Computer Communications 41 (2014) 31–42

Contents lists available at ScienceDirect

Computer Communications

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

Preventing spam in opportunistic networks

Sacha Trifunovic [a,*], Maciej Kurant [b], Karin Anna Hummel [a], Franck Legendre [c]

[a] Computer Engineering and Networks Laboratory, ETH Zurich, Switzerland
[b] Google, Zurich, Switzerland
[c] Uepaa!, Zurich, Switzerland

ARTICLE INFO

Article history:
Received 30 April 2013
Received in revised form 19 December 2013
Accepted 21 December 2013
Available online 24 January 2014

Keywords:
Opportunistic networks
Content dissemination
Spam
Trust
Cold start

ABSTRACT

In case of a network outage or strict censorship, opportunistic networking is an appealing solution to uphold communications. News such as tweets or videos can be disseminated widely in an epidemic and delay-tolerant fashion between mobile phones. Yet, the cooperative nature of such networks can be abused to disseminate unsolicited content at no cost. Aside from harassing other users with spam, this behavior consumes scarce resources such as battery power.

Opportunistic networks' challenging features, such as its highly dynamic node contacts, render traditional decentralized trust and reputation frameworks insufficient. They mainly fail at the 'Cold Start Problem' when mobile users find themselves in a new surrounding without established trust or reputation available, a frequent phenomena in opportunistic networks.

To overcome these challenges we propose *Trust-Based Spreading* (TBS) – a scheme where trusted nodes collaborate and filter spam by opportunistically exchanging assessments to promote or block the spreading of content. TBS copes with the 'Cold Start Problem' by allowing the trust structure to be initialized randomly and being extremely resilient to false positives in the feedback process. We evaluate TBS by replaying a variety of real-world mobility traces and show that TBS disseminates legitimate content almost as effectively as classical epidemic spreading, while significantly limiting the reach of spam.

1. Introduction

The performance of every network infrastructure may be significantly hampered by natural disasters such as earthquakes or floods. Moreover, recent uprisings in Northern Africa [1] have shown that authoritarian governments can arbitrarily *censor* communications. Commonly used techniques range from blocking online social networks [2], e.g., Facebook or Twitter, to enforcing an Internet and mobile phone outage [3].

Fortunately, with increasing penetration of wireless mobile devices, *opportunistic networking* is becoming a feasible and appealing technology to maintain (delay-tolerant) connectivity, even under such harsh conditions and support the freedom of speech. In opportunistic networks [4,5], users/nodes cooperate to distribute content typically over one-hop ad hoc Wi-Fi/Bluetooth links. A user publishing content (e.g., video) or a message (e.g., tweet) shares it with interested users, that, in turn, spread it further. Thanks to the small-world structure of social networks [6] and of human contact graphs [7], such epidemic spreading techniques disseminate the content efficiently in terms of delivery delay [8].

On the downside, epidemic spreading, in its basic form, is indifferent to what is spread. This allows a malicious user (or institution) to disseminate spam, propaganda, or misleading information. Such content, that we henceforth simply refer to as "spam", could supplant valuable information. Moreover, the propagation of spam drains the limited resources of mobile devices such as energy, storage, and network capacity.

This problem is not new and has been extensively studied in related domains, such as P2P and mobile ad hoc networks (MANETs). The traditional solutions in such environments usually rely on some type of distributed trust and reputation management system. While the basic principle of trust and reputation management work in such a disconnected and decentralized setting and can be applied to opportunistic content dissemination, those approaches are unable to cope with the dynamic and sparse nature of opportunistic networks.

One solution to this problem is to resort to closed-group communication where only authorized users are allowed to publish new content or messages. This is, however, a very conservative strategy that lacks the diversity brought by participatory interactions in an open environment. A critical issue is hence to answer the following question:

* Corresponding author.
E-mail address: trifunovic@tik.ee.ethz.ch (S. Trifunovic).

http://dx.doi.org/10.1016/j.comcom.2013.12.003
0140-3664/

"Can we efficiently protect opportunistic networking against spam in an open participatory environment?"

1.1. Challenges and existing approaches

To answer this question we need to be aware of the specific nature of opportunistic networks, i.e., the dynamic, mobility based contacts, the challenges it results in, and how related work targets some of them.

Distributed assessments: Due to the disconnected nature of opportunistic networks no central infrastructure is available to keep track of user behavior and collect or provide assessments. This challenge also exists and has already been thoroughly studied in the related field of MANETs or P2P networks [9,10]. All approaches propose some kind of trust and reputation systems that aim at the following: decide whether or not to interact with a peer, depending on its trust or reputation. Buchegger et al. [11] present a Bayesian reputation system that takes into account the rater's trustfulness to be robust against false ratings. In Walsh and Sirer's Credence [12], ratings about objects are correlated and shared over an overlay network of users with highly correlating ratings. EigenTrust [13] is a reputation system with a global trust value per peer that is calculated in distributed manner. While all these approaches are able to build up trust and reputation values in the environment they are designed for, they do not take into account the sparseness of nodes and their intermittent contacts in opportunistic networks.

Sparse nodes and intermittent contacts: The partial connectedness assumed in MANETs and P2P networks is not given in opportunistic networks. Nodes are much sparser and their contacts intermittent, so no path can be established. While some of the approaches designed for MANETs or P2P networks may be adapted to opportunistic networks and work with reduced performance, there is also work specifically targeting delay-tolerant networks (DTN) which includes opportunistic communication. Ayday and Fekri [14] propose an iterative algorithm for trust and reputation management in DTNs that can deal with the sparseness of nodes and overcome the performance penalties other approaches suffer in this setting. However, their approach does not take into account the challenges arising from a decentralized identity management in such a dynamic environment.

Decentralized identity management: As opportunistic networks are totally disconnected, no central authority is available to manage identities. In such a disconnected and distributed environment, users (nodes) can be authenticated by a self-generated cryptographic ID based on a public/private key pair [15]. As a consequence, content can be signed by the author/publisher, ensuring non-repudiation of published content. It also makes sure that a publisher cannot be imitated, allowing for all the content of the same publisher to be linked. This, however, does not prevent malicious nodes from generating multiple IDs and perform a Sybil attack [16].

While most trust and reputation management approaches in P2P and mobile ad hoc networks consider Sybils only marginally or not at all, there is other related work focusing on Sybil defense. The most prominent defenses are based on the social network structure [17–19]. They all assume in some way or another that Sybil regions are only loosely connected to the main social graph [20]. While this assumption is questionable [21], it also neglects the disconnected and dynamic environment where we do not even have a social network to begin with.

Another related problem is 'Identity Whitewashing', which allows a spammer, whenever he/she is detected, to discard the current identity and create a new one. While most traditional approaches do not consider this problem it is especially important when facing spammers as explained next.

Dynamic environment: All presented approaches fail to prevent the following simple attack on a content dissemination scheme in opportunistic networks: send one spam, change the identity, and repeat. While some related approaches might eventually block the spamming identity, they are all too late, as the ID does not exist anymore. One could argue for the need to set the default policy not to get content from a new author, but this is not possible because nobody could ever publish any content in the first place. Even if we have another way to build up trust or reputation, the environment in opportunistic networks is potentially very dynamic as users might frequently change their location and would need to start from scratch each time they do so. We call this problem the 'Cold Start Problem' and to solve it we need to make sure we can stop spam at the source while letting good content spread as freely as possible.

User interaction: All existing approaches are based on some sort of feedback or rating. To get feedback, direct user interaction is usually required. However, generally further a small fraction of users rate the content they consume [22]. The distributed and disconnected nature of opportunistic networks only reduces the availability of the already scarce feedback. Keeping in mind the 'Cold Start Problem' we need a dissemination scheme that can stop spam at the source with only very little feedback.

1.2. Our contribution

In light of these challenges and keeping in mind the existing approaches we aim at a content spreading scheme that specifically targets the 'Cold Start Problem' taking into account the scarce feedback availability while being resilient to Sybil attacks and identity whitewashing. To accomplish that, we propose a *Trust-Based Spreading* (TBS) mechanism that spreads content based on a user's trust structure. If no real, earned, or otherwise inferred trust structure is available, i.e., due to the 'Cold Start Problem', TBS may be initialized with a random trust structure. For improved performance and Sybil resiliency, we propose to initialize the trust structure with the structure inherent to a user's mobility pattern. This feature of opportunistic networks has been shown to correlate with real trust among users [23]. While the trust structure is expected to improve over time, e.g. by applying any existing trust establishment approach, TBS can cope with inaccurate trust values that naturally result from a random structure.

Initially, the content spread is limited to the trusted surrounding of the publisher. This way, TBS can deal with 'Identity Whitewashing' as it limits every identity's influence from the beginning. Once the content is consumed (i.e., looked at) and assessed (i.e., evaluated as spam or legitimate), it is blocked or promoted for further dissemination. Although some users are required to characterize spam, regular content consumption may be implicitly used as a positive assessment of the many feedback-lazy users [22]. While this improves dissemination speed, TBS is resilient to the many false positives on account of the scarce explicit user feedback.

We show that TBS is able to disseminate legitimate content almost as effectively as classical epidemic spreading, while limiting the reach of spam to a constant amount of nodes. The structured hop-by-hop spreading nature of TBS is additionally inherently resilient to Sybil users.

Note that we do not propose a new trust or reputation management system and many traditional approaches may be used on top of our spreading scheme to improve its accuracy. We especially address the 'Cold Start Problem' and the scarcity of feedback as they are crucial to effectively prevent spam in open participatory content distribution for opportunistic networks. To summarize, our main contributions are:

- We review three classic spreading mechanisms, propose a number of spam-prevention techniques, and discuss potential problems (Section 2).
- We propose *Trust-Based Spreading* (TBS), a content spreading mechanism that effectively prevents spam (Section 3).
- We evaluate the performance of TBS on different real-world mobility traces, across a wide range of parameters such as user behavior, inaccurate assessments, and attacker models. We show that TBS can stop spam at the source, after just reaching a small number of nodes while still achieving an acceptable spreading performance for legitimate content (Sections 4 and 5).

2. Legitimate content dissemination vs. spam

In this section, we first present the targeted scenario, then describe three classic content spreading mechanisms and extend them by a set of methods to fight spam, and finally point out the potential problems with the resulting techniques.

2.1. Scenario

We are primarily interested in *public channels* for content dissemination. This means that every user is completely free to join a channel, and start sending content to all other subscribers of the channel. Each channel is uniquely identified, e.g., by a hashtag in Twitter. Public channels are typically organized around some topic and are very popular in the Internet nowadays. Some examples include Facebook groups, discussion fora, blogs and boards, citizen journalism, public walls, or open publish/subscribe systems [24] in general. Public channels strongly benefit from the diversity brought by participatory interactions in an open environment, but are vulnerable to spam. In the context of opportunistic networks, users subscribe to a channel and receive updates automatically when coming into physical proximity of other users in possession of relevant content and depending on the spreading mechanism.

2.2. Spreading mechanisms in opportunistic networks

Three methods to spread content in opportunistic networks are *Epidemic Spreading* (ES), *Limited Hop Spreading* (LHS), and *Limited Replication Spreading* (LRS). They differ in the range and speed of spreading content. Note that spreading schemes aim at distributing content to all nodes, as opposed to a source to destination communication achieved by routing [25], or anchoring it at a location [26].

Epidemic Spreading (ES) [27]: This is the basic epidemic spreading scheme. A source publishes content and passes it to any subscriber it encounters. Each subscriber holding the content helps in the dissemination process by passing it to further subscribers. This process goes on until every subscriber received the content.

Limited Hop Spreading (LHS) [28]: LHS limits the reach of content in the network. It is essentially ES equipped with a *hop counter* (time to live) assigned to the content. With each hop the content travels, the counter is decreased; when reaching zero, the content is not transmitted further. A hop count of one means a source only uses direct transmission to reach its subscribers; a hop count of infinity makes LHS equivalent to ES.

Limited Replication Spreading (LRS) [29]: LRS limits the speed of spreading the content. Here, each content is assigned a *replication counter*, i.e., a maximal number of neighbors a node can forward the content to. A replication count of one means content can only be passed on from node to node, like a hot potato; a replication count of infinity makes LRS equivalent to ES.

2.3. Introducing spam – the attacker model

We will distinguish between two general classes of attackers. A *Simple Attacker* follows the spreading protocol rules, but misuses the system by spreading spam rather than legitimate content. In contrast, a *Sophisticated Attacker* may disobey any rule or protocol and ignore the hop or replication counter. Sophisticated attackers may also create many identities and change them at will, thus perform a Sybil attack [16]. For both attacker models we assume the worst case of compromised users under full adversary control.

2.4. The first countermeasure – content assessment

The first step in fighting spam is detecting it. We rely on users to classify a given content as legitimate (to *whitelist* it), or spam (to *blacklist* it). Of course, in the real world many assessments may be missing and some may be wrong – we cover such scenarios later in our evaluation (Section 5).

Whitelisting: The purpose of whitelisting is to *promote content* so it may spread further or faster than set by its initial restrictions. A user may whitelist the content explicitly (by directly approving it) or implicitly (e.g., by re-posting, sharing over a different media, or watching say a five minute video until the end). Implicit whitelisting reduces the scarceness of feedback but increases false positives.

Blacklisting: In contrast, the purpose of blacklisting is to *decrease the reach of spammers*. For this reason, all future content of a blacklisted node is automatically discarded. Blacklisting is always an explicit action.

2.5. The second countermeasure – collaboration

To further improve the effect of assessments, we allow nodes to exchange their blacklists and whitelists in all common channels of interest. They do so independently of whether any content is exchanged, and they only share their personal assessments to minimize overhead and cascading effects in assessment. Because an assessment from a single user might be a mistake or a lie, a certain number of assessments have to be received to then actually act upon them. We call this the *threshold of required assessments*.

2.6. Integrating the countermeasures into spreading schemes

ES, LHS, and LRS can be easily enhanced with (collaborative) whitelisting and blacklisting mechanisms. Under LHS, a natural implementation is to set the hop counter to maximum (on whitelisting) or to zero (on blacklisting). Similar changes may be applied to the replication counter under LRS. This strategy should work well against simple attackers (see Section 2.3).

However, a sophisticated attacker may completely ignore the current values of counters making the counter-based methods ineffective. For example, under LRS, the spammer may ignore the replication counter to reach a significant fraction of nodes (using just direct spreading over one hop). Moreover, for every new spam message, a sophisticated attacker may create a new Sybil ID that, by construction, does not appear in any blacklist. Consequently, (collaborative or not) whitelisting and blacklisting are ineffective against sophisticated attacks under ES, LHS, and LRS, and further mechanisms are required.

3. Trust-Based Spreading (TBS)

In this section, we introduce Trust-Based Spreading (TBS). In order to fight spam and promote legitimate content, TBS strongly leverages on the *trust structure* among the nodes.

3.1. Introducing trust – the cornerstone to hampering spam

We assume that every node u assigns a trust value t_{uv} to a node v. Without loss of generality, we assume that t_{uv} can range from 0 ("no trust") to 1 ("full trust"). Trust is not necessarily symmetric: t_{uv} may very well differ from t_{vu}. We use this generic notation so we do not impose any restrictions on how trust can be established. There are many explicit or implicit ways the system can learn the trust values. Reputation systems, such as described in Section 1.1 might be used. Alternatively, there are secure pairing based approaches [30,31] but they as well suffer from long setup times and heavy user interaction, i.e., the 'Cold Start Problem'. While all these approaches work well to improve the trust structure over time, faster approaches, e.g., inferred from the mobility [32], could be used as an initial setup.

Moreover, in Section 5, we show that TBS even performs well under a random trust structure. Some intuition on why, is that TBS just requires part of the structure to be trustworthy as it is very resilient to false positive whitelisting which also includes deliberate false whitelisting.

3.2. TBS dissemination process

The TBS scheme makes use of 'trust' in every interaction between users. The basic operations of the dissemination scheme are summarized in Fig. 1 for the case of legitimate content and in Fig. 2 for the case of spam. Note that TBS makes use of several thresholds summarized in Table 1 for an arbitrary node u. In Section 5.1, we study how these thresholds affect TBS's performance, and how to set them well.

In the first step of TBS, node u accepts content c directly from its *publisher* p only if p is trusted enough, i.e., when

$$t_{up} > \Theta_u^A. \tag{1}$$

Consequently, Θ_u^A controls the initial reach of c (received directly from the publisher). For $\Theta_u^A > 0$, a spammer can thus be already blocked at the first hop.

On consuming content c, the initially reached nodes start assessing and sharing their assessments with other nodes, as described in Section 2.5. Node v accepts the content c from node u (not a direct publisher of c), only once v has received enough whitelist entries for c, weighted by trust, i.e., when

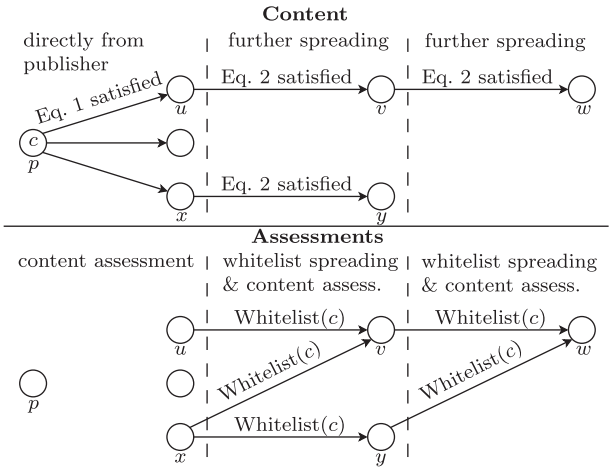

Fig. 1. [TBS dissemination process for legitimate content] Node u receives content c directly from publisher p since Eq. (1) is satisfied (top). Upon consumption, content is assessed and the whitelists are shared with node v (bottom). Content can now spread further to node v if Eq. (2) is satisfied (top). This process is repeated for every hop, e.g. node w.

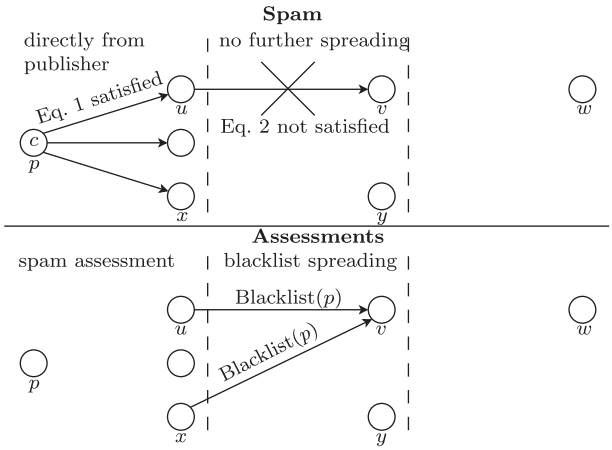

Fig. 2. [TBS dissemination process for spam] Node u receives spam c directly from publisher p since Eq. (1) is satisfied (top). Upon consumption, spam is assessed and the blacklists are shared with node v (bottom). Spam can thus not spread further to node v since Eq. (2) is not satisfied (top). Additionally, no blacklisting nodes (including u) will accept further spam from publisher p. This process usually stops after one hop, thus node w will not even know there was spam present at an earlier hop.

Table 1
Notation used in TBS.

$t_{uv} \in [0,1]$	Trust user u has in v
$\Theta_u^A \in [0,1]$	Threshold of u to accept content from author
$\Theta_u^W \in [0,\infty)$	Threshold of u to accept content whitelisting
$\Theta_u^B \in [0,\infty)$	Threshold of u to accept author blacklisting

$$\sum_w t_{vw} \cdot \mathbb{1}_{\{w \text{ whitelists content } c\}} > \Theta_v^W, \tag{2}$$

where w are nodes encountered by v and $\mathbb{1}_{\{\text{condition}\}}$ is the indicator function returning 1 if the condition is true and 0 otherwise. Consequently, threshold Θ_v^W controls the speed of collaborative whitelisting, and thus content spreading.

Similarly, node v blacklists the publisher p once v has received enough trust-weighted blacklist entries, i.e., when

$$\sum_w t_{vw} \cdot \mathbb{1}_{\{w \text{ blacklists publisher } p\}} > \Theta_v^B. \tag{3}$$

Thus, threshold Θ_v^B controls the speed of collaborative blacklisting. Of course, v may also blacklist p directly, by consuming and assessing the content published by p.

4. Performance evaluation – setup

We evaluate the performance of TBS, ES, LHS, and LRS, by replaying different sets of real-world mobility traces. For the sake of speed it was necessary to write a custom simulator in C. The simulator basically consists of an event queue that replays all connection events from a given trace and executes an interaction among the given nodes according to the selected spreading scheme. In this section we describe the simulation setup; the results are presented in Section 5.

4.1. Performance metrics

We use three main performance metrics to evaluate the performance of the spreading schemes, $\mathcal{R}, \overline{\mathcal{R}}$, and $\mathcal{A}(t)$.

- $\mathcal{R}$ is the *reach of content*, i.e., the number of nodes that receive content, averaged over all contents. Ideally, $\mathcal{R} \to N$ for legitimate content and $\mathcal{R} \to 0$ for spam.

Table 2
Real-world mobility traces used in performance evaluation. Top: general characteristics; bottom: properties of the inferred community structure. (Modularity is a metric describing the quality of the community structure introduced by Newman [33]).

	H06	MIT	DAR
# Nodes N	78	96	1040
Time Period	93 h	14.9 weeks	16.9 weeks
Type	Bluetooth	Bluetooth	AP Assoc.
# Contacts Total	128'979	75'432	4'184'804
# Contacts/Node	1654	786	4024
# Communities	9	9	79
Avg. Community Size	8.67	10.67	13.28
Modularity	0.31	0.53	0.77

$\overline{\mathcal{R}}$ is the *normalized reach of content*. To compare results across different traces, we sometimes normalize $\mathcal{R}$ by $\mathcal{R}^{ES}$ – the reach of legitimate content under ES, as follows: $\overline{\mathcal{R}} = \frac{\mathcal{R}}{\mathcal{R}^{ES}}$.

$\mathcal{A}(t)$ denotes the *availability of content* in the network at time t. $\mathcal{A}(t)$ is here given as the amount of different content possessed by a node at time t, averaged over all nodes. $\mathcal{A}(t)$ is a good metric to describe the *speed* of content dissemination.

4.2. Mobility traces

We use three different sets of real-world mobility traces summarized in Table 2. These particular traces were chosen as they offer a good variety in terms of size, spanned time, the structure of contacts, etc.[1]

Haggle Infocom 2006 (H06): In the scope of the Haggle project, mobility traces were collected during Infocom 2006 [34]. Direct contact of 78 conference attendees were measured with Bluetooth scans every two minutes.

MIT Reality Mining (MIT): The Reality Mining project [35] collected Bluetooth scans of 96 students and staff members with a five minutes interval on the MIT campus. We take the 15 weeks of the trace where mobility was the highest.

Dartmouth (DAR): This trace consists of Wi-Fi access point (AP) associations on the Dartmouth campus [36]. Only users that have an AP association for at least five days a week on average are considered, resulting in 1040 users. Short disconnections (shorter than 60 s) attributed to interference and the well known ping-pong effect (where devices jump back and forth between different APs in less than 60 s) are filtered out. Two nodes are considered in contact when associated with the same AP.

4.3. Trust structure

The trust structure is a critical parameter of TBS. We study possible trust structures by considering two extreme cases: (*i*) strongly correlated with the mobility patterns (best-case), and (*ii*) randomized (worst-case). Both structures can be built automatically and thus mitigate the 'Cold Start Problem', and both structures may further be improved over time by any available trust management framework.

Mobility correlated trust: In this scenario, the trust structure is based on communities existing in the mobility pattern. We find them by applying the Louvain algorithm [37] to the contact graph with edges weighted by their accumulated contact time. To remove isolated nodes, communities smaller than five users are either greedily merged with a community they are well connected to, or discarded. The properties of the resulting *communities* are summarized at the bottom of Table 2.

For each community C, we greedily select three to four *extended communities* among the communities best connected to C. Now, node u assigns trust to node v as follows:

$$t_{uv} = \begin{cases} \mathcal{U}(0.7, 1.0) & \text{if } u \text{ and } v \text{ in same community} \\ \mathcal{U}(0.1, 0.7) & \text{if } v \text{ in extended community of } u \\ 0 & \text{otherwise,} \end{cases} \quad (4)$$

where $\mathcal{U}(l_1, l_2)$ is a random variable uniformly distributed between l_1 and l_2. This is our default trust structure we use with 'TBS'.

Random trust: To generate a trust structure with similar topological properties (e.g., number of nodes and edges, node degree distribution) but not correlated with the mobility pattern, we randomly rewire the contact graph before running the Louvain algorithm (as above). We achieve it by applying double edge swaps [38] to $2N$ randomly selected edge pairs. As a result, the trusted nodes are now randomly selected from the network, independently of the actual contact times, while the node degree, i.e. the connectivity of a node, is preserved. Also, in all traces, the modularity and the community sizes significantly decrease, indicating a much weaker community structure. We refer to TBS using this random trust structure as 'RTBS'.

4.4. Spammers

To stress-test the proposed spam prevention techniques, we make the worst-case assumption that spammers are well integrated in the trust structure (compromised nodes). Additionally, a sophisticated spammer can create and control an arbitrary number of Sybil identities. We make two assumptions about these created Sybils: firstly, any node that comes in contact and interacts with the spammer also interacts with all of its Sybils. Secondly, we assume every Sybil identity to be as well integrated into the trust structure as the spammer is. Both assumptions are hardly realistic but represent the worst possible case. In particular, we assume:

- Every node can become a spammer, i.e., the simulation was repeated with every node being a spammer once.
- The trust of other nodes in node v does not depend on whether v is a regular node, a spammer, or a spammer's Sybil.
- v's Sybils always promote v's spam by whitelisting it upon publication.

4.5. Content generation and consumption

Legitimate content and spam is created regularly defined by the publishing rate $r_p = 1/(24\text{ h})$ for the H06 trace and $r_p = 1/(8\text{ days})$ for the MIT and DART traces (due to their long duration it was not feasible to create content more often). This publishing rate is introduced to see how the system behaves on different days as the mobility pattern of users may vary a lot from one day to another. Because we are interested in the average spreading performance in the network, and because trust is indifferent to a node's status, every node acts both roles, once as a regular node and once as a spammer.

As in the real world, there is a delay between content reception and its consumption. This delay is exponentially distributed with rate $r_c = 1/(2\text{ h})$ for H06 and $r_c = 1/(6\text{ h})$ for MIT and DART. In other words, we assume that a user checks his/her mobile device on average twelve times at a conference and four times during a normal day. This may seem low and research suggests that users actually check their phone up to 150 times a day.[2] However we

[1] The traces are publicly available in the CRAWDAD repository: `http://crawdad.cs.dartmouth.edu/`.

[2] See http://www.kpcb.com/insights/2013-internet-trends or the Tomi Ahonen Almanac 2013.

assume this low consumption rate as a worst case as there might be users that check their phones less often. If we assume a higher consumption rate, i.e., $r_c = 1/(10\text{ min})$, the spreading performance of TBS is practically indistinguishable from ES in our simulations. This is because at low rates the consumption is the main bottleneck. For high consumption rates, the content is already assessed when the next contact occurs, thus nearly all contacts can be used to disseminate content, just as in ES.

Finally, upon consumption, a user assesses the content with probability $p_A \in \{0.1, 0.2, \ldots, 1.0\}$. Assessments might result in false positives (i.e., when spam is whitelisted) with probability $p_F \in \{0, 0.25, 0.5\}$. In principle, also publishers of legitimate content might be blacklisted which refers to the situation of censorship (false negatives). We discuss this issue in Section 6.3.

4.6. Distribution scheme configurations

All spreading schemes use assessment (Section 2.4) and collaboration (Section 2.5) to fight spam. For ES, LHS, and LRS this means that spam may be blacklisted. The blacklists are shared and the amount of blacklists that need to be received in order to block the spam is specified by the *threshold of required assessments* denoted as Θ^N. For LHS and LRS, whitelisting resets the hop and replication counter respectively (hop counter set to 1 and replication counter set to 6). These counters are selected to achieve a good dissemination performance of legitimate content.

For TBS, collaboration is defined by thresholds. We set $\Theta_u^A = 0.7$ for the trusted publishers, i.e., all nodes in the local community described in Section 4.3 are trusted publishers (note, that this includes spammers). The choice of this value depends mainly on how trust is established in the network. In a real system this threshold should be passed from the trust establishment mechanism to TBS as a parameter. For the sake of simplicity, we use the same whitelisting and blacklisting thresholds (see Eqs. (2) and (3)) for all nodes. We denote this threshold as $\Theta^{W/B} = \Theta_u^W \equiv \Theta_u^B$, and we investigate the effect of this threshold in Section 5.1.

5. Performance evaluation – results

In this section, we thoroughly evaluate our schemes under a range of scenarios by replaying the real-world mobility traces described in Section 4.2. First, we study the effect of the assessment threshold choice. Then, we show the influence of the user behavior on the performance of our scheme and discuss possibilities to promote favorable behavior. Next, we show how TBS performs against a sophisticated attacker that injects a number of Sybils into the network. Finally, we compare all spreading schemes in terms of spam reach and content dissemination speed.

5.1. Effects of threshold selection

Selecting the threshold for blacklisting and whitelisting is an important calibration step for LHS, LRS, and TBS. For this reason, we study its effect under either fully correct assessments and assessments with false positives (Fig. 3). There are mainly two findings resulting from this analysis:

- *A low threshold should be selected for all spreading schemes.* This way we maximize the normalized content reach (at least 94% for TBS and slightly more for the other schemes) and minimize the reach of spam. For the following evaluation we use these low threshold values.
- *LHS and LRS are more strongly affected by false positive assessments.* While LHS performs generally bad, with a spam reach of up to 32% for the H06 trace (Fig. 3(a)), the presence of false positives nearly doubles this value (Fig. 3(b)). LRS performs well with correct assessments (Fig. 3(c)), even slightly better than TBS (Fig. 3(f)), with a spam reach of less than 8%. However, introducing false positive assessments, the reach of spam increases drastically; it more than triples to 26% for the H06 trace and even has an eightfold increase from 0.6% to 4.8% for the DAR trace. This means that 49 instead of 6 nodes are reached by every spam. The reach of spam with TBS only doubles in the worst case when false assessments are introduced (Fig. 3(f)). The resilience of TBS against false positives has the following two benefits: (i) even if users provide perfect assessments, an attacker might still produce false assessments, and (ii) the robustness against false positives allows to use implicit assessment, i.e., assessment inferred from the user's consumption behavior, which is not as accurate as explicit assessment.

Conclusion: To achieve maximum content reach and good spam blocking performance, low assessment thresholds should be selected for LHS, LRS, and TBS. This result is valid for all mobility traces and holds further under the existence of false positives.

5.2. Impact of user assessment behavior

It is important to analyze the effect of user behavior on performance for two reasons: (i) to know how sensitive the scheme is to varying behavior, and (ii) to know what behavior is the most desired and should be promoted if possible. Fig. 4 shows the effect of the assessment probability p_A, for the H06 and the DAR trace (MIT trace performs similarly). As expected, all schemes perform better when the assessment probability is high, i.e., when users are more likely assessing content (content reaches at least 94% of the nodes). One way to achieve such high values of p_A in practice is to use implicit rather than explicit assessment (see Section 2.4). This, however, will naturally increase the fraction of incorrect assessments (false positives). Fortunately, as shown in the last section as well as in Fig. 4, TBS is very resilient against false positives (in contrast to other schemes).

Additionally, this user behavior analysis also shows why LHS generally performs much worse than LRS and TBS. In its default behavior – without assessments – content or spam reaches too many nodes, i.e., all the nodes the content producer has come into contact with. This corresponds to 67 users (nearly all) in the H06 trace and 94 in the DAR trace (see 0% assessments in Fig. 4(a) and Fig. 4(b)).

Conclusion: Overall, the performance grows with growing assessment probability, even if some assessments are incorrect, to which TBS is the most resilient scheme.

5.3. TBS under sophisticated attack

One of the most powerful technique a sophisticated spammer can apply is creating an arbitrary number of identities, i.e., Sybils. Sybils allow to multiply the influence of a spammer in the system, e.g., by whitelisting spam. While creating Sybils is fairly straight forward, they are only useful if each of them is integrated into the trust structure. Depending on how trust was established, this is tricky and time consuming to achieve, especially for a larger number of Sybil identities. In our analysis, we assume a sophisticated attacker that is able to integrate its Sybils into the TBS trust structure.

From our analysis we can derive two main conclusions (see Fig. 5):

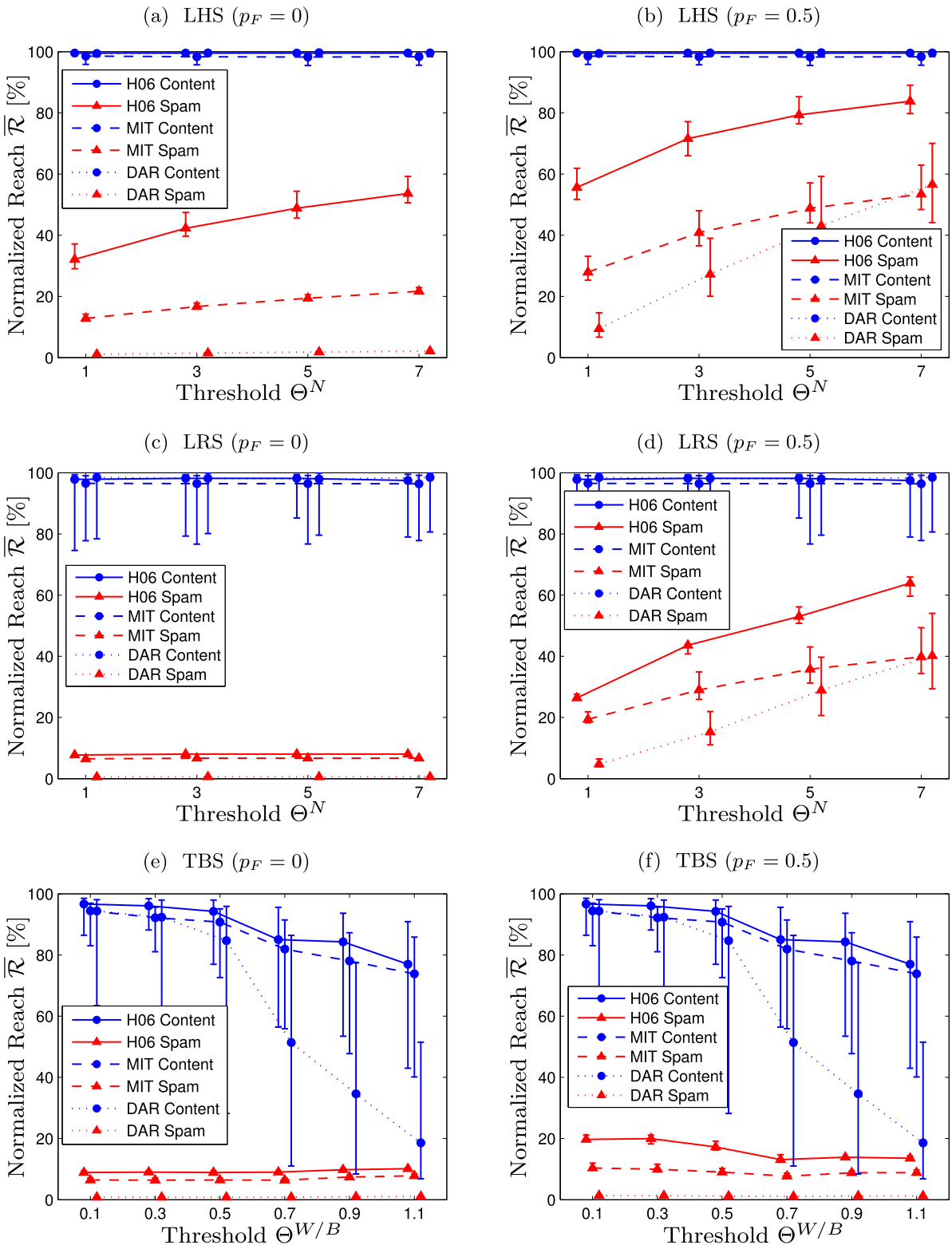

Fig. 3. [The effect of assessment acceptance threshold] The normalized reach $\overline{\mathcal{R}}$ of content (circles) and spam (triangles) as a function of the threshold for accepting white- and blacklists, in case of a simple spammer. The error bars represent the two quartiles around the median and result from the wide range of the assessment probability $p_A \in \{0.1, 0.2, \ldots, 1.0\}$. On the left side, all assessments are correct, i.e, $p_F = 0$, on the right side the probability of incorrect whitelisting is $p_F = 0.5$.

- *TBS is exceptionally robust to Sybils* – the reach $\mathcal{R}$ of spam increases only slightly with increasing number of Sybils s, and quickly flattens out. The reason TBS is so robust against Sybils is based on the same principle as social network based Sybil defenses mentioned in the introduction [17–19], i.e., the Sybil region and the region of honest nodes in the trust structure (graph) are not well connected. Due to the low assessment threshold, one Sybil in the local trust structure is enough to promote spam and creating more does not help since they are connected to the same local trust structure. To be more effective, an attacker would have to be able to infiltrate Sybils into every node's local trust structure.

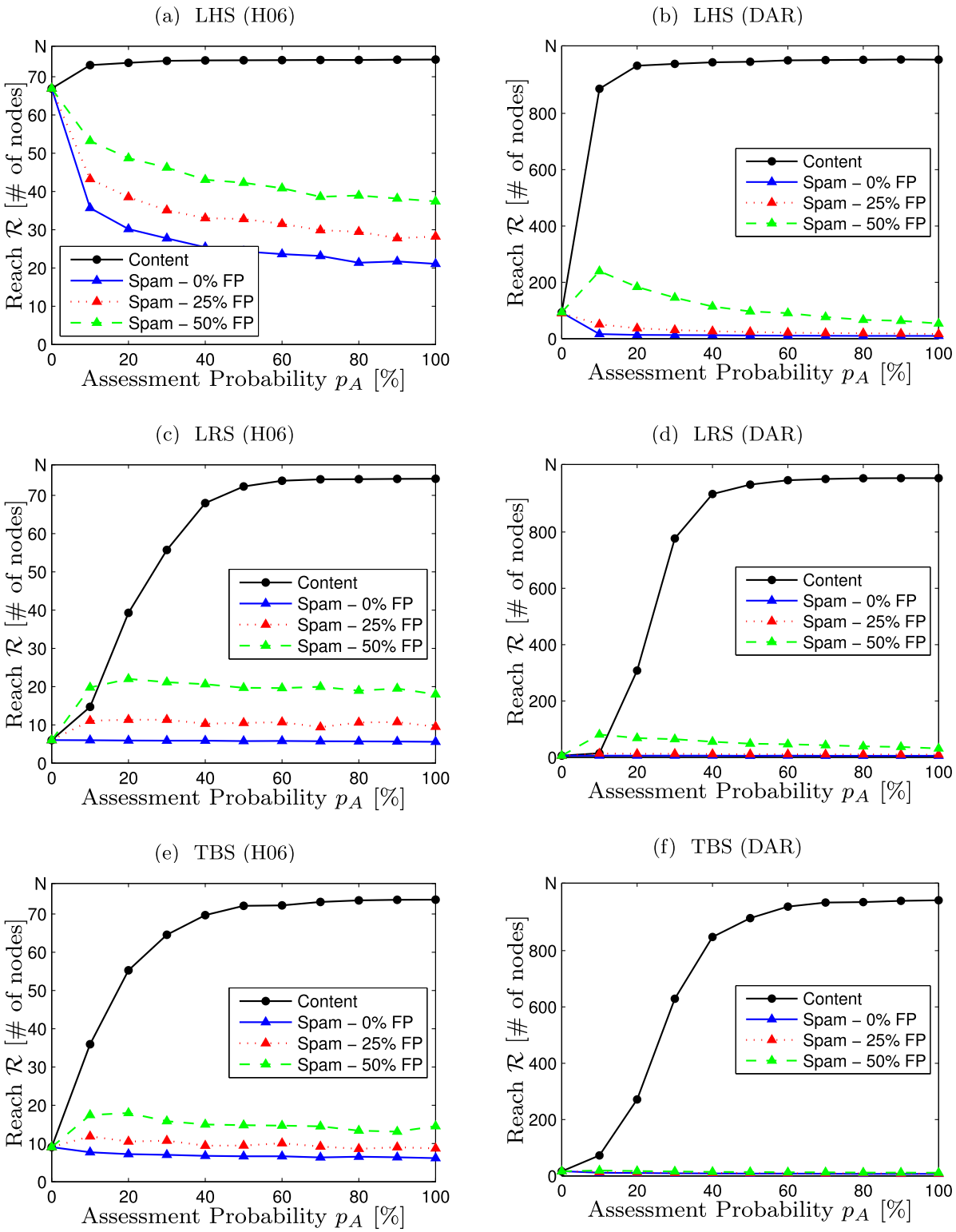

Fig. 4. [The effect of assessment probability p_A] The reach $\mathcal{R}$ for content (circles) and spam (triangles) as a function of the assessment probability p_A for different false positive (FP) probabilities p_F for whitelisting spam. Thresholds are set to $\Theta^{W/B} = 0.1$ and $\Theta^{N} = 1$.

- *The reach of spam is independent of the network size* – it depends only on the number of trusted nodes (i.e., the local trust structure). For all traces the reach of spam is around ten nodes if no Sybils are present and below 20 for a sophisticated attack although the DAR trace has 13 times more nodes than the H06 trace.

Conclusion: Under TBS, even a sophisticated attacker can only spread spam to a constant number of nodes before it gets blocked. The creation of a high number of Sybils is does not increase the attacker's influence.

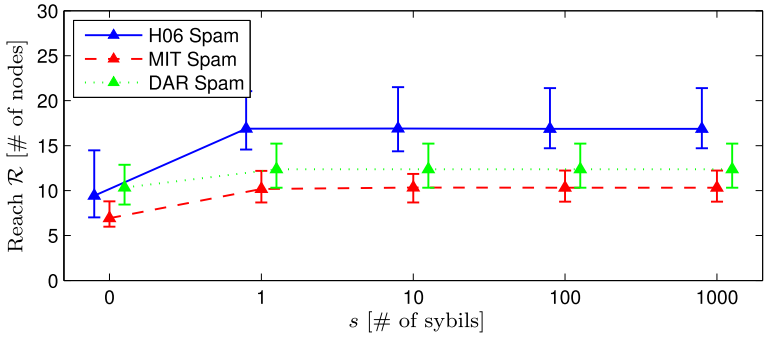

Fig. 5. TBS spam reach $\mathcal{R}$ under Sybil attack, as a function of the number of created Sybils s. The error bars represent the two quartiles around the median and result from the wide range of assessment probability $p_A \in \{0.1, 0.2, \ldots, 1.0\}$ and false positive probability $p_F \in \{0, 0.25, 0.5\}$. The assessment threshold is set to $\Theta^{W/B} = 0.1$.

5.4. TBS vs. classic schemes (overall)

In Fig. 6, we compare the overall performance of the studied spreading schemes for the H06 and the MIT trace (DAR trace performs similarly). We consider the reach of content vs. spam, simple vs. sophisticated attacker, and mobility correlated vs. randomized trust structure (RTBS, see Section 4.3). From this comparison we see that:

- *TBS generally performs best.* While the reach of legitimate content (see Figs. 6(a) and 6(b)) is generally good, i.e., all schemes come close ($> 94\%$) to the best possible ES reach, TBS performs best at stopping spam. In numbers we observe a median spam reach of less than 9 for the simple attacker (see Figs. 6(c) and 6(d)) and less than 17 for a sophisticated attacker (see Figs. 6(e) and 6(f)), independently of the number of nodes in the trace.
- *LRS turns into LHS under a sophisticated attack.* While LRS performs quite well in presence of a simple attacker restraining the median reach of spam to 11 for the H06 trace (Fig. 6(c)) and 9 for the MIT trace (Fig. 6(d)), it performs equally bad as LHS under a sophisticated attack allowing spam to reach 31 nodes for the H06 trace (Fig. 6(e)) and 15 nodes for the MIT trace (Fig. 6(f)). In the real world, a smart attacker could exploit this weakness of LRS additionally by positioning itself at a crowded location, such as a train station, and spread spam to every passing user.
- *RTBS performs nearly as good as TBS.* While this may seem surprising, it stems from the many worst case assumptions we used for the attacker model, the trust structure, and the user consumption and assessment behavior. This result allows us to address the 'Cold Start Problem'. A new user who does not know anybody, can use a random trust structure to bootstrap TBS, and then improve it over time.

Conclusion: Compared to ES, LHS, and LRS, our scheme (TBS) shows the best performance in terms of reach, even under a randomized trust structure.

5.5. Dissemination speed of content

So far we studied only the reach $\mathcal{R}$ of content/spam. However, what counts in practice is not only how far legitimate content reaches, but also *how fast*. Fig. 7 shows the available legitimate content in the network over time, $\mathcal{A}(t)$. Daily patterns such as lower contact rates during several hours (in the night) are the main reason for the slowed down dissemination. This is especially visible for the H06 trace (Fig. 7(a)). Overall, we see that all schemes spread in a qualitatively similar way.

Conclusion: While introducing the spam blocking mechanisms (often based on assessments) slows down the spread of legitimate content its performance is still acceptable and affects TBS only slightly more than LHS and LRS.

5.6. Conclusion of the results

We have demonstrated that TBS effectively blocks spam while maintaining a good dissemination performance of legitimate content. In particular, TBS is (i) very resilient to false positive assessments, i.e., falsely whitelisting spam, allowing for implicit assessment methods, (ii) barely affected by sophisticated attacks with Sybil nodes, stopping spam at the source, and (iii) solves the 'Cold Start Problem' by performing well with a random trust structure. Overall, TBS clearly outperforms all other schemes.

6. Discussion and related work

In this section, we describe related approaches on preventing spam in opportunistic networks. Further, we discuss the potential of a hybrid networking approach and the problem of misusing spam prevention methods for censorship.

6.1. Preventing spam

Spam is a well known problem in the Internet, especially when using Email [39] and social network services [40]. Among the few approaches available for preventing spam in opportunistic networks is 'Hearsay' [41]. 'Hearsay' filters spam by excluding spammers from the trust structure, but performs similar to epidemic spreading (Section 2) with assessments and collaboration [42]. The disadvantages of this approach include unrestricted spreading of spam until the spammer is sufficiently blacklisted, assumption of optimistic spam detection by users (100% assessment probability on average within 15 min after reception), and the requirement of human interaction-based trust establishment (secure pairing).

In the related area of wireless gossiping networks, Gavidia et al. [43] propose a central authority to secure identities and introduce severe restrictions to the amount of data a user may publish. TBS makes a clear contribution to these works as it performs better than epidemic spreading and is a fully distributed solution. Further, TBS supports an open participatory channel deliberately abstaining from restrictions.

6.2. Hybrid networks

In this paper, we exclusively treated totally disconnected opportunistic networks, but TBS can also be integrated into a partially connected hybrid network. On the one hand, opportunistic networks can increase network capacity [44], as well as efficiently

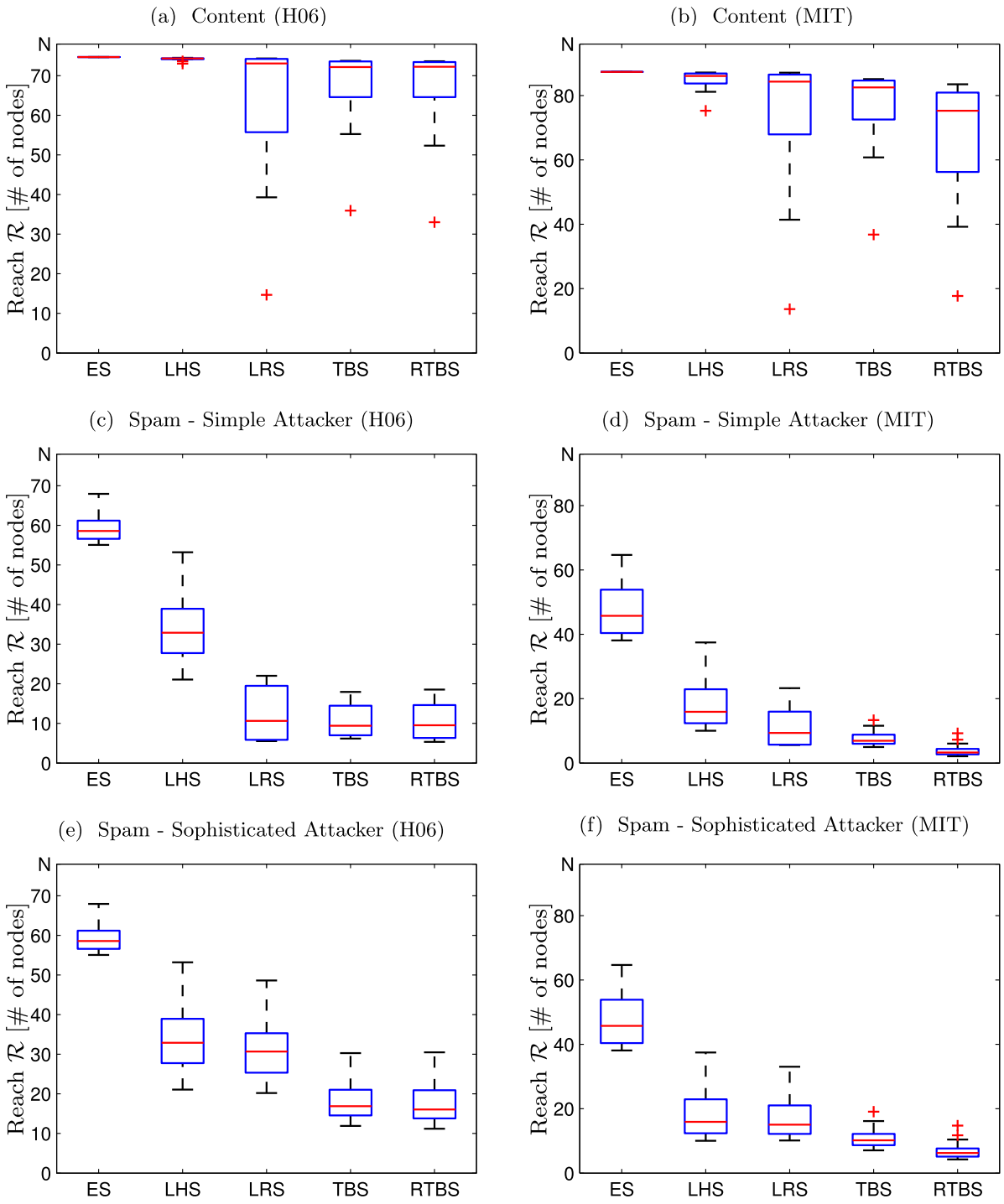

Fig. 6. [Reach of different spreading schemes] For each spreading scheme, a box plot is given where the edges of the box are the quartiles around the median and the whiskers reach the minimum and the maximum of the reach $\mathcal{R}$. The variance of the data results from the wide range of assessment probability $p_A \in \{0.1, 0.2, \ldots, 1.0\}$ and false positive probability $p_F \in \{0, 0.25, 0.5\}$. Thresholds are set to $\Theta^{W/B} = 0.1$ and $\Theta^{N} = 1$.

use the available radio spectrum [45] and offload the 3G infrastructure [46,47]. On the other hand, wireless infrastructure-based networks can speed up the dissemination of content by acting as 'wormholes' between opportunistic domains [48], ease privacy-preserving strategies [49], bootstrap trust via OSN [50], and offer services such as certifications [30] which require a centralized approach. By taking a hybrid approach in TBS, the distribution of assessments (whitelists and blacklists) may be done using the infrastructure network which will increase the performance of TBS even further.

6.3. Censorship

While opportunistic networks easily allow to circumvent censorship [51], by adding collaborative methods to prevent spam, we provide also means to censor solicited content. Users that blacklist providers of legitimate content may thus introduce *false negatives* to the opportunistic network. Hereby, schemes with an absolute threshold, i.e., a constant number of blacklists that are required to block spam are most prone to liars. In contrast, in TBS, these malicious users (liars) have to be trusted peers in all

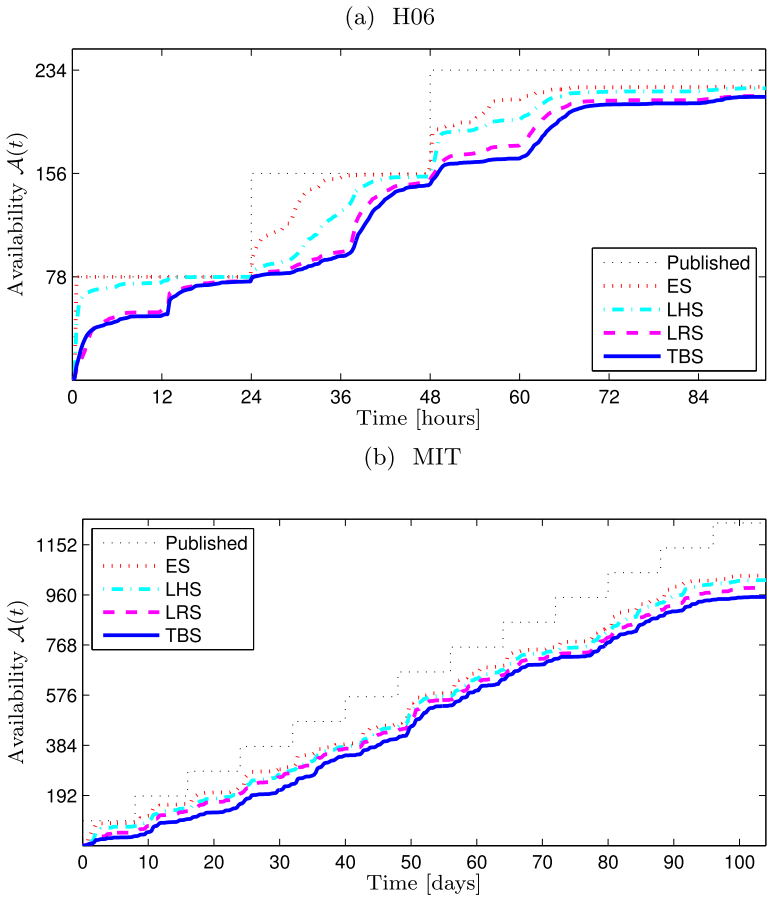

Fig. 7. [Content availability over time] The availability $\mathcal{A}(t)$ of content in the network as a function of time. The black dotted line is the maximum content available published in three rounds. Threshold is set to $\Theta_u^{W/B} = 0.1$ and assessment probability to $p_A \in \{0.1, 0.2, \ldots, 1.0\}$, no false whitelisting, i.e., $p_F = 0$.

communities to effectively censor the whole network. Nevertheless, such liars can still threaten parts of the system and additional countermeasures are needed.

7. Conclusion

In this work, we have focused on protecting opportunistic networks against spam. While existing major classic schemes can be extended by whitelisting and blacklisting, they remain vulnerable to many types of spam attacks. We proposed Trust-based Spreading (TBS) that additionally leverages social trust structures to block spam and promote legitimate content.

By replaying real-world mobility traces, we showed that TBS outperforms the classic epidemic, limited hop, and limited replication spreading. Under TBS, the reach of spam is significantly reduced while legitimate content reaches the availability in the network comparable with that of epidemic spreading. Moreover, TBS is resilient to Sybil attacks, and circumvents the 'Cold Start Problem' by allowing a random trust structure.

While effective against spam, TBS does not solve a complementary yet important problem - censorship in opportunistic networks. We are planning to address this issue in our future work. Furthermore, we want to perform real-world experiments based on the WLAN-Opp framework [52].

Acknowledgment

The work described in this paper is partially supported by the EC's FP7 project SCAMPI and the EC's Marie Curie IEF program contract, PIEF-GA-2010-276336 MOVE-R.

References

[1] S. Cottle, Media and the Arab uprisings of 2011: research notes, Journalism 12 (2011) 647–659.
[2] M. Helft, D. Barboza, Google Shuts China Site in Dispute over Censorship, The New York Times, vol. March, no. 22, 2010.
[3] T.M. Chen, Governments and the Executive 'Internet Kill Switch', IEEE Network 25 (2) (2011) 2–3.
[4] A. Chaintreau, P. Hui, J. Crowcroft, C. Diot, R. Gass, J. Scott, Impact of human mobility on opportunistic forwarding algorithms, IEEE TMC 6 (6) (2007) 606–620.
[5] L. Pelusi, A. Passarella, M. Conti, Opportunistic networking: data forwarding in disconnected mobile ad hoc networks, IEEE Commun. Mag. 44 (Nov. 2006) 134–141.
[6] M. Girvan, M. Newman, Community structure in social and biological networks, PNAS 99 (12) (2002) 7821–7826.
[7] A. Chaintreau, A. Mtibaa, L. Massoulie, C. Diot, The diameter of opportunistic mobile networks, in CoNEXT, 2007.
[8] X. Zhang, G. Neglia, J. Kurose, D. Towsley, Performance modeling of epidemic routing, Comput. Networks 51 (10) (2007) 2867–2891.
[9] K. Hoffman, D. Zage, C. Nita-Rotaru, A survey of attack and defense techniques for reputation systems, ACM Comput. Surv. 42 (1) (2009).
[10] A. Jsang, R. Ismail, C. Boyd, A survey of trust and reputation systems for online service provision, Decision Support Syst. 43 (2) (2007) 618–644.
[11] S. Buchegger, J.-Y. Le Boudec, A robust reputation system for P2P and mobile ad-hoc, Networks, in: P2PEcon, 2004.
[12] K. Walsh, E.G. Sirer, Experience with an object reputation system for peer-to-peer filesharing, in: NSDI, 2006.
[13] S.D. Kamvar, M.T. Schlosser, H. Garcia-Molina, The Eigentrust algorithm for reputation management in P2P networks, in: WWW, pp. 640–651, 2003.
[14] E. Ayday, F. Fekri, An iterative algorithm for trust management and adversary detection for delay-tolerant networks, TMC 11 (2012) 1514–1531.
[15] G. Montenegro, C. Castelluccia, Statistically unique and cryptographically verifiable (SUCV) identifiers and addresses, in: NDSS, 2002.
[16] J.R. Douceur, The Sybil Attack, in: IPTPS, pp. 251–260, 2002.
[17] H. Yu, M. Kaminsky, P.B. Gibbons, A. Flaxman, SybilGuard: defending against sybil attacks via social networks, IEEE/ACM ToN 16 (3) (2008) 576–589.
[18] G. Danezis, P. Mittal, Sybilinfer: detecting sybil nodes using social networks, in: NDSS, 2009.

[19] D. Quercia, S. Hailes, Sybil attacks against mobile users: friends and foes to the rescue, in: INFOCOM, 2010.
[20] B. Viswanath, A. Post, K.P. Gummadi, A. Mislove, An analysis of social network-based Sybil defenses, SIGCOMM Comput. Commun. Rev. 40 (4) (2010) 363–374.
[21] Z. Yang, C. Wilson, X. Wang, T. Gao, B.Y. Zhao, Y. Dai, Uncovering social network sybils in the wild, in: IMC, pp. 259–268, 2011.
[22] G. Chatzopoulou, C. Sheng, M. Faloutsos, A first step towards understanding popularity in YouTube, in: NetSciCom, 2010.
[23] T. Hossmann, G. Nomikos, T. Spyropoulos, F. Legendre, Collection and analysis of multi-dimensional network data for opportunistic networking research, Comput. Commun. 35 (13) (2012) 1613–1625.
[24] V. Lenders, G. Karlsson, M. May, Wireless ad hoc podcasting, SIGMOBILE Mob. Comput. Commun. Rev. 12 (1) (2008) 65–67.
[25] P. Hui, J. Crowcroft, E. Yoneki, BUBBLE rap: social-based forwarding in delay tolerant networks, MobiHoc 10 (2011) 1576–1589.
[26] E. Hyytia, J. Virtamo, P. Lassila, J. Kangasharju, J. Ott, When does content float? characterizing availability of anchored information in opportunistic content sharing, in: INFOCOM, pp. 3137–3145, 2011.
[27] A. Vahdat, D. Becker, Epidemic routing for partially-connected ad hoc networks, Tech. Rep., Duke University, 2000.
[28] M. Vojnovic, A. Proutiere, Hop limited flooding over dynamic networks, in: INFOCOM, pp. 685–693, 2011.
[29] T. Spyropoulos, K. Psounis, C. Raghavendra, Efficient routing in intermittently connected mobile networks: the multiple-copy case, IEEE/ACM ToN 16 (1) (2008) 77–90.
[30] S. Čapkun, J.-P. Hubaux, L. Buttyan, Mobility helps peer-to-peer security, IEEE TMC 5 (2006) 43–51.
[31] Y. Lin, A. Studer, Y. Chen, H. Hsiao, SPATE: Small-group PKI-less authenticated trust establishment, IEEE TMC 9 (12) (2010) 1666–1681.
[32] S. Trifunovic, C. Anastasiades, F. Legendre, Social trust in opportunistic networks, in: NetSciCom, 2010.
[33] M.E.J. Newman, Analysis of weighted networks, Phys. Rev. E 70 (2004).
[34] P. Hui, A. Chaintreau, J. Scott, R. Gass, J. Crowcroft, C. Diot, Pocket switched networks and human mobility in conference environments, in: WDTN, pp. 244–251, 2005.
[35] N. Eagle, A.S. Pentland, Reality mining: sensing complex social systems, Pers. Ubiquitous Comput. 10 (4) (2006) 255–268.
[36] T. Henderson, D. Kotz, I. Abyzov, The changing usage of a mature campus-wide wireless network, Comput. Networks 52 (Oct. 2008) 2690–2712.
[37] V.D. Blondel, J.-L. Guillaume, R. Lambiotte, E. Lefebvre, Fast unfolding of communities in large networks, J. Stat. Mech.: Theory Exp. (2008).
[38] S. Maslov, K. Sneppen, Specificity and stability in topology of protein networks, Science 296 (May 2002) 910–913.
[39] D. Schatzmann, M. Burkhart, T. Spyropoulos, Inferring spammers in the network core, Lecture Notes Comput. Sci. 5448 (2009) 229–238.
[40] H. Gao, J. Hu, Z. Li, Y. Chen, B.Y. Zhao, Detecting and characterizing social spam campaigns, in: IMC, pp. 35–47, 2010.
[41] O. Steenbergen, Hearsay: suppressing spam using trust in mobile social gossiping networks (Master's thesis), Delft University of Technology, 2012.
[42] C. Anastasiades, S. Trifunovic, Secure content dissemination in opportunistic networks (Master's thesis), Swiss Federal Institute of Technology, 2009.
[43] D. Gavidia, G.P. Jesi, C. Gamage, M. van Steen, Canning spam in wireless gossip networks, in: WONS, 2007.
[44] M. Grossglauser, D.N.C.D. Tse, Mobility increases the capacity of ad hoc wireless networks, IEEE/ACM ToN 10 (2002) 477–486.
[45] V. Vukadinović, G. Karlsson, Spectral efficiency of mobility-assisted podcasting in cellular networks, in: MobiOpp, pp. 51–57, 2010.
[46] B. Han, P. Hui, M.V. Marathe, G. Pei, A. Srinivasan, V. Anil, Cellular traffic offloading through opportunistic communications: a case study, in: Chants, pp. 31–38, 2010.
[47] F. Malandrino, M. Kurant, A. Markopoulou, C. Westphal, U.C. Kozat, Proactive seeding for information cascades in cellular networks, in: INFOCOM, pp. 1719–1727, Mar. 2012.
[48] M. Papadopouli, H. Schulzrinne, Seven degrees of separation in mobile ad hoc networks, IEEE Globecom 3 (2000) 1707–1711.
[49] M. Li, K. Sampigethaya, L. Huang, R. Poovendran, Swing & swap: user-centric approaches towards maximizing location privacy, in: ACM WPES, pp. 19–28, 2006.
[50] G. Bigwood, T. Henderson, Bootstrapping opportunistic networks using social roles, in: WoWMoM, 2011.
[51] J. Glanz, J. Markoff, U.S. Underwrites, Internet Detour Around Censors, The New York Times, vol. June, no. 12, 2011.
[52] S. Trifunovic, B. Distl, D. Schatzmann, F. Legendre, WiFi-Opp: ad-hoc-less opportunistic networking, in: Chants, pp. 37–42, 2011.