Multi-Defender Strategic Filtering Against Spear-Phishing Attacks

Aron Laszka[1], **Jian Lou**[2], and Yevgeniy Vorobeychik[2]

1. University of California, Berkeley
2. Vanderbilt University

Laszka@berkeley. edu, {**jian.lou**, yevgeniy.vorobeychik}@vanderbilt.edu

- Mitigate **Spear-Phishing attacks :**
  - e-mail filters which block e-mails with a maliciousness score above a chosen threshold.
- How to choose the threshold? It is tradeoff between False-Positive (FP) and False-Negative (FN).
  - False-Positive (FP) : non-malicious e-mail is filtered out.
  - False-Negative (FN): malicious e-mail is not filtered out.
- Users may be **self-interested** and they may only care about themselves!
- Game Theoretical Approach: A game among Multiple Users and an Attacker.
  - Not only games between users and attacker, also game among users themselves. (They may be self-interested).

Multi-Defender Strategic Filtering Against Spear-Phishing Attacks

Aron Laszka[1], **Jian Lou**[2], and Yevgeniy Vorobeychik[2]

1. University of California, Berkeley
2. Vanderbilt University

Laszka@berkeley. edu, {**jian.lou**, yevgeniy.vorobeychik}@vanderbilt.edu

- How to model the game?
  - Two–stage sequential game (short-term dynamic): all users move first, then the attacker best responds.
  - Simultaneously move game (long-term dynamic): all users and attacker move simultaneously.
- Strategy Space:
  - Users: False Negative ratio (correspondingly get False Positive)
  - Attacker: The set of users to attack.
- Two kinds of equilibrium:
  - Stackelberg Multi-Defender Equilibrium (Short-term)
  - Nash Equilibrium (Long-term)

Want to know more?

Welcome to my poster !☺