# Poster: Detecting Anomalous LAN Activities under Differential Privacy

Norrathep Rattanavipanon[1], Donlapark Ponnoprat[2], Hideya Ochiai[3], Kuljaree Tantayakul[1], Touchai Angchuan[4], and Sinchai Kamolphiwong[4]

[1]College of Computing, Prince of Songkla University, Phuket, 83120, Thailand
{norrathep.r, kuljaree.t}@phuket.psu.ac.th
[2]Data Science Research Center, Department of Statistics, Faculty of Science, Chiang Mai University, Chiang Mai, 50200, Thailand
donlapark.p@cmu.ac.th
[3]Graduate School of Information Science and Technology, The University of Tokyo, Tokyo, 113-8656, Japan
ochiai@elab.ic.i.u-tokyo.ac.jp
[4]Faculty of Engineering, Prince of Songkla University, Songkhla, 90110, Thailand
{touch, ksinchai}@coe.psu.ac.th

**Abstract**

Anomaly detection has emerged as a popular technique for detecting malicious activities in local area networks (LANs). Various aspects of LAN anomaly detection have been widely studied. Nonetheless, the privacy concern about individual users or their relationship in LAN has not been thoroughly explored in the prior work. In some realistic cases, the anomaly detection analysis needs to be carried out by an external party, located outside the LAN. Thus, it is important for the LAN admin to release LAN data to this party in a private way in order to protect privacy of LAN users; at the same time, the released data must also preserve the utility of being able to detect anomalies. This paper investigates the possibility of privately releasing ARP data that can later be used to identify anomalies in LAN. We present four approaches, namely naïve, histogram-based, naïve-$\delta$ and histogram-based-$\delta$, and show that they satisfy different levels of differential privacy – a rigorous and provable notion for quantifying privacy loss in a system. Our real-world experimental results confirm practical feasibility of our approaches. With a proper privacy budget, all of our approaches preserve more than 75% utility of detecting anomalies in the released data.

## Bibliographic Reference

Norrathep Rattanavipanon, Donlapark Ponnoprat, Hideya Ochiai, Kuljaree Tantayakul, Touchai Angchuan, and Sinchai Kamolphiwong. "Detecting Anomalous LAN Activities under Differential Privacy." Security and Communication Networks (2022), https://doi.org/10.1155/2022/1403200.

# Detecting Anomalous LAN Activities under Differential Privacy


Norrathep Rattanavipanon[1], Donlapark Ponnoprat[2], Hideya Ochiai[3], Kuljaree Tantayakul[1], Touchai Angchuan[1], Sinchai Kamolphiwong[1]

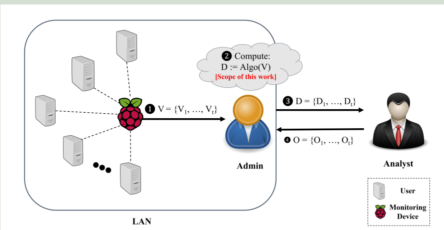
[1] *Prince of Songkla University*, [2] *Chiang Mai University*, [3] *The University of Tokyo*

*norrathep.r@phuket.psu.ac.th*


## Motivation

- No prior work so far has explored the privacy implication of performing LAN anomaly detection, especially in many realistic scenarios where the anomaly detection must be performed by an entity outside LAN or third-party software.
- Simply erasing all users' sensitive information from the output data helps with user anonymization. But it does not provide strong and provable privacy guarantees.
- With this simple technique, a motivated adversary may still be able to deanonymize users through other means, e.g., performing a side-channel analysis or correlating the remaining network traces with the physical world data.
- Hence, there is a need for a technique with *rigorous privacy guarantees, while preserving the utility of detecting anomalies* in the LAN environment.

## System Model




## Privacy Notions

We choose the LAN privacy notions based on differential privacy – a rigorous and provable notion for quantifying privacy loss in a system. The notions use the ARP-Request degree as their underlying data since it has been shown to be a promising feature for detecting anomalies in LAN.

Let $\mathcal{G}$ be the set of graphs between LAN users. We can define 4 privacy notions in LAN anomaly detection as follows:

**Notion 1** ($(\epsilon,\delta)$-edge-DP)**.** An algorithm $\mathcal{M}:\mathcal{G}\to\mathcal{Y}$ satisfies $(\epsilon,\delta)$-edge-DP if, for every pair of edge-neighboring graphs $G$ and $G'$ and every subset $S\subseteq\mathcal{Y}$,

$$\mathbb{P}\left(\mathcal{M}(G)\in S\right)\le e^{\epsilon}\mathbb{P}\left(\mathcal{M}(G')\in S\right)+\delta.$$

**Notion 2** ($(\epsilon,\delta)$-node-DP)**.** $\mathcal{M}:\mathcal{G}\to\mathcal{Y}$ satisfies $(\epsilon,\delta)$-node-DP if, for every pair of node-neighboring graphs $G$ and $G'$ and every subset $S\subseteq\mathcal{Y}$,

$$\mathbb{P}\left(\mathcal{M}(G)\in S\right)\le e^{\epsilon}\mathbb{P}\left(\mathcal{M}(G')\in S\right)+\delta.$$

**Notion 3** ($\epsilon$-edge-DP)**.** $\mathcal{M}:\mathcal{G}\to\mathcal{Y}$ satisfies $(\epsilon,\delta)$-edge-DP if it satisfies $(\epsilon,0)$-edge-DP.

**Notion 4** ($\epsilon$-node-DP)**.** $\mathcal{M}:\mathcal{G}\to\mathcal{Y}$ satisfies $(\epsilon,\delta)$-node-DP if it satisfies $(\epsilon,0)$-node-DP.

## Protection Guarantees

| Notion | # | Protected Info. | Prob. |
|---|---|---|---|
| $(\epsilon,\delta)$-edge-DP | 1 | ARP requests | 1-$\delta$ |
| $(\epsilon,\delta)$-node-DP | 2 | LAN users | 1-$\delta$ |
| $\epsilon$-edge-DP | 3 | ARP requests | 1 |
| $\epsilon$-node-DP | 4 | LAN users | 1 |

## Privacy-preserving algorithms satisfying each notion

Let $V_{jk}$ denote aggregate ARP requests from user $k$, accumulated at interval $j$ and $V_j$ denote the result after appending all ARP requests of all users generated in interval $j$. Algorithms satisfying four privacy notions are shown below:

**Input:** $V=\{V_1,V_2,...,V_t\}$, $t$, $\epsilon$
**Output:** $D=\{D_1,D_2,...,D_t\}$
**for** $j=1$ **to** $t$ **do**
  $D_j\leftarrow\textsc{Sum}(\textsc{Degree}(V_j))$
  $D_j\leftarrow D_j+\text{Laplace}(t/\epsilon)$
  **if** $D_j>0$ **then** $D_j\leftarrow\text{int}(D_j)$
  **else** $D_j\leftarrow 0$
**end**

**Algo 3:** Naïve approach satisfying **Notion-1**

**Input:** $V=\{V_1,V_2,...,V_t\}$, $t$, $\epsilon$, $\delta$
**Output:** $D=\{D_1,D_2,...,D_t\}$
$\rho\leftarrow\left(\sqrt{\log(1/\delta)+\epsilon}-\sqrt{\log(1/\delta)}\right)^2$
**for** $j=1$ **to** $t$ **do**
  $D_j\leftarrow\textsc{Sum}(\textsc{Degree}(V_j))$
  $D_j\leftarrow D_j+N(0,t/2\rho)$
  **if** $D_j>0$ **then** $D_j\leftarrow\text{int}(D_j)$
  **else** $D_j\leftarrow 0$
**end**

**Algo 4:** Naïve-$\delta$ approach satisfying **Notion-2**

---

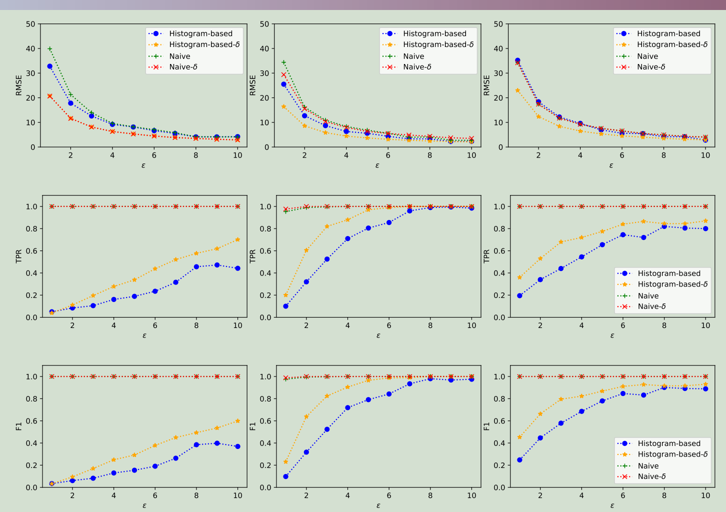
**Input:** $V=\{V_1,V_2,...,V_t\}$, $t$, $\epsilon$
**Output:** $D=\{D_1,D_2,...,D_t\}$
**for** $j=1$ **to** $t$ **do**
  $D_j\leftarrow\textsc{Histogram}(\textsc{Degree}(V_j))$
  **foreach** $bin\in D_j$ **do**
    $bin.count\leftarrow$
      $bin.count+\text{Laplace}(t/\epsilon)$
    **if** $bin.count>0$ **then**
      $bin.count\leftarrow\text{int}(bin.count)$
    **else** $bin.count\leftarrow 0$
  **end**
**end**

**Algo 3:** Histogram-based approach satisfying **Notion-3**

**Input:** $V=\{V_1,V_2,...,V_t\}$, $t$, $\epsilon$, $\delta$
**Output:** $D=\{D_1,D_2,...,D_t\}$
$\rho\leftarrow\left(\sqrt{\log(1/\delta)+\epsilon}-\sqrt{\log(1/\delta)}\right)^2$
**for** $j=1$ **to** $t$ **do**
  $D_j\leftarrow\textsc{Histogram}(\textsc{Degree}(V_j))$
  **foreach** $bin\in D_j$ **do**
    $bin.count\leftarrow$
      $bin.count+N(0,t/2\rho)$
    **if** $bin.count>0$ **then**
      $bin.count\leftarrow\text{int}(bin.count)$
    **else** $bin.count\leftarrow 0$
  **end**
**end**

**Algo 4:** Histogram-based-$\delta$ approach satisfying **Notion-4**

## Utility Evaluation




(a): LAN-1: Tokyo (b): LAN-2: Phuket (c): LAN-3: Hatyai

## Acknowledgements


The ASEAN IVO (http://www.nict.go.jp/en/asean_ivo/index.html) project, ASEAN-Wide Cyber-Security Research Testbed Project, was involved in the production of the contents of this work and financially supported by NICT (http://www.nict.go.jp/en/index.html). This work is also financially supported by Chiang Mai University, Thailand.


**Bibliographic Reference:**


Norrathep Rattanavipanon, Donlapark Ponnoprat, Hideya Ochiai, Kuljaree Tantayakul, Touchai Angchuan, and Sinchai Kamolphiwong. "Detecting Anomalous LAN Activities under Differential Privacy." Security and Communication Networks (2022), https://doi.org/10.1155/2022/1403200.