Protecting Data Privacy in an Increasingly Connected World

Jie Gao

Rutgers University
https://sites.rutgers.edu/jie-gao/

Feb 2nd 2022.
Data Challenge from Ubiquitous Sensing

Enormous amount of inter-connected data collected from everyday living environment.
What privacy models? Cryptographic solutions; statistical solutions.
Protect Data Privacy in Distributed Sensing

- What privacy models? Cryptographic solutions; statistical solutions.
- Where to place a privacy protection module in a system?
- What types of privacy invasion attacks? – advanced machine learning
Protect Data Privacy in Distributed Sensing

- What privacy models? Cryptographic solutions; statistical solutions.
- Where to place a privacy protection module in a system?
  - At destination
  - OR
  - At origin
  
  - External entity perturbs data after collection
  - OR
  - User perturbs data before publishing her data!

- What types of privacy invasion attacks? – advanced machine learning
Data Privacy for Correlated Data

What if users are related in a network?
What if attributes are contagious and correlated?

- Political affiliations, smoking habits, obesity, ...
Run a Survey

Goal: what is the fraction of the population who smoke cigarette?
Run a Survey

Goal: what is the fraction of the population who smoke cigarette?
Individual: may not trust the data collector.

Analysis:

- If A is a smoker, report YES with probability \( \frac{3}{4} \).
- If A is not a smoker, report YES with probability \( \frac{1}{4} \).

The total fraction of YES is \( \frac{p}{2} + \frac{1}{4} \), where \( p \) is the true answer.
Run a Survey

Goal: what is the fraction of the population who smoke cigarette?
Individual: may not trust the data collector.
Random response: flip a coin
  ▶ If HEAD, tell the truth.
  ▶ If TAIL, report YES/NO uniformly randomly.

Analysis:
  ▶ If A is a smoker, report YES with probability $\frac{3}{4}$.
  ▶ If A is not a smoker, report YES with probability $\frac{1}{4}$.
  ▶ The total fraction of YES is $\frac{p}{2} + \frac{1}{4}$, where $p$ is the true answer.
Goal: what is the fraction of the population who smoke cigarette?
Individual: may not trust the data collector.
Random response: flip a coin
  ▶ If HEAD, tell the truth.
  ▶ If TAIL, report YES/NO uniformly randomly.
Analysis:
  ▶ If A is a smoker, report YES with probability $\frac{3}{4}$.
Run a Survey

Goal: what is the fraction of the population who smoke cigarette?

Individual: may not trust the data collector.

Random response: flip a coin
  ▶ If HEAD, tell the truth.
  ▶ If TAIL, report YES/NO uniformly randomly.

Analysis:
  ▶ If A is a smoker, report YES with probability $\frac{3}{4}$.
  ▶ If A is not a smoker, report YES with probability $\frac{1}{4}$. 

The total fraction of YES is $p/2 + 1/4$, where $p$ is the true answer.
Goal: what is the fraction of the population who smoke cigarette?
Individual: may not trust the data collector.

Random response: flip a coin
  ▶ If HEAD, tell the truth.
  ▶ If TAIL, report YES/NO uniformly randomly.

Analysis:
  ▶ If A is a smoker, report YES with probability $3/4$.
  ▶ If A is not a smoker, report YES with probability $1/4$.
  ▶ The total fraction of YES is $p/2 + 1/4$, where $p$ is the true answer.
Run a Survey

Random response: flip a coin

- If HEAD, tell the truth.
- If TAIL, report YES/NO uniformly randomly.

$(\varepsilon, \delta)$-differential privacy:

$$\text{Prob}\{\text{YES} | \text{smoker}\} \leq \text{Prob}\{\text{YES} | \text{nonsmoker}\} \cdot e^\varepsilon + \delta$$
Run a Survey

Random response: flip a coin

▶ If HEAD, tell the truth.
▶ If TAIL, report YES/NO uniformly randomly.

$(\varepsilon, \delta)$-differential privacy:

$$\text{Prob}\{YES|\text{smoker}\} \leq \text{Prob}\{YES|\text{nonsmoker}\} \cdot e^\varepsilon + \delta$$

$$\frac{3}{4} \leq \frac{1}{4} \cdot e^\varepsilon + \delta$$

Random response is $(\ln 3, 0)$-differentially private.
Run a Survey

- What if the collector also knows the social network $G$?
- Smoking is a contagious behavior.

Insight: the correlation structure could be exploited for statistical inference attack.
Run a Survey

- What if the collector also knows the social network $G$?
- Smoking is a contagious behavior.

Insight: the correlation structure could be exploited for statistical inference attack.
Without extra knowledge on social network structures, Bayesian inference is optimal.
Without extra knowledge on social network structures, Bayesian inference is optimal.

Inference with perturbed data with $(\varepsilon, \delta)$-differential privacy has Area Under the ROC Curve (AUC) at most

$$1 - \frac{1 - \delta}{1 + e^{\varepsilon}}$$
Without extra knowledge on social network structures, Bayesian inference is optimal.

Inference with perturbed data with $(\varepsilon, \delta)$-differential privacy has Area Under the ROC Curve (AUC) at most

$$1 - \frac{1 - \delta}{1 + e^\varepsilon}$$

Using contagion models and the social network structure $G$, one can reverse engineer the randomization process and recover the original knowledge more than the designed guarantee.
Contagion model – linear threshold model:

- Edge weight $w_{ij}$ indicates influence from $i$ to $j$;
- Each node has a threshold $\lambda(v) \in [0, 1]$;
- If the sum of influence from all activated neighbors of $v$ goes beyond $\lambda(v)$, $v$ is activated in the next round.
- The process stops when no new node is activated.
Network based inference attack:

- Work the contagion backwards, find the probability $\alpha(v)$ of each node $v$ being initially active.
- With that, run the contagion process to find the probability of a node being active.
We achieve AUC above the theoretical bounds.

<table>
<thead>
<tr>
<th>Network</th>
<th>$\beta$</th>
<th>$\varepsilon$</th>
<th>Upper Bound</th>
<th>Bayesian</th>
<th>CO-DAG</th>
<th>O-DAG</th>
<th>CO-RND</th>
<th>O-RND</th>
<th>Lappas+</th>
</tr>
</thead>
<tbody>
<tr>
<td>Synthetic</td>
<td>0.1</td>
<td>0.201</td>
<td>0.550</td>
<td>0.550</td>
<td>0.575</td>
<td>0.597</td>
<td>0.568</td>
<td>0.587</td>
<td>0.614</td>
</tr>
<tr>
<td></td>
<td>0.3</td>
<td>0.619</td>
<td>0.650</td>
<td>0.648</td>
<td>0.716</td>
<td>0.708</td>
<td>0.689</td>
<td>0.699</td>
<td>0.657</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>1.099</td>
<td>0.750</td>
<td>0.751</td>
<td>0.833</td>
<td>0.803</td>
<td>0.802</td>
<td>0.793</td>
<td>0.695</td>
</tr>
<tr>
<td></td>
<td>0.7</td>
<td>1.735</td>
<td>0.850</td>
<td>0.847</td>
<td>0.904</td>
<td>0.870</td>
<td>0.887</td>
<td>0.868</td>
<td>0.706</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>2.944</td>
<td>0.950</td>
<td>0.951</td>
<td>0.967</td>
<td>0.955</td>
<td>0.965</td>
<td>0.957</td>
<td>0.711</td>
</tr>
<tr>
<td>Core-Periphery</td>
<td>0.1</td>
<td>0.201</td>
<td>0.550</td>
<td>0.545</td>
<td>0.571</td>
<td>0.555</td>
<td>0.553</td>
<td>0.564</td>
<td>0.527</td>
</tr>
<tr>
<td></td>
<td>0.3</td>
<td>0.619</td>
<td>0.650</td>
<td>0.659</td>
<td>0.704</td>
<td>0.699</td>
<td>0.678</td>
<td>0.685</td>
<td>0.571</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>1.099</td>
<td>0.750</td>
<td>0.752</td>
<td>0.806</td>
<td>0.791</td>
<td>0.781</td>
<td>0.781</td>
<td>0.587</td>
</tr>
<tr>
<td></td>
<td>0.7</td>
<td>1.735</td>
<td>0.850</td>
<td>0.851</td>
<td>0.897</td>
<td>0.874</td>
<td>0.876</td>
<td>0.870</td>
<td>0.611</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>2.944</td>
<td>0.950</td>
<td>0.949</td>
<td>0.967</td>
<td>0.948</td>
<td>0.960</td>
<td>0.955</td>
<td>0.620</td>
</tr>
<tr>
<td>Erdos Renyi</td>
<td>0.1</td>
<td>0.201</td>
<td>0.550</td>
<td>0.545</td>
<td>0.590</td>
<td>0.609</td>
<td>0.580</td>
<td>0.587</td>
<td>0.586</td>
</tr>
<tr>
<td></td>
<td>0.3</td>
<td>0.619</td>
<td>0.650</td>
<td>0.646</td>
<td>0.715</td>
<td>0.713</td>
<td>0.689</td>
<td>0.700</td>
<td>0.672</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>1.099</td>
<td>0.750</td>
<td>0.745</td>
<td>0.813</td>
<td>0.805</td>
<td>0.801</td>
<td>0.795</td>
<td>0.701</td>
</tr>
<tr>
<td></td>
<td>0.7</td>
<td>1.735</td>
<td>0.850</td>
<td>0.850</td>
<td>0.890</td>
<td>0.883</td>
<td>0.884</td>
<td>0.876</td>
<td>0.743</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>2.944</td>
<td>0.950</td>
<td>0.949</td>
<td>0.959</td>
<td>0.948</td>
<td>0.939</td>
<td>0.953</td>
<td>0.753</td>
</tr>
<tr>
<td>Power-law Graph</td>
<td>0.1</td>
<td>0.201</td>
<td>0.550</td>
<td>0.540</td>
<td>0.602</td>
<td>0.584</td>
<td>0.577</td>
<td>0.591</td>
<td>0.534</td>
</tr>
<tr>
<td></td>
<td>0.3</td>
<td>0.619</td>
<td>0.650</td>
<td>0.652</td>
<td>0.730</td>
<td>0.711</td>
<td>0.699</td>
<td>0.724</td>
<td>0.573</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>1.099</td>
<td>0.750</td>
<td>0.753</td>
<td>0.821</td>
<td>0.803</td>
<td>0.808</td>
<td>0.796</td>
<td>0.612</td>
</tr>
<tr>
<td></td>
<td>0.7</td>
<td>1.735</td>
<td>0.850</td>
<td>0.852</td>
<td>0.893</td>
<td>0.883</td>
<td>0.887</td>
<td>0.881</td>
<td>0.669</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>2.944</td>
<td>0.950</td>
<td>0.949</td>
<td>0.964</td>
<td>0.961</td>
<td>0.960</td>
<td>0.963</td>
<td>0.756</td>
</tr>
</tbody>
</table>

We surpass the theoretical bound.
Going Beyond the Bounds

Results hold for real world graphs, beating alternative schemes.

<table>
<thead>
<tr>
<th>Network</th>
<th>( \beta )</th>
<th>( \epsilon )</th>
<th>Upper Bound</th>
<th>Bayesian</th>
<th>CO-DAG</th>
<th>O-DAG</th>
<th>CO-RND</th>
<th>O-RND</th>
<th>Lappas+</th>
</tr>
</thead>
<tbody>
<tr>
<td>GrQc</td>
<td>0.1</td>
<td>0.201</td>
<td>0.550</td>
<td>0.554</td>
<td>0.577</td>
<td>0.576</td>
<td>0.564</td>
<td>0.561</td>
<td>N/A*</td>
</tr>
<tr>
<td></td>
<td>0.3</td>
<td>0.619</td>
<td>0.650</td>
<td>0.644</td>
<td>0.720</td>
<td>0.718</td>
<td>0.683</td>
<td>0.685</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>1.099</td>
<td>0.750</td>
<td>0.744</td>
<td>0.833</td>
<td>0.834</td>
<td>0.798</td>
<td>0.797</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>0.7</td>
<td>1.735</td>
<td>0.850</td>
<td>0.849</td>
<td>0.908</td>
<td>0.892</td>
<td>0.892</td>
<td>0.878</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>2.944</td>
<td>0.950</td>
<td>0.951</td>
<td>0.973</td>
<td>0.960</td>
<td>0.971</td>
<td>0.965</td>
<td>N/A</td>
</tr>
<tr>
<td>HepTh</td>
<td>0.1</td>
<td>0.201</td>
<td>0.550</td>
<td>0.553</td>
<td>0.584</td>
<td>0.579</td>
<td>0.568</td>
<td>0.566</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>0.3</td>
<td>0.619</td>
<td>0.650</td>
<td>0.658</td>
<td>0.733</td>
<td>0.716</td>
<td>0.698</td>
<td>0.703</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>1.099</td>
<td>0.750</td>
<td>0.751</td>
<td>0.831</td>
<td>0.825</td>
<td>0.793</td>
<td>0.793</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>0.7</td>
<td>1.735</td>
<td>0.850</td>
<td>0.849</td>
<td>0.917</td>
<td>0.893</td>
<td>0.884</td>
<td>0.876</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>2.944</td>
<td>0.950</td>
<td>0.950</td>
<td>0.972</td>
<td>0.958</td>
<td>0.962</td>
<td>0.962</td>
<td>N/A</td>
</tr>
<tr>
<td>Amazon Videos</td>
<td>0.1</td>
<td>0.201</td>
<td>0.550</td>
<td>0.554</td>
<td>0.598</td>
<td>0.604</td>
<td>0.584</td>
<td>0.585</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>0.3</td>
<td>0.619</td>
<td>0.650</td>
<td>0.649</td>
<td>0.748</td>
<td>0.748</td>
<td>0.706</td>
<td>0.714</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>1.099</td>
<td>0.750</td>
<td>0.747</td>
<td>0.854</td>
<td>0.836</td>
<td>0.828</td>
<td>0.812</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>0.7</td>
<td>1.735</td>
<td>0.850</td>
<td>0.849</td>
<td>0.919</td>
<td>0.899</td>
<td>0.906</td>
<td>0.894</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>2.944</td>
<td>0.950</td>
<td>0.950</td>
<td>0.975</td>
<td>0.962</td>
<td>0.974</td>
<td>0.967</td>
<td>N/A</td>
</tr>
<tr>
<td>Amazon DVDs</td>
<td>0.1</td>
<td>0.201</td>
<td>0.550</td>
<td>0.551</td>
<td>0.576</td>
<td>0.584</td>
<td>0.567</td>
<td>0.570</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>0.3</td>
<td>0.619</td>
<td>0.650</td>
<td>0.652</td>
<td>0.735</td>
<td>0.739</td>
<td>0.699</td>
<td>0.707</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>1.099</td>
<td>0.750</td>
<td>0.750</td>
<td>0.830</td>
<td>0.829</td>
<td>0.806</td>
<td>0.802</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>0.7</td>
<td>1.735</td>
<td>0.850</td>
<td>0.851</td>
<td>0.906</td>
<td>0.897</td>
<td>0.894</td>
<td>0.886</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>2.944</td>
<td>0.950</td>
<td>0.952</td>
<td>0.963</td>
<td>0.957</td>
<td>0.957</td>
<td>0.941</td>
<td>N/A</td>
</tr>
</tbody>
</table>
[RGS SDM’21] “Influencers do not have privacy. ”

- Influencers: nodes whose activation will likely trigger the activation of a ‘giant’ component.
Impossibility Results

[RGS SDM’21] “Influencers do not have privacy. ”

- Influencers: nodes whose activation will likely trigger the activation of a ‘giant’ component.
- Strong trade-off between utility (estimating the number of active nodes) and privacy (the active status) of influencers.
Impossibility Results

[RGS SDM'21] “Influencers do not have privacy. ”

- Influencers: nodes whose activation will likely trigger the activation of a ‘giant’ component.

- Strong trade-off between utility (estimating the number of active nodes) and privacy (the active status) of influencers.

- State-of-the-art mechanisms (Wasserstein mechanisms) add noise with a magnitude $\approx cn$ with a constant $c$, $n$ is the number of nodes.
Privacy Aware Learning
Lessons: rich structures in user data can lead to information leakage that users may not be aware of.
Privacy Aware Learning

- Lessons: rich structures in user data can lead to information leakage that users may not be aware of.
- Be mindful of powerful machine learning algorithms.
Privacy Aware Learning

- Lessons: rich structures in user data can lead to information leakage that users may not be aware of.
- Be mindful of powerful machine learning algorithms.

Q: Can we add carefully designed noises to data such that sensitive attributes cannot be learned while target applications can still work as usual?
Generate **adversarial samples** to fool a classifier, by adding noises to the input.

Generate **artificial samples** from an input data distribution. Two classifiers play in a game.

- A generator trying to generate fake samples.
- A discriminator trying to detect whether a sample is fake or not.

Our Idea: Privacy Aware Generative Noises

Generate tailored privacy aware noises

- Fool a classifier for learning sensitive attributes,
- Does not hurt the performance of target application classifier.
Optimization

Loss function

▶ Adversarial loss: target labels accurately predicted; confusion on sensitive label is maximized.
▶ GAN loss: generated data distribution follows that of real data.
▶ Regularization loss: the magnitude of noise should be small.
OccuTherm occupancy [Munir+, ’19]:

- XBOX Kinect mounted on top of a door
- Getting depth black-and-white images
- Target label: going in/out
- Sensitive label: Identity
### Performance

**Minimum Accuracy for $z$, given prediction accuracy thresholds for $y$**

<table>
<thead>
<tr>
<th>Method</th>
<th>Sensitive Classifier Accuracy (%)</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MNIST $\text{Acc}(\mathcal{D}_y) \geq 0.95$</td>
<td>PubFig $\text{Acc}(\mathcal{D}_y) \geq 0.95$</td>
<td>WiFi $\text{Acc}(\mathcal{D}_y) \geq 0.75$</td>
<td>OccuTherm $\text{Acc}(\mathcal{D}_y) \geq 0.95$</td>
</tr>
<tr>
<td>PR-GAN</td>
<td>65.3</td>
<td>20.6</td>
<td>23.6</td>
<td>3.6</td>
</tr>
<tr>
<td>NGP</td>
<td>67.2</td>
<td>28.1</td>
<td>31.3</td>
<td>7.2</td>
</tr>
<tr>
<td>AP</td>
<td>67.1</td>
<td>61.9</td>
<td>37.7</td>
<td>23.6</td>
</tr>
<tr>
<td>DP</td>
<td>80.6</td>
<td>78.3</td>
<td>50.0</td>
<td>37.6</td>
</tr>
<tr>
<td>Random*</td>
<td>60.0</td>
<td>13.3</td>
<td>12.5</td>
<td>5.0</td>
</tr>
<tr>
<td>Original $\text{Acc}(\mathcal{D}_z)$</td>
<td>98.4</td>
<td>80.7</td>
<td>84.6</td>
<td>99.8</td>
</tr>
<tr>
<td>Original $\text{Acc}(\mathcal{D}_y)$</td>
<td>99.2</td>
<td>98.3</td>
<td>92.6</td>
<td>99.95</td>
</tr>
</tbody>
</table>

* A classifier that outputs a class at random according to $\theta_y(z)$. 
On-going work

- Subspace differential privacy [GGY, AAAI’22]: respect invariant constraints (e.g., county population should be held as invariant in Census data).
- Time-series data [CGSSY, AAMAS’22]: data generated from a Markov chain.
- Binary/integer data with subset sum invariants.
Privacy protection methods should respect the structure in data, or in the process of handling data, to defend against statistical inference attacks.

- On Privacy of Socially Contagious Attributes, Aria Rezaei, Jie Gao, ICDM’19.
- Application-Driven Privacy-Preserving Data Publishing with Correlated Attributes, Aria Rezaei, Chaowei Xiao, Jie Gao, Bo Li, Sirajum Munir, EWSN 2021.

Questions and comments?