Anonymization of Network Traces Using Noise Addition Techniques

By Ahmed AlEroud

Assistant Professor of Computer Information Systems
Yarmouk University, Jordan

Post-doctoral Fellow, Department of Information Systems,
University of Maryland

Agenda

- Data Sharing and Traffic Anonymization
- The Challenge of Anonymizing Network Data
- Objectives
- Sensitive Network Attributes
- Existing Anonymization Techniques
- Noise Addition Techniques
- Experiments and Results
- Conclusions and Future work

Data Sharing: Trace Anonymization

- Why share network data?
  - Collaborative attack detection
  - Advancement of network research
- Any problems with sharing network data?
  - Expose sensitive information
  - Packet header: IP address, service port exposure
  - Packet content: more serious
  - Sharing network trace logs may reveal the network architecture, user identity, and user information
- Solution: anonymization of trace data
  - preserve IP prefix, and change packet content

The Challenge of Anonymizing Network Data

Is it possible to create a technique that detects network threats using shared data with minimal privacy violation?

- In order to answer this question, some sub-questions need to be formulated
  - Which sensitive information is present in network protocols?
  - To what extent will anonymization techniques influence the accuracy of a threat detection system?

Sensitive Network Attributes

Field	Attacks
IP	Adversaries try to identify the mapping of IP addresses in the anonymized dataset to reveal the hosts and the network.
MAC	May be used to uniquely identify an end device. MAC addresses combined with external databases are mappable to device serial numbers and to the organizations or individuals who purchased the devices.
Time-stamps	Time-stamps may be used in trace injection attacks that uses known information about a set of trace generated or otherwise known by an attacker to recover mappings of anonymized fields.
Port Numbers	These fields partially identify the applications that generated the trace in a given trace. This information may be used in fingerprinting attacks to reveal that a certain application with suspected vulnerabilities is running on a network where the trace is collected from.
Counter Anonymization	Counters (such as packet and octet volumes per flow) are subject to fingerprinting and injection attacks.

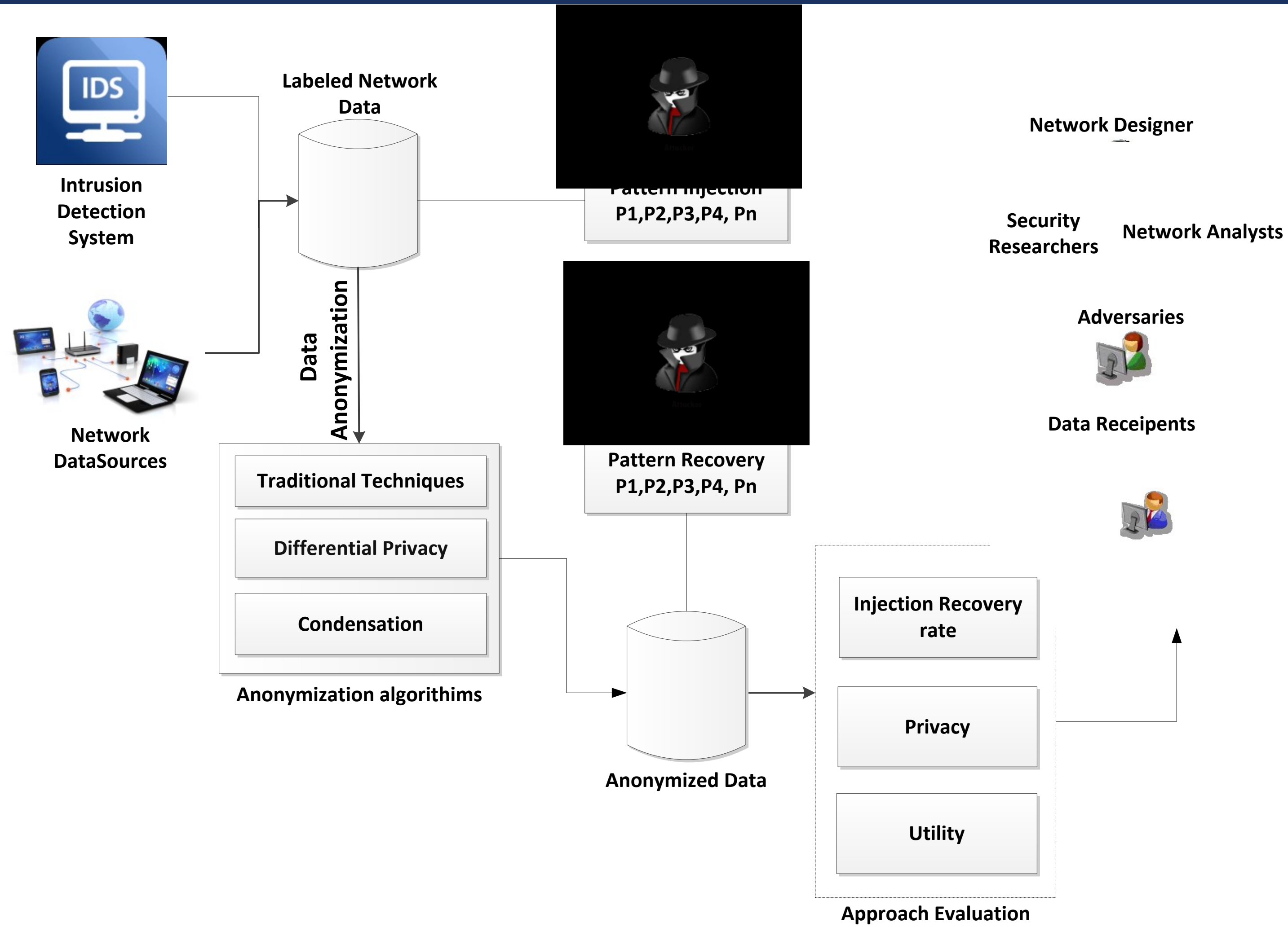
Existing Anonymization Techniques

- *Blackmarking (BM)*
  - Blindly replaces all IP addresses in a trace with a single constant value
- *Truncation (TR{t})*
  - Replaces the t least significant bits of an IP address with 0s
- *Permutation (RP)*
  - Transforms IP addresses using a random permutation (not consistent across IP addresses)
- *Pprefix-preserving permutation (PPP{p})*
  - Permutes the host and network part of IP addresses independently (consistent across IP addresses)

Objectives

- Implement anonymization model for network data, that is strong enough and provides privacy guarantee when sharing network data
- Test various attacking strategies including injection attacks on data anonymized
  - Verify that the approach is more robust guarding against different types of attacks including Fingerprinting attacks on network data

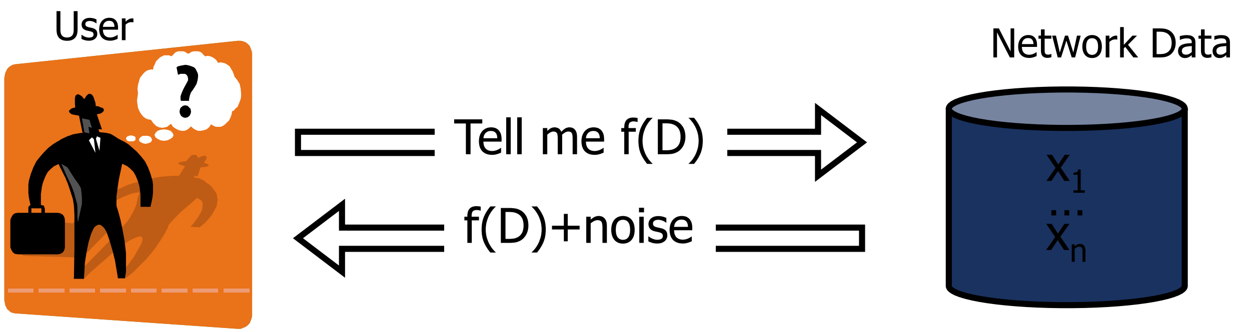
Proposed Solution and Methodology



Differential Privacy

- A privacy model that provides strong privacy guarantee (regardless of what attackers know)
- It works on aggregated values and prevents attackers from inferring the existence of an individual record from the aggregated values (e.g., sum of packet counts)
- The key idea is to add large enough noise (following a specific distribution called Laplace or double exponential) to hide the impact of a single network trace

One Primitive to Satisfy Differential Privacy: Add Noise to Output



- Intuition: f(D) can be released accurately when f is insensitive to individual entries x1, … xn
- Noise generated from Laplace distribution

Differential Privacy Example

Original Data

Packet Size
1024
1234
10240
3333
3456
12340

Average Packet size = 5271

Differential Privacy
(add a noise to average)

Average Packet size =
5271+noise
= 6373

New Data

Packet Size
1024
1234
10240
3333
3456
12340
15000

Average Packet size = 6661

Average Packet size =
6661+noise
= 6175

- Without noise: If the attacker knows the average packet size before the new packet is added, it is easy to figure out the packet's size from the new average.
- With noise: One cannot infer whether the new packet is there.

Differential-Private Anonymization

Original Data

Packet Size
1024
1234
10240
3333
3456
12340

Partitioning into equal-sized Clusters

Packet Size
1024
1234
10240
3333
3456
12340

Compute mean of each column
within each cluster, then
add Laplace noise to the mean
and replace every value with perturbed mean

Packet Size
1099
1099
12221
3217
3217
12221

- The noise added follows Laplace distribution with mean zero and standard deviation = sensitivity / ε.
- Sensitivity = (max value in cluster – min value in cluster) / cluster size
- The larger the cluster size, the smaller the noise
- This method works better for large volume of data

Condensation-based Anonymization of Network Data

- Implemented an algorithm with better utility-privacy tradeoff than existing methods*
- The algorithm consists of two steps:
  - Prefix-preserving clustering and permutation of IP addresses
  - Condensation based anonymization of all other attributes (to prevent injection attacks)

* Ahmed Aleroud, Zhiyuan Chen and George Karabatis. "Network Trace Anonymization Using a Prefix- Preserving Condensation-based Technique". *International Symposium on Secure Virtual Infrastructures: Cloud and Trusted Computing 2016*

IP Anonymization Example

Original IP	Permutation	Clustering	Anonymized IP	
SRC_IP	SRC_IP	SRC_IP	SRC_IP	
10.50.50.12	210.70.70.12	210.70.70.12	210.70.70.17	c1
10.200.21.122	210.160.71.122	210.46.46.20	210.46.46.17	
10.200.21.174	210.160.71.174	210.46.70.20	210.46.70.17	
10.60.60.20	210.46.46.20	210.160.71.122	210.160.71.143	c2
10.200.21.133	210.160.71.133	210.160.71.174	210.160.71.143	
10.60.50.20	210.46.70.20	210.160.71.133	210.160.71.143	

Attributes Anonymized

- The features (attributes) used in network trace data that need to be anonymized and those that are important for intrusion detection are:
  - IP addresses
  - Time-stamps
  - Port Numbers
  - Trace Counters

Experimental Datasets of Network data

Experiments are conducted on

- PREDICT dataset: Protected Repository for the Defense of Infrastructure Against Cyber Threats
- University of Twente dataset: A flow-based dataset containing only attacks
- Since PREDICT mostly has normal flow and Twente mostly has attack flows, we draw a random sample from each and combine them
- The combined data sets:
  - Dataset 1: 70% PREDICT dataset + 30% Twente dataset
  - Dataset 2: 50% PREDICT dataset + 50% Twente dataset
- Metrics:
  - Utility: ROC curve, TP, FP, Precision, Recall, F-measure
  - Average privacy: $2^{h(A|B)}$ where A is original data, B is anonymized, h is conditional entropy (higher is better)

Dataset 1 Experiment: KNN Classification on Anonymized Data

Dataset 1 (70%-30%)

419,666 Total # records

Training set:

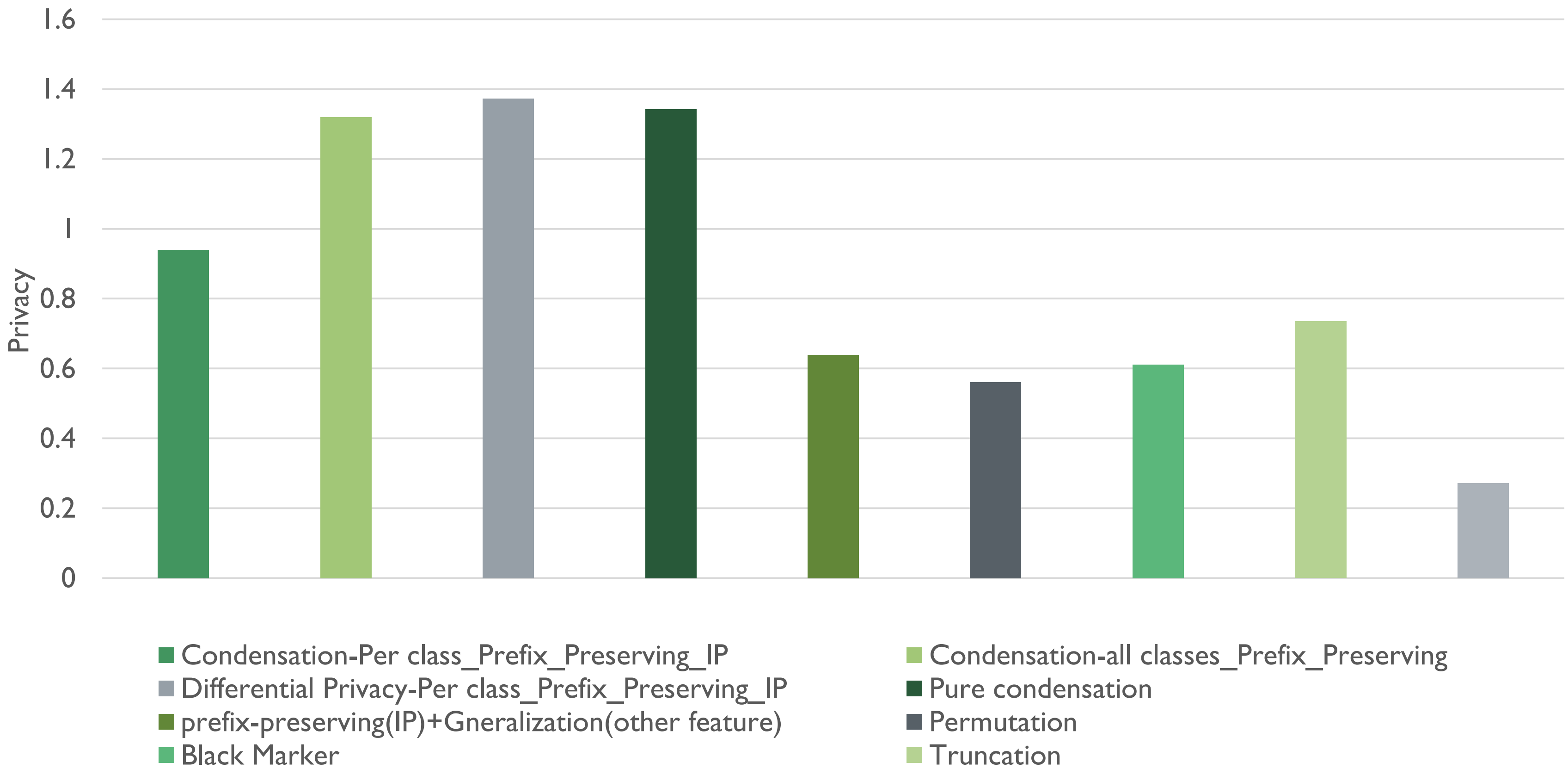
- 177,028 Normal records
- 116,738 Attack records
- 293,766 Total records

Test set:

- 75,862 Normal records
- 50,038 Attack records
- 125,900 Total records

	TP Rate	FP Rate	P	R	F-Measure	ROC Area	Class
Original	0.98	0.013	0.981	0.98	0.98	0.984	Attack
	0.987	0.02	0.987	0.987	0.987	0.984	Normal
	0.984	0.017	0.984	0.984	0.984	0.984	Avg
Condensation-Per class_Prefix_Preserving_IP	0.941	0.059	0.961	0.941	0.951	0.941	Attack
	0.941	0.059	0.913	0.941	0.927	0.941	Normal
	0.941	0.059	0.942	0.941	0.941	0.941	Avg
Condensation-all classes_Prefix_Preserving_IP	0.628	0.582	0.62	0.628	0.624	0.523	Attack
	0.418	0.372	0.426	0.418	0.422	0.523	Normal
	0.545	0.498	0.543	0.545	0.544	0.523	Avg
Differential Privacy-Per class_Prefix_Preserving_IP	0.941	0.059	0.96	0.941	0.95	0.94	Attack
	0.941	0.059	0.913	0.941	0.927	0.94	Normal
	0.941	0.059	0.941	0.941	0.941	0.94	Avg
Pure condensation	0.691	0.612	0.631	0.691	0.66	0.54	Attack
	0.388	0.309	0.454	0.388	0.418	0.54	Normal
	0.571	0.491	0.56	0.571	0.564	0.54	Avg
prefix-preserving(IP)+ Generalization(other feature)	1	1	0.602	1	0.752	0.5	Attack
	0	0	0	0	0	0.5	Normal
	0.602	0.602	0.362	0.602	0.452	0.5	Avg
Permutation	0.999	1	0.602	0.999	0.751	0.5	Attack
	0	0.001	0.048	0	0	0.5	Normal
	0.602	0.602	0.381	0.602	0.452	0.5	Avg
Black Marker	1	1	0.602	1	0.752	0.5	Attack
	0	0	0	0	0	0.5	Normal
	0.602	0.602	0.362	0.602	0.452	0.5	Avg
Truncation	0.983	0.999	0.598	0.983	0.744	0.196	Attack
	0.001	0.017	0.034	0.001	0.002	0.196	Normal
	0.592	0.608	0.374	0.592	0.448	0.196	Avg
Reverse Truncation	0.082	0.163	0.432	0.082	0.137	0.46	Attack
	0.837	0.918	0.376	0.837	0.519	0.46	Normal
	0.382	0.463	0.41	0.382	0.289	0.46	Avg

Dataset 1 Privacy Results



Dataset 2 Experiment: KNN Classification on Anonymized Data

Dataset 2 (50%-50%)

278,067 Total # of records

Training set:

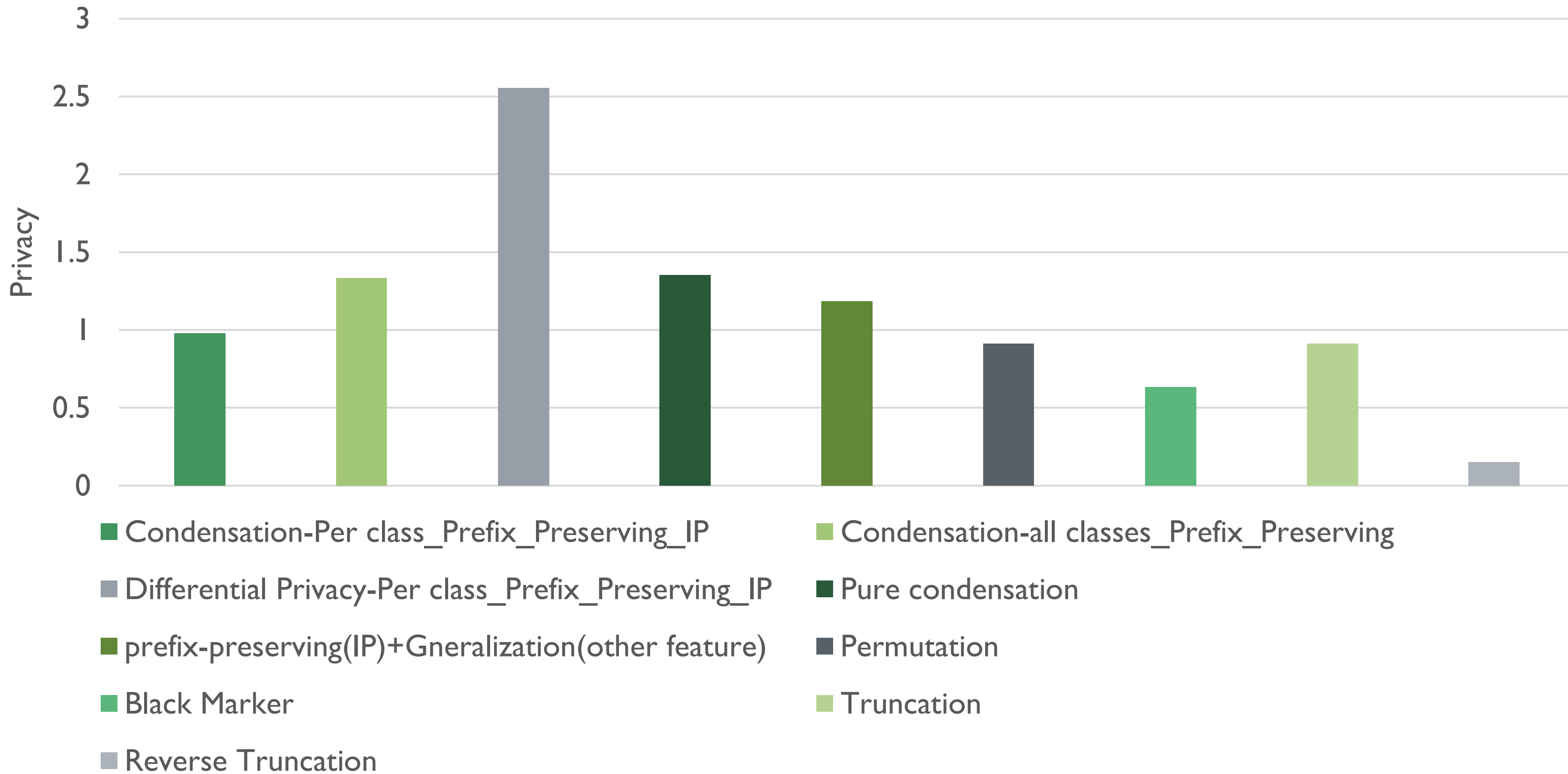
- 81,386 Normal records
- 113,260 Attack records
- 194,646 Total records

Test set:

- 35,153 Normal records
- 48,268 Attack records
- 83,421 Total records

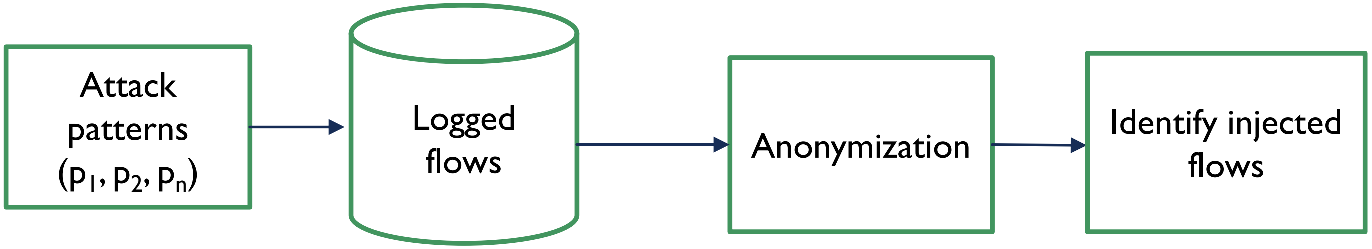
	TP Rate	FP Rate	P	R	F-Measure	ROC Area	Class
Original	0.991	0.013	0.991	0.991	0.991	0.989	Attack
	0.987	0.009	0.987	0.987	0.987	0.989	Normal
	0.989	0.011	0.989	0.989	0.989	0.989	Avg
Condensation-Per class_Prefix_Preserving_IP	0.954	0.118	0.917	0.954	0.935	0.918	Attack
	0.882	0.046	0.934	0.882	0.907	0.918	Normal
	0.924	0.088	0.924	0.924	0.923	0.918	Avg
Condensation-all classes_Prefix_Preserving_IP	0.553	0.562	0.575	0.553	0.564	0.495	Attack
	0.438	0.447	0.416	0.438	0.427	0.495	Normal
	0.504	0.514	0.508	0.504	0.506	0.495	Avg
Differential Privacy-Per class_Prefix_Preserving_IP	0.975	0.125	0.915	0.975	0.944	0.945	Attack
	0.875	0.025	0.962	0.875	0.916	0.945	Normal
	0.933	0.083	0.935	0.933	0.932	0.945	Avg
Pure condensation	0.662	0.597	0.603	0.662	0.631	0.532	Attack
	0.403	0.338	0.464	0.403	0.431	0.532	Normal
	0.553	0.488	0.545	0.553	0.547	0.532	Avg
prefix-preserving(IP)+ Generalization(other feature)	1	1	0.579	1	0.733	0.67	Attack
	0	0	0	0	0	0.67	Normal
	0.579	0.579	0.335	0.579	0.424	0.67	Avg
Permutation	0.083	0.31	0.27	0.083	0.127	0.387	Attack
	0.69	0.917	0.354	0.69	0.468	0.387	Normal
	0.339	0.566	0.305	0.339	0.271	0.387	Avg
Black Marker	0	0	0	0	0	0.5	Attack
	1	1	0.421	1	0.593	0.5	Normal
	0.421	0.421	0.178	0.421	0.25	0.5	Avg
Truncation	0	0	0.25	0	0	0.25	Attack
	1	1	0.421	1	0.593	0.25	Normal
	0.421	0.422	0.322	0.421	0.25	0.25	Avg
Reverse Truncation	0.906	0.9	0.58	0.906	0.708	0.503	Attack
	0.1	0.094	0.437	0.1	0.163	0.503	Normal
	0.567	0.56	0.52	0.567	0.478	0.503	Avg

Dataset 2 Privacy Results



Anonymization under Injection Attacks

- Test injection attacks on data anonymized by our algorithms
  - Are the datasets anonymized with differential privacy robust enough against Injection Attacks?
- Flows with specific and unique characteristics are prepared by possible intruders and injected in traces before anonymization
- Can one identify injected patterns from anonymized data?



Injected Patterns *

	Packets	Source port	Destination port	Duration	Octets
P_1	1	Fixed	80	-	160
P_2	5	R(65k)	R(65k)	200	256
P_3	110	Fixed	80	200	480[+32]
P_4	10	R(65k)	R(65k)	200	832[+32]
P_5	50	R(65k)	R(65k)	150+R(300)	1208[+R(8)]

- Values in square brackets denote the field evolution between flows.
- R(x): random number between 1 and x.
- Total number of injected flows is 650 (130 flows from each pattern)

* Martin Burkhart, Dominik Schatzmann, Brian Trammell, Elisa Boschi, and Bernhard Plattner. 2010. The role of network trace anonymization under attack. *SIGCOMM Comput. Commun. Rev.* 40, 1 (January 2010), 5-11.

Anonymization Policies

	IP Addr.	Ports	Time [S]	Packets	Octets
A_1	Permutation	-	-	-	-
A_2	Permutation	-	-	O(5)	O(50)
A_3	Permutation	B(8)	O(30)	-	-
A_4	Permutation	B(2)	O(60)	-	-
A_5	Permutation	B(8)	O(30)	O(5)	O(50)
A6: Condensation	-	-	-	-	-
Differential Privacy	-	-	-	-	-

- B(x): bucketized in x buckets,
- O(x): Added a uniform random offset between −x and +x,

Successful Injection Attack Example (oops!)

Injection Pattern

	Packets	Source port	Destination port	Duration	Octets
P_2	5	R(65k)	R(65k)	200	256

Injected record

ID	SRC_IP	DST_IP	PACKETS	OCTETS	START_TIME	START_MSEC	END_TIME	END_MSEC	SRC_PORT	DST_PORT	TCP_FLAGS	DST_PORT	DURATION	TYPE
92144	172.16.50.201	10.220.223.10	5	256	1.39835E+12	940	1.39835E+12	940	36717	61768	0	1	200	1
155653	192.168.51.68	172.16.90.3	5	256	1.39835E+12	665	1.39835E+12	659	3245	35037	0	1	200	1
242622	10.60.60.20	10.150.200.200	5	256	1.39835E+12	44	1.39835E+12	59	36290	31465	0	1	200	1

Anonymization method

	IP Addr.	Ports	Time [S]	Packets	Octets
A_2	Perm.	-	-	O(5)	O(50)

ID	SRC_IP	DST_IP	PACKETS	OCTETS	START_TIME	START_MSEC	END_TIME	END_MSEC	SRC_PORT	DST_PORT	TCP_FLAGS	DST_PORT	DURATION	TYPE
155648	116.251.19.176	98.162.247.69	616	45	40345	0	40345	0	4530	80	2	1	0	1
155649	108.239.60.192	83.39.140.125	4	83	1222259989	507	1222259989	507	113	59346	20	1	0	2
155650	113.69.150.12	7.6.81.7	4	67	1222262255	227	1222262255	227	113	58085	20	1	0	2
155651	240.54.249.20	65.78.151.232	2	89	1.39835E+12	699	1.39835E+12	699	56876	6666	0	1	0	1
155652	72.159.16.47	17.130.149.225	6	49	1222260518	262	1222260518	262	113	42461	20	1	0	2
155653	206.36.9.209	44.200.197.229	8	260	1.39835E+12	665	1.39835E+12	659	3245	35037	0	1	200	1
155654	59.100.174.176	86.185.155.99	6	79	1.39835E+12	562	1.39835E+12	562	56878	2007	0	1	0	1
155655	225.101.113.49	165.132.147.120	4	75	1222260753	724	1222260753	724	64221	113	2	1	0	2
155656	30.190.69.221	119.82.22.111	4	103	1.39835E+12	878	1.39835E+12	878	53816	3828	0	1	0	1
155657	12.160.24.12	29.107.15.54	3069	57	40345	0	40345	0	53152	80	2	1	0	1
155658	148.67.0.23	43.48.244.67	14	2021	1222187543	237	1222187543	647	22	1454	27	1	0	2
155659	244.144.214.239	49.129.28.253	1	56	1222260095	941	1222260095	941	113	51192	20	1	0	2
155660	191.147.42.21	210.28.99.211	5	91	1.39835E+12	675	1.39835E+12	675	58035	1058	0	1	0	1
155661	28.215.221.239	221.17.46.73	5	280	1.39835E+12	356	1.39835E+12	356	49545	8080	0	3	0	1
155662	41.183.63.15	112.34.162.148	4	139	1.39835E+12	916	1.39835E+12	916	1497	80	0	1	0	1

Injected Patterns discovered using K-NN search

	1
	15130
2	75070
3	190667
4	220870
5	41106
6	92144
7	155653
8	242622
9	275461
10	273329
11	276004
12	253237
13	203653
14	20768
15	236750
16	237633
17	3267
18	77141
19	32392
20	194177
21	255112
22	240982
23	178214

Failed Injection Attack Example (YES!)

Injection Pattern

	Packets	Source port	Destination port	Duration	Octets
P_2	5	R(65k)	R(65k)	200	256

Injected record

ID	SRC_IP	DST_IP	PACKETS	OCTETS	START_TIME	START_MSEC	END_TIME	END_MSEC	SRC_PORT	DST_PORT	TCP_FLAGS	DST_PORT	DURATION	TYPE
92144	172.16.50.201	10.220.223.10	5	256	1.39835E+12	940	1.39835E+12	940	36717	61768	0	1	200	1
155653	192.168.51.68	172.16.90.3	5	256	1.39835E+12	665	1.39835E+12	659	3245	35037	0	1	200	1
242622	10.60.60.20	10.150.200.200	5	256	1.39835E+12	44	1.39835E+12	59	36290	31465	0	1	200	1

Anonymization using Differential Privacy

ID	SRC_IP	DST_IP	PACKETS	OCTETS	START_TIME	START_MSEC	END_TIME	END_MSEC	SRC_PORT	DST_PORT	TCP_FLAGS	DST_PORT	DURATION	TYPE
155648	1.92E+02	2.08E+02	8.91E+01	1.04E+04	1.43E+12	1.93E+02	1.43E+12	4.07E+02	-4.86E+03	8.20E+04	-4.34E-02	-2.98E+00	1.23E+03	1.00E+00
155649	2.46E+02	2.45E+02	2.46E+02	1.61E+01	1.49E+11	4.75E+02	1.49E+11	5.12E+02	-3.64E+04	7.53E+03	2.54E+01	1.81E+00	-6.61E+02	2.00E+00
155650	2.46E+02	2.45E+02	2.16E+02	9.06E+00	3.58E+11	7.42E+02	3.58E+11	5.73E+02	7.60E+03	1.70E+04	1.13E+01	1.94E+00	4.28E+02	2.00E+00
155651	1.92E+02	1.08E+01	1.15E+02	1.02E+05	6.70E+11	4.77E+02	6.70E+11	4.61E+02	-1.11E+04	1.74E+04	1.04E+00	7.64E+00	-6.39E+03	1.00E+00
155652	2.46E+02	2.45E+02	2.95E+02	2.51E+01	-2.99E+11	5.28E+02	-2.99E+11	4.36E+02	-3.88E+04	6.20E+03	3.75E+01	7.73E-01	-2.98E+02	2.00E+00
155653	1.92E+02	1.72E+02	1.81E+02	5.16E+04	4.94E+11	3.32E+02	4.94E+11	2.25E+02	3.71E+04	1.44E+04	1.29E+00	-1.20E+00	-7.16E+03	1.00E+00
155654	1.02E+01	1.05E+01	7.16E+01	-7.21E+04	1.64E+12	8.31E+02	1.64E+12	8.10E+02	3.45E+04	5.26E+04	-3.47E-01	1.66E+00	-3.11E+03	1.00E+00
155655	2.45E+02	2.46E+02	4.25E+02	7.58E+00	-3.02E+11	5.40E+02	-3.02E+11	8.53E+02	4.21E+03	4.88E+04	3.20E+01	-6.12E-01	5.94E+01	2.00E+00
155656	1.02E+01	1.72E+02	-6.58E+01	-2.34E+04	1.23E+12	4.27E+02	1.23E+12	6.29E+02	1.00E+04	3.73E+04	2.44E-01	2.18E+00	-7.77E+03	1.00E+00
155657	1.92E+02	2.09E+02	1.47E+02	-8.68E+03	1.16E+12	6.75E+02	1.16E+12	6.16E+02	3.81E+04	1.18E+03	3.39E-01	1.06E-01	6.12E+03	1.00E+00
155658	1.76E+02	2.46E+02	2.98E+02	1.73E+01	-1.07E+11	5.98E+02	-1.07E+11	4.76E+02	-2.37E+04	4.66E+04	3.24E+01	-1.41E-01	-2.29E+02	2.00E+00
155659	2.46E+02	2.45E+02	2.10E+02	1.97E+01	4.22E+10	4.99E+02	4.22E+10	4.10E+02	-3.21E+04	2.71E+04	2.80E+01	1.27E+00	-1.98E+02	2.00E+00
155660	1.02E+01	1.05E+01	5.45E+01	3.88E+04	1.37E+12	5.14E+02	1.37E+12	4.77E+02	3.79E+04	2.02E+04	3.98E-02	3.48E+00	9.18E+03	1.00E+00
155661	1.08E+01	1.01E+01	2.31E+02	1.33E+05	8.33E+11	2.37E+02	8.33E+11	4.89E+02	3.01E+04	-4.17E+03	8.08E-01	3.13E+00	1.06E+04	1.00E+00
155662	1.02E+01	1.72E+02	4.10E+01	-4.58E+04	2.02E+12	1.07E+03	2.02E+12	1.04E+03	4.20E+04	4.10E+03	-8.95E-01	3.23E+00	2.42E+03	1.00E+00

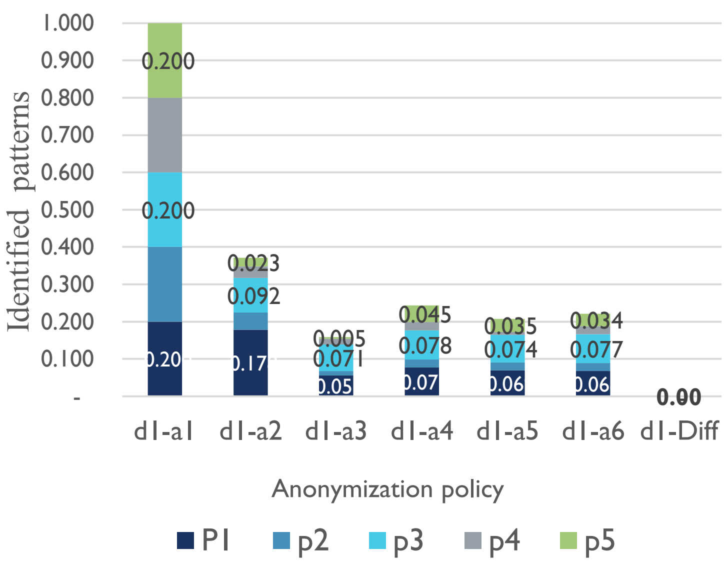
No Injected Patterns discovered using K-NN search

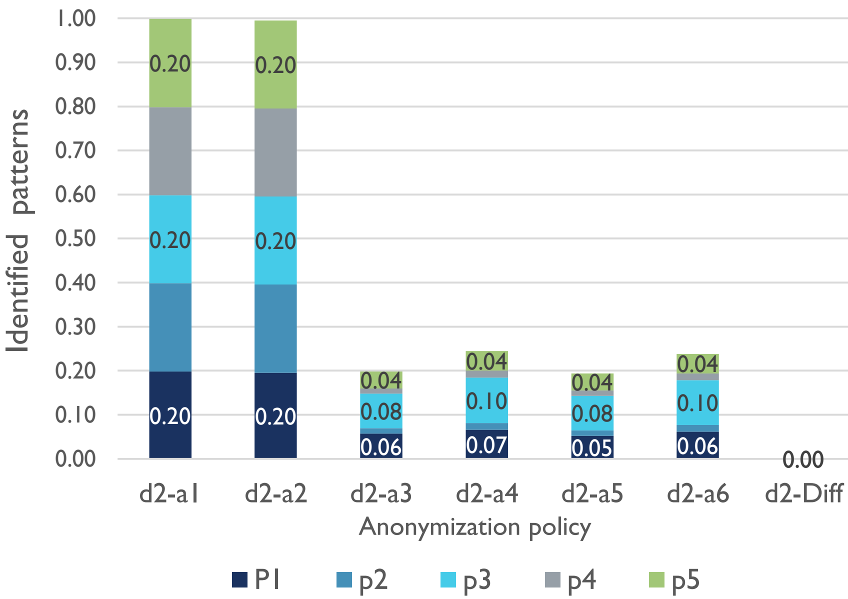
	1
	276016
	270833
3	268886
4	262293
5	270747
6	263262
7	266447
8	259187
9	266129
10	264681
11	259738
12	276016
13	267364
14	26196
15	269328
16	261915
17	277229
18	262534
19	270489
20	271405
21	271947

Experiments on Pattern Injection

- 130 records from each pattern are injected in each dataset before anonymization (total 650 injection attempts)
- The data is anonymized using 7 anonymization policies including Differential Privacy
- K-NN search is used to recover the injected patterns
- The number of identified injected patterns using each anonymization policy is reported

Robustness Against Data Injection Attacks





Findings

- We proposed a method to anonymize network traces that:
    1. Utilizes Differential Privacy providing a very strong privacy guarantee
    2. Is robust against injection attacks
    3. Has negligible impact (less than 2%) when anonymized data are fed to intrusion detection systems
    4. Achieves better privacy-utility tradeoff than existing techniques

Future Work

- Testing if the utility of the proposed method is affected when the number of the injected patterns increases
- Creating a GUI interface to automatically perform all anonymization procedures
- Big-data environment
  - Conduct experiments in big-data test-bed
  - Exploit parallelism for big-data
  - Investigate scalability of proposed techniques in big-data platforms
- Explore additional domains within cybersecurity (e.g. logs)

References

1. A. Aleroud, G. Karabatis, Queryable Semantics for the Detection of Cyber Attacks
2. A. Aleroud, G. Karabatis, P. Sharma, and P. He, "Context and Semantics for Detection of Cyber Attacks," ***Int. J. Information and Computer Security***., vol. 6, no. 1, pp. 63-92, 2014.
3. A. AlEroud and G. Karabatis, "Context Infusion in Semantic Link Networks to Detect Cyber-attacks: A Flow-based Detection Approach," in ***Eighth IEEE International Conference on Semantic Computing***, Newport Beach, California, USA, 2014.
4. A. AlEroud and G. Karabatis, "**A Contextual Anomaly Detection Approach to Discover Zero-day Attacks**," in ***ASE International Conference on Cyber Security***, Washington, D.C., USA 2012, pp. 383-388.
5. A. AlEroud and G. Karabatis, "Detecting Zero-day Attacks using Contextual Relations," in ***Ninth International Knowledge Management in Organizations Conference***, Santiago, Chile, 2014
6. A. AlEroud and G. Karabatis, "Toward Zero-Day Attack Identification Using Linear Data Transformation Techniques," in ***IEEE 7th International Conference on Software Security and Reliability (SERE'13)***, Washington, D.C., 2013, pp. 159-168.