Bayesian inference to evaluate information leakage in complex scenarios

Carmela Troncoso
Gradiant, Spain
17th July 2013
Privacy beyond encryption

Common belief: “if I encrypt my data, then the data is private”
- Encryption works and gets more and more efficient!
- But does not hide all data
  - Origin and destination
  - Timing
  - Frequency
  - Location
  - ...

These data contain a lot of information
- WWII: The English recognized German Morse code operators
Easy, let’s hide this information!

- Delay messages to change frequency and timing patterns
  - Messages cannot be delayed for too long
- Add dummy events to confuse the adversary
- Pad packets to hide their length
  - Bandwidth is in general limited
- Reroute messages to hide origin and destination
  - Delays messages
  - Needs of collaboration or dedicated infrastructure
- Obfuscate the location
  - Obfuscation must not prevent usability
Maybe is not that easy...

Design decisions to:
- Balance available resources and privacy
- Balance usability and privacy

And do not forget there is an adversary
- not only observes public input/outputs of the system...
- ... also *knows* the privacy-preserving mechanism operation
  - e.g., ISP providers, system administrator, Data Retention, ...

**Information will leak!!**

How to quantify the information leaked?
This is a problem we all have
Given an observation...

Anonymous communications
Who speaks with whom?

Location privacy mechanisms
Which is the real location?

Image forensics
Was the image tampered?

Source identification
What device originated the image?
Case study
Anonymous communications
Anonymous communications

- Hide who speaks to whom
  - sender, receiver, type of service, network address, friendship network, frequency, relationship status.

- Main building block for privacy-preserving applications
  - Desirable privacy (comms, surveys,…)
  - Mandatory privacy (eVoting,)

- Subject to constraints (bandwidth, delay,…) They must leak information!
Traffic analysis of Anonymous Communications

- Systems are evaluated against one attack at a time
  - Network constraints
  - Users knowledge
  - Persistent communications
  - ...

- Based on heuristics and simplified models
  - Exact calculation of probability distributions in complex systems was considered as an intractable problem
Mix networks as an example

- Mixes hide relations between inputs and outputs
- Mixes are combined in networks in order to
  - Distribute trust (one good mix is enough)
  - Load balancing (no mix is big enough)
The traffic analysis game

Who speaks to whom?
Routing constraints

Max Length = 2 hops

Non trivial given the observation!!
Routing constraints

Really, non-trivial!

(we could think about user knowledge in the same way)
(Re)Defining Traffic analysis

Find hidden state of mixes
(Re)Defining Traffic analysis

Find hidden state of mixes

\[
\Pr[HS \mid O, C] = \frac{\Pr[O \mid HS, C] \Pr[HS \mid C]}{\sum_{HS} \Pr[O \mid HS, C]}
\]
(Re)Defining Traffic analysis

Find hidden state of mixes

\[
\text{Pr}[\text{HS} \mid O, C]
\]

Too large to enumerate

\[
\text{Pr}[\text{HS} \mid O, C] = \frac{\text{Pr}[O \mid \text{HS}, C] \text{Pr}[\text{HS} \mid C]}{\sum_{\text{HS}} \text{Pr}[O \mid \text{HS}, C]} = \frac{\text{Pr}[O \mid \text{HS}, C] K}{Z}
\]
Sampling to get probabilities

- Computing $\Pr[HS|O,C]$ infeasible: too many HS
  - ... but we only care about marginal distributions
  - Is Alice speaking to Bob?

- If we had many samples of HS according to $\Pr[HS|O,C]$
  - We could simply count how many times Alice speaks to Bob

- Markov Chain Monte Carlo methods
  - Sample from a distribution difficult to sample from directly
Metropolis Hastings

**Simple**

1. Given $HS_0$ (an internal configuration of the mixes)
2. Propose a new state $HS_1$
3. Accept with probability $\min(1, \alpha)$, reject otherwise

$$\alpha = \frac{\Pr[HS_1 \mid O, C] \cdot Q(HS_0 \mid HS_1)}{\Pr[HS_0 \mid O, C] \cdot Q(HS_1 \mid HS_0)} = \frac{\Pr[O \mid HS_1, C]^K}{Z} \cdot \frac{Q(HS_0 \mid HS_1)}{Z} \cdot Q(HS_1 \mid HS_0)$$

**Pr[O|HS,C]** is a generative model (in general simple)

**Q()** is a proposal function
- e.g., swap two links in a mix

The stationary distribution corresponds to $Pr[HS \mid O,S]$

We can sample!

The bayesian traffic analysis of mix networks, C. Troncoso and G. Danezis, 16th on Computer and Communications Security (CCS 2009)
Why is this useful?

- **Evaluation information theoretic metrics for anonymity**
  
  \[ H = \sum_{R_i} \Pr[A \rightarrow R_i | O, C] \log(\Pr[A \rightarrow R_i | O, C]) \]

  - e.g., comparison of network topologies

- **Estimating probability of arbitrary events**
  
  - Input message to output message?
  - Alice speaking to Bob ever?
  - Two messages having the same sender?

- **Accommodate new constraints**
  
  - Key to evaluate new mix network proposals

---

Persistent communications

Perfect!

Anonymity set size = 6
Entropy metric $H_A = \log 6$
Persistent communications

- Rounds in which Alice participates output a message to her friends
  - Her friends appear more often
  - We can infer set of friends!
## Statistical Disclosure Attacks

- Statistically finds frequent receivers
- Count & Substract “noise”
- 20 users, 5 msgs/batch
- Alice’s friends [0,13,19]

### Example

<table>
<thead>
<tr>
<th>Round</th>
<th>Receivers</th>
<th>SDA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>[15, 13, 14, 5, 9]</td>
<td>[13, 14, 15]</td>
</tr>
<tr>
<td>2</td>
<td>[19, 10, 17, 13, 8]</td>
<td>[13, 17, 19]</td>
</tr>
<tr>
<td>3</td>
<td>[0, 7, 0, 13, 5]</td>
<td>[0, 5, 13]</td>
</tr>
<tr>
<td>4</td>
<td>[16, 18, 6, 13, 10]</td>
<td>[5, 10, 13]</td>
</tr>
<tr>
<td>5</td>
<td>[1, 17, 1, 13, 6]</td>
<td>[10, 13, 17]</td>
</tr>
<tr>
<td>6</td>
<td>[18, 15, 17, 13, 17]</td>
<td>[13, 17, 18]</td>
</tr>
<tr>
<td>7</td>
<td>[0, 13, 11, 8, 4]</td>
<td>[0, 13, 17]</td>
</tr>
<tr>
<td>8</td>
<td>[15, 18, 0, 8, 12]</td>
<td>[0, 13, 17]</td>
</tr>
<tr>
<td>9</td>
<td>[15, 18, 15, 19, 14]</td>
<td>[13, 15, 18]</td>
</tr>
<tr>
<td>10</td>
<td>[0, 12, 4, 2, 8]</td>
<td>[0, 13, 15]</td>
</tr>
<tr>
<td>11</td>
<td>[9, 13, 14, 19, 15]</td>
<td>[0, 13, 15]</td>
</tr>
<tr>
<td>12</td>
<td>[13, 6, 2, 16, 0]</td>
<td>[0, 13, 15]</td>
</tr>
<tr>
<td>13</td>
<td>[1, 0, 3, 5, 1]</td>
<td>[0, 13, 15]</td>
</tr>
<tr>
<td>14</td>
<td>[17, 10, 14, 11, 19]</td>
<td>[0, 13, 15]</td>
</tr>
<tr>
<td>15</td>
<td>[12, 14, 17, 13, 0]</td>
<td>[0, 13, 17]</td>
</tr>
</tbody>
</table>
Statistical Disclosure Attacks

- Statistically finds frequent receivers
- Count & Substract “noise”
- 20 users, 5 msgs/batch
- Alice’s friends [0,13,19]

#### Efficient

- Needs a lot of data for reliability
- More complex models (replies, pool mixes)

<table>
<thead>
<tr>
<th>Round</th>
<th>Receivers</th>
<th>SDA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>[15, 13, 14, 5, 9]</td>
<td>[13, 14, 15]</td>
</tr>
<tr>
<td>2</td>
<td>[19, 10, 17, 13, 8]</td>
<td>[13, 17, 19]</td>
</tr>
<tr>
<td>3</td>
<td>[0, 7, 0, 13, 5]</td>
<td>[0, 5, 13]</td>
</tr>
<tr>
<td>4</td>
<td>[16, 18, 6, 13, 10]</td>
<td>[5, 10, 13]</td>
</tr>
<tr>
<td>5</td>
<td>[1, 17, 1, 13, 6]</td>
<td>[10, 13, 17]</td>
</tr>
<tr>
<td>6</td>
<td>[18, 15, 17, 13, 17]</td>
<td>[13, 17, 18]</td>
</tr>
<tr>
<td>7</td>
<td>[0, 13, 11, 8, 4]</td>
<td>[0, 13, 17]</td>
</tr>
<tr>
<td>8</td>
<td>[15, 18, 0, 8, 12]</td>
<td>[0, 13, 17]</td>
</tr>
<tr>
<td>9</td>
<td>[15, 18, 15, 19, 14]</td>
<td>[13, 15, 18]</td>
</tr>
<tr>
<td>10</td>
<td>[0, 12, 4, 2, 8]</td>
<td>[0, 13, 15]</td>
</tr>
<tr>
<td>11</td>
<td>[9, 13, 19, 19, 15]</td>
<td>[0, 13, 15]</td>
</tr>
<tr>
<td>12</td>
<td>[13, 6, 2, 16, 0]</td>
<td>[0, 13, 15]</td>
</tr>
<tr>
<td>13</td>
<td>[1, 0, 3, 5, 1]</td>
<td>[0, 13, 15]</td>
</tr>
<tr>
<td>14</td>
<td>[17, 10, 14, 11, 19]</td>
<td>[0, 13, 15]</td>
</tr>
<tr>
<td>15</td>
<td>[12, 14, 17, 13, 17]</td>
<td>[0, 13, 17]</td>
</tr>
</tbody>
</table>
Co-inferring routing and profiles

A simple approach

- Iterate profile and routing
- Introduces systematic errors if done naively

Actually we want to find $Pr[M, \Psi | O, C]$

- $M$ is the routing, $\Psi$ are the profiles (multinomial distribution)
- Sounds familiar...

Gibbs sampling

MCMC to sample from a joint distributions

Iterate

- $X \leftarrow Pr[X | Y, O, C]$
- $Y \leftarrow Pr[Y | X, O, C]$
Gibbs sampling for anonymity systems

From matching to profiles

\[ \Pr[\Psi | M, O, C] \]

Observation

\[ V_{AB} = 1 \]
\[ V_{AO} = 3 \]
\[ V_{OB} = 3 \]
\[ V_{OO} = 17 \]

Count messages and use the multinomial prior

\[ \Psi = \text{Dirichlet} \left( V_{AB}, V_{AO} \right) \]
Gibbs sampling for anonymity systems

From profiles to matchings

\[ \Pr[M \mid \Psi, O, C] \]

\[ \Psi_{Alice} = \{ \Pr[A \rightarrow B], \Pr[A \rightarrow O]\} \]
\[ \Psi_{Others} = \{ \Pr[O \rightarrow B], \Pr[O \rightarrow O]\} \]

Sadly not as simple...

1. If possible analytical
2. Use MCMC-MH
3. Other alternatives?
And if profiles are dynamic?

- Previous methods work for static behavior
  - But this does not seem very realistic...

- The Bayesian approach: Particle filtering
  - Sequential Monte Carlo
  - Infer dynamic hidden variables when the state space is intractable analytically

- The adversary observes volumes of communication and wants to infer poisson rates that generates them
  \[
  \Pr[\lambda_{AB_t} | \lambda_{AB_{t-1}}, O, C]
  \]
Particle filtering

1. Start with some particles $\lambda_{AB_t}^1, \lambda_{AB_t}^2, \ldots, \lambda_{AB_t}^N$

2. Evolve particles according to model

3. Compute their likelihood according to the current and previous observation

$$L[\lambda_{AB_{t+1}}^1 | \lambda_{AB_t}^1, O] = p_1$$
$$L[\lambda_{AB_{t+1}}^2 | \lambda_{AB_t}^2, O] = p_2$$
$$\vdots$$
$$L[\lambda_{AB_{t+1}}^N | \lambda_{AB_t}^N, O] = p_N$$

4. Resample $N$ particles according to probabilities: “best” particles
   e.g.,

$$\lambda_{AB_{t+1}}^1 = \lambda_{AB_t}^1, \lambda_{AB_{t+1}}^2 = \lambda_{AB_t}^2, \ldots, \lambda_{AB_{t+1}}^N = \lambda_{AB_t}^N$$

5. Back to 2
Particle filtering for anonymity systems

**Observation**

Input and output volume

$t$: \( V_A = 2, V_O = 4, V_B = 1, V_{OO} = 5 \)

$t+1$: \( V_A = 1, V_O = 5, V_B = 2, V_{OO} = 4 \)
Particle filtering for anonymity systems

\[ \lambda_{AB_t}^{1}, \lambda_{AB_t}^{2}, \lambda_{AB_t}^{3} \]

Propose new rates
\[ \lambda_{AB_{t+1}}^{1}, \lambda_{AB_{t+1}}^{2}, \lambda_{AB_{t+1}}^{3} \]

Resample
- Probability of generating observation
- Likelihood of evolution
- Trained (loose) with real data

Start with some rates

You cannot hide for long: De-anonymization of real-world dynamic behaviour, G. Danezis and C. Troncoso, Under submission (ask me!)
Results

Enron dataset (http://www.cs.cmu.edu/~enron/)

![Graph showing the Enron dataset results with various markers and lines representing different metrics such as $\lambda_{AB}$, $\hat{\lambda}_{SSD}$, $\hat{\lambda}_{SDA}$, and $\hat{\lambda}_{SDA}$. The graph plots messages/week against epochs (t in weeks) with a probability $\Pr[\lambda_{AB} < 0.1]$ on the lower axis.]
Advantages

- **Systematic**
  - Generative model tends to be easy

- **Return probability distributions**
  - More informative than ML
  - Allows for multiple inferences

- **Confidence estimates**
  - Key in real analysis!

**What I did not say**

- I have avoided all the scary details
- Getting the model correctly is non-trivial
Applications

We have seen three Bayesian methods

- Metropolis Hastings sampling \( \Pr[HS|O,C] \)
  - Location privacy - tracking
  - Differential privacy
- Gibbs sampling \( \Pr[X,Y|O,C] \)
  - Location privacy – de-anonymization
- Particle filtering \( \Pr[\lambda_t|\lambda_{t+1},O,C] \)
  - Privacy-preserving video surveillance

Lots to do

- Tor: website fingerprinting, flow correlation, flow watermarking, routing,…
- Location privacy: dynamic behaviour
- Cloud computing: side channels
The message I wanted to convey

- We are solving the same problem again and again
- Privacy and forensics are not that far
- Privacy research can be a source for inspiration
- And the other way around! Come apply your methods to our systems!
- LSDA with Fernando Pérez-Gonzalez (UVigo)

Bayesian inference as systematic approach
- Allows to tackle complex scenarios
- Sampling reduces computational requirements
Thanks!

I hope I have awaken your curiosity 😊

I’ll be around, come talk to me!
Write to me at ctroncoso@gradiant.org