Online Social Networks and Media 2 (2017) 19–31

Contents lists available at ScienceDirect

Online Social Networks and Media

journal homepage: www.elsevier.com/locate/osnem

Efficient and timely misinformation blocking under varying cost constraints

Iouliana Litou[a,*], Vana Kalogeraki[a], Ioannis Katakis[b], Dimitrios Gunopulos[b]

[a] Athens University of Economics and Business, Athens, Greece
[b] University of Athens, Athens, Greece

ARTICLE INFO

Article history:
Received 20 February 2017
Revised 11 July 2017
Accepted 11 July 2017
Available online 29 July 2017

Keywords:
Misinformation blocking
Social networks
Emergency events

ABSTRACT

Online Social Networks (OSNs) constitute one of the most important communication channels and are widely utilized as news sources. Information spreads widely and rapidly in OSNs through the word-of-mouth effect. However, it is not uncommon for misinformation to propagate in the network. Misinformation dissemination may lead to undesirable effects, especially in cases where the non-credible information concerns emergency events. Therefore, it is essential to timely limit the propagation of misinformation. Towards this goal, we suggest a novel propagation model, namely the Dynamic Linear Threshold (DLT) model, that effectively captures the way contradictory information, i.e., misinformation and credible information, propagates in the network. The DLT model considers the probability of a user alternating between competing beliefs, assisting in either the propagation of misinformation or credible news. Based on the DLT model, we formulate an optimization problem that under cost constraints aims in identifying the most appropriate subset of users to limit the spread of misinformation by initiating the propagation of credible information. We prove that our suggested approach achieves an approximation ratio of $1 - 1/e$ and demonstrate by experimental evaluation that it outperforms its competitors.

1. Introduction

The massive use of Online Social Networks (OSNs), that enumerate millions of daily active users [1,2], have led to drastic changes on the communication and information sharing among users. Many people have integrated popular online social networks in their everyday lives and rely on them as one of their major news sources, with over 28% of people claiming to get their news from social media and even local news media exploit OSNs to report news.[1] For instance, during the Marathon Bombing in Boston on April 15, 2013, the news broke initially on Twitter and local news media actively exploited the platform to report news regarding the event.[2] However, there are differences in terms of the triggers, actions and news values that are prevalent in news media and in general in public shares in social media [3].

An intriguing behavior of the users of an OSNs is information sharing. Aside from being informed through the network, users further propagate information of interest to their friends. An interesting study reveals that more than 45.1% of the messages published by a user are further propagated by his/her followers and over 37.1% are propagated by followers of up to 4 hops away from the original publisher [4]. However, it is not uncommon for false news to propagate through social networks, causing significant repercussion outside the network, e.g., a tweet stating that there was an explosion in the White House caused stocks to temporarily plung.[3] It becomes therefore clear that it is vital to detect and timely block the propagation of deceptive information, since dissemination of false news in OSNs can have undesirable effects with a disastrous impact outside the network. Especially in cases where the news concern emergency related events, limiting the spread of misinformation becomes imperative.

Misinformation is defined as any malicious, deceptive or irrelevant information regarding an event, that is spread either deliberately or unintentionally [5,6]. Manifold factors contribute to the complexity of the task of misinformation blocking. Sources of misinformation are multiple and varying, as is the users' susceptibility

* Corresponding author
E-mail addresses: litou@aueb.gr (I. Litou), vana@aueb.gr (V. Kalogeraki), katak@di.uoa.gr (I. Katakis), dg@di.uoa.gr (D. Gunopulos).

[1] http://www.mushroomnetworks.com/infographics/social-media-news-outlets-vs.-traditional-news-infographic.

[2] https://media.twitter.com/success/bostonglobe-uses-twitter-to-source-news-and-keep-the-public-informed-during-the-boston.

[3] http://www.cnbc.com/id/100646197.

http://dx.doi.org/10.1016/j.osnem.2017.07.001
2468-6964/

to the news they are exposed. Furthermore, the longer the deceptive information propagates in the network without contradiction from a reputable source, the greater is the effect the misleading information can have, and it is thus crucial to timely notify users about the credible information. To effectively accomplish the limitation of misinformation, understanding the way news propagate and how contradictory information affect users' decision to accept and share content is required. Moreover, it is desirable to avoid unnecessary broadcasting of messages, since this causes excess load to the infrastructures that might be already over-utilized during emergencies [7].

For the effective and efficient solution of the problem, a set of interesting research questions should be addressed: (i) how can a user's susceptibility be estimated, so as to efficiently approximate the degree to which the user is willing to adopt or renounce an idea, (ii) how can we identify the sources of misinformation, (iii) how can we estimate the the influence of the users across the network, and finally (iv) given cost constraint, expressed as the number of sources to contradict misinformation, which is the most appropriate subset of users to initiate the cascade in order to timely limit the dissemination of misinformation?

The problem of limiting misinformation or rumors in OSNs, commonly referred as misinformation or rumor blocking, is the subject of study in [8] where the authors discover the minimal subset of users that minimize the propagation achieved by rumor originators. Authors in [5] also study the problem of identifying the smallest set of users to act as protectors against the dissemination of misinformation. Their goal is to find the smallest set of highly influential users whose decontamination with good information assists in limiting the spread of misinformation to a desired ratio within a given number of steps. In [9] authors consider the evolving propagation of competitive campaigns and aim at selecting the appropriate subset of nodes so that the eventual influence of the misinformation campaign is minimized. Albeit the aforementioned works address the problem of misinformation blocking, there are significant limitations: First, none of the above schemes considers propagation times and therefore, fail to capture any time constraints during the dissemination of the information in their models. Furthermore, the susceptibility of a user related to the propagation of content is ignored. Finally, they assume that whenever a user is convinced about certain information, the user will remain loyal to this belief as the propagation unfolds, regardless of the state of its neighbors. However, this assumption is not always correct. There are multiple examples of users that have unintentionally contributed to the spread of false news and they later apologize for their mistakes.[4]

The objective of this work is to minimize the misinformation spread while overcoming the above limitations. Our approach is two-fold: First, we propose a novel propagation model, namely the Dynamic Linear Threshold (DLT) model, where each user is associated with a *renouncement* threshold that expresses her susceptibility to adopt an idea. Instead of considering that the threshold remains static throughout the propagation process, we define it to be dynamically adjusted after a user adopts an idea. The adjustment of the threshold denotes that a user has the ability to renounce a previously adopted belief. However, adopting a different opinion regarding the validity of the information is less likely after the user is convinced on a specific piece of information. Furthermore, the proposed model considers the impact of competing ideas to the users' choice of adopting an idea. Finally, contrary to existing models such as the LT model, we consider that the influence of a user over any other user in the network varies over time and there is a certain time frame that a user is more likely to be influenced by a specific neighbor.

To limit the spread of misinformation, in the second part of our approach, we propose an algorithm that, given a budget, aims at selecting an appropriate subset of users to initiate the propagation of credible information. Our goal is to identify a subset of users to act as seeds for the dissemination of credible news, so that the number of users *infected* during the spread of a non-credible information is minimized. We define as infected the users that contribute to the propagation of the misinformation. We formulate the problem as an optimization task and propose a greedy approach for selecting the subset of users to contradict the misinformation. We require that the size of the subset does not exceed the budget, satisfying therefore the cost constraints. The suggested technique exploits simulated annealing to determine the credible seed set, i.e., the users that initiate the propagation of credible information and we formally prove that our approach achieves an approximation ratio of $1 - 1/e$.

1.1. Contributions

The contribution of this work is summarized as follows:

- We introduce a novel propagation model where users' susceptibility to news dynamically adapts over time. The model captures the hesitance of users to renounce their beliefs, making it appropriate to realistically describe the spread of information in OSNs. Furthermore, while previous works ignore time constraints during the propagation process, we capture time as an actual unit. We argue that propagation time differs among any two users and exploit the Poisson Distribution to capture the probability of timely delivery of messages.
- We suggest a greedy approach that efficiently solves the problem of misinformation limitation by selecting an appropriate subset of users to decontaminate the network. Our proposed model for seed selection is efficient and independent of the structure of the underlying network. We prove that it approximates the best solution with an approximation ratio of $1 - 1/e$ and it is suitable for real-time misinformation containment, as it requires less than 23 min to identify up to 1000 seeds in dense networks.
- We illustrate through extensive experimental evaluation that our approach achieves notably better results than its competitors with respect to misinformation limitation with a remarkable reduction in resource costs. The suggested approach requires in all cases a significantly lower amount of messages exchange in the network, while achieving better results in misinformation blocking under different cost constraints, making it thus efficient and cost-effective.

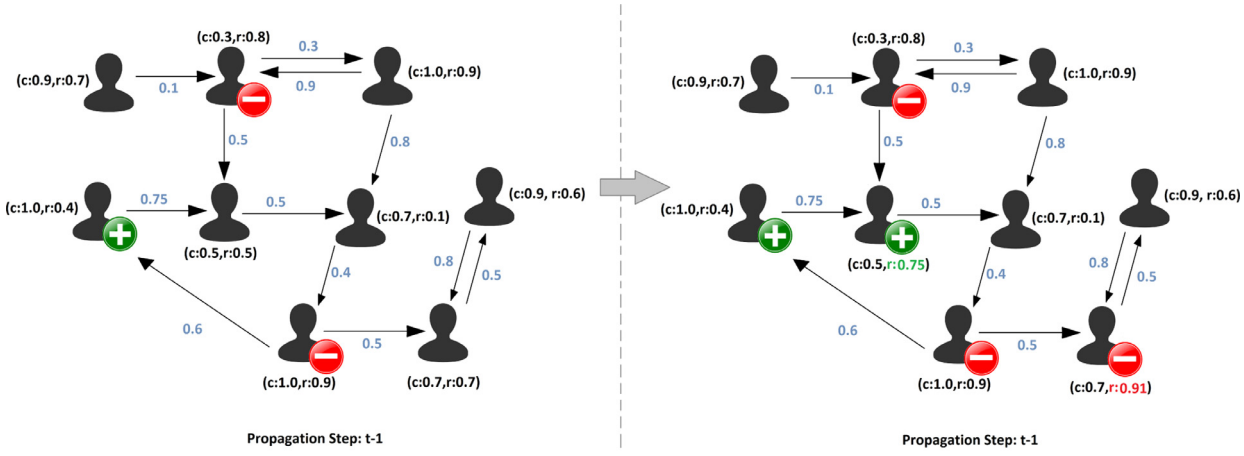
2. Model and problem definition

In this section we describe the network model and formally define the problem of misinformation blocking. We further prove that identifying the appropriate subset of users to limit the spread of misinformation is NP-complete.

2.1. Network model

A social network is commonly represented as a directed weighted graph $G(V, E)$. Users of the network constitute the nodes and a directed edge from user u to v denotes the flow of information. For example, in the Twitter network, the edge $u \rightarrow v$ denotes that user v has replied to or retweeted messages published by u. We associate each edge $u \rightarrow v$ with a weight w_{uv}, which expresses

[4] http://twitchy.com/2015/05/04/we-screwed-up-shep-smith-apologizes-for-false-report-on-baltimore-shooting-video/.



(a) Network snapshot at step $t-1$ (b) Network snapshot at step t

Fig. 1. DLT Propagation Example.

the influence of u over v and may be estimated based on prior user interactions.

Each node u is assigned a *credibility score* c_u and a *renouncement threshold* r_u. The credibility score c_u expresses the trustworthiness of the user u. Effectively estimating the users' credibility is out of the scope of this work, however, there are approaches suggested by former studies on how to identify sources of misinformation. For example, in [5] authors exploit the Reverse Diffusion Process to detect possible sources of misinformation given a snapshot of the diffusion. Prakash et al. [10] aim in identifying the initiators of a spread given an epidemic propagation model. Techniques that aim in real time identification of non-credible information, as those suggested in [11], may also be exploited to characterize the trustworthiness of a user. The study in [12], where the authors estimate the validity of an information considering the mood sensitivity factor, may also assist towards the identification of non-credible users. Another interesting work towards this direction is the work of Wu et al. [13] where based on past rumors circulated in the network a framework for early rumor detection is developed, that can be exploited in order to infer users' credibility.

The renouncement r_u denotes the incredulity of the user u in renouncing an already adopted opinion, either credible or misinformation. Users that present high renouncement are characterized as skeptical in adopting and propagating information, while users with low renouncement are less reluctant. We assume that users have an initial reluctance regarding information sharing, but after they are convinced on disseminating the information, their renouncement threshold increases, as they become hesitant in later switching their beliefs.

Following the terminology of [8], we distinguish between three different user roles, namely *infected, protected* and *inactive*.

Infected users (I): Non-credible users or users that adopt and propagate false news. We do not distinguish between users that intentionally spread false news from those that are influenced by the wrong news, since the user contribution towards the spread of misinformation is independent of this fact. Infected users are considered *negatively influenced* users.

Protected users (C): Users that adopt the credible information and therefore are able to further propagate it to the network are referred to as Protected users or *positively influenced*. Unlike the work in [8] that assumes that the protected users remain protected throughout the propagation process, we define them to be susceptible to misinformation.

Inactive users (R): Users that are unaffected by the propagation of either credible or false news. It holds that $R = V \setminus (I \cup C)$, where V is the set of all users.

An example of the social graph is presented in Fig. 1a, where negative, positive and non-signed users represent the infected, protected and inactive users, respectively. The values on the edges denote the weights, while the c and r values represent the renouncement threshold and the credibility scores respectively.

2.2. Problem definition

We define the problem addressed in this as follows: Given (i) a social graph $G(V, E)$, (ii) the weight of the edges w_{uv}, $\forall u \rightarrow v \in E$ among users of the network, denoting the influence of user u over v, (iii) the set of initially non-credible users before the propagation unfolds (this constitutes the misinformation originators I_0), and (iv) a parameter k denoting the cost constraints and expressed as the desired number of user in the network to act as initiators of credible news propagation, our goal is to select a subset of users S with $|S| \leq k$, so that the number of infected users I during the propagation is minimized. We aim at a cost-effective solution where we define the *cost* for the seed selection process as the amount of messages exchanged among users. Thus, it could be monetary (for an SMS) or resource allocation cost. Note that the parameter k may be adjusted in order to reflect the available budget, i.e., the amount of messages to be sent in the network, and hence reduce costs. I_0 constitutes a subset of the infected users and the seed subset S of the credible users, hence $I_0 \subset I$ and $S \subset C$. The study in [14] reveals that users tend to be very polarized regarding certain topics, e.g., conspiracy, and are more susceptible regarding the diffuse of false information based on their polarization tendencies. The homophily of the users also plays an important role to the misinformation diffusion. We note that although in this work we assume absence of user polarization or confirmation bias, these aspects can be captured by appropriately defining the subset I_0 and S. That is, users that are considered malevolent can be included in I_0, and users that confirm facts before posting in Social Media can be included in S a priori. Although we assume S to be initially empty, relaxing this assumption does not impose any restrictions to our approach. We refer to users in S as *seeds*. The role of the seeds is to decontaminate the infected users by propagating the credible information. We assume that seeds are convinced about the credible information, regardless of their re-

nouncement threshold. We show in our experimental evaluation how this assumption impacts the performance of our approach.

The aforementioned problem is referred in the bibliography as the *Influence Blocking* problem [8,15] and can be proven to be NP-Complete.

Theorem 1. *The influence blocking problem is NP-complete*

Proof. Consider the special case where all users have the same probability to influence their neighbors, regardless of the time one user is exposed to another user's influence, and that probability equals 1, i.e., $w_{uv} = 1, \forall u \rightarrow v \in E$. This implies that if user u adopts a belief, all out-neighbors of u (i.e., users influenced by u) will adopt the same belief. Given the set of users V, we construct a set of subsets of V, $SG = \{V_1, V_2, \ldots, V_m\}$, so that if a user in V_i, $\forall i \in [1, m]$ is infected, all users in the subset V_i are infected. In order to reverse the belief adopted by a subset V_i, we have to convince at least one node $u \in V_i$ of the credible information. In order to leave no node infected we have to select a subset of users S, $|S| \leq k$, so that $S \cup V_i \neq \emptyset$, $\forall i \in [1, m]$. The above problem as described, constitutes the *Hitting-Set Problem*, a well know NP-complete problem. Since the special case is NP-complete, we conclude the NP-Completeness of the original problem. □

3. Propagation model

In this section we present the first part of our approach for solving the problem of misinformation propagation. We introduce our propagation model (DLT) and then present our methodology for computing the influence when users are exposed to contradictory information and the updated renouncement scores.

3.1. The dynamic linear threshold model

The most common propagation models in Social Networks are the Independent Cascade (IC) model and the Linear Threshold (LT) model [16]. Users that adopt a belief and further assist in the propagation of the belief are referred to as *active*. In the IC model each active user has a single chance of activating his/her inactive neighbors and the probability to succeed is expressed as a function of the weight of the corresponding edges. The propagation unfolds in discreet steps, with the users activated in the later step having the chance to influence their currently inactive neighbors. In the LT model, each user is associated with a threshold θ. Unlike the IC model, a user is activated whenever the sum of the weights of the incoming edges from the currently active neighbors, i.e., the incoming influence, exceeds the threshold θ of the user.

Both the IC and the LT propagation models express the information dissemination of a single campaign and hence ignore the way conflicting propagation may impact one another. Furthermore, traditional IC and LT models ignore propagation times. Variations of the above models consider time limitations [17–19], yet none of them takes into account opposing campaigns that may be propagated simultaneously.

To overcome the above limitations, we propose the Dynamic Linear Threshold (DLT). Our suggested propagation model differs from the traditional LT model in the following ways:

- The LT model assumes the propagation of a single idea in the network, while DLT considers competing ideas that simultaneously propagate and evolve over time.
- We estimate the influence of a user u to a neighbor v not solely based on the weight of the edge $u \rightarrow v$, but also based on the time frame, i.e., there is a time window that a neighbor is most likely to be influenced.
- In the LT model, whenever a user adopts an idea, she remains thereafter loyal during the propagation process. On the contrary, we assume that the user may renounce an idea, based on the input influence of the neighbors.
- Finally, the threshold of user v, denoted as renouncement r_v, can be dynamically updated over time after user v adopts an idea.

Table 1
Parameters table.

G(V,E)	Weighted social graph
V	Nodes of the graph denoting users of the Social Network
E	Edges on the graph denoting the flow of information
$u \rightarrow v$	Edge between users u and v
w_{uv}	Weight of edge $u \rightarrow v$, denoting the influence of user u over v
c_u	Credibility score of user u
r_u	Renouncement threshold user u
I	The set of infected users
R	The set of inactive users
C	The set of protected users
S	The set of seed users selected to propagate the credible information
T_{uv}	Set of time intervals t_i between subsequent interactions of u with v
λ	The variance $Var(T_{uv})$
$p_{uv}(t; \lambda)$	Probability of user u influencing v at time t
$IF(v\|t)$	Influence exceeded to user v by the neighbors at time t

Similarly to the LT model, we consider that a user v adopts a belief when the influence from the incoming neighbors exceeds the renouncement threshold r_v. To estimate the probability that a user u influences a user v at time t we exploit the Poisson Distribution, taking into account past interactions of user u and v. User v adopts either the misinformation or the credible information, depending on the beliefs of the incoming neighbors. The propagation process unfolds at discrete steps as follows: (i) At step $\tau - 1$ each user u that is either negatively or positively influenced (infected or credible respectively), influences the currently inactive neighbors $v \in out(u)$ with a *probability* $IF(v|t)$ that is estimated based on the weight of the edge and the *time window* t. (ii) A user v adopts a belief based on the influence of the incoming neighbors $u \in in(v)$. In case that the negative/positive influence exceeds the renouncement threshold r_v, then user v becomes infected/credible. We assume that, if both conflicting propagation reach the renouncement threshold r_v simultaneously, user v adopts neither of them. This is based on the fact that when a user is convinced on both conflicting information, due to the high ambiguity presented for the specific information, he remains hesitant concerning the adoption of either. (iii) At step τ, each node that is either positively or negatively influenced at step $\tau - 1$ is added to the credible or infected set and the renouncement thresholds are updated.

The aforementioned propagation steps are repeated until no more users can be influenced, i.e., the credible and infected sets remain unchanged. The existence of conflicting influences at node v is taken into consideration when deciding whether v will adopt a specific belief, as explained in Section 3.2. In Fig. 1 we present an example of the propagation, where newly influenced users at step τ have their renouncement thresholds updated.

3.2. Influence and renouncement computation

In this section, we formulate the influence channeled to the users from their neighborhood during the propagation of conflicting information. We also define how the renouncement threshold of a user is updated after a belief is adopted.

In the DLT model we consider that the propagation unfolds over time and the transmission times between any two users vary. To estimate the probability of user u influencing user v at time t we

assume a Poisson Distribution, that is,

$$p_{uv}(t;\lambda) = \frac{\lambda^t e^{-\lambda}}{t!}$$
$$\text{where} \quad \lambda = Var(T_{uv}) = \frac{1}{n}\sum_{1}^{n}(t_i - \mu) \tag{1}$$

In the above equation $T_{uv} = \{t_1, t_2, \ldots, t_n\}$, is the set of time intervals between any two subsequent interactions of user u with v, t_i is the ith interval, n is the total number of subsequent interactions and μ is the mean of all values in T_{uv}. By exploiting the above distribution to model the propagation times between nodes, each neighbor $u \in in(v)$ is given a different probability for influencing user v, which evolves over time. Hence, the influence of v from his neighbors changes dynamically over time. This is a reasonable assumption considering the way updates are presented to users in Online Social Networks, i.e., a timeline presentation with the more recent update on top. As time elapses, the influence of user u on user v gradually fades, since the probability of the update of u presented in the timeline of v decreases. Each neighbor presents a different decreasing ratio. However, updates may not necessarily appear in the *most recent* way in the news feed of the user [20,21], hence the Poisson Distribution captures the fact that between each two users u and v, with edge $u \rightarrow v$, there is a different time window within which the probability of u influencing v increases. Note that more sophisticated methods that predict whether a user will view a message within a certain time window from its publication may be exploited [22].

We consider that a user v at time t is influenced by his/her neighbors according to the following Influence Function:

$$IF(v|t) = \sum_{u \in in(v)} B(u|t) \cdot p_{uv}(t;\lambda) \cdot w_{uv} \tag{2}$$

where $in(v)$ denotes the incoming neighbors of user v, $B(u|t)$ equals -1 if user u has adopted the non-accurate belief at time t, $B(u|t)$ equals 1 if user u is convinced on the credible information at time t, or is 0 otherwise (i.e. if no belief has been adopted). In the above equation $p_{uv}(t;\lambda)$ denotes the probability that user u influences v at time t and w_{uv} denotes the influence of u over v as expressed by the weight of the edge $u \rightarrow v$. User v adopts a belief if $|IF(v|t)| \geq r_v$ and $B(v|t)$ at time t is set to 1 if $IF(v|t) > 0$, or -1 if $IF(v|t) < 0$. In case $|IF(v|t)| < r_v$ we assume that no belief can be adopted either because none of the credible or non-credible exceeds the threshold or both exceed it, thus causing confusion to the user and therefore leading her to dismiss both ($B(v|t) = 0$).

The above function considers the co-existence of conflicting cascades, since the misinformation decreases the sum while the credible information increases it. Therefore, unless either the positive or negative influence is strong enough, the user becomes hesitant on which one to adopt. Intuitively, when a user is exposed to both the negative and the positive propagation, the influence of the one is counterbalanced by the influence of the other.

Whenever user v adopts a belief B_i, then the renouncement score $r_v(t)$ of user v at time t is updated as follows:

$$r_v(t) = 1 - (1 - r_v(0))^{y+1} \tag{3}$$

where $1 - r_v(0)$ denotes the inherited reluctance of user v regarding the adoption of a cascade, and y the number of times a user switched believes. The above equation intuitively expresses that the renouncement of user v increases whenever he/she adopts a belief, hence, the user becomes more reluctant renouncing an adopted belief.

4. Misinformation blocking

In order to limit the spread of false news, and therefore the number of infected users, we suggest the selection of a limited subset of users that will assist the decontamination of infected users by cascading the credible news. Our approach aims primarily in preventing the misinformation from reaching the users in the network rather than convincing them to switch their opinion after they are contaminated. However, even in the case of decontamination, the difficulty of the decontamination process once a user is convinced on a specific belief is captured by the propagation model, as the thresholds of the users increase after a belief is adopted, denoting that it becomes more difficult to debunk to people who are already infected. Since the problem of finding the appropriate subset of users for misinformation blocking is NP-Complete, we develop a greedy algorithm that iteratively adds nodes to the seed set S, until the desired number k of seeds is reached, or no more seeds can be added. The algorithm exploits simulated annealing at each iteration in order to determine the most appropriate seed.

Based on Eq. (2), for a node v not to be infected, it should hold $IF(v|t) \geq 0$ or $|IF(v|t)| < r_v$ otherwise, that is:

$$\sum_{u \in in(v)} B(u|t) \cdot p_{uv}(t;\lambda) \cdot w_{uv} \geq 0, \text{ or}$$
$$\left| \sum_{u \in in(v)} B(u|t) \cdot p_{uv}(t;\lambda) \cdot w_{uv} \right| < r_v \tag{4}$$

In order to identify the misinformation originators, we exploit the credibility scores of the users, denoting as misinformation originators the users with the lower credibility scores. We argue that simply displaying a warning next to any communication from unreliable individuals is not sufficient, as the credibility of a user may vary based on variety of factors, e.g., proximity to the event, thus it should be identified in real-time and differ based on the specific information that is spread. Additionally, other factors such as the users' credulity that may impact on their credibility, could be considered. Given the set of misinformation originators I_0, the nodes likely to be infected at the next step belong to the neighborhood of I_0. Therefore, for nodes with incoming edges from I_0, i.e., $v \in out(I_0)$ we have to maximize $IF(v|t)$ of user v, while considering the number of nodes I_k infected after k seeds are selected. Hence, we define the function $g(S_k)$ as follows:

$$g(S_k) = \sum_{v \in out(I_0)} IF(v|t) - (|I_k| - |I_0|)$$
$$= \sum_{v \in out(I_0)} \sum_{u \in in(v)} B(u|t) \cdot p_{uv}(t;\lambda) \cdot w_{uv} - (|I_k| - |I_0|) \tag{5}$$

In the above equation $(|I_k| - |I_0|)$ denotes the additional nodes infected given the seed set S_k. We need not only to maximize the positive influence, but also minimize the set of infected nodes. Therefore, by maximizing $g(S_k)$ we either increase positive influence by maximizing $\sum_{v \in out(I_0)} \sum_{u \in in(v)} B(u|t) \cdot p_{uv}(t;\lambda) \cdot w_{uv}$ or decrease the set of infected nodes at time t by minimizing $(|I_k| - |I_0|)$. Since $B(u|t)$ equals 1 when the node u is credible, -1 if u is infected or 0 otherwise, the above equation may be written as follows:

$$g(S_k) = \sum_{v \in out(I_0)} \left(\sum_{u \in W_k} p_{uv}(t;\lambda) \cdot w_{uv} - \sum_{u \in I_0} p_{uv}(t;\lambda) \cdot w_{uv} \right)$$
$$-(|I_k| - |I_0|) \tag{6}$$

where $W_k = C \cup S_k$. W_k denotes the neighbors of node v that are protected and therefore have a positive influence and I_0 are the infected nodes and have negative impact. In Algorithm 1 we present the simulated annealing approach for achieving the maximization of $g(S_k)$. Parameters α, T and T_{min} are tunable, depending on the refinement required in the solution. However, these can effect the execution times. For the experiments the values are set as in Algorithm 1.

Algorithm 1: REACT.

Data: $G(V, E)$, I_0, C, $seedsSize$, t
$S \leftarrow \emptyset$;
while $|S| < seedsSize$ **do**
  $T \leftarrow 1.0$;
  $T\_min \leftarrow 0.0001$;
  $\alpha \leftarrow 0.7$;
  $g(S_k) \leftarrow null$;
  while $T > T\_min$ **do**
    $u_k \leftarrow randomSeed(V)$;
    $S_k \leftarrow S \cup u_k$;
    $I_k \leftarrow getInfected(I_0, C, S_k)$;
    $g_{new}(S_k) = \sum_{v \in out(I_0)} \Big(\sum_{u \in C \cup S_k} p_{uv}(t;\lambda) \cdot w_{uv} - \sum_{u \in I_0} p_{uv}(t;\lambda) \cdot w_{uv} \Big) - (|I_k| - |I_0|)$;
    if $g(S_k) == null$ **then** $g(S_k) \leftarrow g_{new}(S_k)$;
    $ap \leftarrow e^{\frac{g_{new}(S_k) - g(S_k)}{T}}$;
    if $ap \geq random()$ **then**
      $u \leftarrow u_k$;
      $g(S_k) \leftarrow g_{new}(S_k)$;
    end
    $T = T \cdot \alpha$;
  end
  $S \leftarrow S \cup u$
end

The algorithm computes the seed set S as follows:

1: Initially a random seed u_k is selected at step k.
2.1: Given u_k, the seed set S and the misinformation originators I_0, we estimate the set of users I_k infected when u_k is added to the seed set $S_k = S \cup u_k$ as well as the new value of the $g(S_k)$ function. I_k is computed by simulating the propagation under the DLT model given the seed set S_k. For the simulated annealing iteration we store the value of the prior best seed as well as seed u. Hence, $g(S_k)$ expresses the value of the best candidate seed at step k and u the best candidate, and $g_{new}(S_k)$ the value of the u_k candidate investigated at the current iteration of the simulated annealing process.
2.2: After the simulated annealing of the kth iteration completes the best candidate u is added to the seed set S, and the set is updated as $S = S \cup u$.
3: Steps 1 through 2 are repeated until the required number of seeds is selected or no more nodes may be protected by the selection of additional seeds.

We do not consider streaming data in order to estimate the evolution of the propagation process after the seeds are selected. Albeit, it is possible that the propagation may not evolve as predicted. In that cases the ability to adapt to such concept drifts by exploiting adaptive learning techniques [23] may be desirable.

5. Approximation ratio

We now prove that the maximization of the $g(S)$ function, by iteratively adding seeds in a greedy manner, leads to a selection of a seed set S that approximates the optimal solution within a factor of $1 - 1/e$ [16]. In order to achieve the above approximation rate the function generating the seed set has to be non-negative, monotone and sub-modular. In contrast to the Linear Threshold model, we note that there is a negative part presented in the $g(S)$ function, i.e., $\sum_{u \in I_0} p_{uv}(t;\lambda) \cdot w_{uv}$, and hence the proof of monotonicity and non-negativity slightly differs from the LT model.

Non-negativity: Since we maximize $g(S_k)$, we have that node u_k selected at step k adds value, therefore, by definition $\sigma(u_k) = g(S_k \cup u_k) - g(S_k)$ is non-negative.
Monotonicity: $g(S_k) \geq g(S_{k-1}) \forall S_k = S_{k-1} \cup u_k$.

Proof. Given the set of infected nodes I_{k-1} when $k-1$ seeds are selected and the set of nodes P_{u_k} that are infected when u_k is not a seed, i.e., P_{u_k} are nodes sheltered by u_k, based on Eq. (6), at the kth iteration we have:

$$\mathbf{g(S_k)} = \sum_{v \in out(I_0)} \Big(\sum_{u \in W_k} p_{uv}(t;\lambda) \cdot w_{uv} - \sum_{u \in I_0} p_{uv}(t;\lambda) \cdot w_{uv} \Big) - (|I_k| - |I_0|)$$

$$= \sum_{v \in out(I_0)} \Big(\sum_{u \in (C \cup S_{k-1} \cup u_k)} p_{uv}(t;\lambda) \cdot w_{uv} - \sum_{u \in I_0} p_{uv}(t;\lambda) \cdot w_{uv} \Big) - ((|I_{k-1}| - |P_{u_k}|) - |I_0|)$$

$$= \sum_{v \in out(I_0)} \Big(p_{u_k v}(t;\lambda) \cdot w_{u_k v} + \sum_{u \in (C \cup S_{k-1})} p_{uv}(t;\lambda) \cdot w_{uv} - \sum_{u \in I_0} p_{uv}(t;\lambda) \cdot w_{uv} \Big) - ((|I_{k-1}| - |P_{u_k}|) - |I_0|)$$

$$= |P_{u_k}| + \sum_{v \in out(I_0)} p_{u_k v}(t;\lambda) \cdot w_{u_k v} + \sum_{v \in out(I_0)} \Big(\sum_{u \in W_{k-1}} p_{uv}(t;\lambda) \cdot w_{uv} - \sum_{u \in I_0} p_{uv}(t;\lambda) \cdot w_{uv} \Big) - (|I_{k-1}| - |I_0|)$$

$$= |P_{u_k}| + \sum_{v \in out(I_0)} p_{u_k v}(t;\lambda) \cdot w_{u_k v} + g(S_{k-1})$$

We note that if u_k has no incoming edges to v, then $P_{u_k} + \sum_{v \in out(I_0)} p_{u_k v}(t;\lambda) \cdot w_{u_k v} = 0$. Therefore,

$$g(S_k) = |P_{u_k}| + \sum_{v \in out(I_0)} p_{u_k v}(t;\lambda) \cdot w_{u_k v} + g(S_{k-1}) \quad (7)$$

$$\iff g(S_k) \geq g(S_{k-1})$$

□

Submodularity: In order to prove that a function is submodular, for two sets $M^1 \subseteq M^2$ and a node u_k, it should hold that $g(M^1 \cup u_k) - g(M^1) \geq g(M^2 \cup u_k) - g(M^2)$, i.e., node u_k should add greater value when added to the subset.

Proof. From Eq. (7) we get that

$$g(M^1 \cup u_k) = |P^{M^1}_{u_k}| + \sum_{v \in out(I_0)} p_{u_k v}(t;\lambda) \cdot w_{u_k,v} + g(M^1) \quad (8)$$

where $P^{M^1}_{u_k}$ is the set of nodes protected when u_k is added to M^1. Subsequently,

$$g(M^1 \cup u_k) - g(M^1) = |P^{M^1}_{u_k}| + \sum_{v \in out(l_0)} p_{u_k v}(t;\lambda) \cdot w_{u_k v} \quad (9)$$

Similarly for M^2, with $P^{M^2}_{u_k}$ denoting the set of nodes protected when we add u_k to M^2, it holds that

$$g(M^2 \cup u_k) - g(M^2) = |P^{M^2}_{u_k}| + \sum_{v \in out(l_0)} p_{u_k v}(t;\lambda) \cdot w_{u_k v} \quad (10)$$

Based on Eqs 9 and 10, in order to prove that the function is submodular, we have to prove that $|P^{M^1}_{u_k}| \geq |P^{M^2}_{u_k}|$, i.e., the nodes protected when u_k is added to $M_1 \subseteq M^2$, equal or exceed the number of nodes protected by u_k when the node is added to M^2. For the nodes P^{M^2} protected from the seed set $M^2 \supseteq M^1$, it holds that $P^{M^2} = P^{M^1} \cup P^{M^2 \setminus M^1}$, i.e., the set of nodes protected from M^2 equals the set of nodes protected from M^1, augmented by the protected nodes due to the additive seeds belonging in the relative complement of M^2 with respect to M^1. Denoting as P_{u_k} the set of nodes protected when u_k is the only seed, we get that $P^{M^1}_{u_k} = P_{u_k} \setminus P^{M^1}$ and $P^{M^2}_{u_k} = P_{u_k} \setminus P^{M^2}$. By further analysing the equation of $P^{M^2}_{u_k}$, we get $P^{M^2}_{u_k} = P_{u_k} \setminus P^{M^2} = P_{u_k} \setminus (P^{M^1} \cup P^{M^2 \setminus M^1}) = (P_{u_k} \setminus P^{M^1}) \cap (P_{u_k} \setminus P^{M^2 \setminus M^1}) = P^{M^1}_{u_k} \cap (P_{u_k} \setminus P^{M^2 \setminus M^1})$. Hence $|P^{M^2}_{u_k}| \leq |P^{M^1}_{u_k}|$ and consequently the function is submodular. □

6. Experimental evaluation

We conducted a set of experiments on two real world datasets, namely the Sandy Dataset and the NetHepPh dataset. The Sandy Dataset consists of tweets related to the hurricane Sandy, a large scale emergency event. We choose this dataset as it is indicative of the way emergency related information flows in the network. The second dataset is a collaborative network that covers scientific collaborations between authors of papers submitted to High Energy Physics - Phenomenology category in the e-print arXiv and it is widely used for testing information diffusion purposes [5,15,18]. In order to assess the performance of our approach, we implemented and compared three seed selection techniques.

Degree: Under the Degree seed selection technique, nodes are added to the seed set based on their out-Degree, i.e., nodes with the highest number of outgoing edges are selected as seeds. Based on our definition of edges in Section 2.2, nodes with high out-Degree values are highly influential, as an edge from node u towards v, i.e., $e_{u \to v}$ denotes that v is influenced by u. Therefore, the highest the degree of a user, the greater the number of users influenced by this node. Although this technique reaches many users, the users most likely to be infected may never be reached as the degree metric does not consider misinformation originators and their proximity to the rest of the users in the network.

Greedy viral stopper (GVS): The GVS technique constitutes a modification of the algorithm suggested by Nguyen et al. [5]. Instead of searching the minimum seed set for the decontamination of nodes, seeds are selected until a desired number of seeds is reached. The GVS model iteratively selects seeds. At each step k a node u that results in the greatest number of users protected from seed set $S_k \cup u$ is selected. However, this approach ignores that users may later switch beliefs regarding the information they propagate.

REACT: REACT (REal-time And CosT-effective misinformation blocking) constitutes the implementation of our approach, that exploits simulated annealing to determine the appropriate set of seeds to decontaminate the network.

Concerning propagation times between users, since these are not provided, we randomly generate a set of 10 time intervals $t_i \in T_{uv}$ between any two user u and v, where $u, v \in V$ and $e_{u \to v} \in E$ in the social graph $G(V, E)$ of the network. Every t_i ranges between 0 and 5 minutes, i.e., 0–300 s. For the experimental evaluation, we assign the initial renouncement thresholds and credibility scores of the users uniformly at random. We evaluate the performance of our approach under varying cost constraints, by tuning the parameter k that expresses the number of users to act as seeds for the dissemination of credible news.

6.1. Decontamination performance

In the first set of experiments, our goal is to estimate the number of nodes decontaminated under varying sizes of misinformation originators. For these experiments we ignore any time constraints that may be required on the misinformation blocking. This is achieved by setting $p(t \leq 5min; \lambda)$ for all nodes, that results to 1, since all times are generated between 0 and 5 min, as mentioned above. The misinformation originators are set to 10% and 15% of the nodes with the lower credibility values of all nodes in the social graph. We choose 10% in accordance to [8] and 15% as in [5]. Although misinformation originators may be more densely connected to one another, rather than randomly placed and connected on the network, e.g., due to homophily, our approach for seed selection is independent of the structure of the network and the connectivity of the users and therefore performance is not expected to significantly vary in such cases.

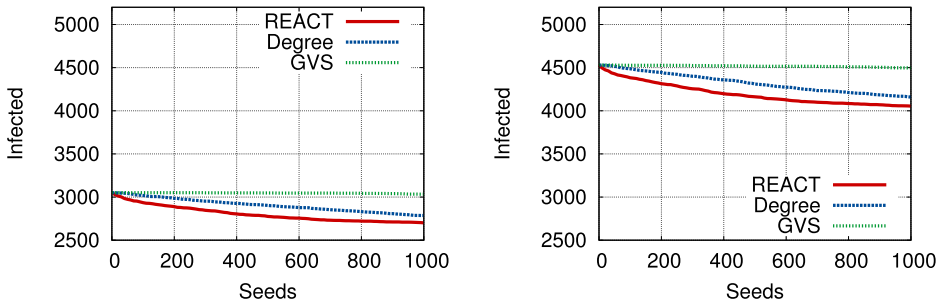
6.1.1. Sandy dataset

The first dataset is related to tweets regarding the Sandy Hurricane, a major emergency event that unfolded in 2012, from October 22 to November 2, and severely affected the area of New York City. The dataset is derived by the work in [24] and tweets related to the event were collected based on the keywords “sandy” and “hurricane”. In order to form the graph, we used the replies a tweet received. For each reply an edge is formed between the user that published the tweet and the user that replied. The influence flows from the original publisher to the responder, since the users that reply are affected from the tweet and therefore influenced by the publisher of the tweet. We refer to the user of the original tweet as *source*. Users that presented no interactions are excluded from the social graph. The final graph $G(V, E)$ consists of 25838 users and 23913 edges. The influence of a source u to a user v, i.e., the weight of the edge, is calculated as follows:

$$influence_{u \to v} = \frac{R_v(u)}{max\{R_v(u') : \forall u' \in V\}} \quad (11)$$

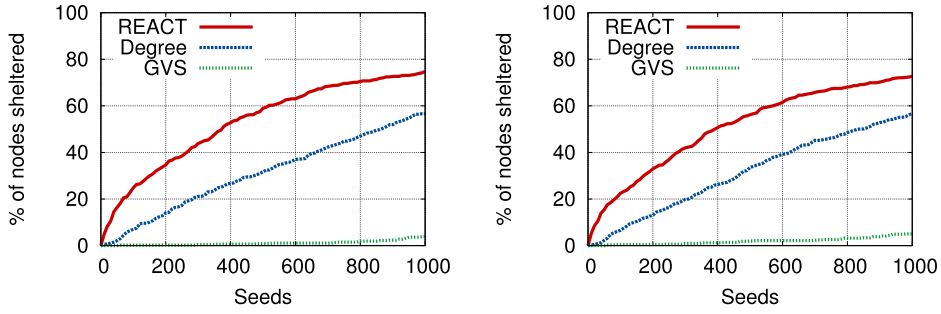
where $R_v(u)$ is the total number of replies of user v to user u. Intuitively, the influence of u to v is estimated based on the maximum number of interactions of v with any other user $u' \in V$ in the network. We chose to express the weight of the edges based on the replies, rather than the retweets, since the dataset presented a limited number of retweets.

In Fig. 2 we present the number of infected nodes after the propagation process is completed, given seed sets of different sizes, and in Fig. 3 we present the percentage of nodes *sheltered* by the selected seeds. We define as sheltered the nodes that are infected during the propagation of misinformation when protectors (i.e. seeds) are absent. Overall, REACT outperforms the Degree approach for seed selection, by managing to decontaminate a larger portion of users in the network and at a significant redundancy of load to the network due to excess messages, as observed in Fig. 4. The fact that REACT performs better than the Degree approach, while not overloading the network with excess traffic of messages,



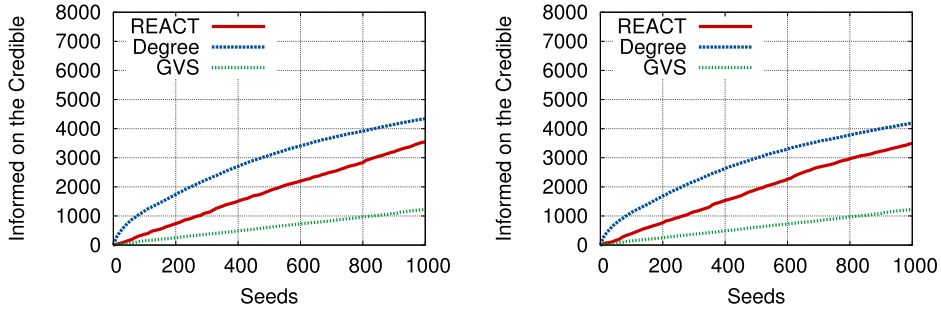
(a) 10% misinformation originators (b) 15% misinformation originators

Fig. 2. Number of Infected nodes (Sandy).



(a) 10% misinformation originators (b) 15% misinformation originators

Fig. 3. Percentage of nodes sheltered (Sandy).



(a) 10% misinformation originators (b) 15% misinformation originators

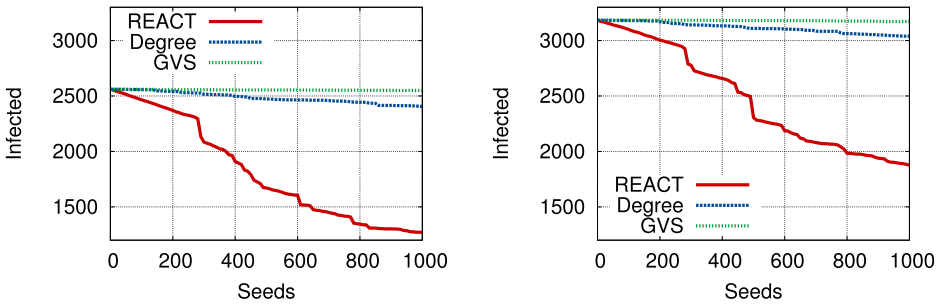
Fig. 4. Cost-Nodes receiving additional messages (Sandy).

is an important advantage of the suggested technique, as it demonstrates that it is cost effective in terms of resources consumption. This is particularly useful in cases that the network may be already over-utilized, as the cases of emergency events [25]. GVS approach seems to perform poorly under all cases, as it only considers the number of nodes informed on the credible information rather the minimization of nodes infected. Same trends are observed when we set the misinformation originators to either 10% or 15%, with a slightly drop at the number of nodes protected when the misinformation originators increase.

6.1.2. High-energy physics citation network - NetHepPh

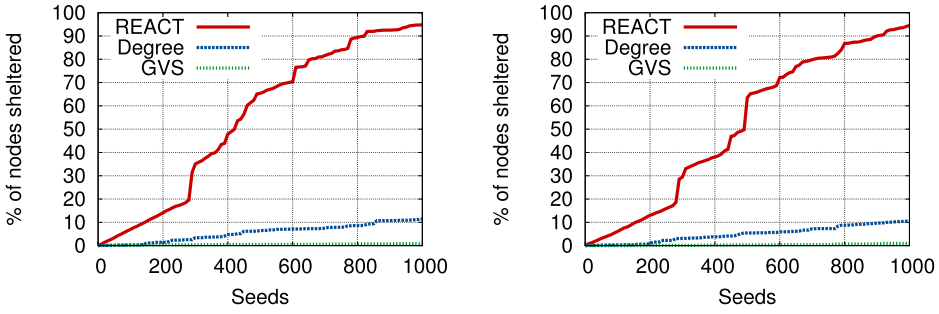
This dataset is a citation graph provided as part of the 2003 KDD Cup [26] and information on the construction of the original graph can be found in [27]. As the graph contains only undirected edges while we require a directed graph, we alter the original dataset so that when an author u co-authored a paper with author v, the graph contains two directed edges, from u to v and from v to u. The resulting graph consists of 12008 nodes and 237010 directed edges. Since the dataset contains unweighted edges, we assign weights uniformly at random, similarly to [5]. Contrary to the Sandy Dataset, NetHepPh presents high connectivity among nodes and since all nodes are connected bi-directionally, there is a higher chance that a node influences another, making it difficult to reach at a steady state during propagation.

Results for the NetHepPh dataset are presented in Figs. 5 through 7. We observe that the advantage of REACT over its competitors becomes clearer in this dataset. Unlike the Degree or the GVS approach, REACT considers the probability of a user renouncing an adopted cascade during the propagation propagation. Since the users are tightly connected it is more likely that a user renounces an adopted belief, as the influence exceeded by her neighbors is greater. Similarly to the Sandy dataset, same trends are observed in the number of nodes sheltered regardless of the number of misinformation originators. We observe in Fig. 7 that more messages concerning the credible information are propagated in the network using REACT, which is expected, as the propagation of the credible information from seeds with the Degree and GVS is restricted, as suggested by Figs. 5 and 6.



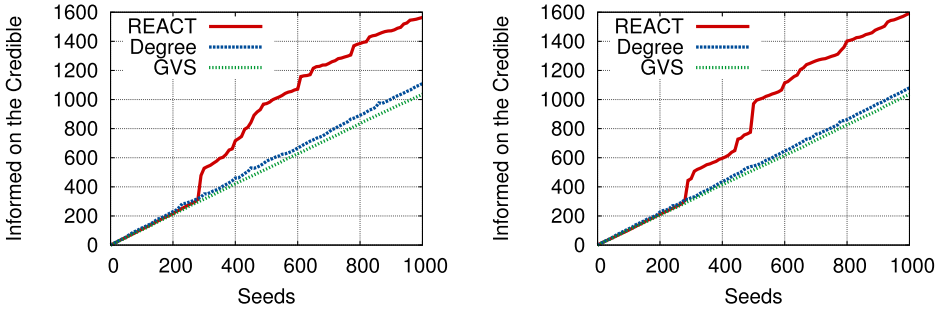
(a) 10% misinformation originators (b) 15% misinformation originators

Fig. 5. Number of Infected nodes (NetHepPh).



(a) 10% misinformation originators (b) 15% misinformation originators

Fig. 6. Percentage of nodes sheltered (NetHepPh).



(a) 10% misinformation originators (b) 15% misinformation originators

Fig. 7. Cost-Nodes receiving additional messages (NetHepPh).

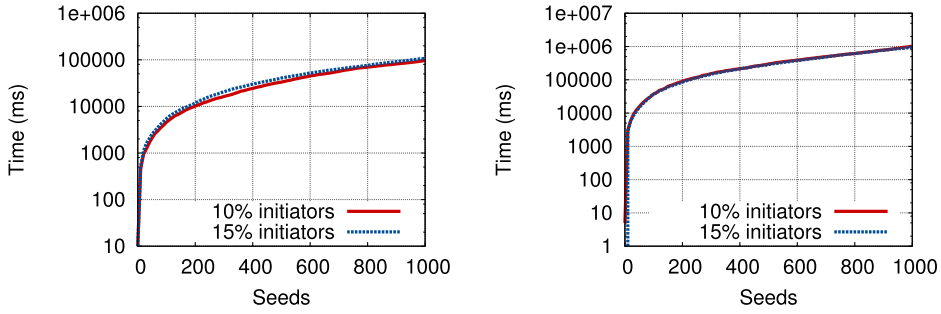
6.2. Execution times and time constraints

In Fig. 8a we present the execution times for the REACT approach in the Sandy dataset. We note that as the number of misinformation originators increases, so does the execution time, but only slightly. Overall, the process requires less than 15 min. In Figure 8b we present the corresponding results for the NetHepPh dataset. We observe that in the NetHepPh dataset the execution time slightly increases, yet less than 23 min are required to estimate a seed set of size $k = 1000$ with 15% of the nodes as misinformation originators.

In Fig. 9a and b we present the percentage of nodes sheltered in the Sandy and NetHephPh network respectively, under different propagation time requirements. Results are similar for both 10% and 15% misinformation originators, hence we only present the results for 15%. We set the propagation time between any two nodes in the network to be either tight, i.e., 2 min or relaxed, i.e., 5 min. Constraining the time requirements between any two users affects the adoption probability (Eq. (2)) and the time required for the information to propagate, i.e., the constraint of 2 min expresses the emergency of a user informing her neighbors, as opposed to 5. The results suggest that when time requirements are relaxed, then the number of users sheltered increases. The above observation becomes obvious in networks that present higher connectivity and longer paths between users, as is the case with the NetHephPh network.

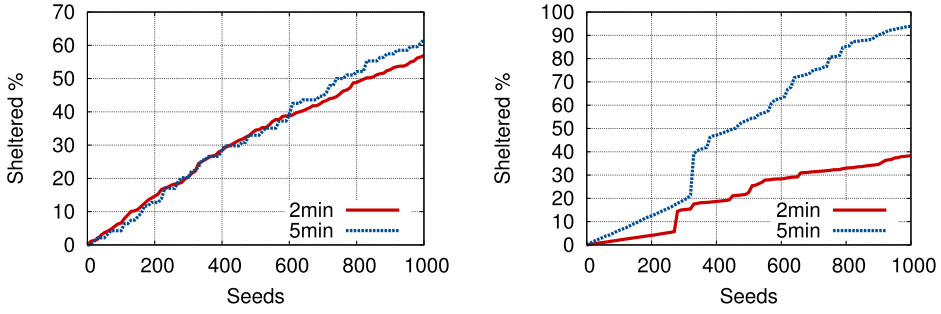
6.3. Seeds unwilling to share the credible information

In Fig. 10 we present the percentage of nodes sheltered in case seeds selected are unwilling to propagate the credible information. For this set of experiments we set the time requirements to 5 min delay at each edge. The probability that a user declines to initiate the propagation of the credible news is expressed as the initial renouncement threshold of the user. A random probability r is generated for each seed u and if $r_u \leq r$, u participates in the credible news propagation. Results are similar for 10% and 15% misinformation originators on both datasets, hence we only present results for 15% misinformation originators. On average 51% of the seeds



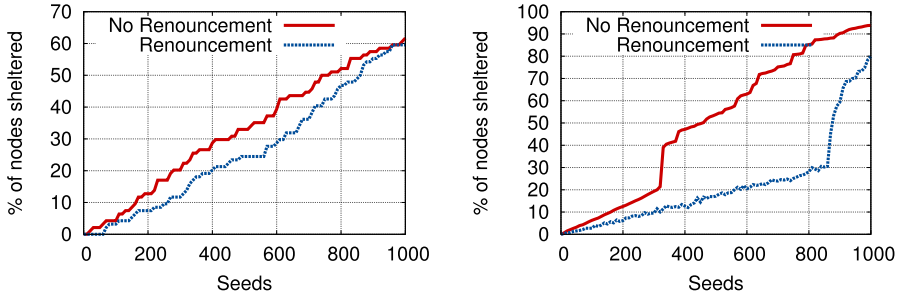
(a) Sandy Dataset (b) NetHepPh Dataset

Fig. 8. Time required for seed set selection.



(a) Sandy(15% misinformation originators) (b) NetHepPh(15% misinformation originators)

Fig. 9. Nodes sheltered under different time requirements.



(a) Sandy(15% misinformation originators) (b) NetHepPh (15% misinformation originators)

Fig. 10. Impact of seeds unwillingness to share the information.

accept to propagate the information on the Sandy dataset and 49% on NetHepPh. In the Sandy dataset (Fig. 10a) we note that there is a slight drop on the number of sheltered nodes, yet the trend remains similar. As more seeds are added, the number of sheltered nodes increases with at almost the same rate. However, we observe that in the network of NetHepPh the impact of the seeds denial to propagate the information creates an interesting behavior. From 0 to 850 selected seeds, with an average of 49% participating in the propagation, i.e., 0–425 seeds, the number of sheltered nodes increases at a steady rate. However, when more than 425 seeds participate, there is a sharp increase in the number of sheltered nodes. In particular, when a certain number of nodes is sheltered, and herein block the misinformation propagation, the more limited seems to be the spread of misinformation. This may be due the tight connectivity of the nodes in the network. It is interesting to note that this behavior is presented even when all seeds accept to propagate the credible news, although more smoothly, as observed for the values between 250 and 400 seeds.

7. Related work

Epidemic propagation models. Beutel et al. [28] study the problem of competing ideas/viruses spreading in the network. In their work they examine the validity of the assumption the "winner takes it all" made by most works that examine competitive models. Given a variation of the SIS model that considers that a user may be infected on either propagation campaigns, they aim in identifying the fixed points for the system, in which both viruses can survive. Rapti et al. [29] consider the propagation of viruses in multiple profile networks. In their work, users have different affinity to a particular piece of information. By exploiting the SIS propagation model, they aim at determining the necessary conditions un-

der which the virus reaches a particular equilibrium state. They examine special network topologies, i.e., clique and arbitrary graphs with low and high connectivity, to extract the necessary conditions. In this work, users affinity remains fixed, while in our work we consider it to be varying over time and it is presented through the renouncement value. Moreover, we consider the spread of conflicted information through the LT model, i.e, users may not be susceptible on a single information, but on both competitive information diffused in the network. Moreover, in this work we aim at identifying the subset of users to initiate a propagation so that the spread of the conflicted information is minimized. Myers et al. [30] examine how different contagions interact with each other as they spread through the network. A statistical model is proposed that allows for competition as well as cooperation of different contagions in information diffusion. A mixing model is developed, where the probability of a user adopting a piece of content is based on what other content the users was previously exposed. It is assumed that the infection probability does not change from user to user.

Competition using independent cascading and linear thresholds. Clark et al. [31] exploit a Game-Theoretic approach to estimate the best strategy *players* should deploy in competitive environments, so as to maximize their influence in the network. They propose the Dynamic Influence in Competitive Environments (DICE) propagation model, where each user has the ability to hold multiple ideas with different probability. The DICE model uses Markov processes to model the propagation of ideas through a social network. LT and IC model constitute special cases of the DICE, where users hold a single idea. For the special case of social graphs with Strongly Connected Components, the suggested objective function of the players is submodular. In their work, competing players may choose to add seeds to maximize influence, while our objective is to block the influence of a specific idea. Moreover, we consider that misinformation spread is unintentional and therefore not strategic, thus, no utility function can be determined for the seeds of the misinformation. He et al. [15] study the Influence Blocking Maximization (IBM) problem. They suggest the Competitive Linear Threshold (CLT) propagation model. Each vertex in the social graph is assigned a threshold and each edge is associated with a negative and a positive weight. Users may be positively or negatively activated, depending on which activation is triggered first. In cases both activations reach the threshold simultaneously, the negative prevails over the positive. In their model they assume that there no independent paths may exist between nodes of the seed set to other nodes, i.e., if there are multiple paths between any two users, then the one is sub-path of the other. Moreover, time is only considered in the seed selection process, while during propagation it is expressed as the number of hops in the path. Lin et al. [32] study the problem of influence maximization under the Competitive Linear Threshold (CLT) model, but define CLT slightly differently than the work [15]. A node is activated by the party that has the highest overall influence and exceeds the threshold of the user. They assume that if a node is activated by a party, it cannot be activated again by another party. Players select seeds interchangeably. That is, they define a multi-round multi-party influence maximization problem.

Fan et al. [8] study the rumor blocking problem in OSNs. Their goal is to select the minimal subset of users, referred as *Protectors*, in order to minimize the propagation achieved by *Rumor Originators*. They define the One-Activate-One and One-Activate-Many diffusion models and deploy Breadth First Search (BFS) to Rumor Forward Search Trees to locate bridge ends. In their model protectors influence has priority over rumors when both reach a user in the network simultaneously. In contrary to their work, we consider the number of protectors given and aim at locating the more appropriate seeds. Similar to the work in [8], Nguyen et al. [5] aim at identifying the smallest set of protectors to contain the spread of misinformation to a desired ratio $1-\beta$ in T steps. As in [8] they assume that credible information takes over when a node is simultaneously informed about both pieces of information. Their goal is to find the smallest set of highly influential nodes whose decontamination with good information assists in limiting the spread of misinformation to a desired ration $1-\beta$ in T steps. To achieve the above objective they propose a Greedy Viral Stopper algorithm that utilizes a modification of the hill-climbing algorithm and a community bases approach that greedily selects nodes form each community. Butak et al. [9] aim in the limitation of misinformation in OSNs under the Multi-Campaign Independent Cascade model. Considering the evolving propagation of competitive campaigns, their goal is to select the appropriate subset of nodes so that the eventual influence of the misinformation campaign is minimized. They refer to this problem as Eventual Influence Limitation. In their model, each of the competing cascades follows the Independent Cascade model and whenever a node is influenced simultaneously from both campaigns, then it is assumed that the "good" campaign prevails. Moreover, after a node adopts a belief, its state remains constant throughout the propagation process. They further prove the performance of their approach in the presence of missing data, where states of nodes are know with a certain probability. Zhang et al. [33] suggest an approach to limit misinformation while maximizing the spread of good information in the network. They initially identify a set of *Getaway Nodes*, that are the nodes that contribute to further propagation of misinformation and offer the greatest marginal gain to the spread. The gain is estimated as the number of incident nodes activated by the Getaway Nodes. Given the Getaway nodes, their goal is to select a subset of users to initiate the credible propagation so that the Gateway nodes are positively activated before the misinformation reaches them.

Borotin et al. in their work [34] propose extensions of the Linear Threshold propagation model to formulate competitive influence cascades in the network. They define a competitive model that a user adopts a technology if the the collective weight of neighbors that adopted the technology exceeds the users threshold. Time-constraints are ignored in their model and a user that adopts a specific technology may not switch state afterwards. They further define a variation of LT model where users have different thresholds for each technology propagated in the network and an alternative of the latter, where at the end of each propagation all inactive users are randomly choosing a technology and adopt it. Pathak et al. [35] suggest a Generalized Linear Threshold Model for multiple cascades in social networks that allows users to switch states regarding the adopted cascade. The subject of their study is to estimate the equilibrium of the different cascades. Users reluctance in adopting a new idea remains constant, regardless of the users prior believes alternation. Finally, the way opposing campaigns correlate to the users conviction is ignored.

Influence maximization. In our previous work [19] we study the problem of efficient information dissemination in location-based social networks under time constraints. The objective is to identify a subset of individuals to propagate the information and make intelligent route selection that can result in maximizing the reach within a time window. The underlying propagation model is the Independent Cascade propagation model and any contradictory propagation is ignored. Li et al. [36] address the problem of real time targeted influence maximization for online advertisements. Their goal is to find a seed set that maximizes the expected influence over users who are relevant to a given advertisement. The Independent Cascade model is exploited to characterize the dissemination of the information to the network and user profiles express user preferences. The preferences of the users are considered in order to estimate the probability of a user being effectively influenced on a specific topic that is propagated on the network.

Gayraud et al. [37] study the problem of influence maximization in evolving Social Networks. They show that when considering evolving networks the diffusion function is no longer submodular for the transient models and for the transient Independent Cascade model is neither monotone. They assume that the entire graph sequence is known in advance and aim in selecting a subset of users to initiate the propagation so that the final number of users activated under any diffusion model is maximized. Chou and Chen[38] propose a Multiple Factors-Aware Diffusion model to determine the diffusion of a content over Social Networks. For each user they exploit a probabilistic classifier that considers multiple predefined features that affect the tendency o user adopting adopting an item after she is exposed to it. Chen et al. [39] develop a community-based influence maximization (CIM) framework, to tackle the problem of influence maximization by exploiting the graph structure. Specifically, they discover the community structure of the network and exploit the detected communities to narrow down the possible seed candidates and later select the best seed nodes from the candidate set. To capture the temporal process of information diffusion they consider the heat diffusion model (HDM).

8. Conclusions

In this paper, we propose an approach for misinformation limitation that is aware of two important dimensions, currently neglected by the related work. The first one is that propagation time differs among the users of a social network. The second is that users' susceptibility to new information should dynamically adapt over time. Considering the aforementioned dimensions in propagation unfolding in the social networks, we design a propagation model that manages to capture the way contradictory information unfolds in the network, namely the Dynamic Linear Threshold (DLT) model. DLT, unlike existing models, considers the aspect of users' hesitation in adopting and further propagating a specific piece of information when they are exposed to directly opposing information. Given DLT, we develop REACT (REal-time And CosT-effective misinformation blocking), a greedy seed selection algorithm that identifies the appropriate subset of users contributing in the minimization of misinformation spread. We provide theoretical and experimental results that prove and demonstrate the efficiency and effectiveness of our algorithms. Specifically, we formally prove that REACT achieves a constant approximation ration of $1 - 1/e$ to the best solution under the DLT propagation and we contact a set of experiments in two real-world networks that showcase the superiority of REACT to state-of-the-art approaches for misinformation blocking under various scenarios.

Our research agenda includes a number of challenging problems to be addressed in future work. An interesting problem to consider is the case of the dynamic nature of the social graph. Links between users in the network are constantly generated or disappear and even the influence probabilities vary over time, leading to constant updates on the structure of the graph that should be consider during the misinformation blocking process. Another interesting issue is the case of non-credible users identification in real time, which requires the analysis of both information and users' validity.

Acknowledgments

This research has been financed by the European Union through the FP7 ERC IDEAS 308019 NGHCS project, the Horizon2020 688380 VaVeL project and a Google 2017 Faculty Award.

References

[1] Twitter company - about,(https://about.twitter.com/company).
[2] Company info - facebook, (http://newsroom.fb.com/company-info/).
[3] A. Olteanu, C. Castillo, N. Diakopoulos, K. Aberer, Comparing events coverage in online news and social media: the case of climate change, in: Proceedings of the 9th ICWSM, 2015, pp. 288–297.
[4] S. Ye, S.F. Wu, Measuring message propagation and social influence on twitter.com, in: Proceedings of the 2nd SocInfo, in: SocInfo'10, Springer-Verlag, Berlin, Heidelberg, 2010, pp. 216–231.
[5] N.P. Nguyen, G. Yan, M.T. Thai, S. Eidenbenz, Containment of misinformation spread in online social networks, in: Proceedings of the 4th Annual ACM Web Science Conference, in: WebSci '12, ACM, New York, NY, USA, 2012, pp. 213–222.
[6] S. Antoniadis, I. Litou, V. Kalogeraki, A model for identifying misinformation in online social networks, in: Proceedings of the On the Move to Meaningful Internet Systems: OTM 2015 Conferences, 2015, pp. 473–482.
[7] M. Guy, P. Earle, C. Ostrum, K. Gruchalla, S. Horvath, Integration and dissemination of citizen reported and seismically derived earthquake information via social network technologies, in: Proceedings of the 9th IDA, in: IDA'10, Springer-Verlag, Berlin, Heidelberg, 2010, pp. 42–53.
[8] L. Fan, Z. Lu, W. Wu, B. Thuraisingham, H. Ma, Y. Bi, Least cost rumor blocking in social networks, in: Proceedings of the 33rd IEEE ICDCS, in: ICDCS '13, IEEE Computer Society, Washington, DC, USA, 2013, pp. 540–549.
[9] C. Budak, D. Agrawal, A. El Abbadi, Limiting the spread of misinformation in social networks, in: Proceedings of the 20th WWW, in: WWW '11, ACM, 2011, pp. 665–674.
[10] B.A. Prakash, J. Vreeken, C. Faloutsos, Spotting culprits in epidemics: how many and which ones? in: Proceedings of the 12th IEEE ICDM, in: ICDM '12, IEEE Computer Society, Washington, DC, USA, 2012, pp. 11–20.
[11] C. Castillo, M. Mendoza, B. Poblete, Predicting information credibility in time-sensitive social media., Internet Res. 23 (2013) 560–588.
[12] J. Marshall, D. Wang, Mood-sensitive truth discovery for reliable recommendation systems in social sensing, in: Proceedings of the 10th ACM Conference on Recommender Systems, in: RecSys '16, 2016, pp. 167–174.
[13] L. Wu, J. Li, X. Hu, H. Liu, Gleaning wisdom from the past: early detection of emerging rumors in social media, Proceedings of the 2017 SIAM International Conference on Data Mining, pp. 99–107.
[14] A. Bessi, F. Petroni, M.D. Vicario, F. Zollo, A. Anagnostopoulos, A. Scala, G. Caldarelli, W. Quattrociocchi, Homophily and polarization in the age of misinformation, Eur. Phys. J. Spec. Top. 225 (10) (2016) 2047–2059.
[15] X. He, G. Song, W. Chen, Q. Jiang, Influence blocking maximization in social networks under the competitive linear threshold model., in: SDM, SIAM / Omnipress, 2012, pp. 463–474.
[16] D. Kempe, J. Kleinberg, E. Tardos, Maximizing the spread of influence through a social network, in: Proceedings of the 9th ACM SIGKDD, in: KDD '03, ACM, New York, NY, USA, 2003, pp. 137–146.
[17] B. Liu, G. Cong, D. Xu, Y. Zeng, Time constrained influence maximization in social networks, in: Proceedings of the 12th IEEE ICDM, 2012, pp. 439–448.
[18] W. Chen, W. Lu, N. Zhang, Time-critical influence maximization in social networks with time-delayed diffusion process, CoRR abs/1204.3074 (2012).
[19] I. Litou, I. Boutsis, V. Kalogeraki, Using location-based social networks for time-constrained information dissemination, in: Proceedings of the 15th IEEE MDM, 1, 2014, pp. 162–171.
[20] Facebook, How does news feed decide which stories to show?, 2015,(https://www.facebook.com/help/166738576721085). Accessed: 2015-07-31.
[21] P. Rosania, While you were away..., 2015,(https://blog.twitter.com/2015/while-you-were-away-0). Accessed: 2015-07-31.
[22] M. Pielot, R. de Oliveira, H. Kwak, N. Oliver, Didn't you see my message?: Predicting attentiveness to mobile instant messages, in: Proceedings of the 32nd ACM CHI, in: CHI '14, ACM, New York, NY, USA, 2014, pp. 3319–3328.
[23] J.a. Gama, I. Žliobaitė, A. Bifet, M. Pechenizkiy, A. Bouchachia, A survey on concept drift adaptation, ACM Comput. Surv. 46 (4) (2014) 44:1–44:37.
[24] A. Zubiaga, H. Ji, Tweet, but verify: epistemic study of information verification on twitter, Soc. Netw. Anal. Min. 4 (1) (2014).
[25] A. Hughes, L. Palen, Twitter adoption and use in mass convergence and emergency events, Int. J. Emerg. Manag. 6 (3) (2009) 248–260.
[26] J. Gehrke, P. Ginsparg, J. Kleinberg, Overview of the 2003 kdd cup, SIGKDD Explor. Newsl. 5 (2) (2003) 149–151.
[27] J. Leskovec, J. Kleinberg, C. Faloutsos, Graphs over time: densification laws, shrinking diameters and possible explanations, in: Proceedings of the 11th ACM SIGKDD, in: KDD '05, ACM, New York, NY, USA, 2005, pp. 177–187.
[28] A. Beutel, B.A. Prakash, R. Rosenfeld, C. Faloutsos, Interacting viruses in networks: can both survive? in: KDD, ACM, 2012, pp. 426–434.
[29] A. Rapti, K. Tsichlas, S. Sioutas, G. Tzimas, Virus propagation in multiple profile networks, CoRR abs/1504.03306 (2015).
[30] S.A. Myers, J. Leskovec, Clash of the contagions: Cooperation and competition in information diffusion, in: Proceedings of the 12th IEEE ICDM 2012, Brussels, Belgium, 2012, pp. 539–548.
[31] A. Clark, R. Poovendran, Maximizing influence in competitive environments: a game-theoretic approach, in: Decision and Game Theory for Security, in: Lecture Notes in Computer Science, vol. 7037, Springer Berlin Heidelberg, 2011, pp. 151–162.
[32] S.-C. Lin, S.-D. Lin, M.-S. Chen, A learning-based framework to handle multi-round multi-party influence maximization on social networks, in: Proceedings of the 21th ACM SIGKDD, in: KDD '15, ACM, New York, NY, USA, 2015, pp. 695–704.
[33] H. Zhang, H. Zhang, X. Li, M. Thai, Limiting the spread of misinformation while effectively raising awareness in social networks, in: Computational Social Networks, in: Lecture Notes in Computer Science, vol. 9197, Springer International Publishing, 2015, pp. 35–47.

[34] A. Borodin, Y. Filmus, J. Oren, Threshold models for competitive influence in social networks, in: Proceedings of the 6th WINE, in: WINE'10, Springer-Verlag, Berlin, Heidelberg, 2010, pp. 539–550.
[35] N. Pathak, A. Banerjee, J. Srivastava, A generalized linear threshold model for multiple cascades, in: Proceedings of the 10th IEEE ICDM, 2010, pp. 965–970.
[36] Y. Li, D. Zhang, K.-L. Tan, Real-time targeted influence maximization for online advertisements, Proc. VLDB Endow. 8 (10) (2015) 1070–1081.
[37] N.T. Gayraud, E. Pitoura, P. Tsaparas, Diffusion maximization in evolving social networks, in: Proceedings of the 2015 ACM COSN, in: COSN '15, ACM, New York, NY, USA, 2015, pp. 125–135.
[38] C.-K. Chou, M.-S. Chen, Multiple factors-aware diffusion in social networks, in: T. Cao, E.-P. Lim, Z.-H. Zhou, T.-B. Ho, D. Cheung, H. Motoda (Eds.), Advances in Knowledge Discovery and Data Mining, Lecture Notes in Computer Science, vol. 9077, Springer, 2015, pp. 70–81.
[39] Y.-C. Chen, W.-Y. Zhu, W.-C. Peng, W.-C. Lee, S.-Y. Lee, Cim: Community-based influence maximization in social networks, ACM Trans. Intell. Syst. Technol. 5 (2) (2014) 25:1–25:31.

Iouliana Litou is a PhD candidate in Computer Science at Athens University of Economics and Business under the supervision of the Professor Vana Kalogeraki. She holds a M.Sc. degree in Computer Science and a B.Sc. degree in Informatics from the same department. She is a member of the Real Time Distributed Systems Group and works as a Research Assistant on EU-funded research projects. Her research interests focus on Social Networks Analytics, Social Campaigns Influence Maximization, Misinformation Discovery and Propagation Limitation. She has published papers in International Conferences and Scientific Journals related to her area of expertise. More information can be found on her personal web page: http://www2.aueb.gr/users/litou/.

Vana Kalogeraki is an Associate Professor leading the Distributed and Real-Time Systems research at Athens University of Economics and Business. Previously she has held positions as an Associate and Assistant Professor at the Department of Computer Science at the University of California, Riverside and as a Research Scientist at Hewlett-Packard Labs in Palo Alto, CA. She received her PhD from the University of California, Santa Barbara in 2000. Prof. Vana Kalogeraki has been working in the field of distributed and real-time systems, participatory sensing systems, peer-to-peer systems, crowdsourcing, mobility, resource management and fault-tolerance for over 20 years and has authored and co-authored over 150 papers in journals and conferences proceedings, including co-authoring the OMG CORBA Dynamic Scheduling Standard. Prof. Kalogeraki was invited to give keynote talks at MoVid2015, DNCMS 2012, SN2AE 2012, PETRA 2011, DBISP2P 2006 and MLSN 2006 in the areas of participatory sensing systems and sensor network middleware and delivered tutorials and seminars on peer-to-peer computing. She has served as the General co-Chair of SEUS 2009, the General co-Chair of WPDRTS 2006 and as a Program co-Chair of MDM 2017, DEBS 2016, MDM 2011, ISORC 2009, ISORC 2007, ICPS 2005, WPDRTS 2005 and DBISP2P 2003, a Tutorial Chair for ACM DEBS 2015, a Workshops Chair for IEEE SRDS 2015, a Demo Chair for IEEE MDM 2012, in addition to other roles such as Area Chair (IEEE ICDCS 2016, 2012) and as program committee member on over 200 conferences. She was also awarded a Marie Curie Fellowship, three best paper awards at the 11th ACM International Conference on Distributed Event-Based Systems (DEBS 2016), 24th IEEE International Parallel and Distributed Processing Symposium (IPDPS 2009) and the 9th IEEE Annual International Symposium on Applications and the Internet (SAINT 2008), a Best Student Paper Award at the 11th IEEE/IPSJ International Symposium on Applications and the Internet (SAINT 2011), a UC Regents Fellowship Award, UC Academic Senate Research Awards and a research award from HP Labs. Her research has been supported by an ERC Starting Independent Researcher Grant, the European Union, joint EU/Greek "Aristeia" grant, a joint EU/Greek "Thalis" grant, NSF and gifts from SUN and Nokia.

Dr Ioannis Katakis is a senior researcher at the National and Kapodistrian University of Athens working on EU-funded research projects as a Quality Assurance Coordinator. His research interests include Mining Social and Web Data, Knowledge Discovery from Text and Urban Data Streams, Multi-label Learning, Adaptive and Recommender Systems. He has published papers in International Conferences and Scientific Journals related to his area of expertise, organized three workshops, edited four special issues and is an Editor at the journal Information Systems. More information can be found at his personal web page: http://www.katakis.eu.

Dimitrios Gunopulos is a Professor at the Department of Informatics and Telecommunications, at the National and Kapodistrian University of Athens. He got his PhD from Princeton University in 1995. He has held positions as a Postoctoral Fellow at the Max-Planck-Institut for Informatics, Research Associate at the IBM Almaden Research Center, Visiting Researcher at the University of Helsinki, Assistant, Associate, and Full Professor at the Department of Computer Science and Engineering in the University of California Riverside, and Visiting Researcher in Microsoft Research, Silicon Valley. His research is in the areas of Data Mining, Knowledge Discovery in Databases, Databases, Sensor Networks, Peer-to-Peer systems, and Algorithms. He has co-authored over a hundred journal and conference papers that have been widely cited and a book. He has 11 Ph.D. students that have joined industry labs or have taken academic positions. His research has been supported by NSF (including an NSF CAREER award), the DoD, the Institute of Museum and Library Services, the Tobacco Related Disease Research Program, the European Commission, the General Secretariat of Research and Technology, AT&T, Nokia, the Stavros Niarchos Foundation, a Yahoo Faculty Award and a Google Faculty Award. He has served as a General co-Chair in SIAM SDM 2018, SIAM SDM 2017, HDMS 2011 and IEEE ICDM 2010, as a PC co-Chair in ECML/PKDD 2011, in IEEE ICDM 2008, in ACM SIGKDD 2006, in SSDBM 2003, and in DMKD 2000.