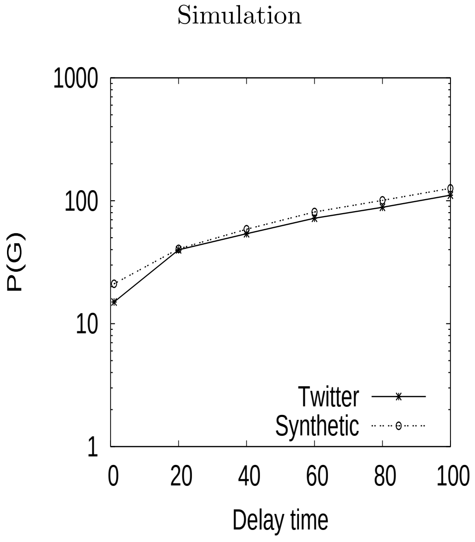
(a) Avg infected time



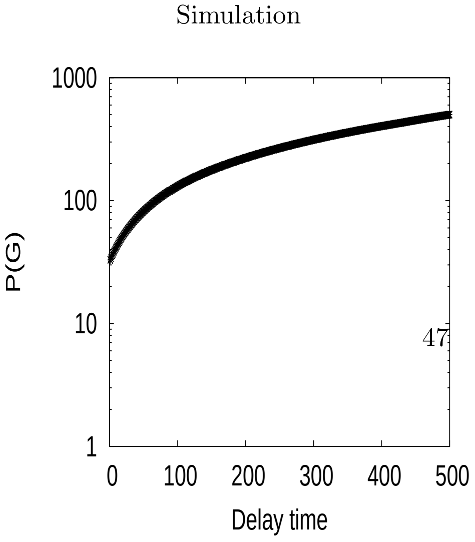
(b) Max infected time



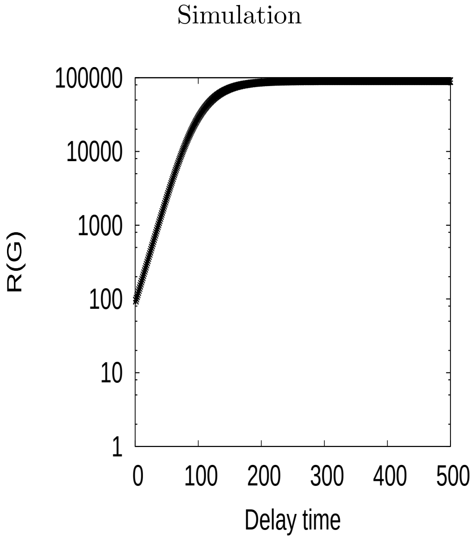
(c) Outbreak size



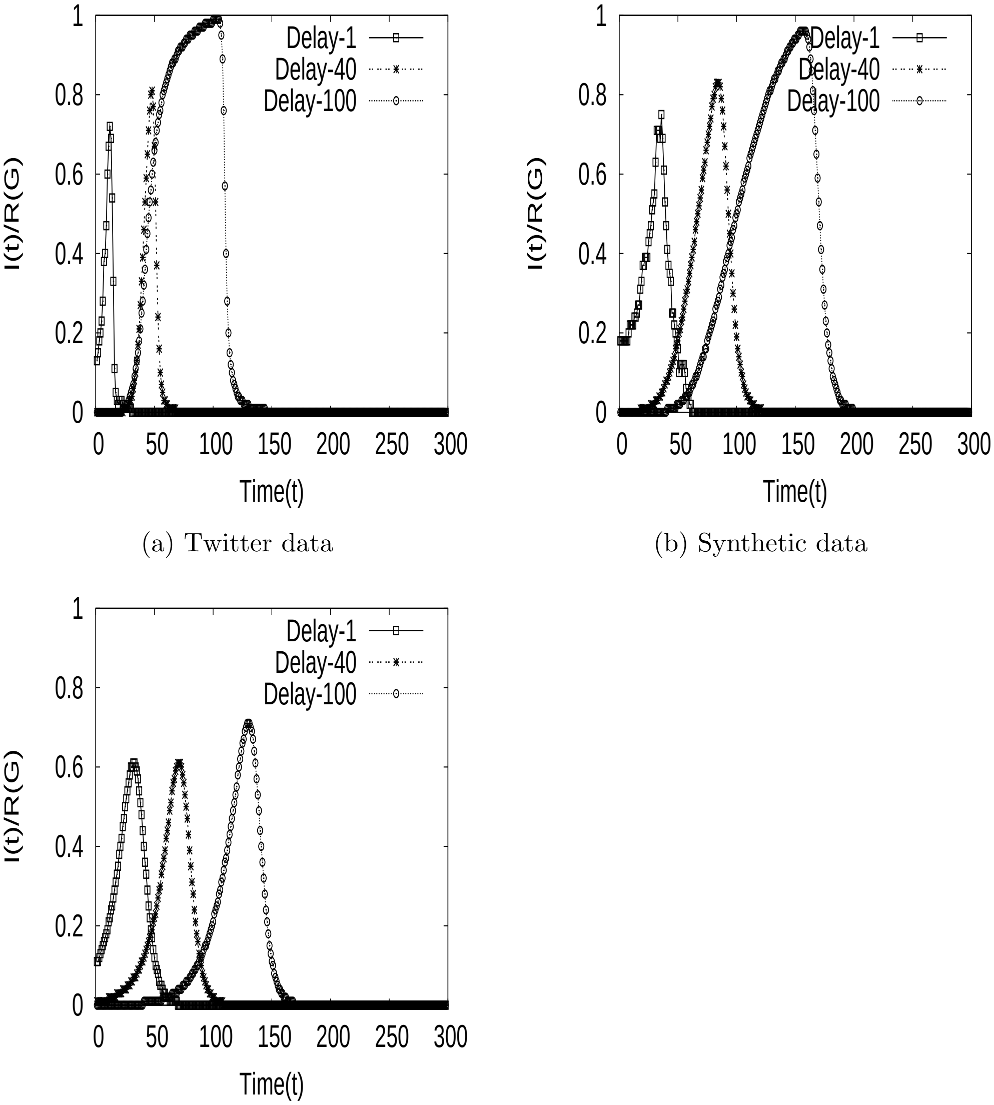
(d) Point of decline



e) Point of decline



(f) Outbreak size



(a) Twitter data

(b) Synthetic data

(c) Mean-field solution

Figure 8: Growth of Infected vs. Out-break size: Delayed start model

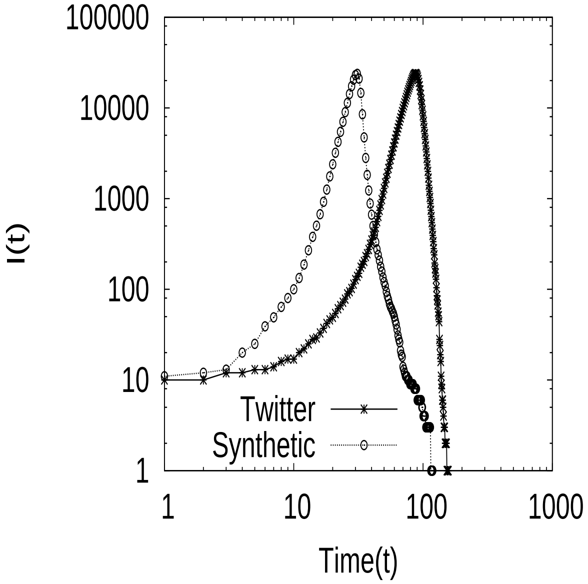
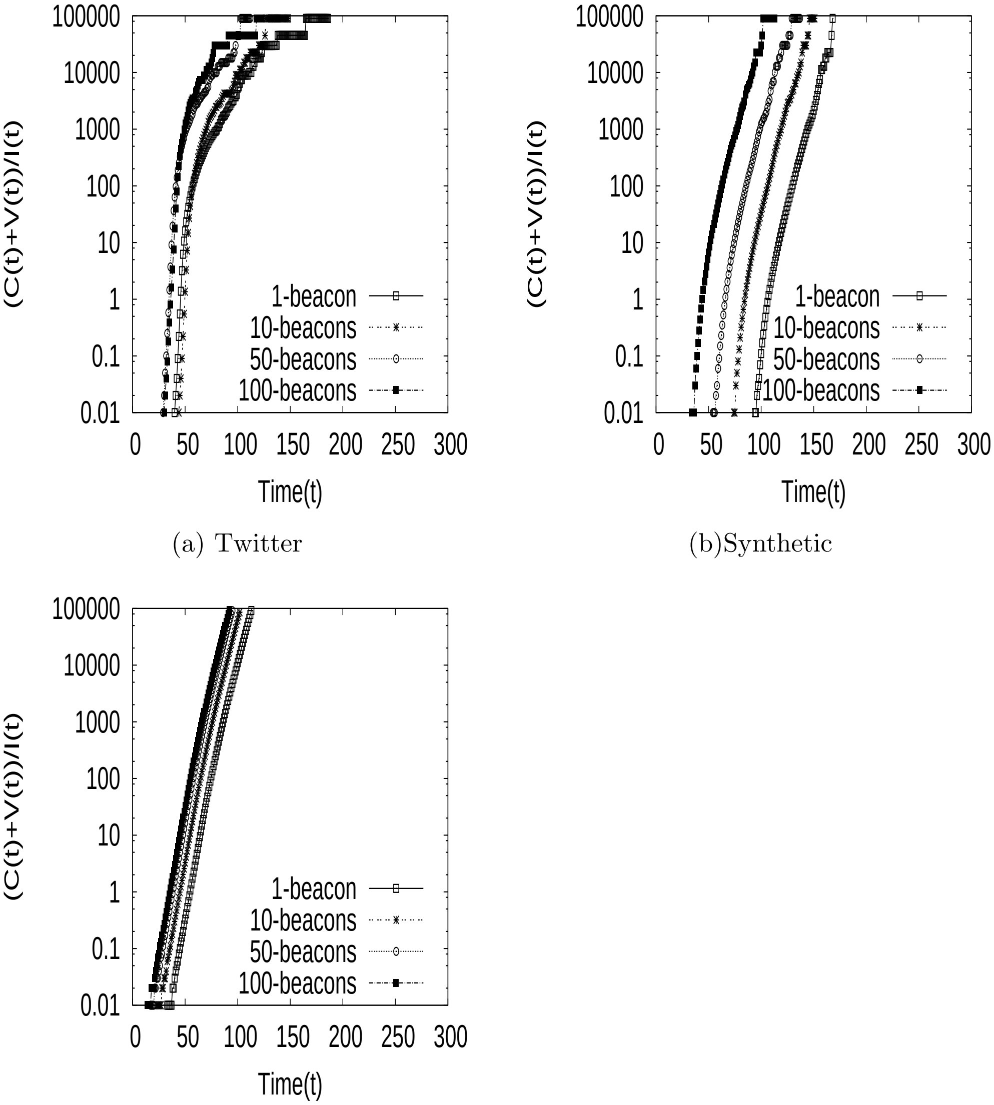


Figure 9: Rumor growth: Beacon Model

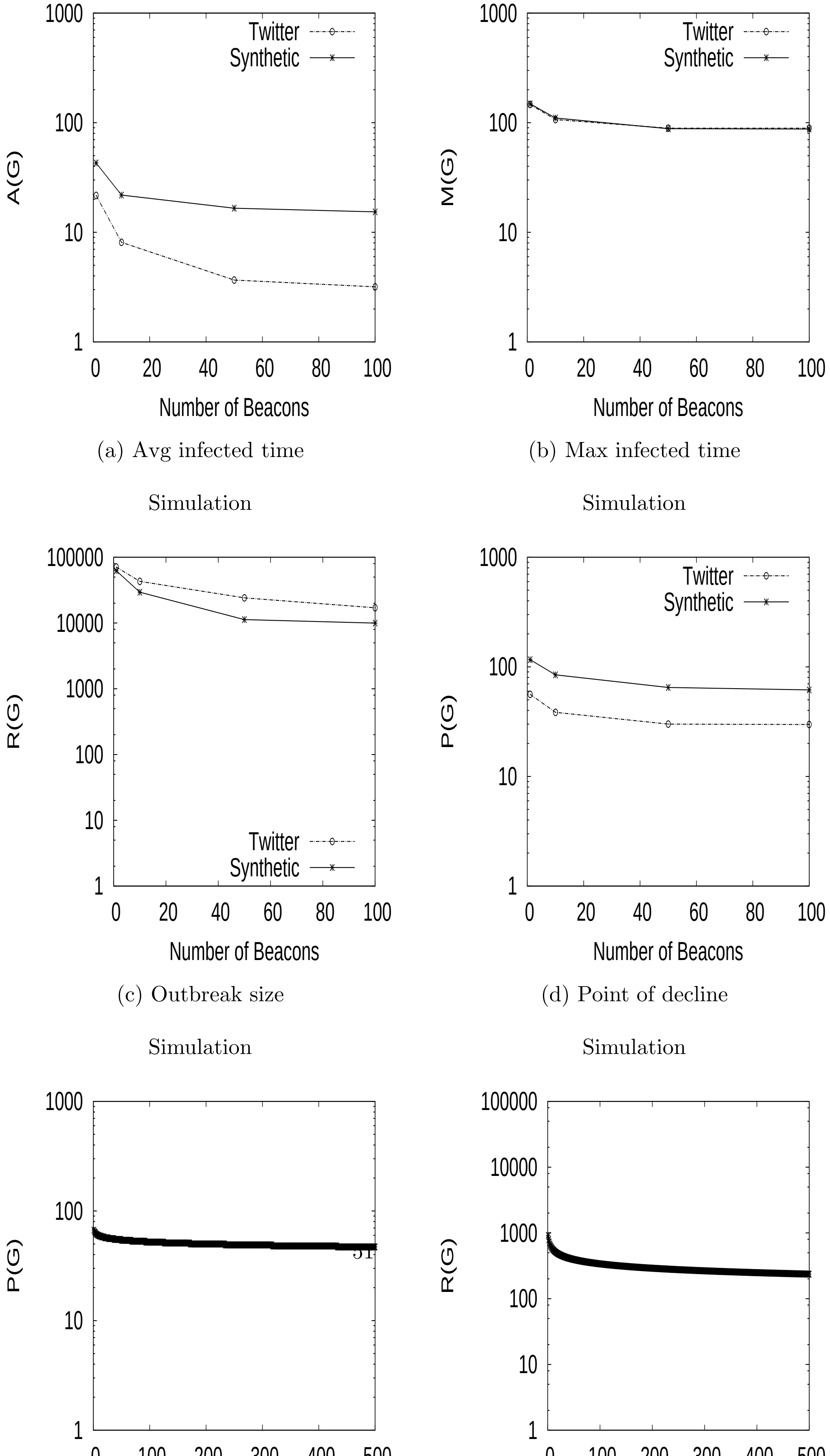


(a) Twitter

(b)Synthetic

(c) Mean-field

Figure 10: Behavior of time varying metrics: Beacon Model

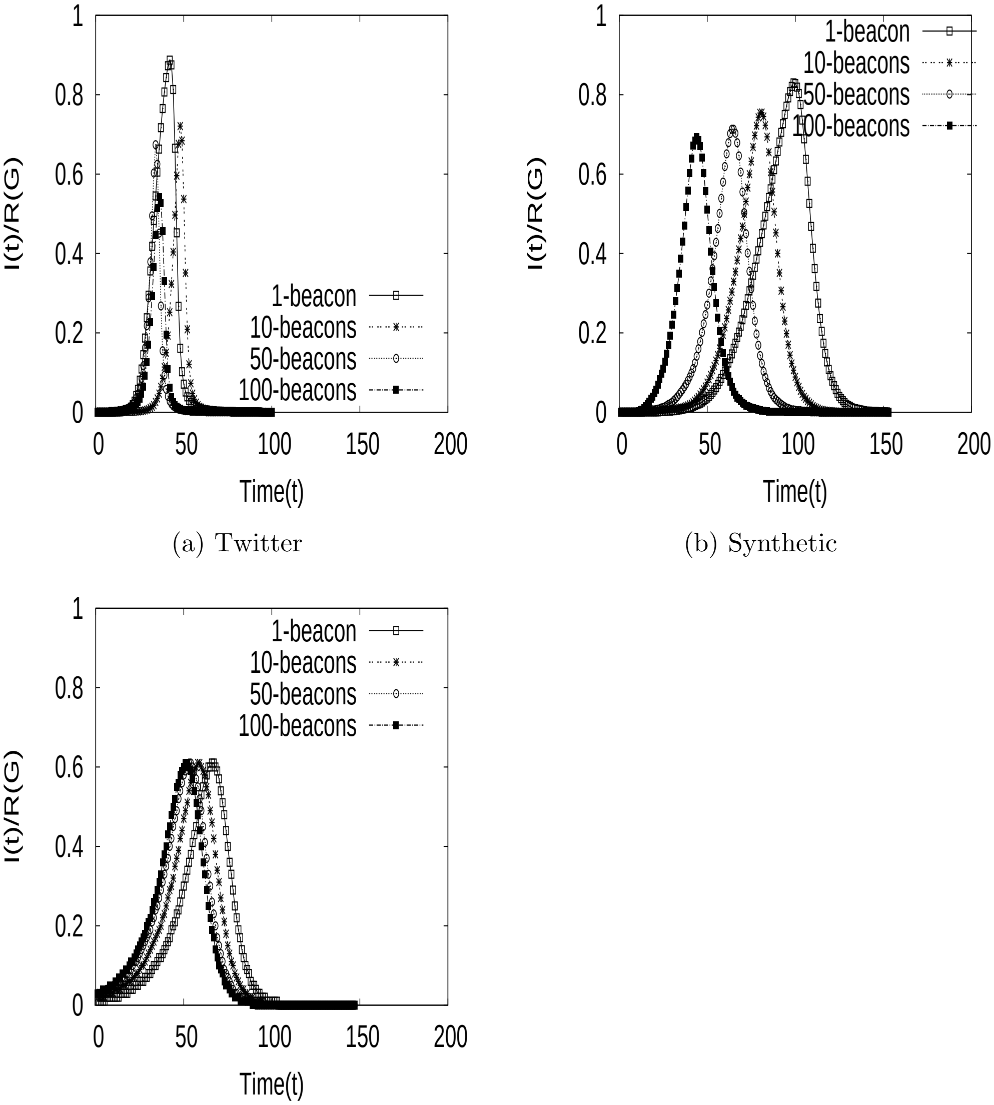


(a) Avg infected time

(b) Max infected time

(c) Outbreak size

(d) Point of decline



(a) Twitter

(b) Synthetic

(c) Mean field

Figure 12: Growth of Infected vs. Out-break size: Beacon model

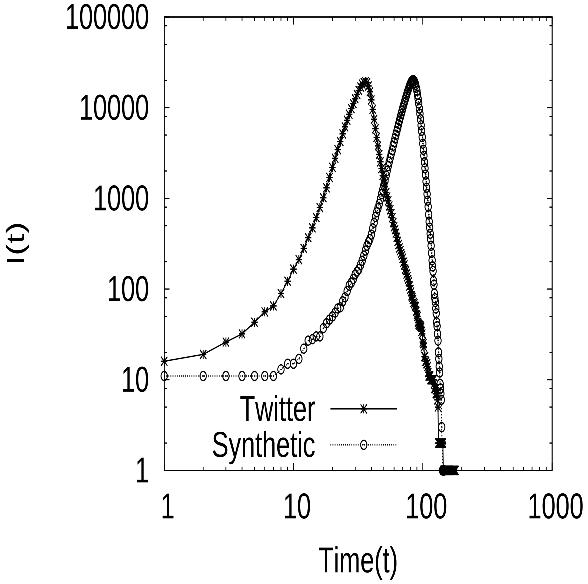
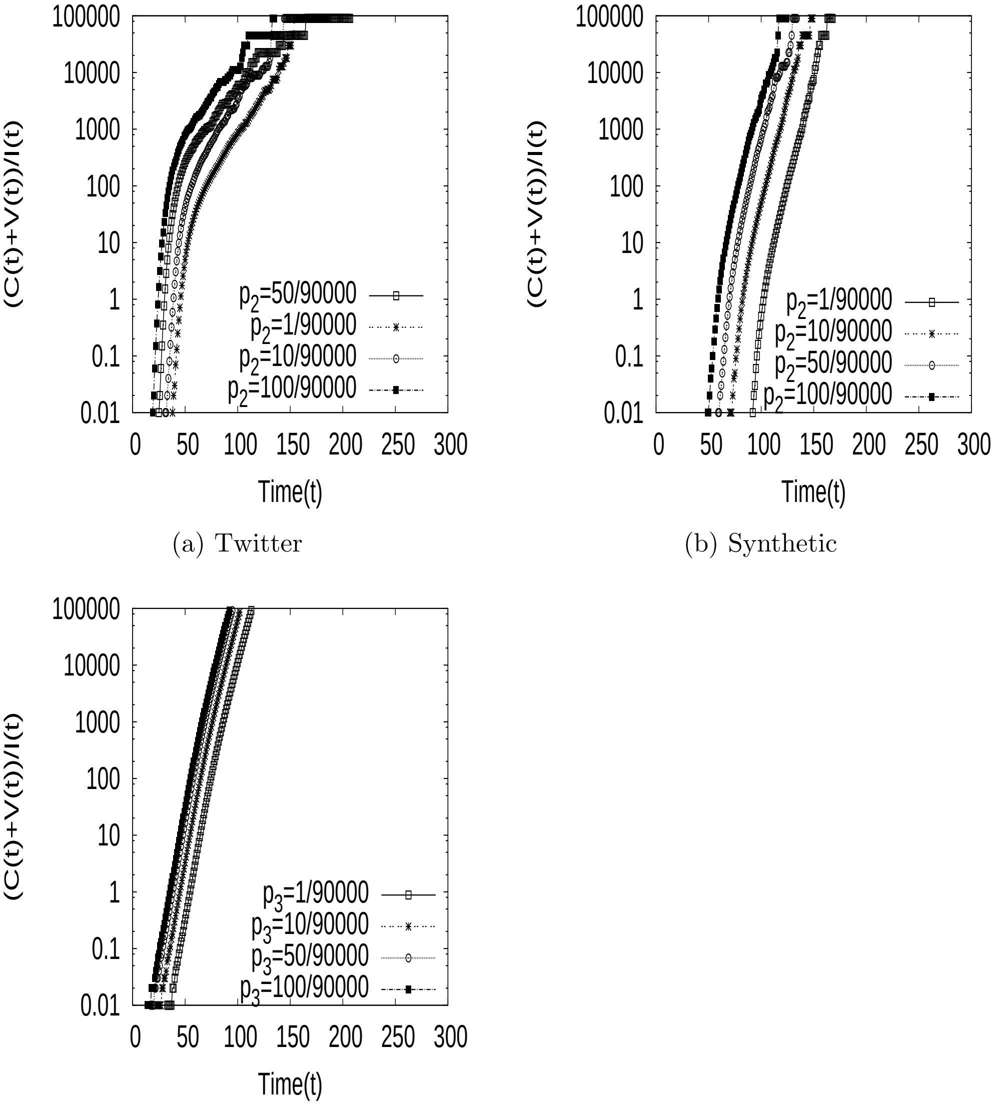


Figure 13: Rumor growth: Neighborhood Model

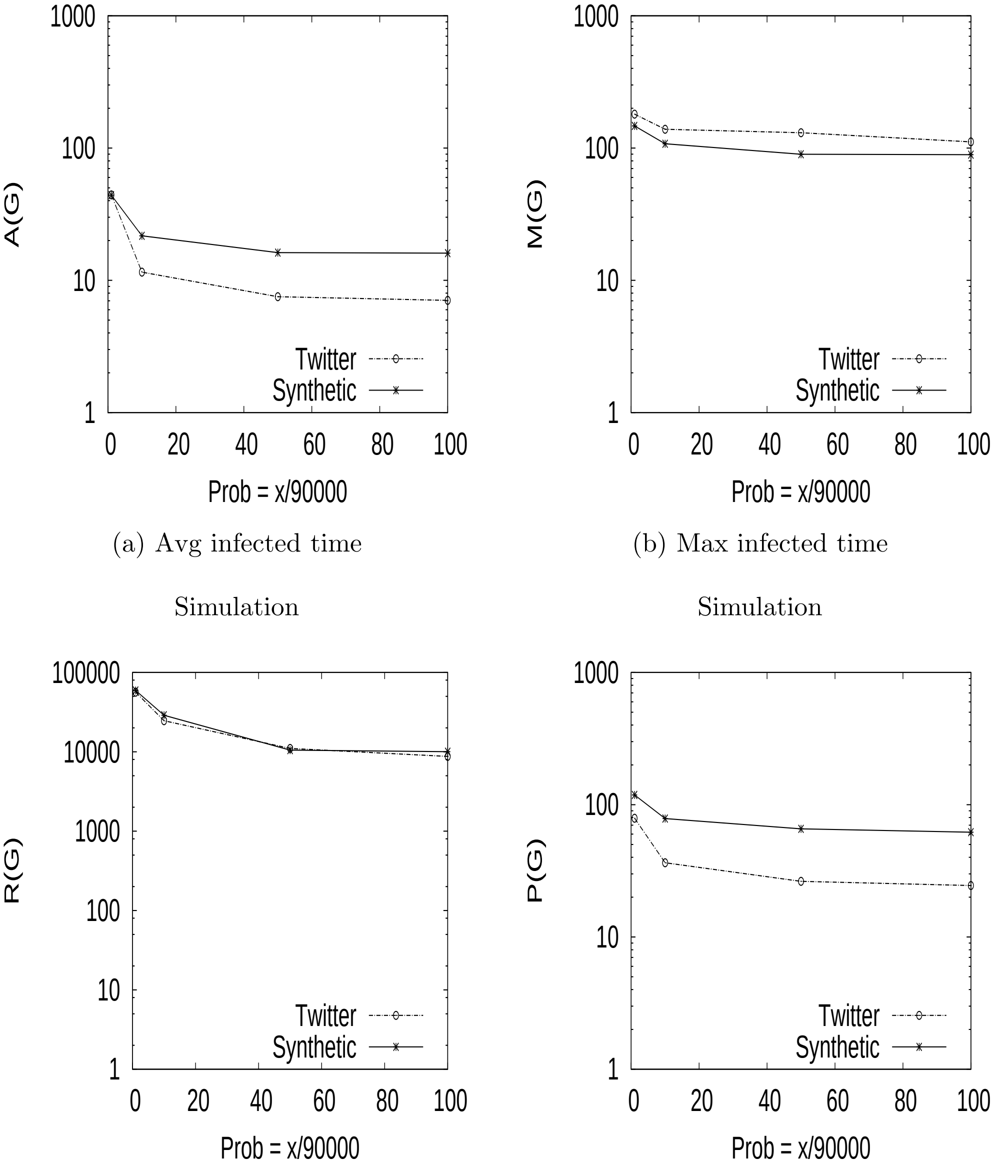


(a) Twitter

(b) Synthetic

(c) Mean-field

Figure 14: Behavior of time varying metrics: Neighborhood Model

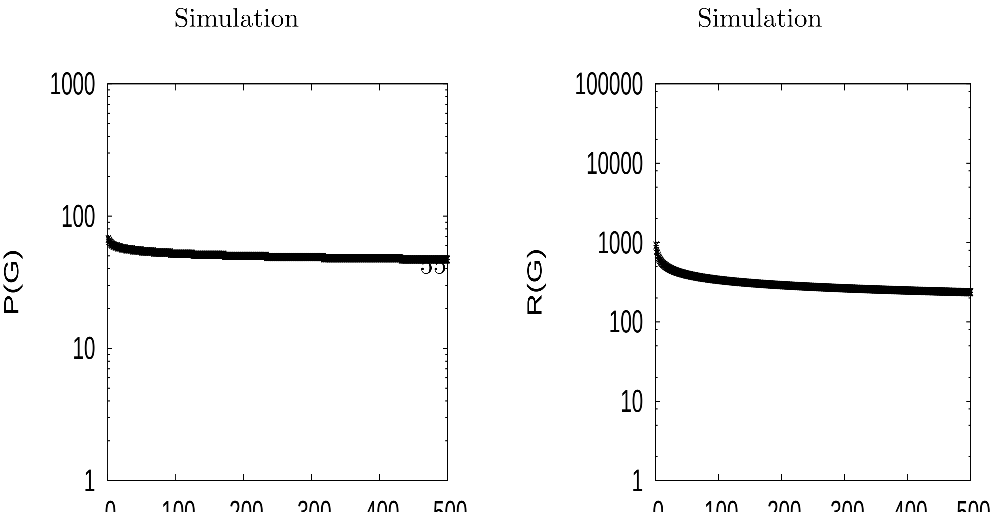


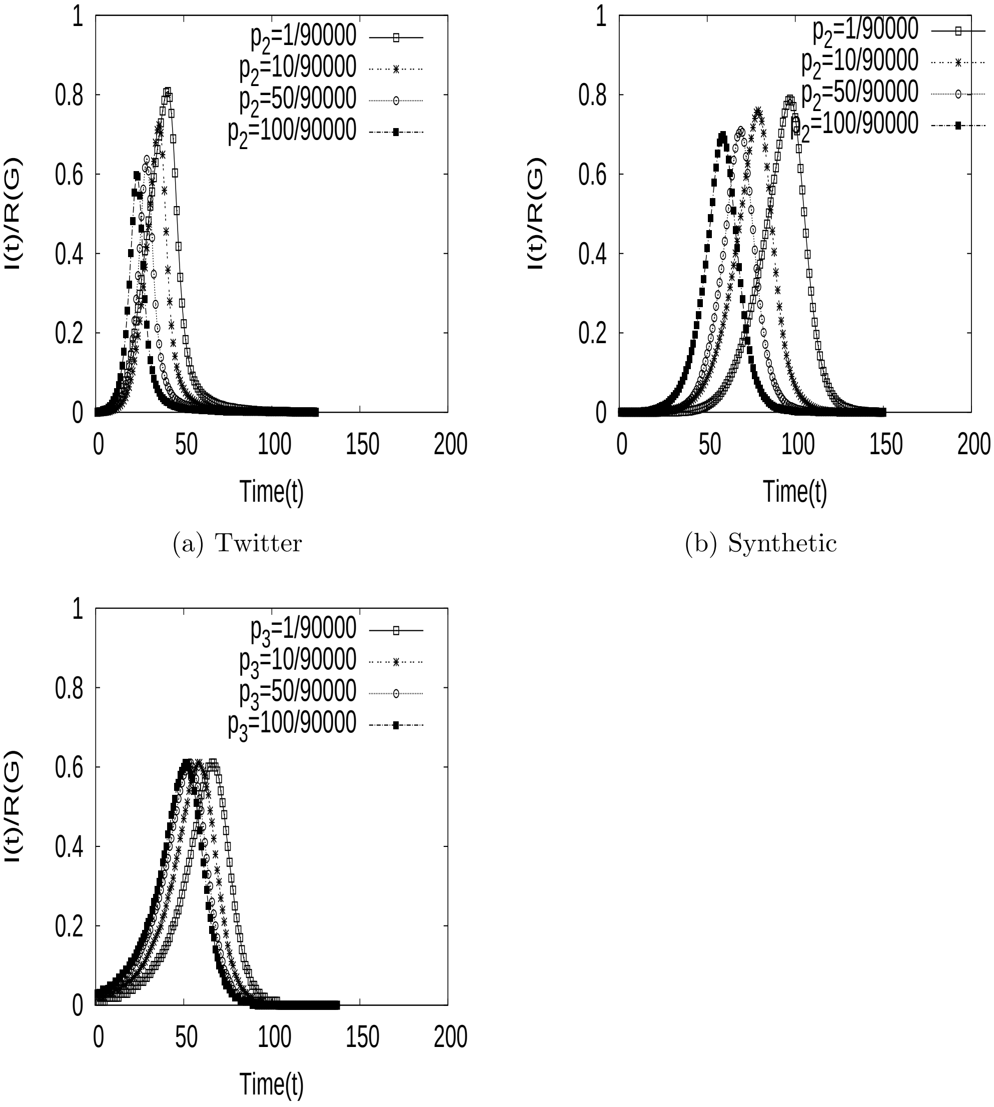
(a) Avg infected time

(b) Max infected time

(c) Outbreak size

(d) Point of decline





(a) Twitter

(b) Synthetic

(c) Mean-field

Figure 16: Growth of Infected vs. Out-break size: Neighborhood model