Protecting Data Privacy in an Increasingly Connected World

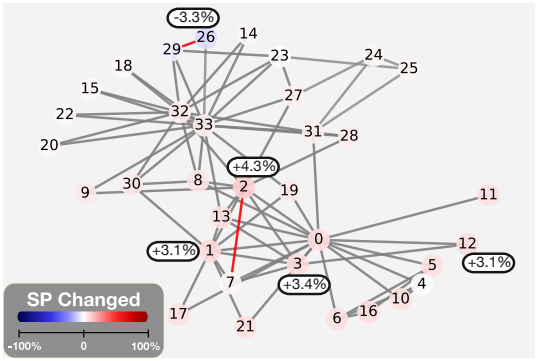
Jie Gao

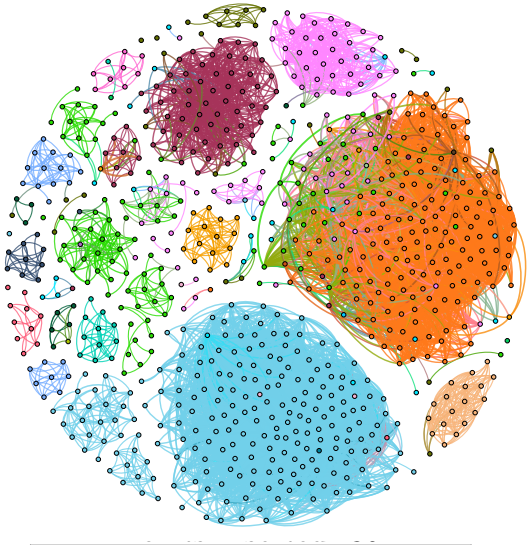
Rutgers University
https://sites.rutgers.edu/jie-gao/

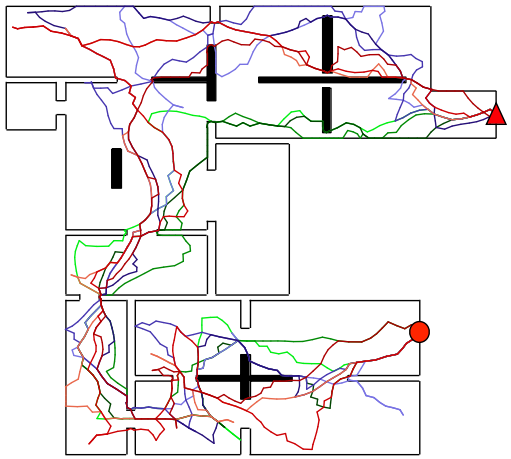
Feb 2nd 2022.

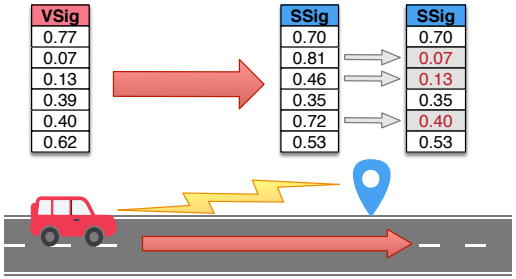
Data Challenge from Ubiquitous Sensing

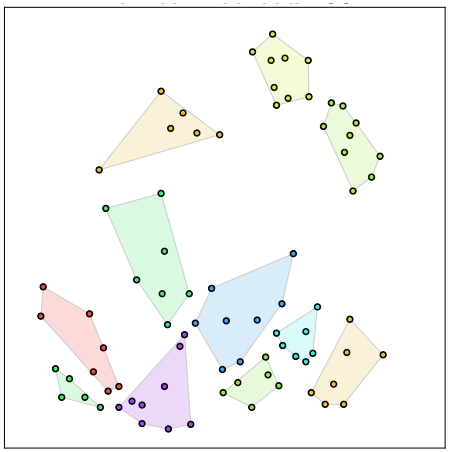
Enormous amount of inter-connected data collected from everyday living environment.

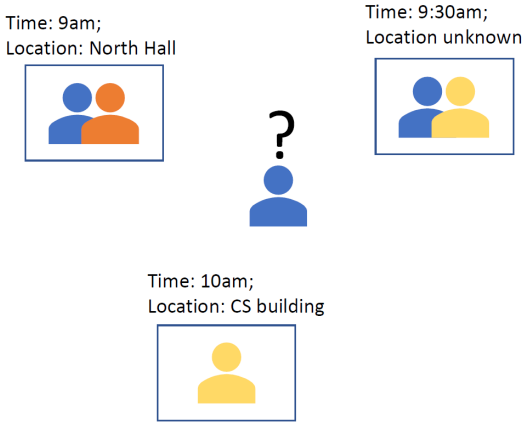












Protect Data Privacy in Distributed Sensing

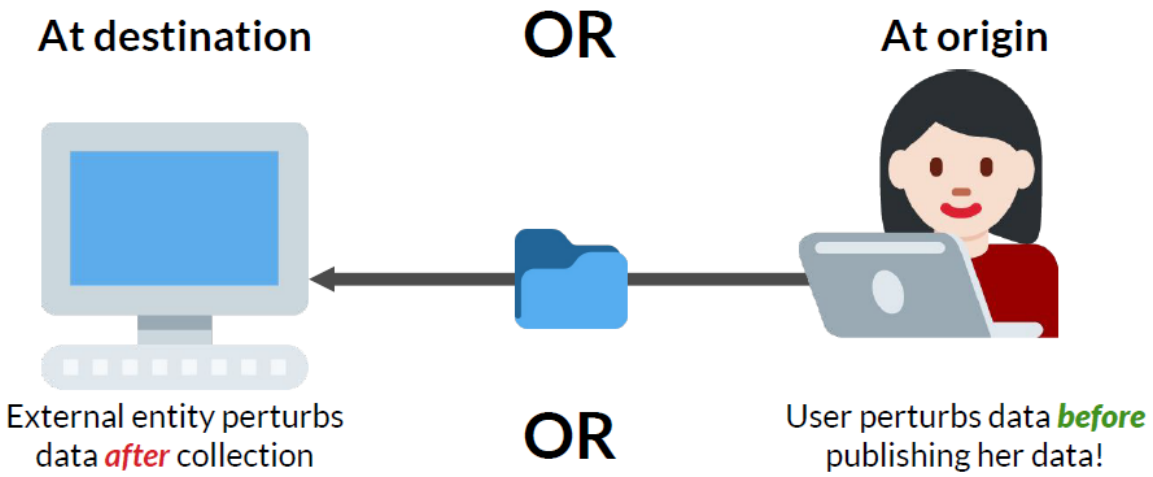
- What privacy models? Cryptographic solutions; statistical solutions.

Protect Data Privacy in Distributed Sensing

- What privacy models? Cryptographic solutions; statistical solutions.
- Where to place a privacy protection module in a system?

Protect Data Privacy in Distributed Sensing

- What privacy models? Cryptographic solutions; statistical solutions.
- Where to place a privacy protection module in a system?

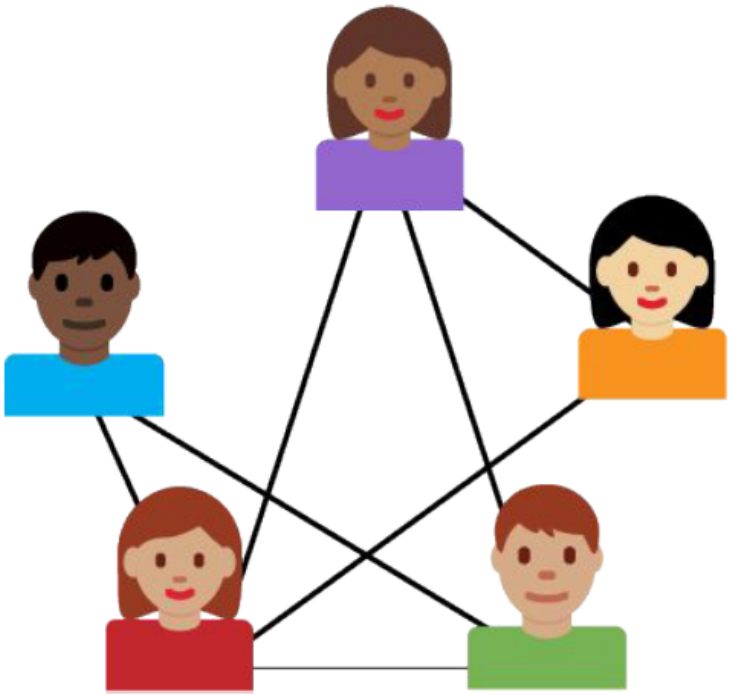


- What types of privacy invasion attacks? – advanced machine learning

Data Privacy for Correlated Data

What if users are 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.

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.

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 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 3/4.
- If A is not a smoker, report YES with probability 1/4.

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 $3/4$.
- 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.

(ε, δ)-differential privacy:

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

Run a Survey

Random response: flip a coin

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

(ε, δ)-differential privacy:

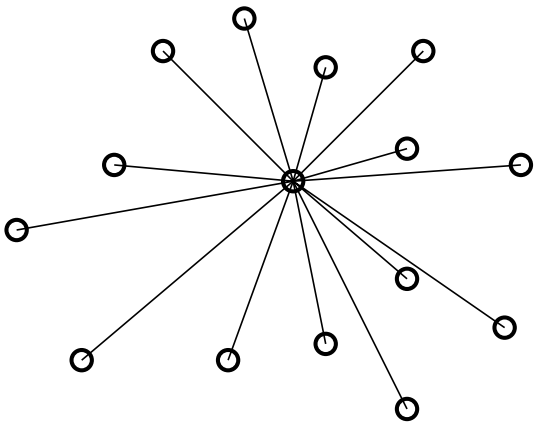
$$\text{Prob}\{YES|smoker\} \leq \text{Prob}\{YES|nonsmoker\} \cdot e^{\varepsilon} + \delta$$

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

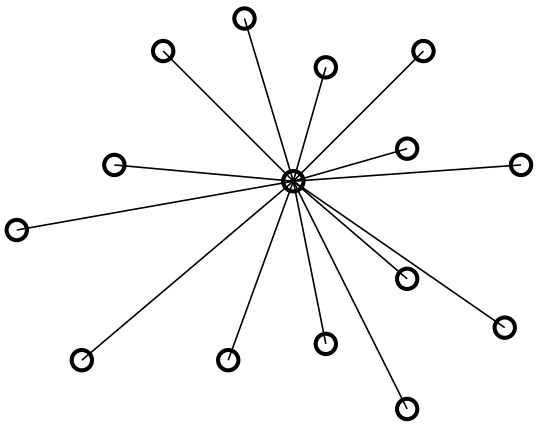
Random response is $(\ln 3, 0)$-differentially private.

Run a Survey

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



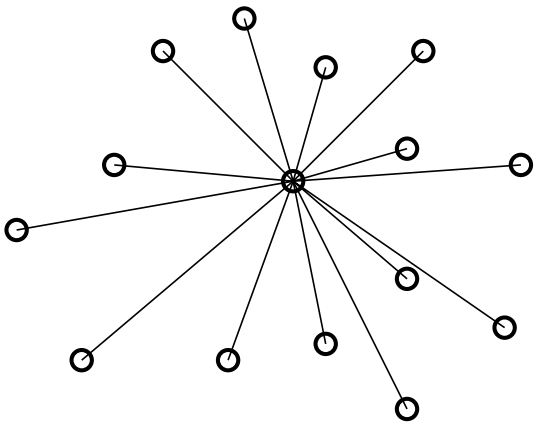
3/4 of all friends report YES



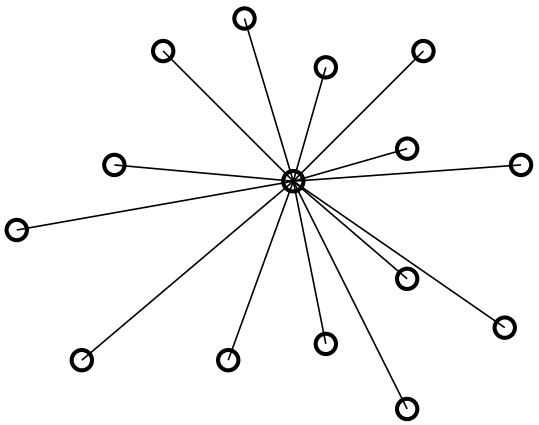
1/4 of all friends report YES

Run a Survey

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



3/4 of all friends report YES



1/4 of all friends report YES

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

Social Network Based Attack

[RG ICDM'19]

- Without extra knowledge on social network structures, Bayesian inference is optimal.

Social Network Based Attack

[RG ICDM'19]

- Without extra knowledge on social network structures, Bayesian inference is optimal.
- Inference with perturbed data with (ε, δ)-differential privacy has Area Under the ROC Curve (AUC) at most

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

Social Network Based Attack

[RG ICDM'19]

- Without extra knowledge on social network structures, Bayesian inference is optimal.
- Inference with perturbed data with (ε, δ)-differential privacy has Area Under the ROC Curve (AUC) at most

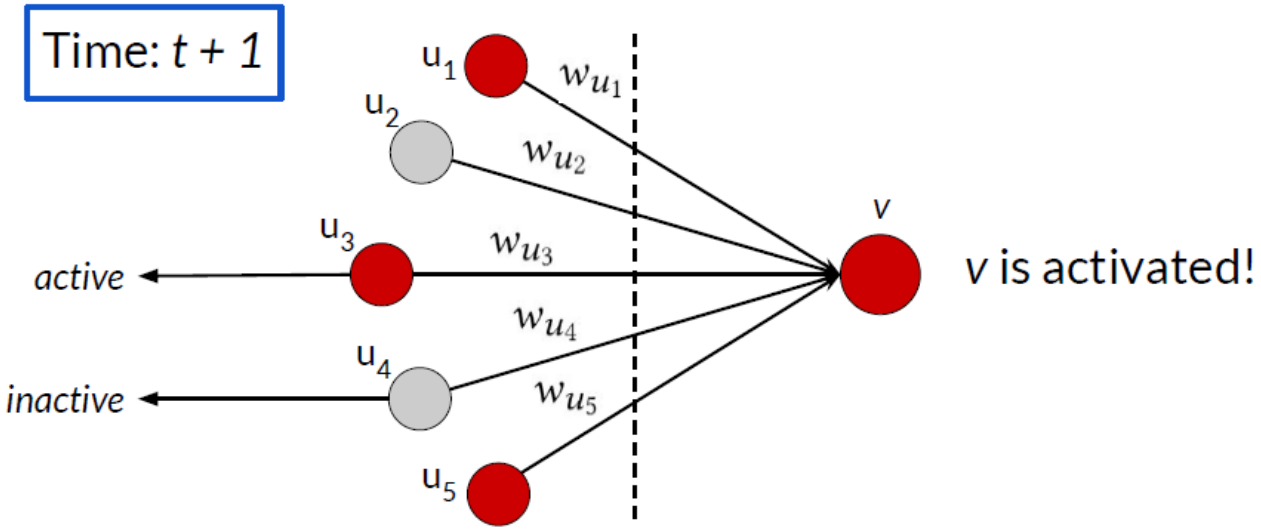
$$1 - \frac{1-\delta}{1+e^{\varepsilon}}$$

- 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.

Social Network Based Attack

Contagion model – linear threshold model:

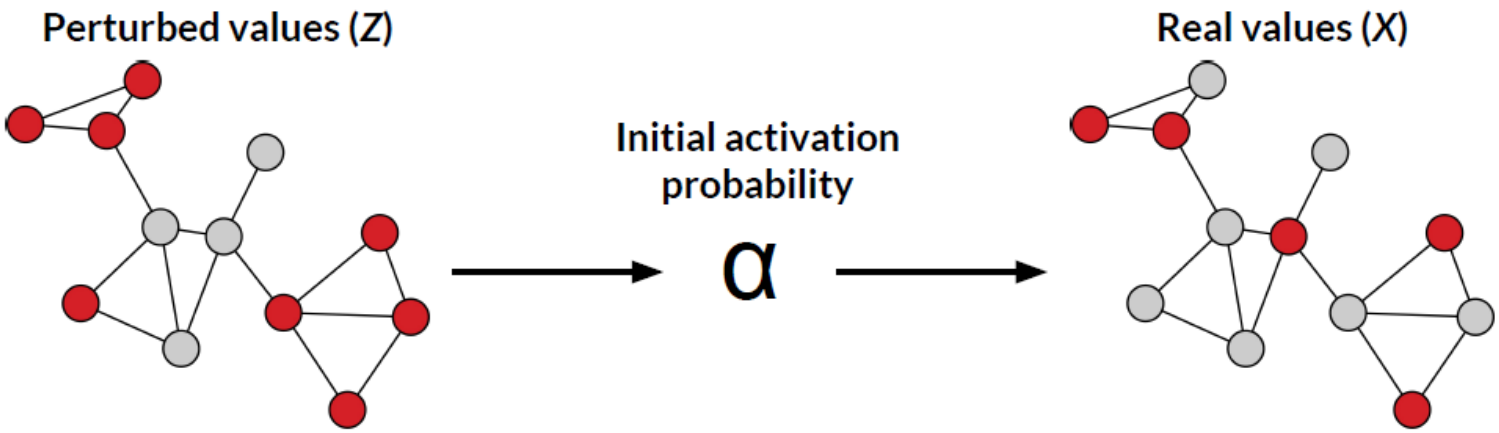
- Edge weight w_{ij} indicates influence from i to j;
- Each node has a threshold $\lambda(v) \in [0, 1]$;
- If the sum of influence from all activated neighbors of v goes beyond $\lambda(v)$, v is activated in the next round.
- The process stops when no new node is activated.



Social Network Based Attack

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.



Going Beyond the Bounds

We achieve AUC above the theoretical bounds.

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

We surpass the ***theoretical bound***

Going Beyond the Bounds

Results hold for real world graphs, beating alternative schemes.

	Network	β	ε	Upper Bound	Bayesian	CO-DAG	O-DAG	CO-RND	O-RND	Lappas+
Real-world	GrQc	0.1	0.201	0.550	0.554	**0.577**	0.576	0.564	0.561	N/A*
		0.3	0.619	0.650	0.644	**0.720**	0.718	0.683	0.685	N/A
		0.5	1.099	0.750	0.744	0.833	**0.834**	0.798	0.797	N/A
		0.7	1.735	0.850	0.849	**0.908**	0.892	0.892	0.878	N/A
		0.9	2.944	0.950	0.951	**0.973**	0.960	0.971	0.965	N/A
	HepTh	0.1	0.201	0.550	0.553	**0.584**	0.579	0.568	0.566	N/A
		0.3	0.619	0.650	0.658	**0.733**	0.716	0.698	0.703	N/A
		0.5	1.099	0.750	0.751	**0.831**	0.825	0.793	0.793	N/A
		0.7	1.735	0.850	0.849	**0.917**	0.893	0.884	0.876	N/A
		0.9	2.944	0.950	0.950	**0.972**	0.958	0.962	0.962	N/A
	Amazon Videos	0.1	0.201	0.550	0.554	0.598	**0.604**	0.584	0.585	N/A
		0.3	0.619	0.650	0.649	**0.748**	0.748	0.706	0.714	N/A
		0.5	1.099	0.750	0.747	**0.854**	0.836	0.828	0.812	N/A
		0.7	1.735	0.850	0.849	**0.919**	0.899	0.906	0.894	N/A
		0.9	2.944	0.950	0.950	**0.975**	0.962	0.974	0.967	N/A
	Amazon DVDs	0.1	0.201	0.550	0.551	0.576	**0.584**	0.567	0.570	N/A
		0.3	0.619	0.650	0.652	0.735	**0.739**	0.699	0.707	N/A
		0.5	1.099	0.750	0.750	**0.830**	0.829	0.806	0.802	N/A
		0.7	1.735	0.850	0.851	**0.906**	0.897	0.894	0.886	N/A
		0.9	2.944	0.950	0.952	**0.963**	0.957	0.957	0.941	N/A

Impossibility Results

[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

Privacy Aware Learning

- Lessons: rich structures in user data can lead to information leakage that users may not be ware of.

Privacy Aware Learning

- Lessons: rich structures in user data can lead to information leakage that users may not be ware 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 ware 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?

Background: Adversarial Learning

Generate adversarial samples to fool a classifier, by adding noises to the input.

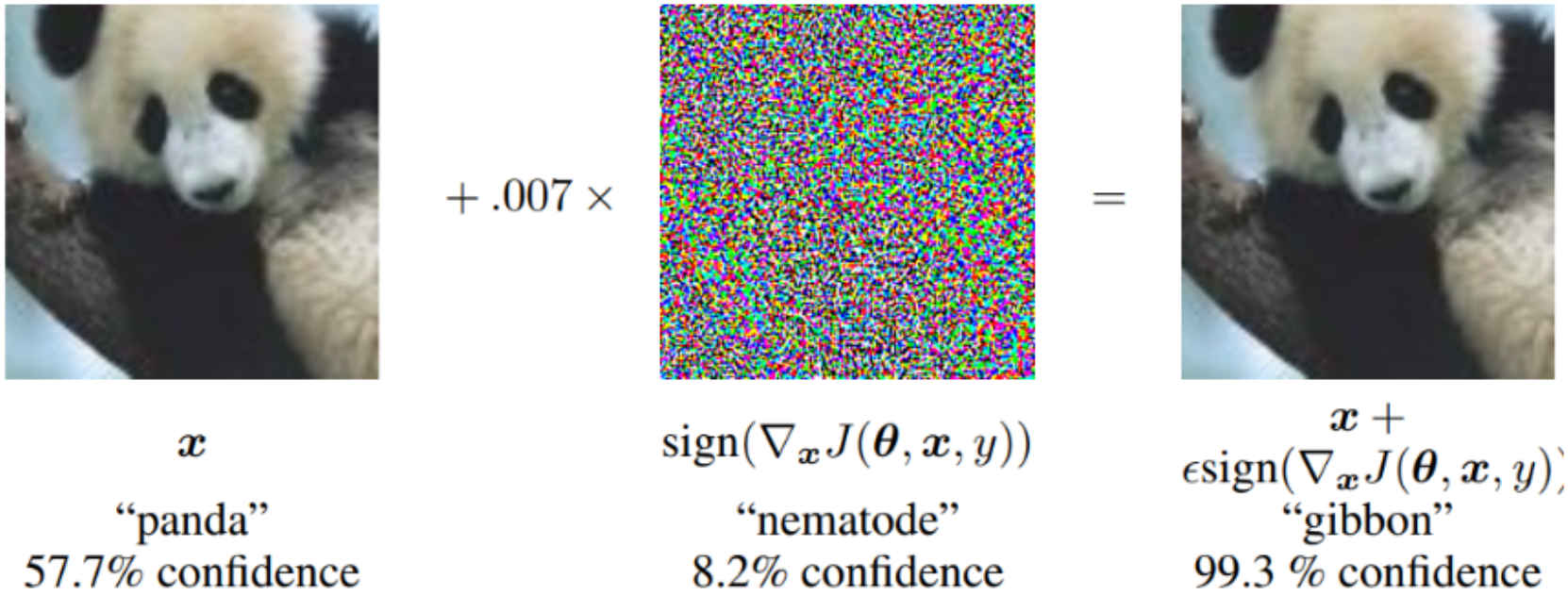


Image taken from https://arxiv.org/pdf/1412.6572.pdf

Background: Generative Adversarial Networks

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

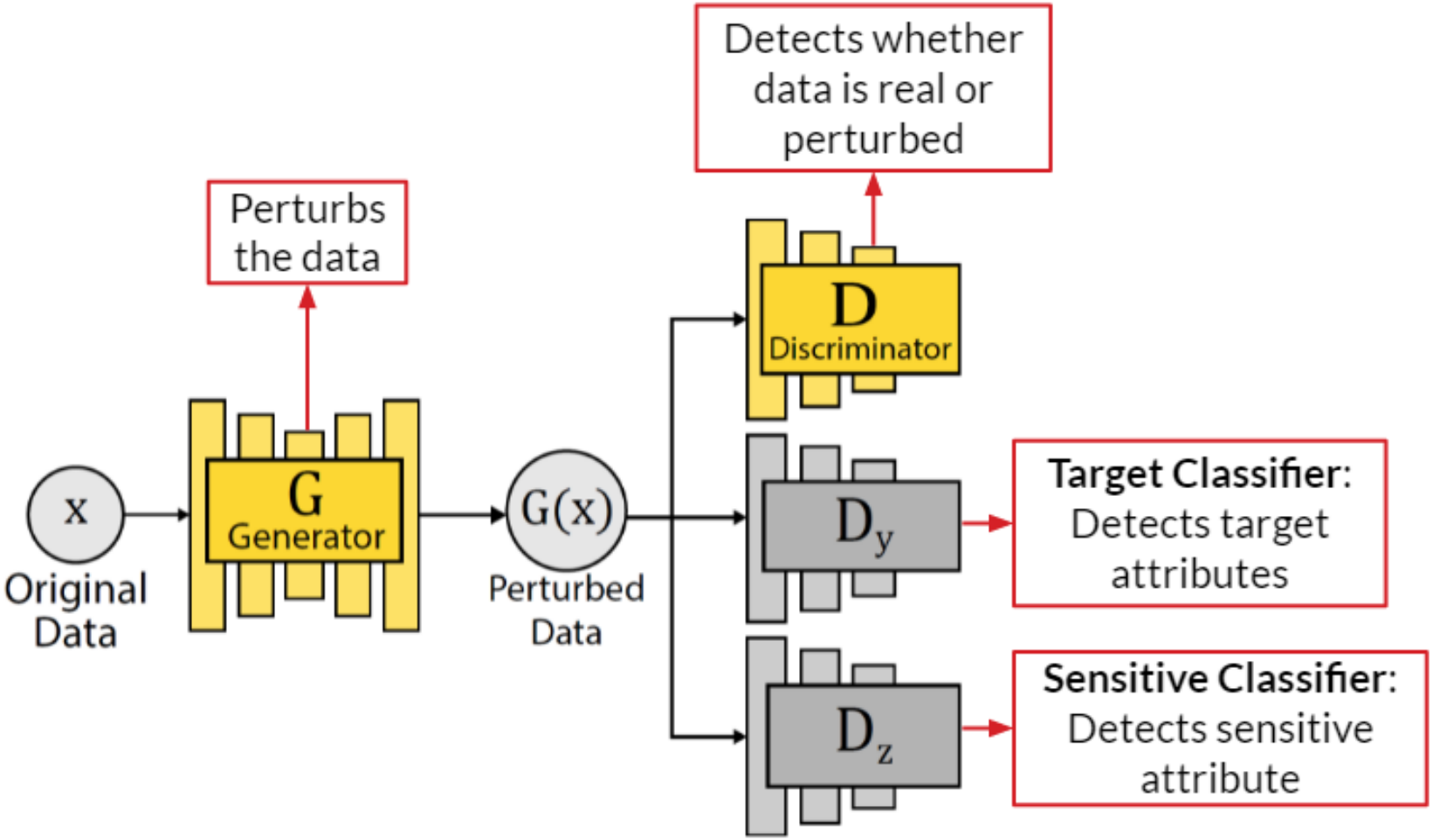
- A generator trying to generate fake samples.
- A discriminator trying to detect whether a sample is fake or not.

Image taken from https://arxiv.org/pdf/1907.05600v2.pdf

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.

Example Application: Smart Building Occupancy

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***

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

* 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.

Summary and Acknowledgement

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.
- Influencers and the Giant Component: the Fundamental Hardness in Privacy Protection for Socially Contagious Attributes, Aria Rezaei, Jie Gao, Anand D. Sarwate, SDM'2021.
- Application-Driven Privacy-Preserving Data Publishing with Correlated Attributes, Aria Rezaei, Chaowei Xiao, Jie Gao, Bo Li, Sirajum Munir, EWSN 2021.

Questions and comments?