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: