Using Mechanism Design to Incentivize Spam Detection and Prevention in Mobile Social Networks

Abstract

TBA

1 Introduction

Over the past few years, online social networking has become a popular way for people to communicate and get to know new friends. While users are sharing their personal information like friends, hobbies and pictures, social networks also attracts malicious parties. One of the most severe problem is spam, which a shown in a Sophos's report [1] that there was a 70% rise in the proportion of firms that report encountering spam and malware attacks via social networks during 2009.

The mobile social network, much like web based social network, includes users connect with each other via mobile devices in a wireless way *e.g.*, Bluetooth. In such a scenario, any two agents can establish a connection in an ad hoc fashion, which makes a spammer more convenient to spread spam. The malicious action is easy to be recognized by users but hard to prevent, because it costs energy of an agent to send the spammer's information to others, so it might prefer just deleting the spam message.

While working in isolation may cause an agent in a mobile social network easily compromised by new spammer, a collaborative spam reporting mechanism can overcome this weakness by allowing mobile agents share collective knowledge and experience, hence improve the overall accuracy and efficiency of spam detection.

2 Approach

We first design an incentive model using game theory for mobile agents to collaborate truthfully without free-riding.

Then, we show the existence of a Nash equilibrium under which agents can report spam in an incentive compatible manner. We will also illustrate an iterative algorithm that converges to the equilibrium.

At last, we use discrete event simulation to demonstrate the convergence to Nash equilibrium.

3 Modeling and Equilibrium Analysis

The set of agents $\mathcal{N} = \{1, 2, \cdots, N\}$

The distance between two agents: $d \in \mathbb{R}^+$, and distance function D: $\mathcal{N} \times \mathcal{N} \rightarrow \mathbb{R}^+$

The set of neighbors of agent i: $\mathcal{N}_i^d$

The utility function (to be designed).

Question.1.: Whether to introduce "trust management"?

4 Algorithm

Question.2.: Whether to discuss the methods of detecting spam?

1. an intuitive iterative algorithm
2. primal / dual iterative algorithm

5 Conclusion

References

[1] Sophos, "Security threat report: 2010," Sophos, Tech. Rep., 2010. [Online]. Available: http://www.sophos.com/sophos/docs/eng/papers/sophos-security-threat-report-jan-2010-wpna.pdf