Key Challenges in Defending Against Malicious Socialbots

Position Paper

Yazan Boshmaf, Ildar Muslukhov, Konstantin Beznosov, Matei Ripeanu
Outline

Problem
Motivation

Socialbots

Challenges

OSN Security
Problem Motivation
Reaching Out to Millions

Obama Raised Half a Billion Online in 2008

(Source: Jose Vargas, Voices on The Washington Post, November, 2008)
Mobilizing the Masses

The Arab Spring, January 2011 - Now

Photo credit: Peter Macdiarmid, Getty Images

Photo credit: Steve Crisp, Reuters

Salem et al. Civil movements: The impact of Facebook and Twitter. The Arab Social Media Report, 2011
Predicting the Future: Elections

Twitter elections predictions (*Tweetminster*) outperform market research (*YouGov*)

(Source: Jemima Koss, The Guardian, May 2010)
Predicting the Future: Markets

Twitter mood (Calm) predicts Dow Jones Industrial Average (DJIA)

Day-to-day Overlap

Calm lagged by 3 days

Socialbots
Bots and Socialbots

Computer program used to perform highly repetitive operations (AI?)

Automation software (to pass off as human) + Social media account

Socialbot
Rise of the Socialbots

Zack Coburn and Greg Marra, Olin College, 2010

The Web Ecology Project (Social Engineering), 2011

ACM Interactions Magazine Cover Story, April 2012
Misusing Socialbots on a Large Scale?

An automated social engineering tool for:

- Infiltration
- Misinformation
- Data collection

Boshmaf et al. The Socialbot Network: When Bots Socialize for Fame and Money. ACSAC’11
OSN Security
Tolerate Socialbots
Adversarial Machine Learning

Adversarial learning is a cyclical process shown in Figure 1.

1. **Begin Attack**: The attacker initiates the attack.
2. **Attack**: The attacker controls this phase to maximize the damage caused.
3. **Detect**: The defender detects the attack and responds.
4. **Defender Responds**: The defender controls this phase to mitigate the attack.
5. **Mutate**: The attacker mutates the attack strategy to evade detection.
6. **Attacker Detects**: The attacker detects the defense mechanism.
7. **Defense**: The defender defends against the mutated attack.
8. **Defender Controls**: The defender controls this phase to refine the defense.

This cycle continues indefinitely, with each phase influencing the outcome of the next.

**Figure 1. The adversarial cycle.**
Graph-theoretic Defense Techniques

Sybil detection via social networks\(^1\)

With adversary running large-scale infiltration\(^2\)

---

1 Haifeng Yu. Sybil Defenses via Social Networks: A Tutorial and Survey. ACM SIGACT News’11
2 Boshmaf et al. The Socialbot Network: When Bots Socialize for Fame and Money. ACSAC’11
Prevent Socialbots
Observation: It’s all about automation

Prevent it and the socialbot threat will go away (almost surely)

Not an easy job!
Challenges

Solve at least one
OSN Vulnerabilities: Ineffective CAPTCHAs
OSN Vulnerabilities: Ineffective CAPTCHAs

CAPTCHA-solving businesses

Koobface Botnet

OSN Vulnerabilities: Ineffective CAPTCHAs

CAPTCHA-solving businesses

Koobface Botnet

Design a reverse Turing test that is \textit{usable} and \textit{effective} even against "illegitimate" human solvers.
How about Social Authentication?

Use “personal” social knowledge to challenge users.

Kim et al. Social authentication: Harder than it looks. FC’12
Histogram of Attack Advantage

When the number of challenge images is 1,

many people are vulnerable to impersonation.

Even for 5 challenge images,

some people can be impersonated with probability 100%.

Kim et al. Social authentication: Harder than it looks. FC’12
OSN Vulnerabilities: Fake (Sybil) User Accounts and Profiles
Guarantee an anonymous, yet credible, online-offline identity binding in online and open-access systems
How can we deal with Sybils?

- Centralized trusted authority
- Tie identities to resources
- Use external information
OSN Vulnerabilities: Large-Scale Network Crawler
Effectively limit large-scale Sybil crawls of OSNs without restricting users’ social experience.
How about using a credit network?

Viswanath et al. Exploring the design space of social network-based Sybil defenses. COMSNETS’12
C. Challenges building credit network-based Sybil tolerance

The key challenge in building credit network-based Sybil tolerance is to ensure that the attacker cannot increase the credit available to her from the rest of the network. This is because the attacker's Sybil identities have exactly the same available credit as the rest of the network, and therefore, the attacker cannot use her Sybil identities to obtain credit from the network.

Assumption #1: Small edge cut

Assumption #2: Available credit on each directed edge

As shown in Figure 6, the total amount of credit available within the network is 31. This includes the initial credit of 3, the 2 credits each Sybil identity has, and the 24 credits that the legitimate nodes have. The total available credit within the rest of the network is also 31, as no credit has been transferred to or from the attacker.

Figure 7 shows the edge cut between the well-behaved nodes (hollow) and the misbehaving nodes (solid). The credit available on each edge is shown in the figure. The total amount of credit available on the edge cut is 10, which is less than the total available credit within the network.

In summary, the key challenges in building credit network-based Sybil tolerance are:

1. Small edge cut: The edge cut between the well-behaved nodes and the misbehaving nodes should be small.
2. Available credit on each directed edge: The available credit on each directed edge should be sufficient to prevent the attacker from obtaining additional credit from the network.

Further, each directed edge, ab, represents the initial credit allocation that represents the amount of unconsumed credit by the destination node ab (i.e., the credit available on each directed edge).

Transactions between two nodes in a credit network are contingent upon the availability of credit along network paths that traverse over links while extending credit to each other. For this purpose, nodes can route credit to a node via network paths that traverse over links within the network. Any credit available on edges between the attacker's Sybil identities and the rest of the network cannot be used for payments between nodes that do not directly connect to the attacker with just one identity; in this case, the relevant set of edges is the cut between the subgraph consisting of the attacker's Sybil identities and the rest of the network.

Note that for simplicity, the links are not along the edge cut are shown as dashed lines, for simplicity: the set of edges is the cut between the subgraph consisting of the attacker's Sybil identities and the rest of the network. The total credit available to the misbehaving user accounts does not by itself allow the attacker to obtain additional available credit with other users (though she can use her Sybil identities to purchase from legitimate nodes. Thus, available credit is only shown as dashed lines.

We now discuss the key challenges associated with building credit network-based Sybil tolerance.

Viswanath et al. Exploring the design space of social network-based Sybil defenses. COMSNETS’12
OSN Vulnerabilities: Exploitable Platforms and APIs
Challenge #4

Detect abusive and automated usage of OSN platforms and their social APIs across the Internet
OSN Vulnerabilities: Poorly Designed Privacy/Security Controls
Challenge #5

Develop usable OSN security and privacy controls that help users make more informed decisions
Take-home message(s)

• Large-scale infiltration is feasible
  – has serious privacy and security implications

• Socialbots make it difficult for OSN security defenses and their users to detect their true nature
  – defending against such bots raises a set of unique challenges

• Effective, socio-technical defenses less vulnerable to both human and technical exploits are needed
Key Challenges in Defending Against Malicious Socialbots

Yazan Boshmaf
Ildar Muslukhov
Konstantin Beznosov
Matei Ripeanu

Funded by:
Backup
Socialbot Network: Concept

Botmaster

C&C Channel

SocialBots

Online Social Network

SocialBot

Infiltrated user (randomly picked)

Infiltrated user (with mutual friends)

Boshmaf et al. The Socialbot Network: When Bots Socialize for Fame and Money. ACSAC’11
Prototype Architecture

Our Machine

- BotUpdater
- Master Controller
- C&C Engine
- Botmaster
- Updates
- Commands, Botcargo

Facebook Servers

- 3rd Party Websites & APIs (Access Tokens)
- Blurbs, Tokens
- Graph API + HTTP
- b_i

API Wrapper
HTTP Scraper
Native Controller
Socialbot

Our Machine

- C&C Engine, Botcargo
- Botmaster
- Updates
- Graph API + HTTP
- b_i

Facebook Servers
Methodology

- Prototype on Facebook
- 102 Socialbots, single Botmaster
- Operated for 8 weeks (Spring 2011)
- Single machine
  - Different IPs
  - HTTP proxy emulating different browsers and OSs
- Approved by UBC ethics board
Most Users Decide Within Three Days
Too Many Friends: Too Many Bots?

**f-socialbots**

![Graph of f-socialbots]

**m-socialbots**

![Graph of m-socialbots]
Mutual Friends Matter

![Graph](image)

- **Figure 4:** Degree distribution of the generated random sample of Facebook user profiles during the bootstrapping phase, with a sample size of 5,053 valid profile identities.

- **Figure 5:** Cumulative distribution function of number of days it took to observe a fraction of the overall accepted friendship requests during the bootstrapping phase.

- **Figure 6:** Average acceptance rate of the resulted infiltration as a function of the number of mutual friends the socialbots has with the infiltrated users. (95% conf.)

- **Figure 7:** Average acceptance rate as a function of the number of friends a user profile has during the bootstrapping phase. (for the requests sent by m-socialbots, 95% conf.)

- **Figure 8:** Average acceptance rate as a function of the number of friends a user profile has during the bootstrapping phase. (for the requests sent by f-socialbots, 95% conf.)

- **Figure 9:** Distribution of the overall infiltration among the socialbots. A point in the figure represents how many socialbots infiltrated the corresponding number of user profiles.

95% conf.
Successful Infiltration is Team Work

- 88 bots infiltrated 10-40 profiles
- 10 bots infiltrated 60-80 profiles

Number of socialbots vs. Number of infiltrated profiles

- 70% of infiltrations involved 10-40 profiles
- 23% of infiltrations involved 60-80 profiles
Private Data Exposed

Socialbots: **102**, their friends: **3,055**, their friends’ friends: **1,085,785**

Birth dates: 48,810 before ➔ 580,649 after (11.9x more)
Web-based Botnet Integration

Infected Machines

C&C Channel

Botherder

Botnet + Socialbot Network

Botmaster

Socialbots

OSN Channel

Online Social Network
Advice: OSNs and Security Research

Read it!

GOT ETHICS?