Novel Approaches to Preserving Utility in Privacy Enhancing Technologies

Meisam Mohammady

Research Scientist, Data61

CSIRO, Australia

June 2022

CYBER SECURITY COOPERATIVE RESEARCH CENTRE

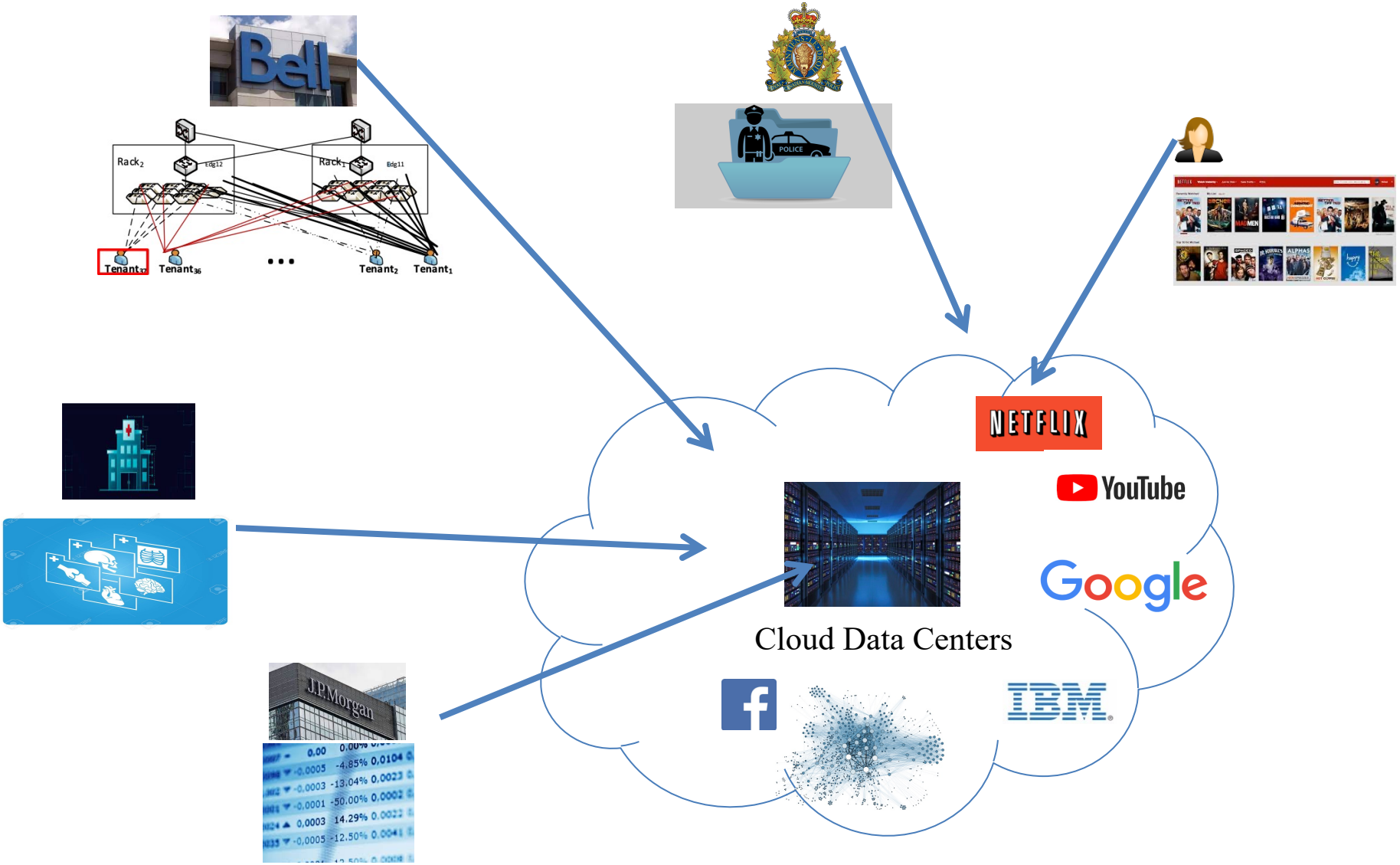
Outline

- Introduction
- Novel Approaches to Preserving Utility in Privacy Enhancing Technologies
  - I. *Multi-view*: Preserving Utility in Network Trace Anonymization
  - II. *DPOAD*: ***D***ifferentially ***P***rivate ***O***utsourcing of ***A***nomaly ***D***etection
- Ongoing Research

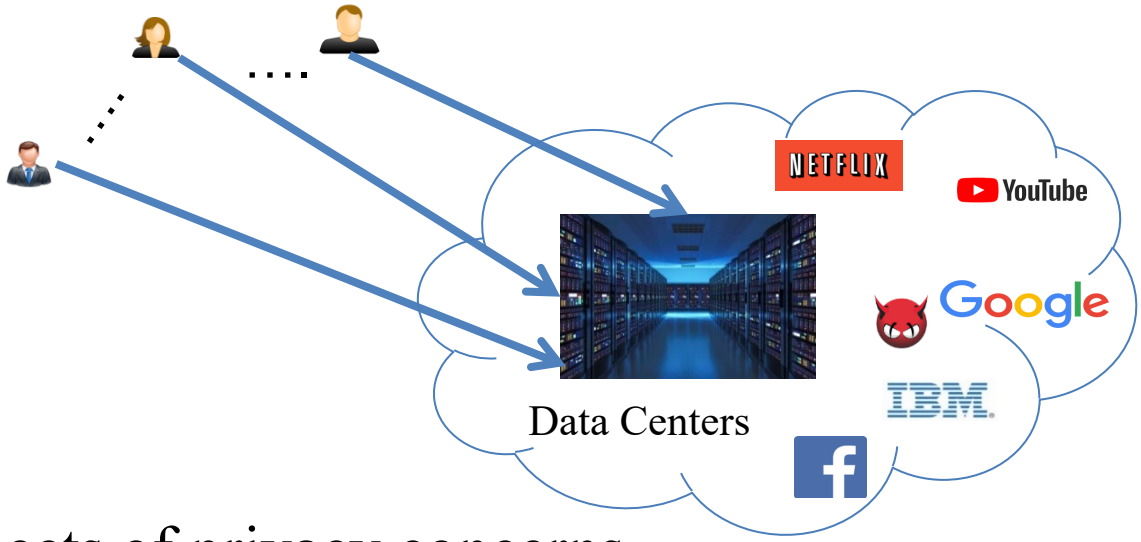
Outline

- Introduction
- Novel Approaches to Preserving Utility in Privacy Enhancing Technologies
  - I. *Multi-view*: Preserving Utility in Network Trace Anonymization
  - II. *DPOAD*: ***D***ifferentially ***P***rivate ***O***utsourcing of ***A***nomaly ***D***etection
- Ongoing Research

The Need for Data Privacy



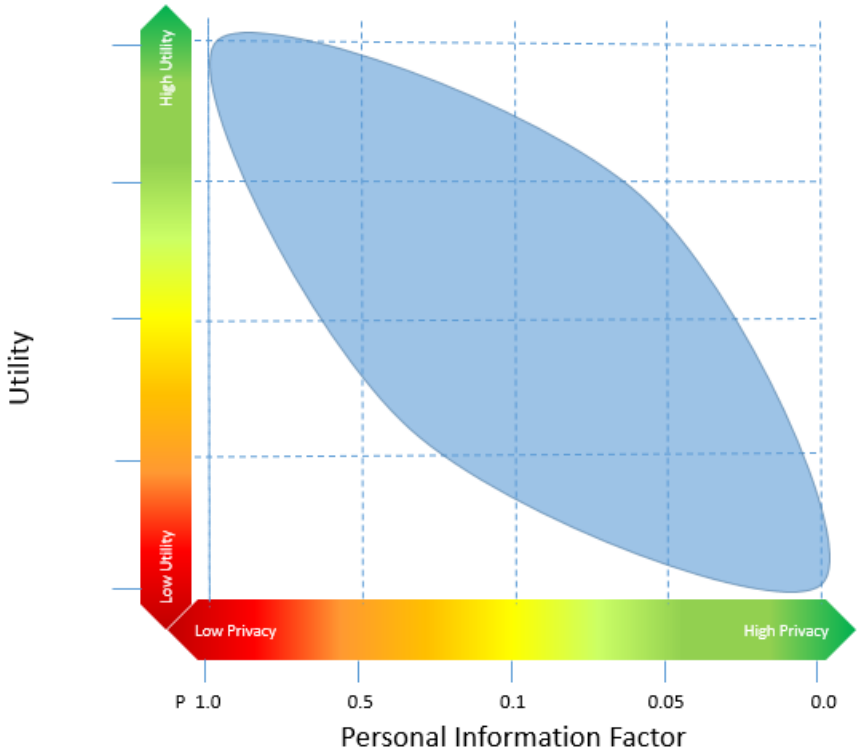
Privacy Issues and Regulations



Main aspects of privacy concerns

- Customers' activities can be exploited by unauthorized parties through targeting them with Ads, black mailing, etc. [USENIX 06]
- Leaked network topology information may cause other attacks, e.g., DoS [INFOCOM 12]
- According to GDPR Article 28, providers cannot share tenants' data without protection while acquiring the services of third parties

Privacy Degrades Utility (Trade-off)

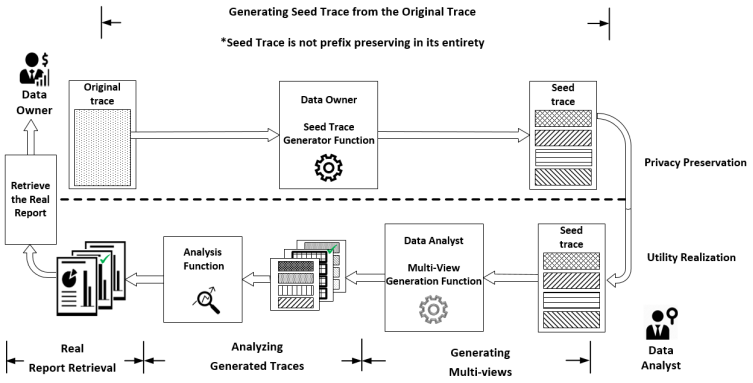


Finding the optimal point for each application is important and challenging.

Highlight of the Proposed Systems

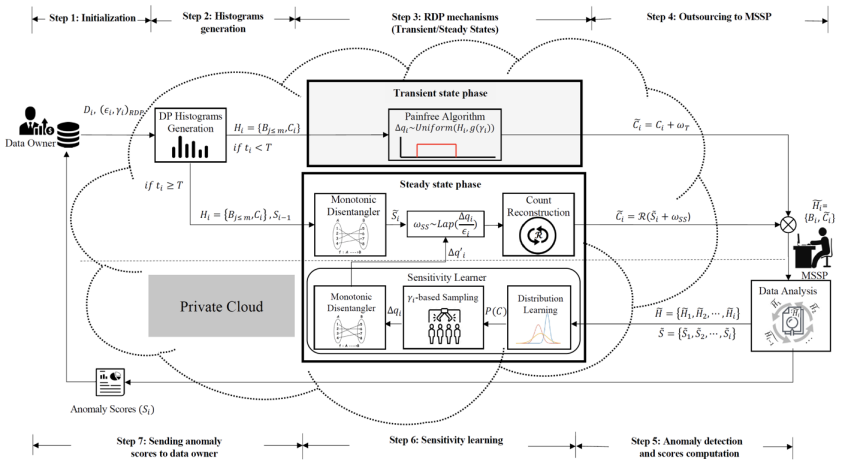
The *Multi-view* Approach [CCS' 18]:

PP Outsourcing of Network Security Analysis



The *DPOAD* Approach [Submitted to ICDE' 22]:

Outsourcing Anomaly Detection with Differential Privacy



Outline

- Introduction
- Novel Approaches to Preserving Utility in Privacy Enhancing Technologies
  - *I.* *Multi-view*: Preserving Utility in Network Trace Anonymization
  - *II.* *DPOAD*: ***D***ifferentially ***P***rivate ***O***utsourcing of *A*nomaly ***D***etection
- Ongoing Research

Privacy and Utility Requirements

ISP

Original Trace						
Start Time	Src Port	Dst Port	P	Pkts	ScrIPaddr	DstIPaddr
10:23:42:50	902	600	TCP	6	10.1.1.0	194.20.2.2
10:23:42:53	901	2000	TCP	8	10.1.1.1	194.20.52.32
10:23:43:54	900	63	TCP	10	128.0.1.12	74.60.2.2
10:53:42:54	800	1900	TCP	2	190.21.1.52	84.12.2.2

Prefix Preserving Anonymizer (CryptoPAn)

Anonymization Function

CryptoPAn Output						
Start Time	Src Port	Dst Port	P	Pkts	SrcIPaddr	DstIPaddr
10:23:42:50	902	600	TCP	6	117.14.242.125	108.0.15.22
10:23:42:53	901	2000	TCP	8	117.14.242.124	108.0.47.22
10:23:43:54	900	63	TCP	10	37.64.22.54	78.10.26.22
10:53:42:54	800	1900	TCP	2	67.02.02.11	98.02.30.22

The data owner needs to be convinced that the outsourced traces are anonymized (**privacy of the data**)

The anonymized data needs to be accurate enough to allow for a valid analysis (**utility of the data**)

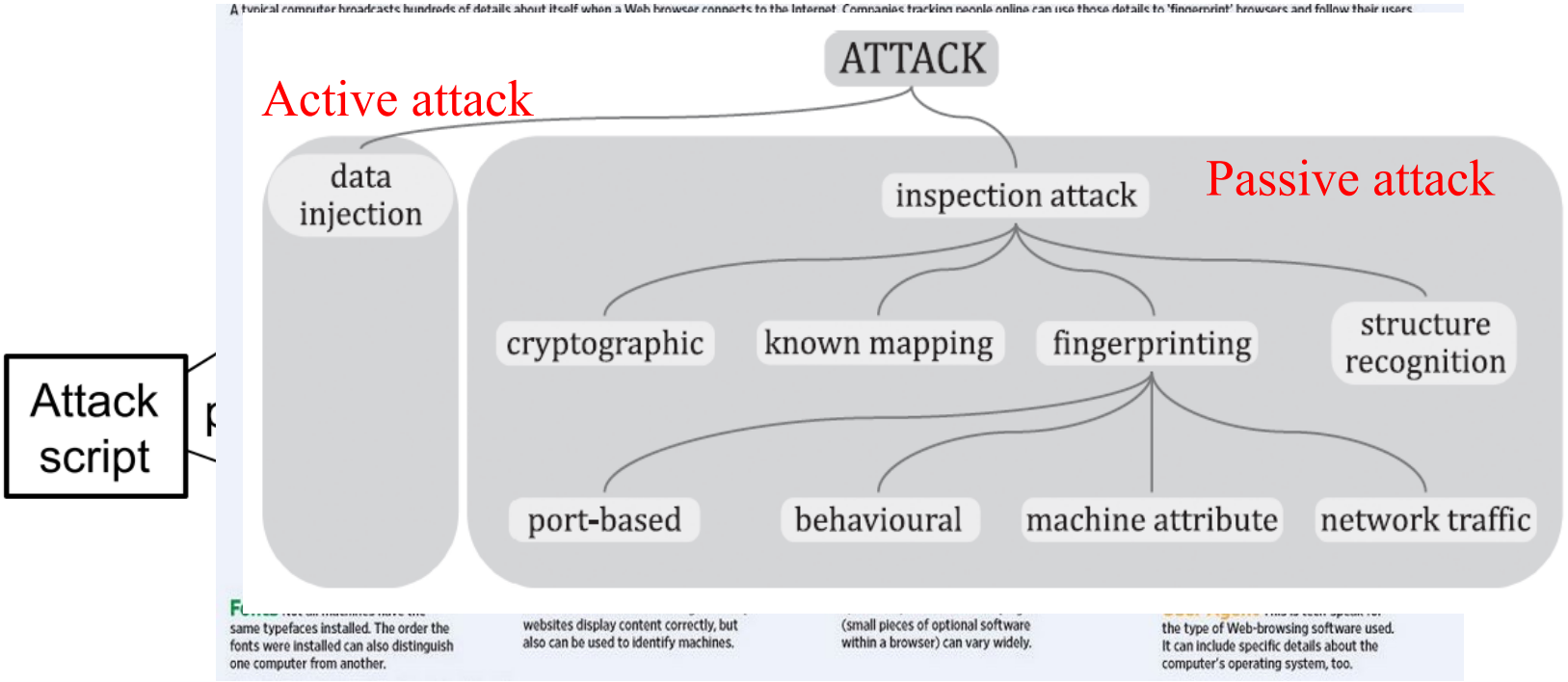
PP(-,K):

If two real addresses share first X bits, then the same two anonymized addresses

share first X bits

Taxonomy of attacks Against Anonymized Traces

Data injection [2]



Semantic Attacks on Prefix Preserving Anonymization (CryptoPAn)

Original Trace

Start Time	Src Port	Dst Port	P	Pkts	ScrIPaddr
10:23:42:50	902	600	UDP	6	10.1.1.0
10:23:42:53	901	2000	UDP	8	150.10.10.1
10:23:43:54	900	63	UDP	10	128.10.10.1
10:53:42:54	800	1900	TCP	2	10.1.1.0
10:53:42:55	750	2330	TCP	1	150.10.1.0
10:53:42:56	220	591	TCP	1	150.10.20.0
10:53:42:57	22	2600	TCP	1	10.1.1.0

CryptoPAn Output

Start Time	Src Port	Dst Port	P	Pkts	ScrIPaddr
10:23:42:50	902	600	UDP	6	117.14.242.125
10:23:42:53	901	2000	UDP	8	159.61.5.252
10:23:43:54	900	63	UDP	10	135.243.4.124
10:53:42:54	800	1900	TCP	2	117.14.242.125
10:53:42:55	750	2330	TCP	1	159.61.13.126
10:53:42:56	220	591	TCP	1	159.61.20.124
10:53:42:57	22	2600	TCP	1	117.14.242.125

Step1: Injecting flows *(injection)*

Step2: Recognizing injected flows via *Start Time* and *Src Port*

(Fingerprinting)

Step3: Identifying more flows via shared Prefixes

(Structure Recognition)

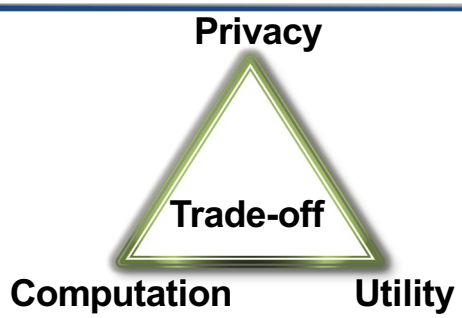
Step4: De-anonymizing IPs or prefixes

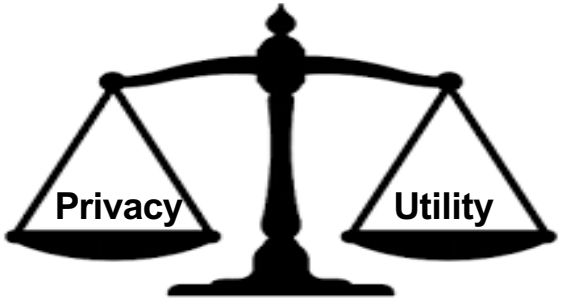
(Known Mapping)

Existing Anonymizations Techniques

Tool name	Anonymized Fields					Anonymization method					Weaknesses	
	Netflow fields	IP address	Port	Header	Payload	Prefix-preserving	Permutation	Truncation	Hashing	Shifting	Highly sanitized	Semantic attacks
Anontool [ICC 06]	●	●				●	●		●		●	●
CANINE [ICTSMA 05]	●	●	●			●	●			●	●	●
CoralReef [USENIX 01]	●	●	●			●	●	●				●
Flaim [USENIX 06]	●	●	●			●	●		●	●	●	●
IPsumdump		●		●		●						●
NFDUMP	●	●			●	●						●
SCRUB [CORR 07]		●		●	●		●				●	
TCPanon					●						●	
tcpdpriv		●		●	●	●	●	●			●	●
tcpmkpub [SIGCOMM 06]		●		●	●	●		●				●
tcpurify		●			●		●	●				●

Main Idea





Again, the famous privacy-utility trade-off!

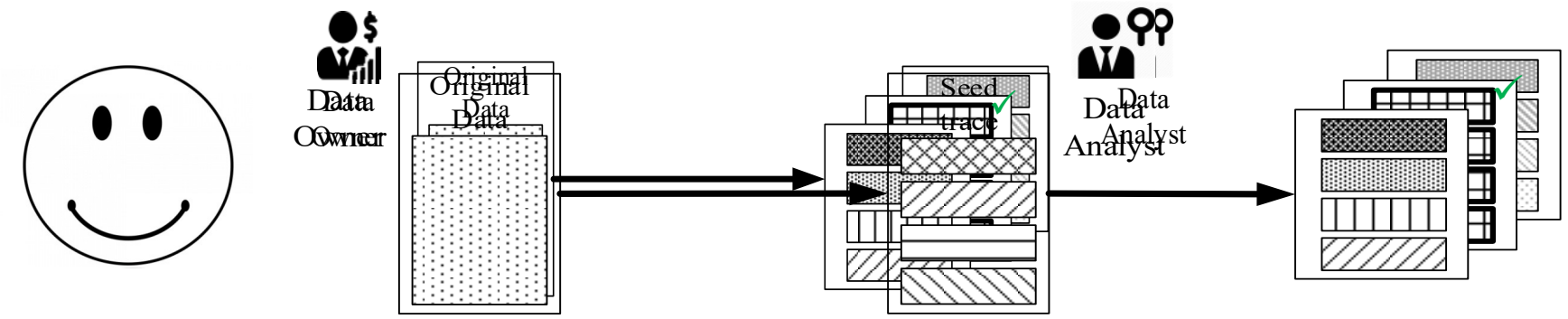
Can we have the best of both worlds by sacrificing something else (more expendable)?

Answer: Preserve both privacy/utility with more computations (today's computation is cheaper esp. with clouds)

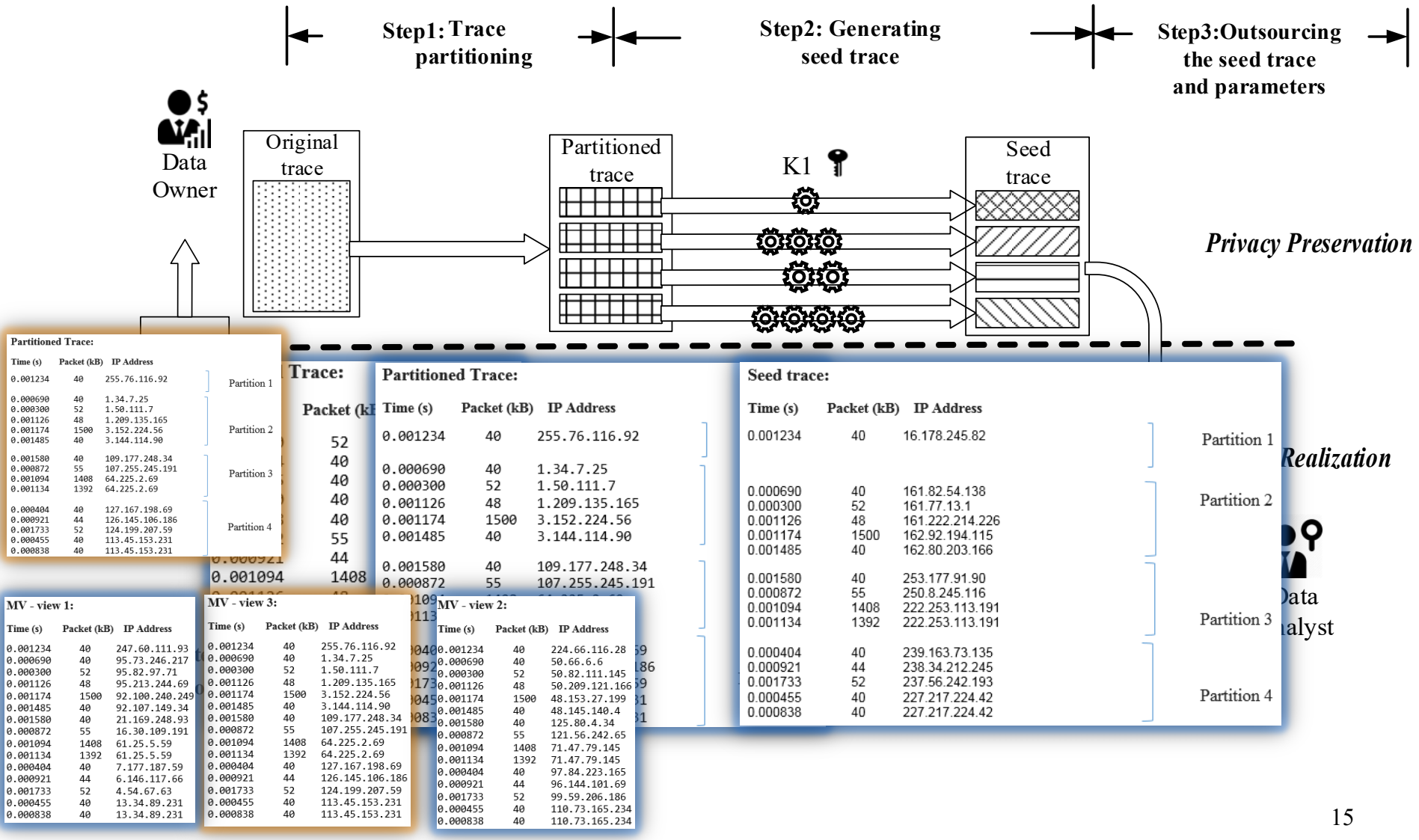
Multi-view Approach in a Nutshell

- Data owner asks analyst to analyze multiple views of the original data
  - Privacy: The real view is hidden among many fake views
  - Utility: Data owner secretly retrieves the analysis results of the real view

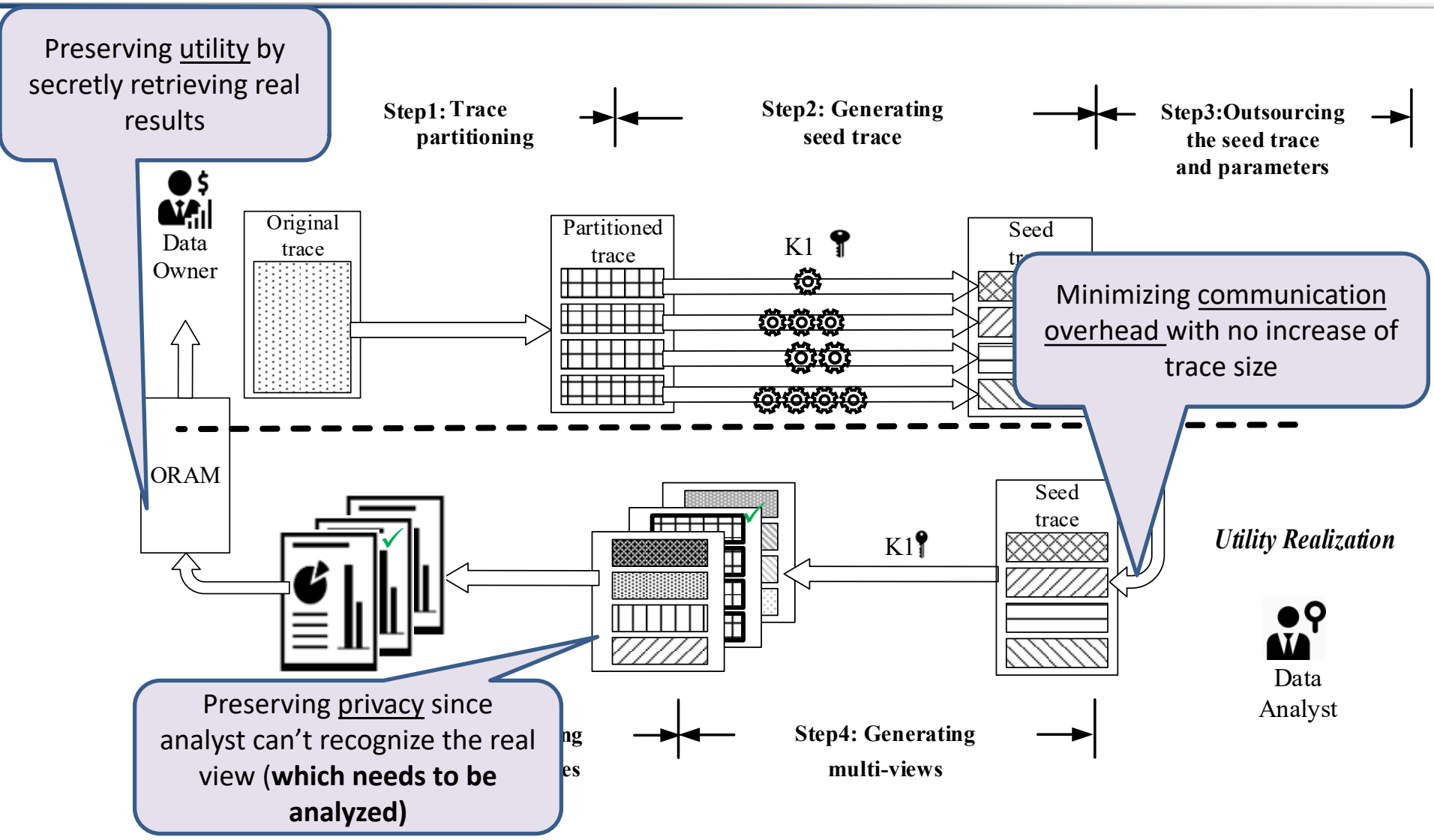
- **Key challenge:** How to minimize the communication overhead considering the sheer size of network traces?



The Multi-view Approach



The Multi-view Approach Benefits

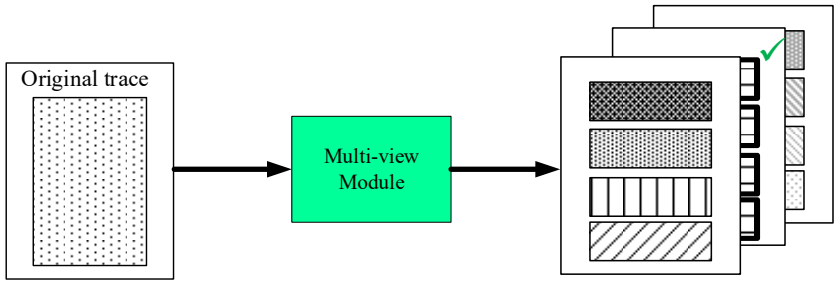


Adversary Model [Sigcomm 10]

- The adversary is an **honest-but-curious** analyst
- The goal of the adversary is to find all possible ***matches*** between the anonymized and original traces
- Suppose the trace consists of ***d* groups** (e.g., those in the same subset), and among these an ***α*-knowledge** adversary can successfully **inject** or **fingerprint** α ($\leq d$) groups.

Quantifying View Indistinguishability

Fake views must be generated such that the adversary cannot **distinguish** them from the real view!



α-knowledge adversary

Definition. A multi-view solution is said to satisfy ϵ −Indistinguishablity against an α-knowledge adversaryif and only if

$$e^{-\epsilon} \leq \frac{\Pr(\text{view i may be the real view})}{\Pr(\text{view r may be the real view})} \leq e^{\epsilon}$$

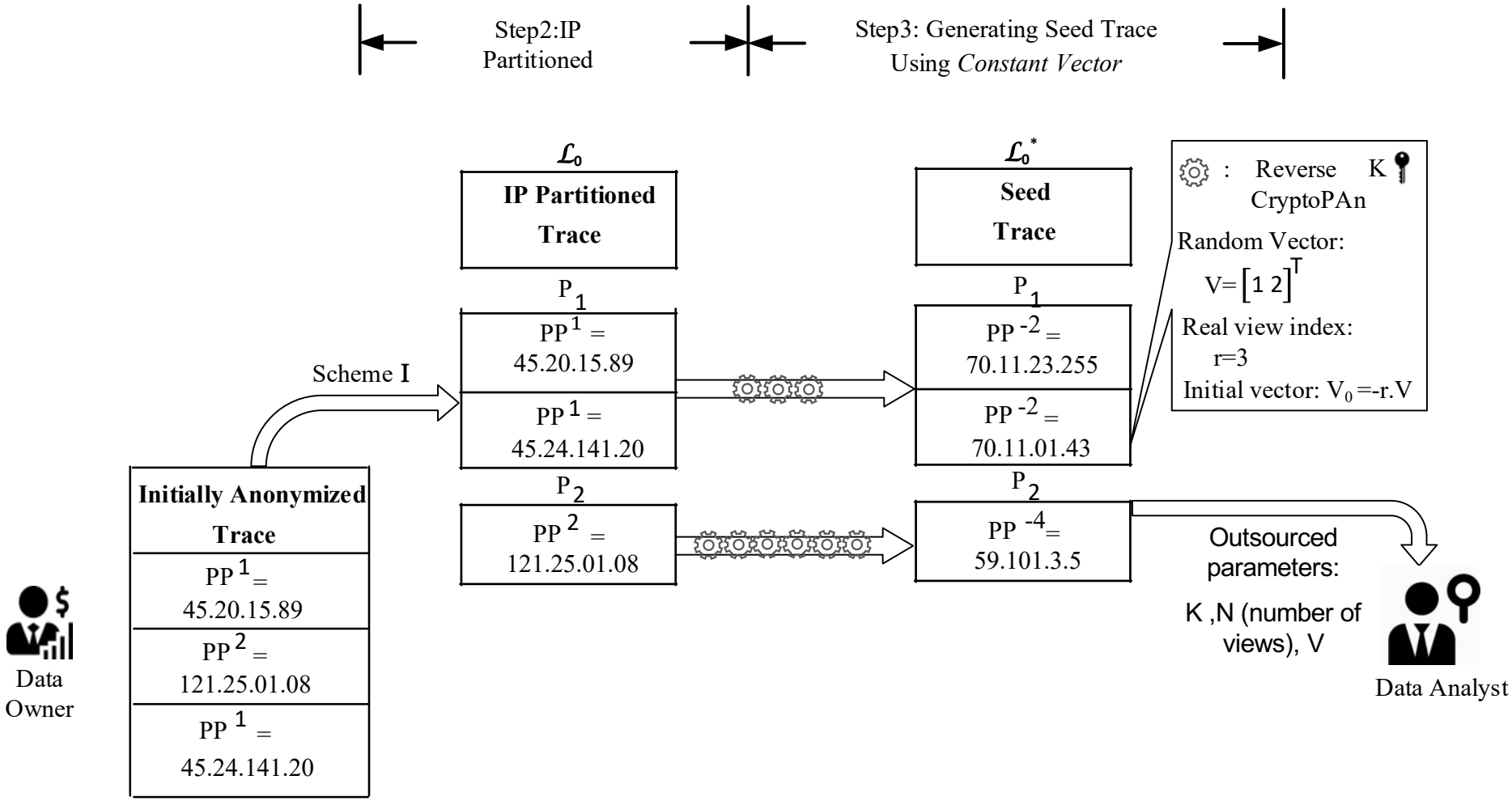
ϵ–Indistinguishable Multi-view Mechanisms

ϵ would depend on the specific design of a multi-view solution.

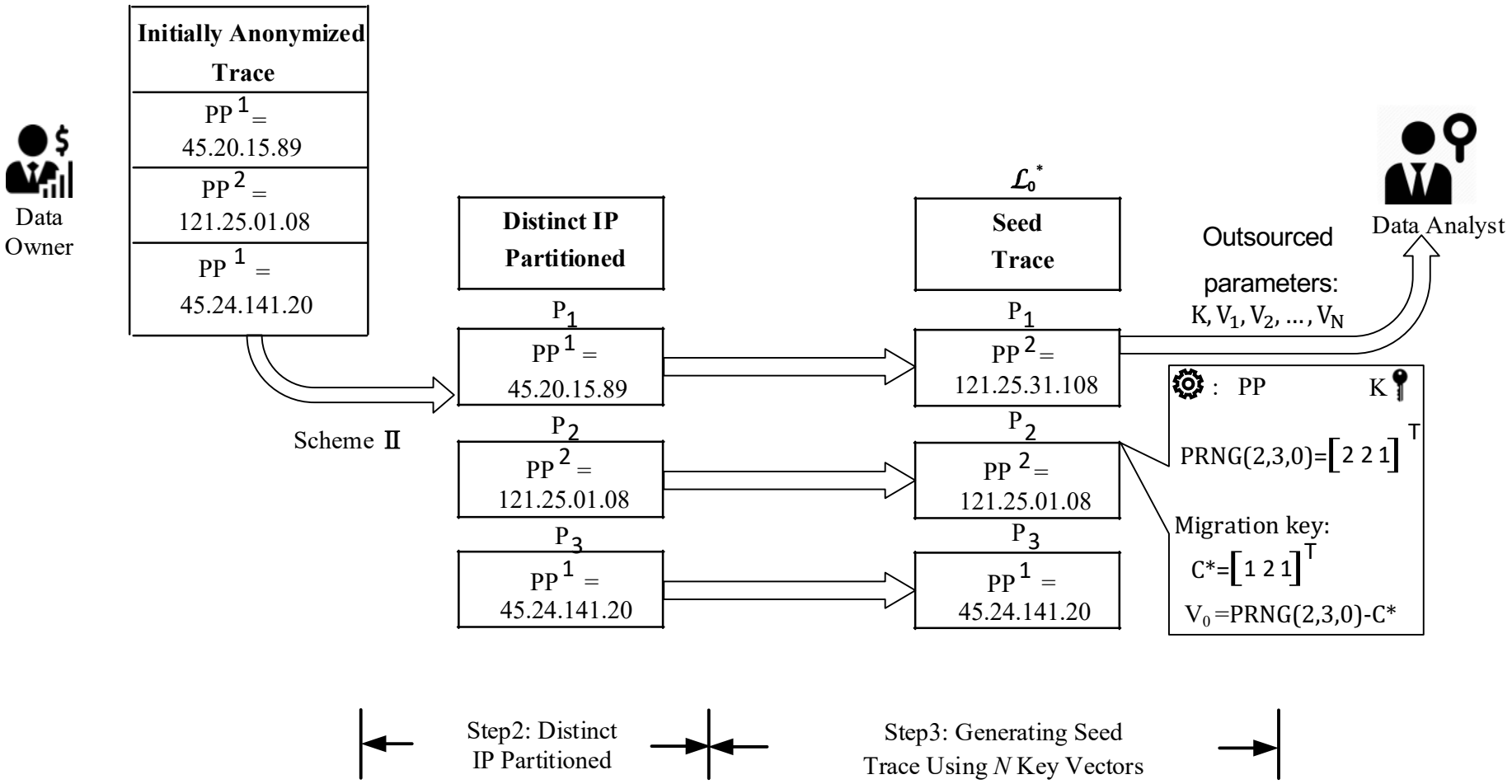
- Scheme I: Perfect Indistinguishability (ϵ=0) with less protected partitions
  (Fake views still contain a lot of sensitive information)

- Scheme II: Sacrifices some indistinguishability to achieve better protected partitions (in the sense of slightly less real view candidates)

Scheme I: Subnet-based Partitioning Approach



Scheme II: IP-based Partitioning Approach



Scheme II:Indistinguishability Analysis

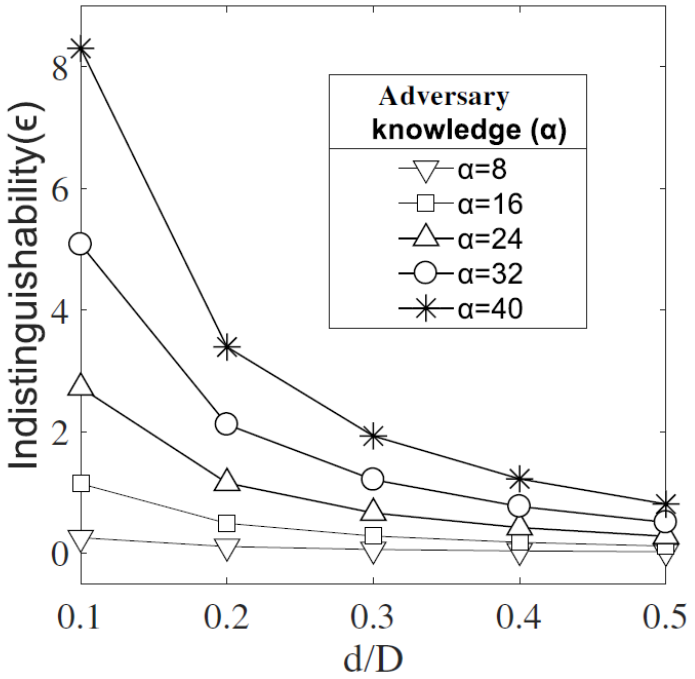
Theorem. The indistinguishability parameter ϵ of the generated views in scheme II is lower-bounded by

$$ln\left[\frac{D^{\alpha}}{d^{\alpha}} \cdot \prod_{i=0}^{\alpha-1} \frac{d-i}{D-i}\right]$$

D: Number of distinct addresses

d: Number of prefix groups (subnets)

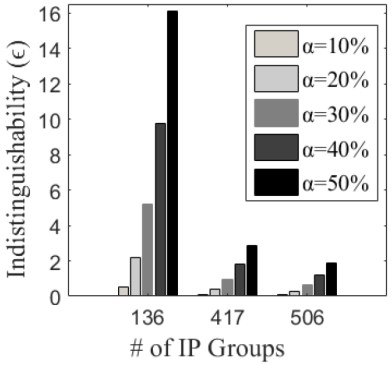
α: Adversary's knowledge



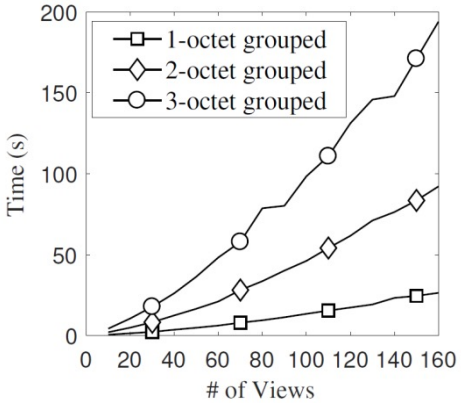
Experiments

Dataset

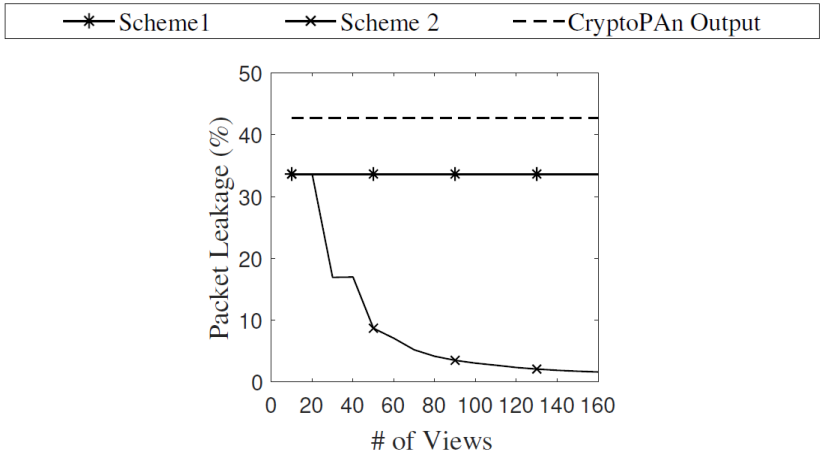
Attribute	Value
Type Header	Header
Start date	August 1st, 2016
End date	October 21st, 2016
Size	70 GB
Format	pcap and TXT
Distinct IP addresses	883
Number of packets	1M



Computational overhead

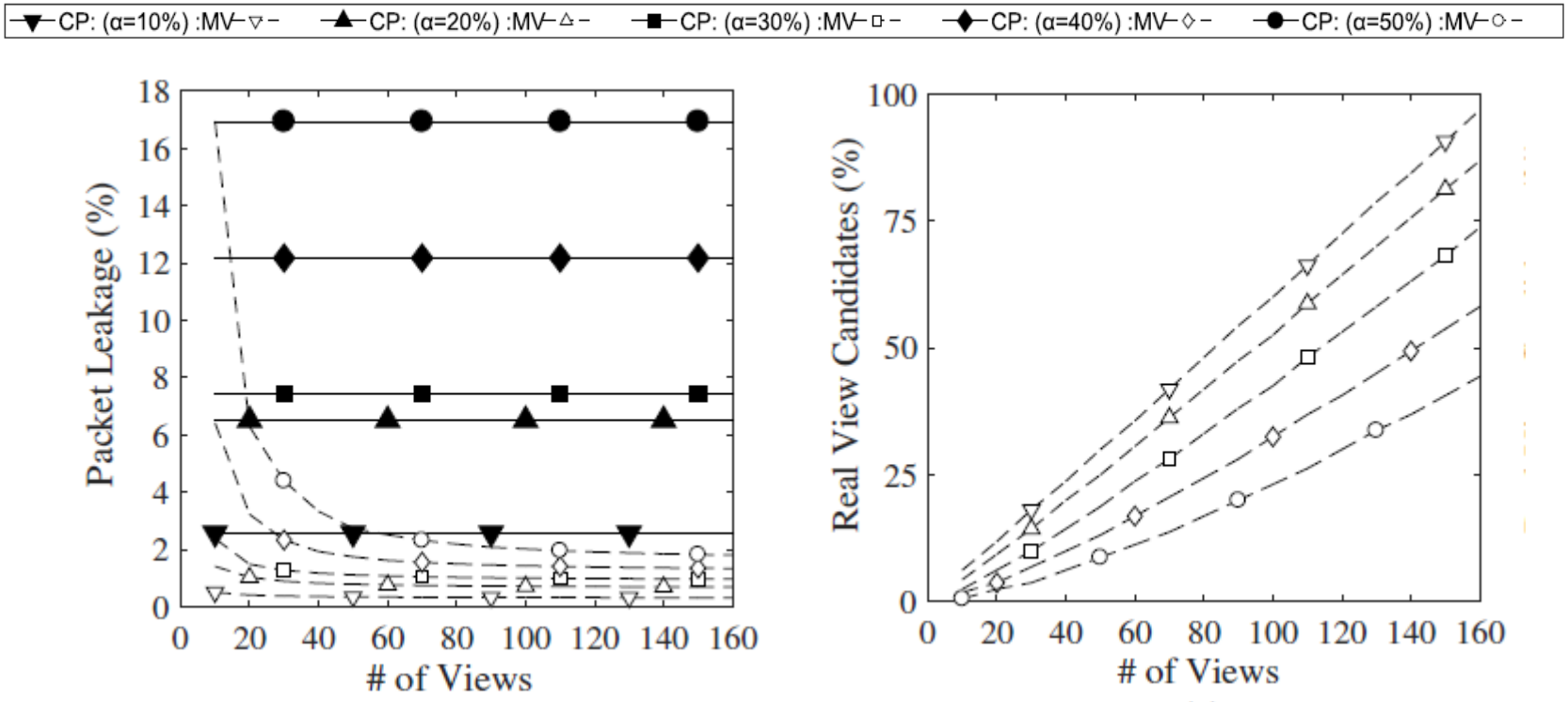


Comparison between the two schemes



Experiments

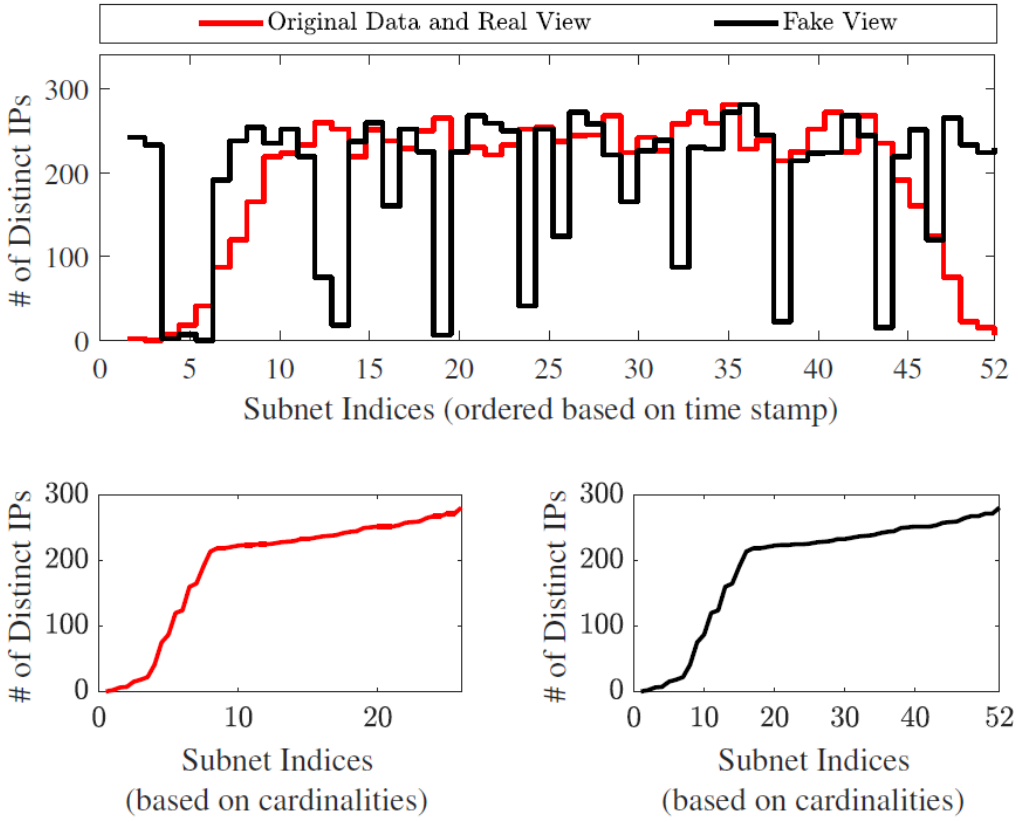
Privacy Evaluation



Time Octets Grouping (1377 Groups)

Experiments

Utility Evaluation



Conclusion

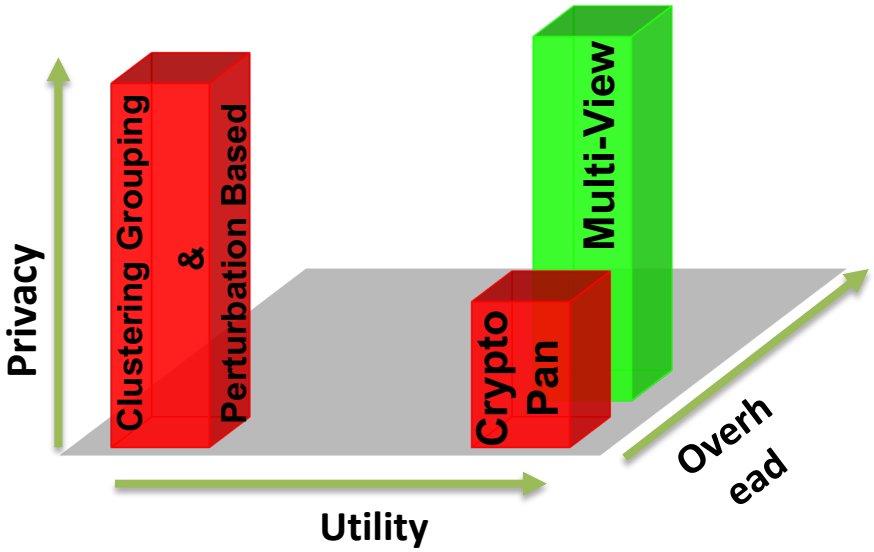
1) Multi-view approach offers the following features:

   a) Protects sensitive information in network traces

   b) Preserves utility by providing a higher ratio of privacy to utility than the state of the art does

   c) Minimizes communication overhead

2) Tradeoff is shifted from privacy-utility to privacy-computation cost, where the cost can be adjusted depending on the desired level of protection



Outline

- Introduction
- Novel Approaches to Preserving Utility in Privacy Enhancing Technologies
  - I. *Multi-view*: Preserving Utility in Network Trace Anonymization
  - II. *DPOAD*: ***D***ifferentially ***P***rivate ***O***utsourcing of ***A***nomaly ***D***etection
- Ongoing Research

Outline

① Introduction

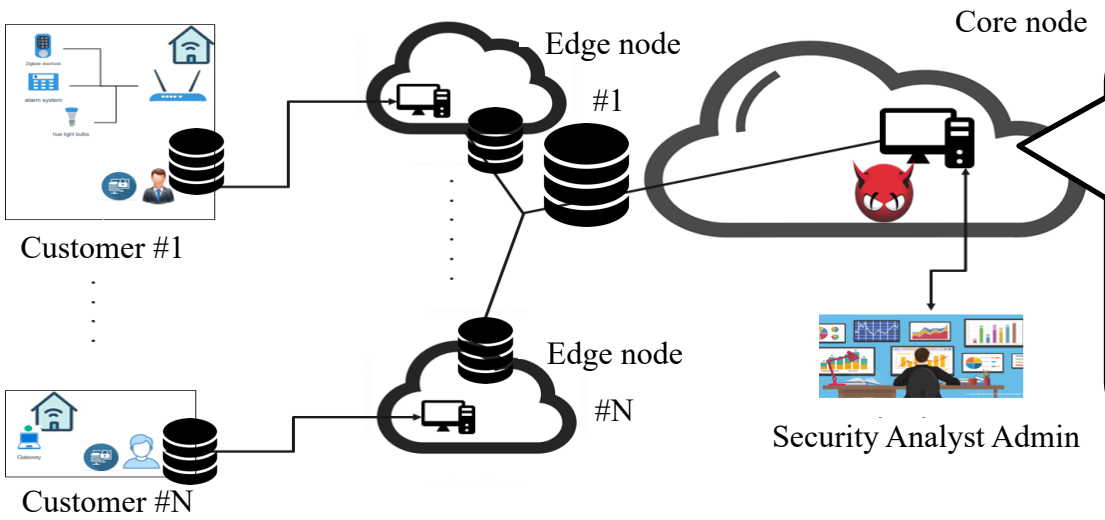
② Novel Approaches to Preserving Utility in Privacy Enhancing Technologies

- ① Multi-view: Preserving Utility in Network Trace Anonymization
- ② DPOAD: Differentially Private Outsourcing of Anomaly Detection

③ Ongoing Research

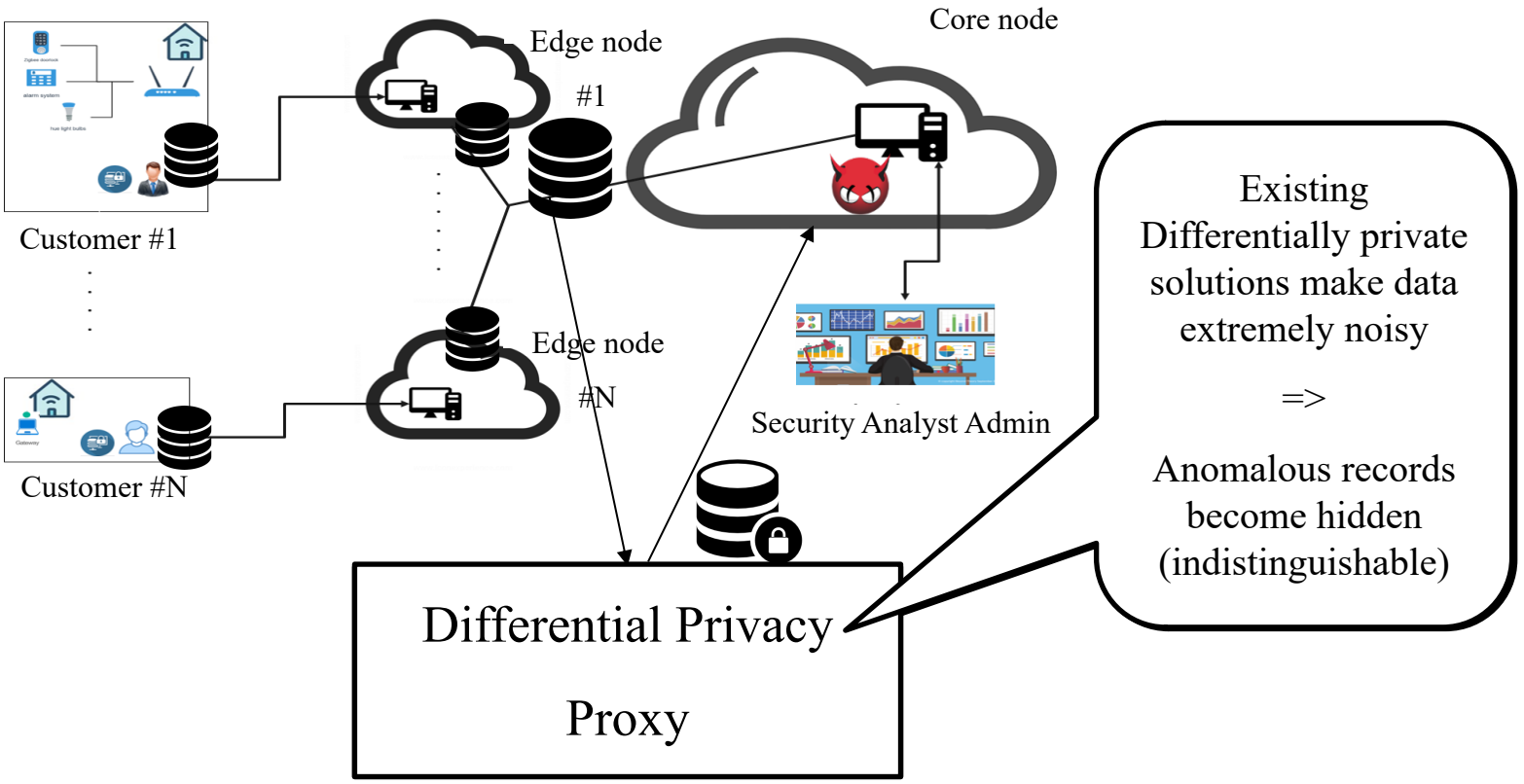
Motivation: Different Network Slices

Local decision making may lack the global view necessary to thwart large-scale attacks

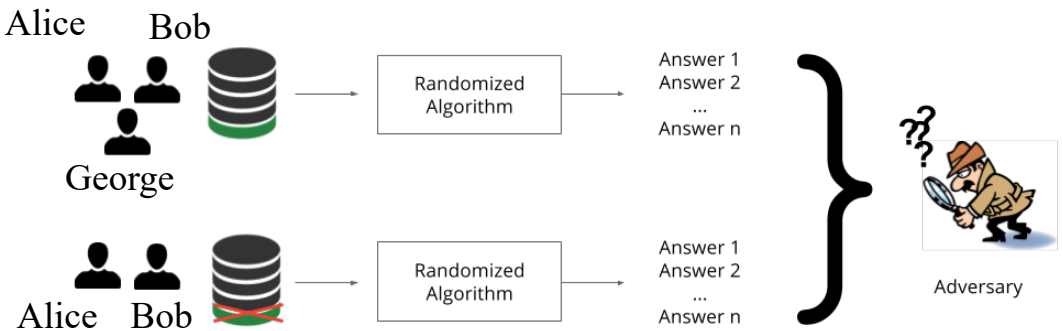


Aggregation of the personal data requires GDPR compliance

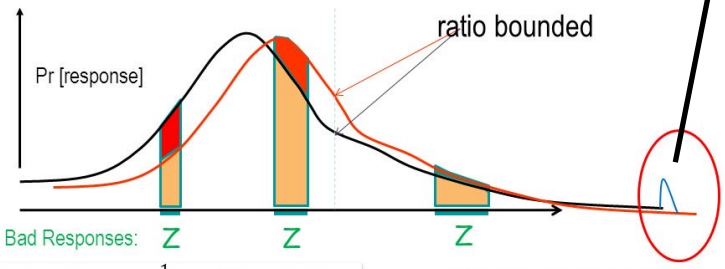
Motivation: Network Slices with Privacy Proxy



DP Contradicts Anomaly Detection



Sanitizer M gives (ε, δ) -differential privacy if:
for all **adjacent** D_1 and D_2, and all $A \mu$ range(M):
$\Pr[M(D_1) 2 A] \leq e^{\varepsilon} \Pr[M(D_2) 2 A] + \delta$



Typical setting $\epsilon = \frac{1}{10}$ and δ negligible

This course: δ negligible

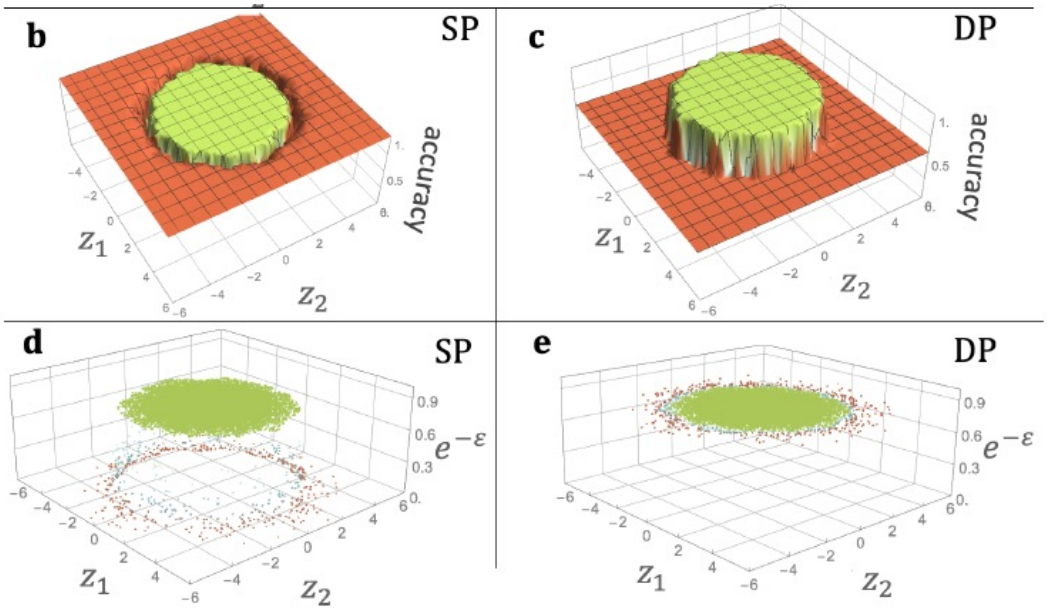
Also anomalous records are here!

Differential Privacy and Anomaly Detection in Home

Asif et al. [CCS'19]

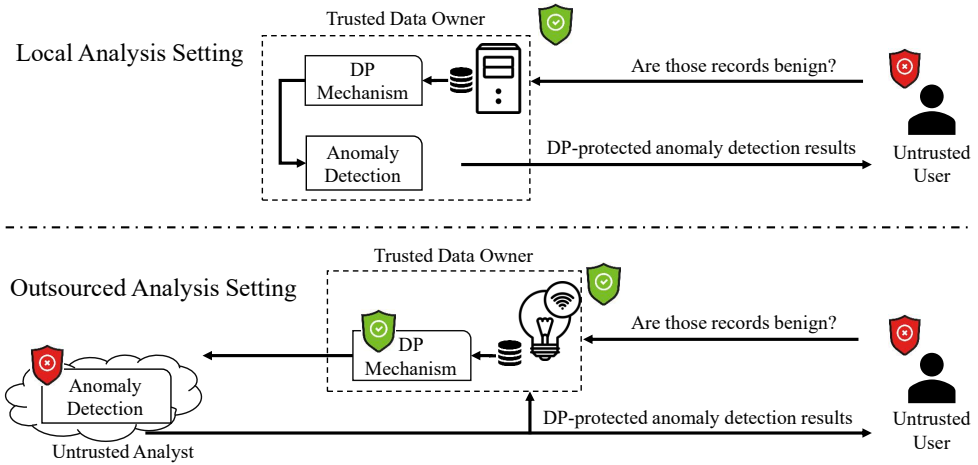
Anomaly-Restricted DP

Local Setting

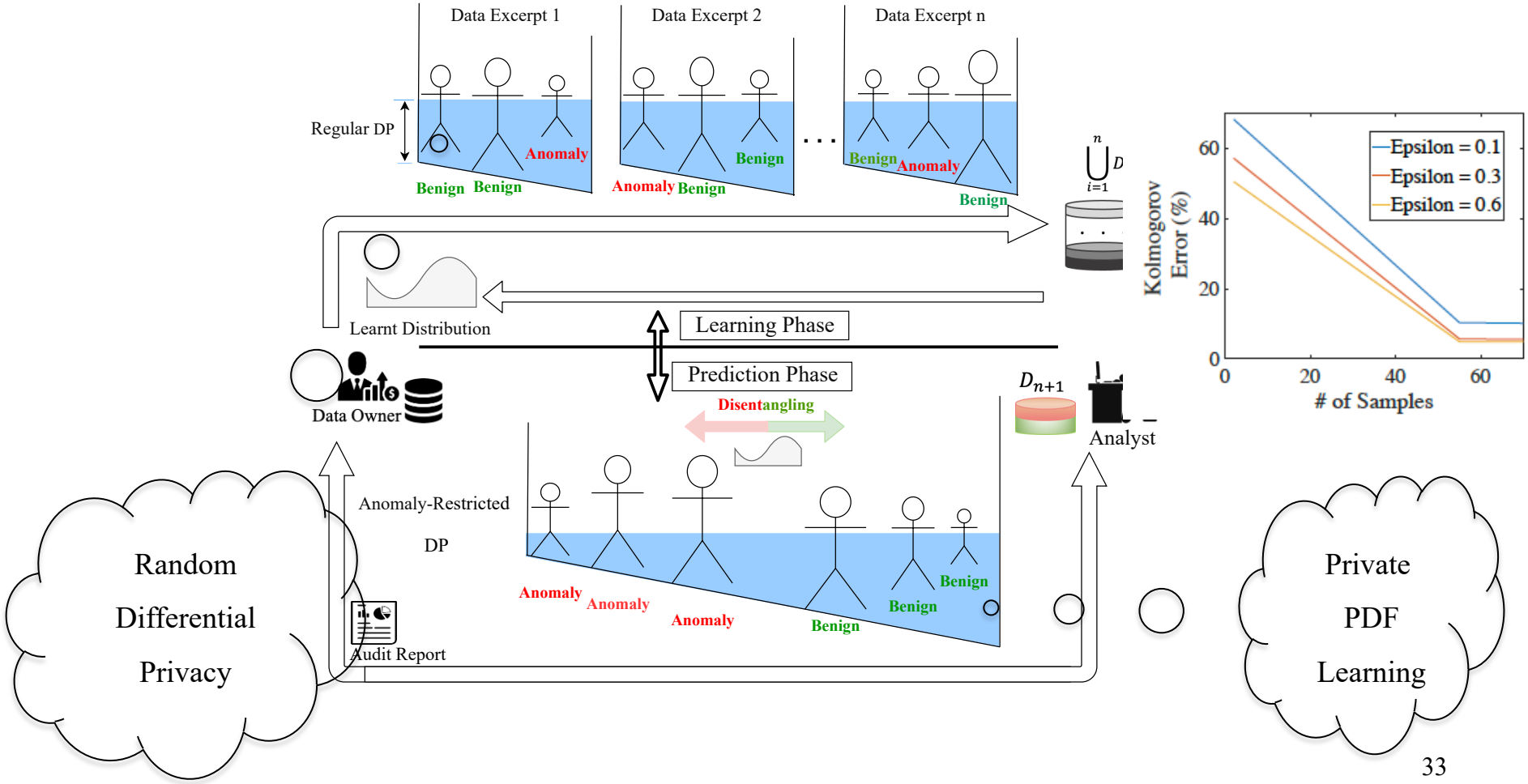


How about

Oustsourcing setting?



DPOAD Overview (Intuition)



DPOAD: Main Idea

Implement the "outsourcing" under continuous interactions.

Sensitivity Update: is **to estimate a "sensitivity"** value which **protects the privacy of normal users** but **sacrifices the privacy of anomalous** records

DPOAD

(Similar to Kalman Filtering approach)

Update

Optimal multiplicative factor

Updated (*a posteriori*) state estimate $\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k \tilde{\mathbf{y}}_k$

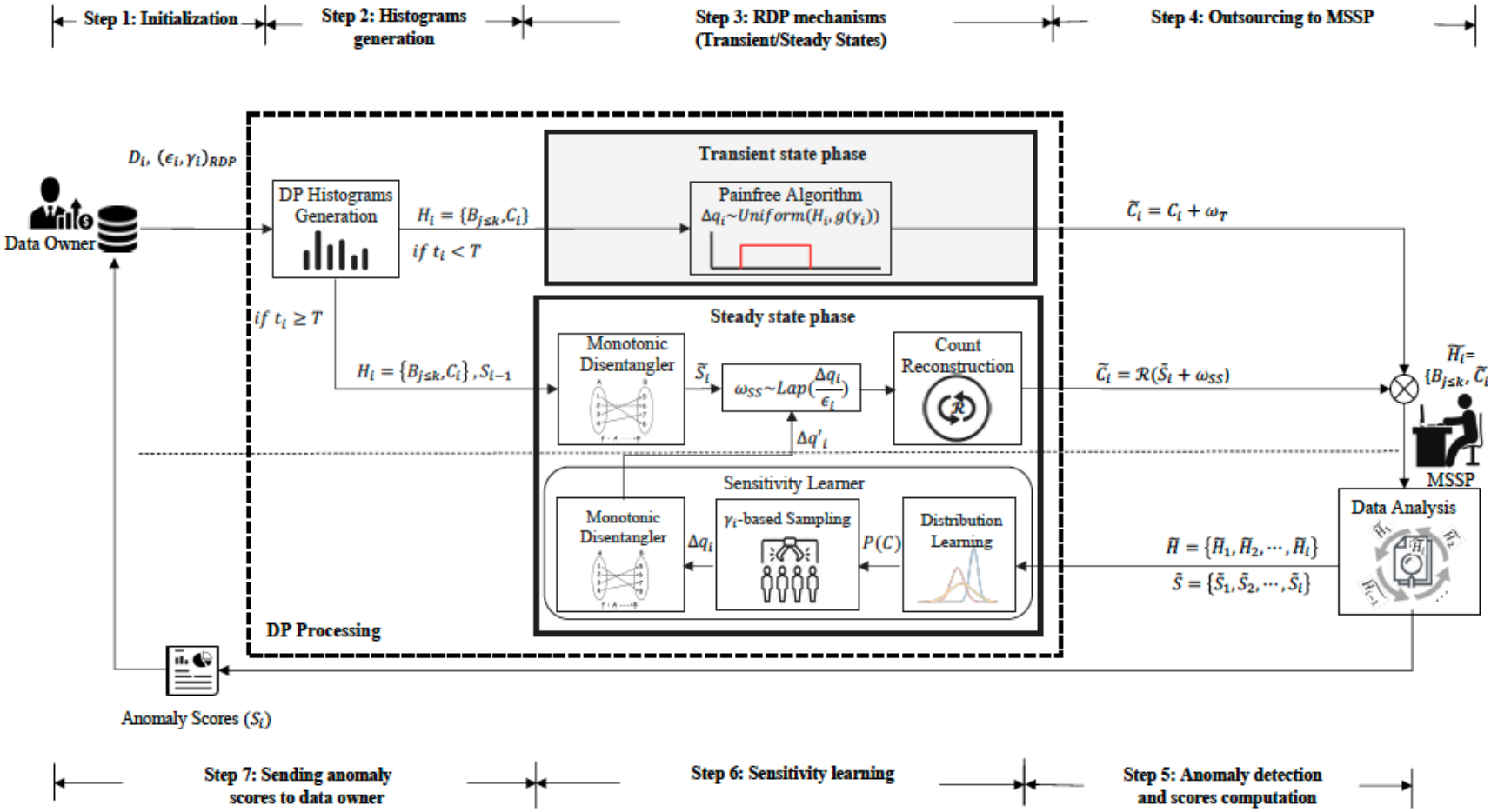
Predict

Optimal multiplicative factor

Predicted (*a priori*) state estimate $\hat{\mathbf{x}}_{k|k-1} = \mathbf{F}_k \hat{\mathbf{x}}_{k-1|k-1}$

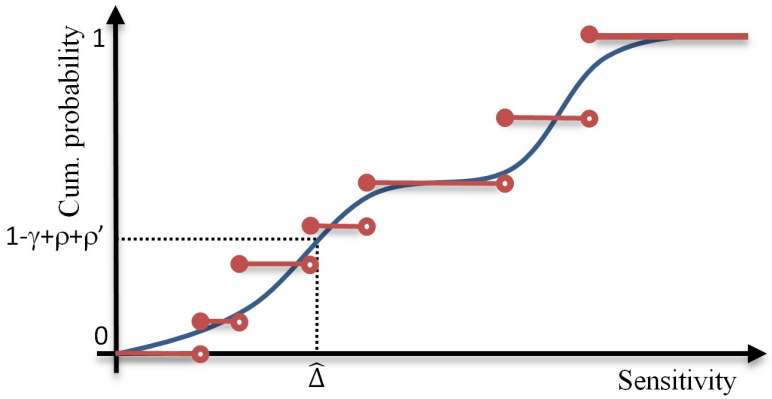
Outsource: is to **enforce** the estimated sensitivity value (way smaller and hence less noise to be added)

DPOAD Overview (System Design)



Transient State Phase

Pain-free Algorithm [ICML'17]:



Algorithm 1 SENSITIVITYSAMPLER

Input: database size n, target mapping $f : \mathcal{D}^n \to \mathcal{B}$, sample size m, order statistic index k, distribution P

for $i = 1$ **to** m **do**

Sample $D \sim P^{n+1}$

Set $G_i = \|f(D_{1\ldots n}) - f(D_{1\ldots n-1,n+1})\|_{\mathcal{B}}$

end for

Sort $G_1, \ldots, G_m$ as $G_{(1)} \leq \ldots \leq G_{(m)}$

return $\hat{\Delta} = G_{(k)}$

Laplace mechanism with the sensitivity computed using Algorithm 1 is $\epsilon(m,k), \gamma(m,k)$-RDP.

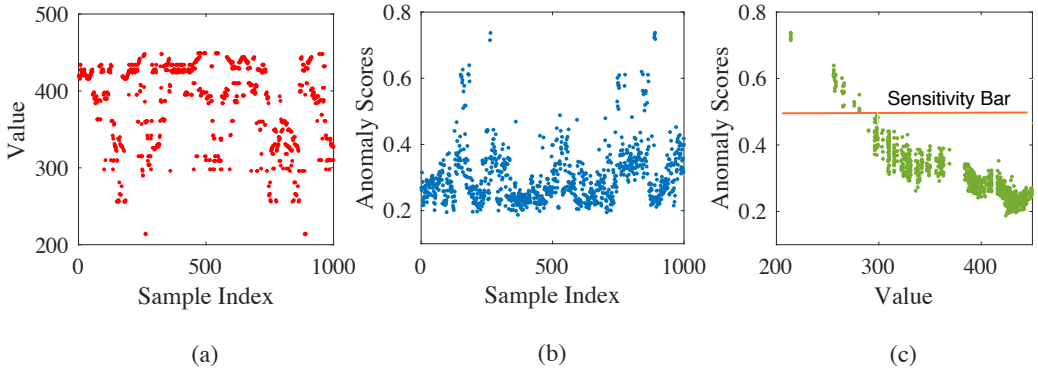
Steady State Phase

Elements of effective differentially private sensitivity learning:

1. Monotonic Disentangler:

The process of **mapping** the *data* to *their anomaly scores* to output a *monotonic* version in terms of outlierness.

The process of **scaling** the anomaly scores back to the histogram count to preserve accuracy.



Steady State Phase

Elements of effective differentially private sensitivity learning:

2. PDF Learning: The process of approximating the PDF of dataset using the noisy anomaly scores

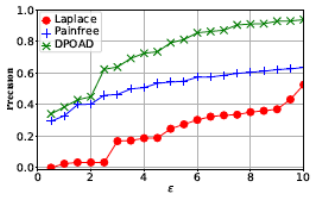
Theorem 3. *There is a computationally efficient* $(\epsilon, 0)$*-differentially private* (α, β)*-learning algorithm for* $\mathcal{C}_N$ *that uses* $n = O((N + \log(1/\beta))/\alpha^2 + N\log(1/\beta)/(\epsilon\alpha))$ *samples.*

[NIPS'15]

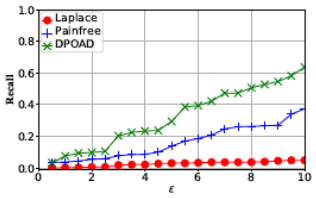
Experiments

Databset	Size	# of Attributes
IoT	-	12 events + 10 sensors
Parking	35,718	4
Electric consumption	2M	9
Breast cancer	286	9
Credit card	30,000	23
KDD	494,021	42

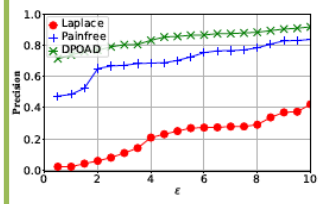
IOT Data



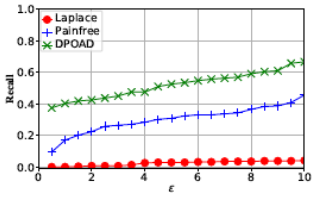
(a) Precision vs ϵ



(b) Recall vs ϵ



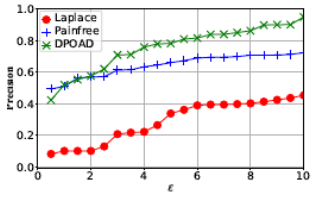
(c) Precision vs ϵ



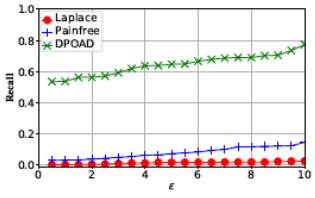
(d) Recall vs ϵ

Electric Consumption Data

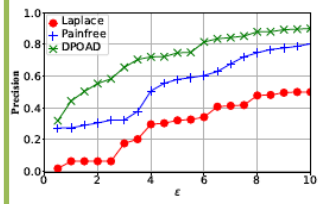
Credit Card Data



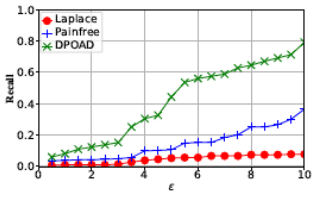
(e) Precision vs ϵ



(f) Recall vs ϵ



(g) Precision vs ϵ



(h) Recall vs ϵ

Breast Cancer Data

Conclusion

1) DPOAD provides the first practical differentially private anomaly detection in outsourcing setting.

2) We formally benchmark DPOAD under the Laplace mechanism for network, IoT and credit card anomaly (fraudulent) detection.

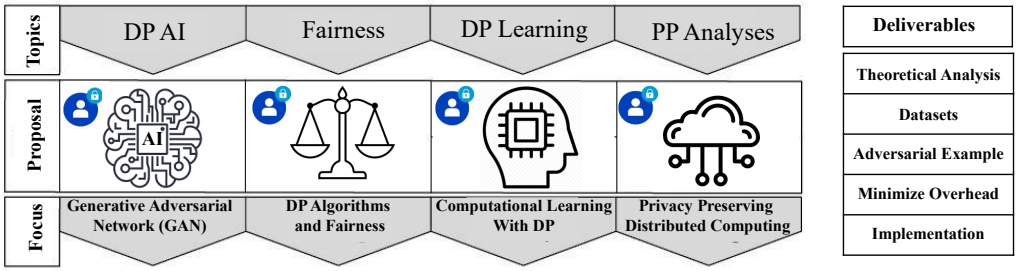
3) Our experimental results demonstrate that DPOAD significantly improves the accuracy of the anomaly detection compared to the baselines.

40

Outline

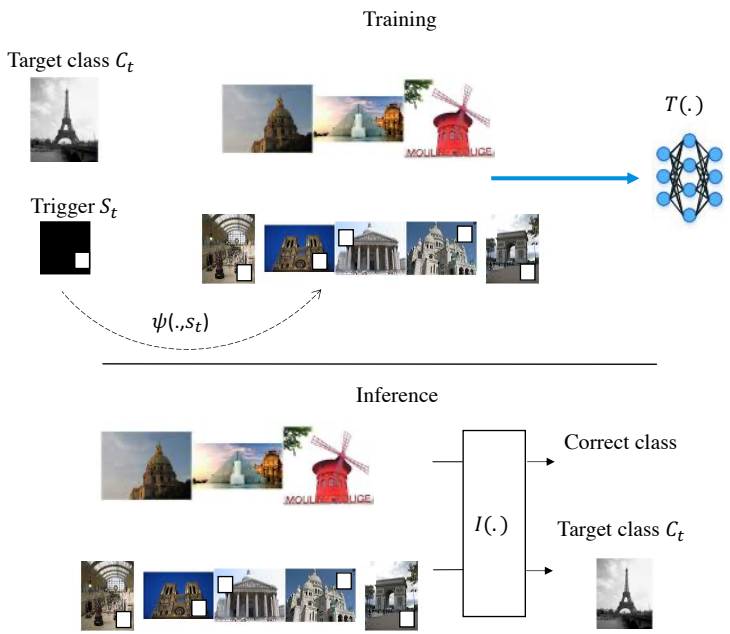
Future Research Directions

- Privacy preserving, fair and accountable algorithms
  - Tools:
  - Deep learning
  - Computational learning theory
  - Cybersecurity
  - Applications:
  - Health data monitoring and analysis
  - Cloud computing
  - Safe networking
- Secure distributed computation for IoT and cyber-physical systems
  - Federated Learning
  - Hybrid models like Secure multiparty computation (SMC) + DP

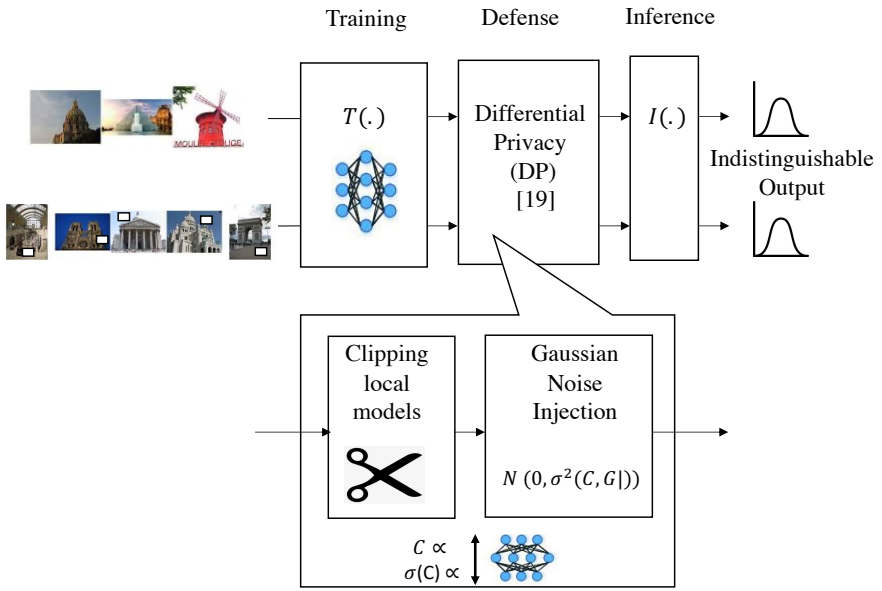


Example: DP for FL Security

Attack Overview



System Overview



Publication: Kane Walter, Meisam Mohammady, Surya Nepal, Salil Kanhere. Optimally Closing the backdoor in Federated Learning According to the Model Size. Resubmitted to Privacy Enhancing Technologies (Submitted to TDSC'22).

Concluding Remarks

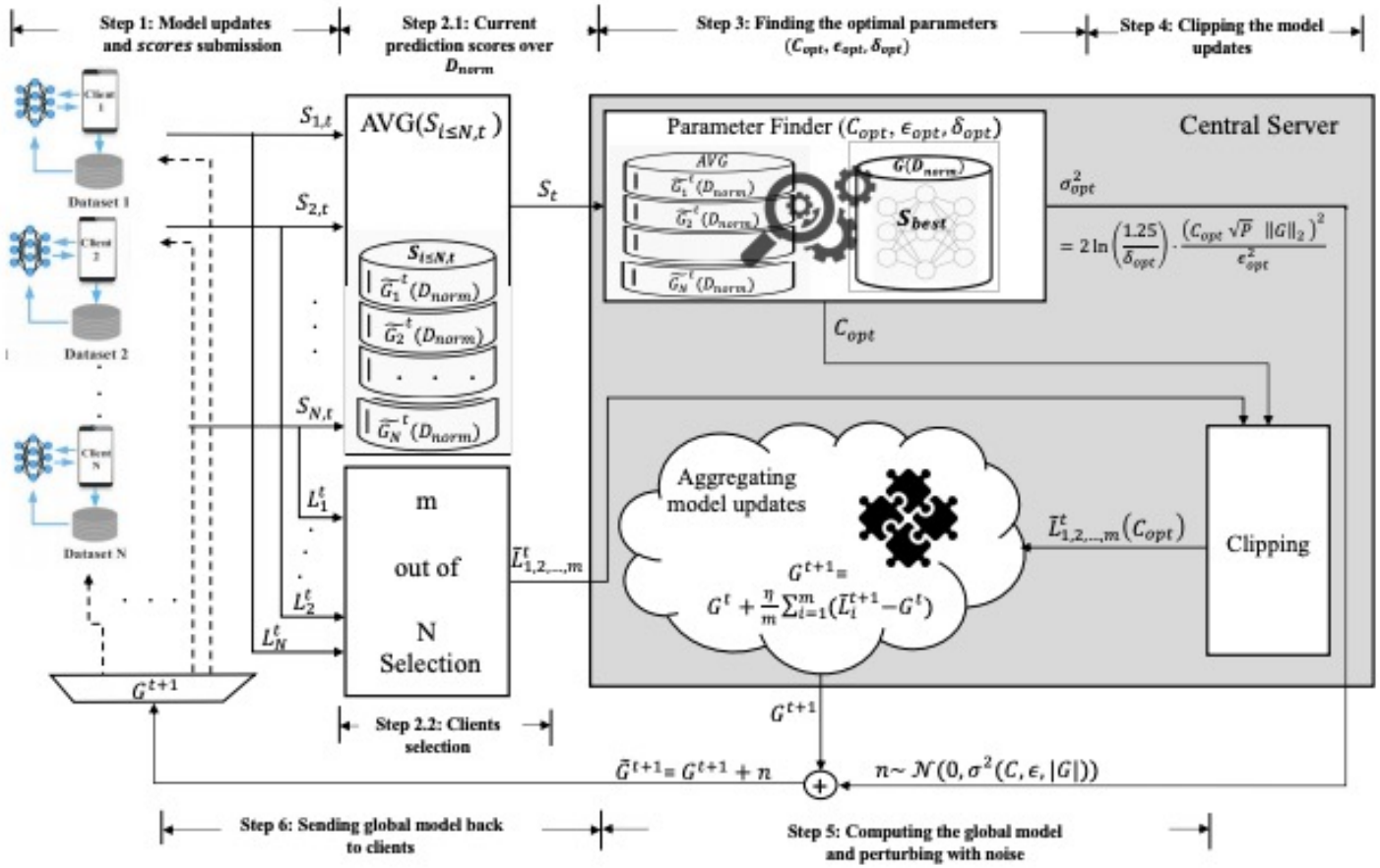
- Privacy Enhancing Technologies with Optimal Utility
- Ongoing and Future Research
- Results of Active Collaboration:

Email: meisam.mohammady@csiro.au

Thank you

Backups

Application of DP in Secure Federated Learning



Indistinguishability Analysis

- The statement inside the probability is the adversary's decision on a view, declaring it as a fake view or a real view candidate, using his/her α −knowledge

- Moreover, we note that generated views differ only in their IP values (fp-QI attributes are similar for all the views)

- Hence, the adversary's decision can only be based on the published set of IPs in each view through comparing shared prefixes among those IP addresses which he/she already know (α)

$$e^{-\epsilon} \leq \frac{\Pr(\text{view i may be the real view})}{\Pr(\text{view r may be the real view})} \leq e^{\epsilon}$$

Migration Technique

DEFINITION. ***Migration Function***: *Let $\mathcal{S}$ be a set of IP addresses consisting of d groups of IPs $S_1, S_2, \cdots, S_d$ with distinct prefixes $s_1, s_2, \cdots, s_d$ respectively, and K be a random CryptoPAn key. Migration function $M : \mathcal{S} \times C(\text{set of positive integers}) \rightarrow \mathcal{S}^*$ is defined as*

$$\mathcal{S}^* = M(\mathcal{S}) = \{S_i^* | \forall i \in \{1, 2, \cdots, d\}\}$$

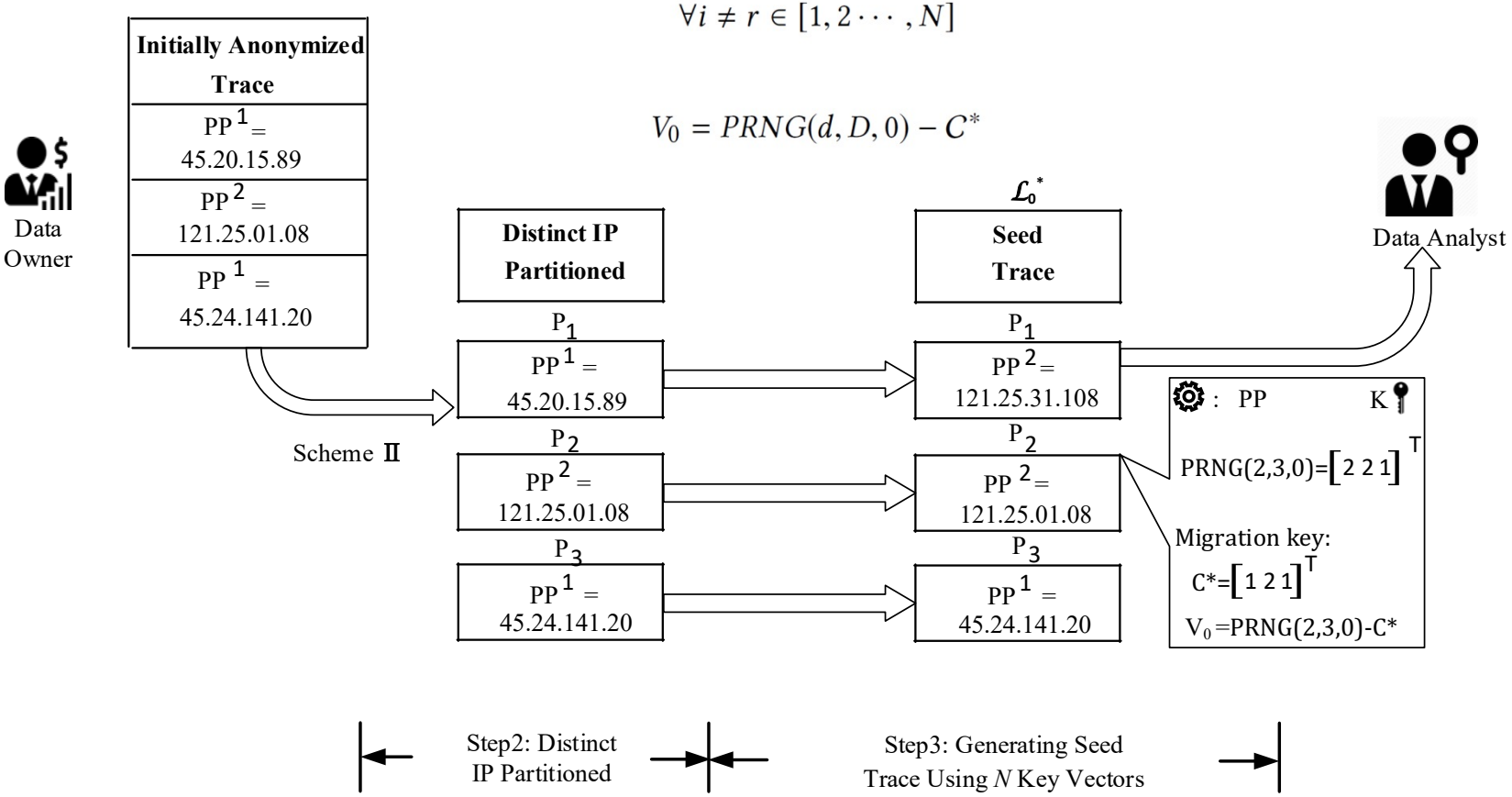
$$\text{where } S_i^* = \{PP^{c_i}(s_i \oplus a_j, K), \forall a_j \in S_i\}$$

where $C = PRNG(d, d) = \{c_1, c_2, \cdots, c_d\}$ is the set of d non-repeating random key indices generated between $[1, d]$ using a cryptographically secure pseudo random number generator.

Scheme II: Key Generation

$$V_i = PRNG(d, D, i) - PRNG(d, D, i-1),$$
$$\forall i \neq r \in [1, 2 \cdots, N]$$

$$V_0 = PRNG(d, D, 0) - C^*$$



Background (Utility Metrics)

- **Usefulness.** Database mechanism $M_{q,f}$ is (γ, ζ)-useful if with probability $1 - \zeta$, for every database $d \subseteq D, \left|M_q(d) - q(d)\right| \leq \gamma$.

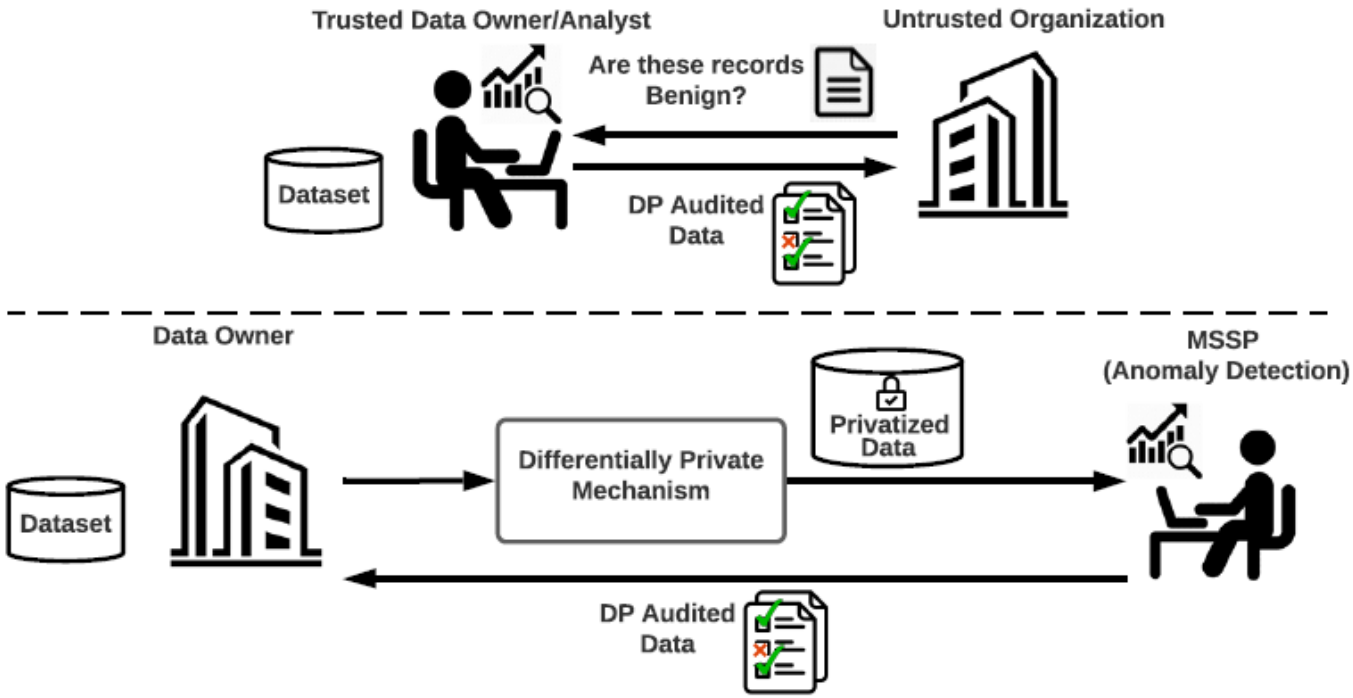
- **Entropy.** Let x be a random variable on $\mathbb{R}$ with PDF $f(x)$. The entropy of $x$is defined as $H(x) = -\int_0^\infty f(x)\, log\big(f(x)\big)$.

Related Works

- Several existing works
  - Smooth sensitivity [ECML PKDD 15]
  - Anomaly Exclusion [AISec@CCS 16]
  - ML approaches (Stochastic Gradient Descent (SGD)) [CSCML 18]
  - Anomaly Exclusion [CCS' 19]

- Drawbacks of existing solutions
  - Most such approaches assume the data-owners are able to run anomaly detection by themselves (Publishing framework vs Outsourcing) which is not the case in our motivation examples
  - Outlier detection and privacy protection are intrinsically conflicting tasks. This seemingly impossible problem has not been properly addressed.
    - In case of outsourcing, since the analysis is done on DP-results this contradiction is even more challenging
  - Analysis using global sensitivity of outlier counts makes the outputs too noisy

DPOD vs. Existing Works



Pain free RDP functions

Budgeted	Optimise	ρ	γ	m	k
$\gamma \in (0,1)$	m	$\exp\left(W_{-1}\left(-\frac{\gamma}{2\sqrt{e}}\right)+\frac{1}{2}\right)$	$\bullet$	$\left\lceil \frac{\log\left(\frac{1}{\rho}\right)}{2(\gamma-\rho)^2} \right\rceil$	$\left\lceil m\left(1-\gamma+\rho+\sqrt{\frac{\log\left(\frac{1}{\rho}\right)}{2m}}\right)\right\rceil$
$m \in \mathbb{N}, \gamma$	k	$\exp\left(\frac{1}{2}W_{-1}\left(-\frac{1}{4m}\right)\right)$	$\geq \rho+\sqrt{\frac{\log\left(\frac{1}{\rho}\right)}{2m}}$	$\bullet$	$\left\lceil m\left(1-\gamma+\rho+\sqrt{\frac{\log\left(\frac{1}{\rho}\right)}{2m}}\right)\right\rceil$
$m \in \mathbb{N}$	γ	$\exp\left(\frac{1}{2}W_{-1}\left(-\frac{1}{4m}\right)\right)$	$\rho+\sqrt{\frac{\log\left(\frac{1}{\rho}\right)}{2m}}$	$\bullet$	m