The Economics of Mass Surveillance and The Questionable Value of Anonymous Communications

George Danezis and Bettina Wittneben

Presented By: Warren Powers
Overview

- Introduction
- Model
- Obtaining network information
- Surveillance
- Discussion of effectiveness
- Conclusion
Introduction

- Participants belong to clubs
  - If one participant is under surveillance, all information shared and membership is revealed

- Questions
  - How many need to be under surveillance?
  - Who do we put under surveillance?
  - How does anonymity affect target selection?
Model

- People and spaces (or clubs)
- Relationships
  - When people belong to spaces
  - No links between people
  - Relationships have strength
    - Symbolizes degree of association between person and space
- Graph from set of people to set of spaces.
Extracting the Network from Data

- Used data from mailing list archives
- Mapped email address to space
- Mapped email to person
- Relations were created from messages to lists
Effectiveness of Partial Surveillance

- What is revealed?
  - Observing one member of a space
    - All relationships associated with that space

- Choice of target
  - Those with highest degree
    - Among spaces not under surveillance
  - Repeated as budget allows
Data Uncovered – Full Info

Spaces, People and Relations Uncovered – Full Information

Fraction Uncovered

Fraction of People Under Direct Surveillance
Partial Information

- We can monitor volume of messages
  - Not degree or correspondents
- Target selection more difficult
  - Lower return on investment
Spaces, People and Relations Uncovered – Volume Information

Fraction of People Under Direct Surveillance

Fraction Uncovered
Discussion

- The first model could represent no anonymity
  - Can obtain much information with little surveillance

- The second model represents some anonymity
  - No cover traffic

- Anonymized communication is helpful but not perfect
Diminishing Returns

- Initial investment provides great return
- As budget is increased, marginal returns decrease.
  - Cost per unit of intelligence rapidly increases
- Useful information may be very costly
- Privacy violation is high
Interception Figures

- Warrants issued vs. number under surveillance
- UK population
  - Full information graph used
  - Formula for those under surveillance
    - \( \frac{0.5}{0.01} \times X \)
  - \( \frac{0.5}{0.01} \times 1849 = 92000 \) people
  - Info on 50 people revealed for each one monitored
Failure of Adaptive Target Selection

- Adaptive strategies are inferior to volume selection

- Adaptive Strategies
  - High known degree
    - Likely to have links to undiscovered spaces
  - Structural equivalence
    - High know degree and few nodes sharing its position
Strategy Comparison

**People Uncovered – Two Adaptive Strategies**

- **Non-adaptive**
- **Known degree**
- **Structural uniqueness**

**Fraction Uncovered**

**Fraction of People Under Direct Surveillance**
**Target Selection for Disruption**

- What if goal is to disrupt network?
  - Remove nodes with highest degree
  - Remove nodes with high volume

- Selection results not very different
  - Need to remove twice as many using volume info.
Size of largest component and node deletion

- Size by degree
- Size by volume
- Size by random
- Number by degree
- Number by volume

Size of largest component

Number of nodes deleted
Conclusion

- Information is leaked through third parties
- A small carefully selected set of nodes reveal a large amount of information
- Unlinkability is not sufficient, unobservability is necessary
- Surveillance will violate privacy of innocent parties
- Finding guilty parties will be costly