

Social Spam, Campaigns, Misinformation and Crowdturfing

Kyumin Lee
Department of Computer
Science
Utah State University
kyumin.lee@usu.edu

James Caverlee
Department of Computer
Science and Engineering
Texas A&M University
caverlee@cse.tamu.edu

Calton Pu
College of Computing
Georgia Institute of
Technology
calton@cc.gatech.edu

ABSTRACT

This tutorial will introduce peer-reviewed research work on information quality on social systems. Specifically, we will address new threats such as social spam, campaigns, misinformation and crowdturfing, and overview modern techniques to improve information quality by revealing and detecting malicious participants (e.g., social spammers, content polluters and crowdturfers) and low quality contents.

Categories and Subject Descriptors: H.3.5 [Online Information Services]: Web-based services; J.4 [Computer Applications]: Social and behavioral sciences

General Terms: Design, Experimentation, Security

Keywords: social spam, campaign, misinformation, crowdturfing, social media

1. OUTLINE

The past few years have seen the rapid rise of many successful social systems - from Web-based social networks (e.g., Facebook, LinkedIn) to online social media sites (e.g., Twitter, YouTube) to large-scale information sharing communities (e.g., reddit, Yahoo! Answers) to crowd-based funding services (e.g., Kickstarter, IndieGoGo) to Web-scale crowdsourcing systems (e.g., Amazon MTurk, Crowdflower). However, with this success has come a commensurate wave of new threats, including bot-controlled accounts in social media systems for disseminating malware and commercial spam messages, adversarial propaganda campaigns designed to sway public opinion, collective attention spam targeting popular topics and memes, and propagate manipulated contents.

This tutorial will introduce peer-reviewed research work and techniques on information quality on social systems. It will be organized in four parts with a focus on (i) social spam, (ii) campaigns, (iii) misinformation, and (iv) crowdturfing. In the end of the tutorial, we will address challenges, opportunities and tools in the research areas.

Social Spam. In this session, we will overview various social spam detection approaches:

- How to detect suspicious URLs.
- Discussing a relationship between social capitalists and spammers, and how to give penalty to not only the spammers but also these social capitalists.
- Supervised and unsupervised spam detection approaches.
- Social Honeypot was proposed to monitor spammers' behaviors and collect their information.
- Using the crowd wisdom to identify social spammers.

Campaigns. We will introduce how these malicious participants form groups and run campaigns to target social systems more effectively, and overview campaign detection approaches:

- Graph-based social spam campaign detection.
- Content-driven campaign detection.
- Detect and track political campaigns in social media by using a classification approach.
- Frequent itemset mining method with behavioral models to detect fake reviewer groups.

Misinformation. This session will introduce what kind of misinformation exist on social systems, and survey possible approaches to detect the misinformation:

- Measure information credibility on social media by using classification approaches with the crowd power.
- Automatic rumor detection approach on Sina Weibo, China's leading micro-blogging service provider.
- Identify fake images on Twitter during Hurricane Sandy.
- Methods for the information credibility in emergency situation. The methods consist of an unsupervised approach and a supervised approach to detect message credibility.

Crowdturfing. This session will overview real examples of weaponizing crowdsourcing and techniques to identify these manipulated contents and crowd workers who spread manipulated contents on behalf of requesters:

- Introduce real examples reported by the news media.
- Understand what kind of crowdturfing tasks are available on crowdsourcing sites.
- Understand a crowdturfing market size in both eastern and western crowdsourcing sites.
- Track and reveal crowdsourced manipulation of social media. Especially, focus on the western crowdsourcing sites and overview how to detect crowdturfers on social media.