Leveraging cybersecurity with machine-learning

Takeshi Takahashi, Ph.D., CISSP, PMP

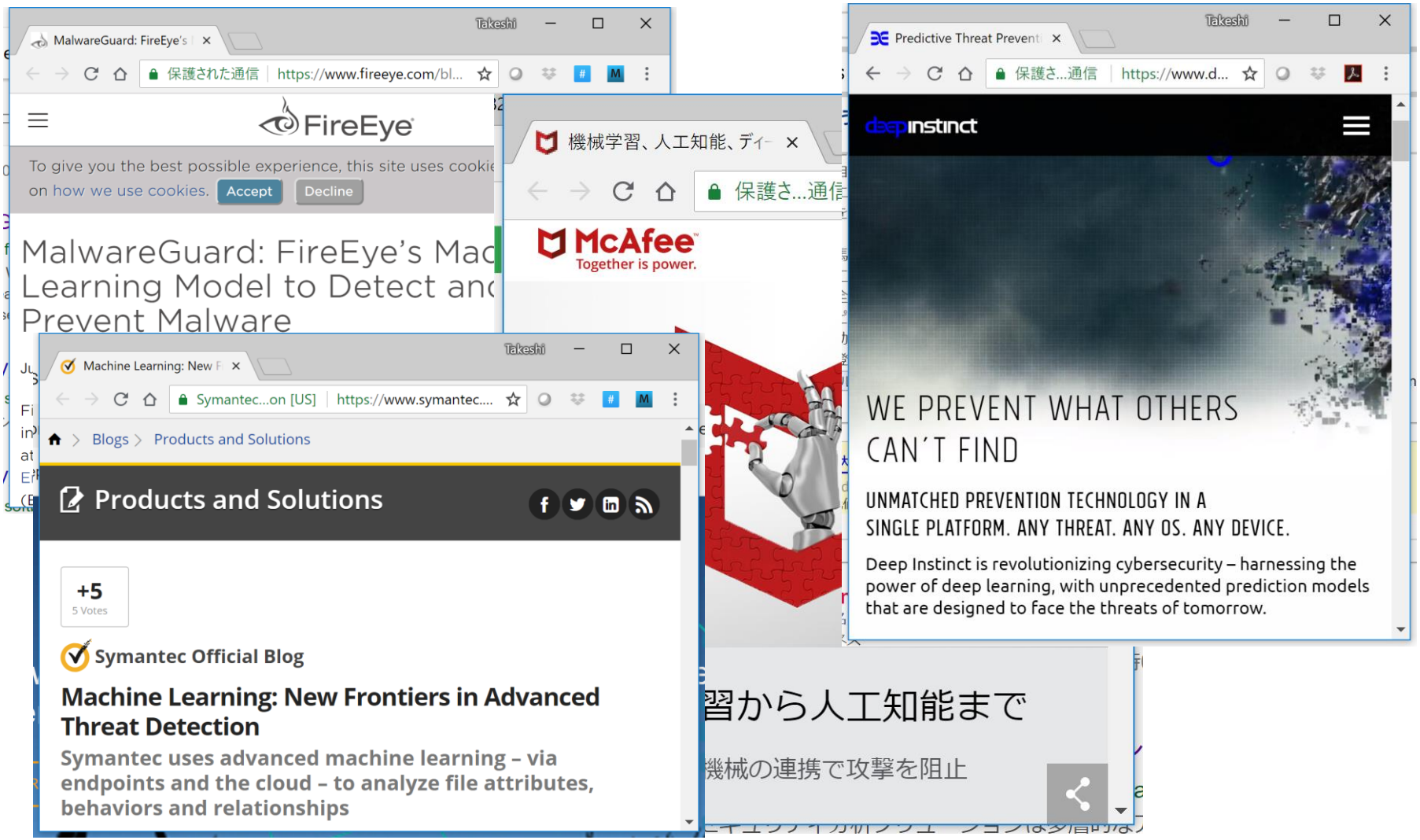
Research Manager
Cybersecurity Laboratory / AI Science R&D Promotion Center
NICT

Agenda

1. The use of AI for cybersecurity in these days

2. Our research activities in a nutshell

AI techniques are said to be used in industries



Anti-virus vendors claim that they use deep learnings, but the algorithms or the specs they use were non-disclosed thus the details were unknown.

The use of AI for cybersecurity is a hot topic now

Major research organizations work on the applicability of AI techniques for cybersecurity in these days. Here are the list of organizations that presented AI-related papers in USENIX Security 2018.

Europe
- EPFL
- Frauhofer FKIE
- Max Planck Institute for Informatics
- RWTH Aachen University
- Siemens CERT
- Universidade de Lisboa

Israel
- Bar-Ilan Uniersity

Asia
- Chinese Academy of Science
- Beijing Jiaotong University

United States
- Boston University
- Columbia University
- Florida Institute of Technology
- Google Inc
- Indiana University
- Iowa State University
- MIT
- UC Santa Barbara
- University of Chicago
- University of Delaware
- University of Illinois
- University of Maryland
- Virginia Tech

The use of AI for cybersecurity is a hot topic now

Major research organizations work on the applicability of AI techniques for cybersecurity in these days. Here are the list of organizations that presented AI-related papers in CSS 2018.

Europe
- Lancaster University
- University College London

Asia
- Inha University
- Peking University
- Zhejiang University
- The Hong Kong Polytechnic University
- Chinese Academy of Sciences
- Hanyang University
- National University of Singapore

United States
- University of Central Florida
- Florida International University
- Northwest University
- Lehigh University
- The Pennsylvania State University
- Virginia Tech
- University of Pennsylvania
- Symantec
- UC Riverside
- UC Berkeley
- University of Illinois at Urbana-Champaign
- University of Massachusetts

A few example topics on ML researches

Traffic anomaly detection & malware detection (long standing area)

- Explainable system
- Performance improvements /real-time operations

Attacks on computing systems

- Solving captcha
- Malfunctioning voice recognition systems

Deanomysation (attacks against privacy)

- Code Authorship Identification
- Document author attribute classification
- Identification of account pertaining review comments

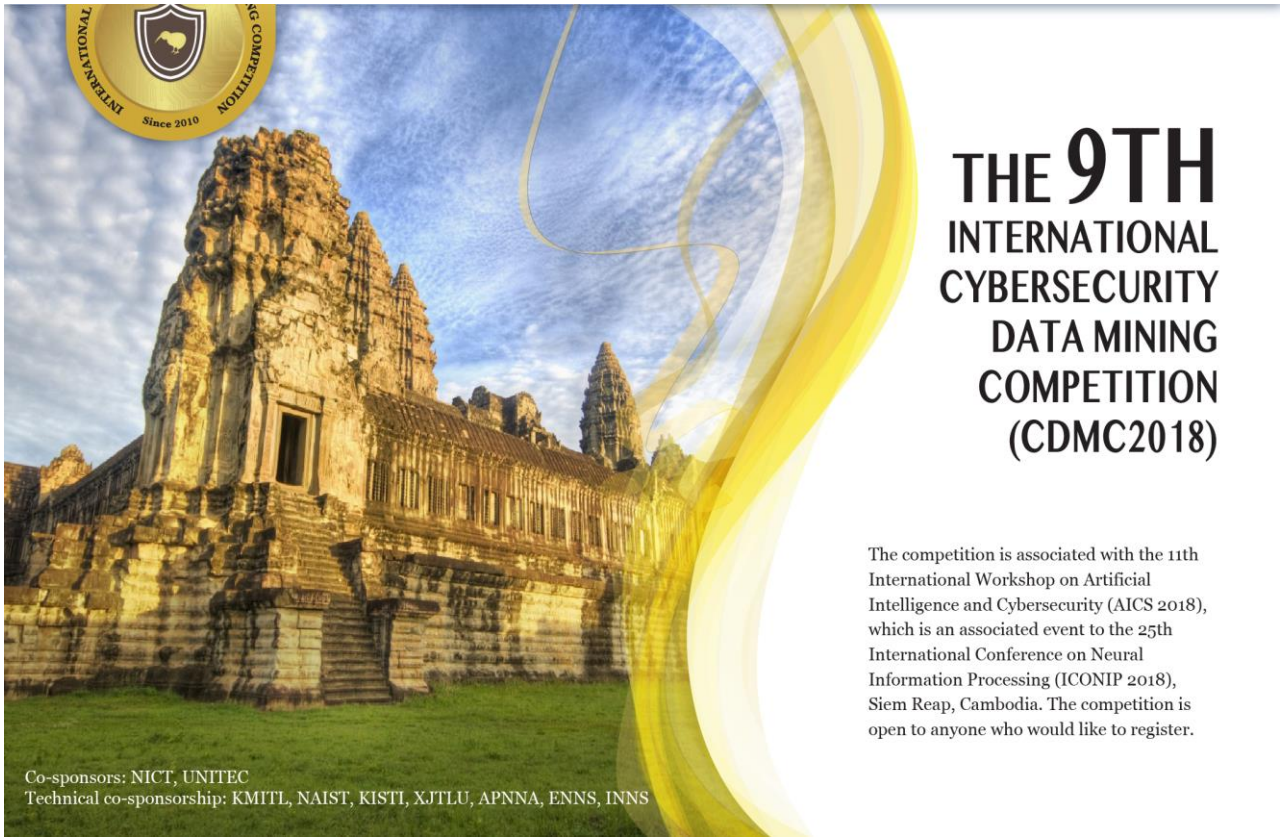
Proactive defense techniques

- Program debloating (minimize vulnerabilities)
- Watermarking DNN
- Event prediction

Vulnerabilities of ML

- Poisoning attacks
- Vulnerabilities of transfer learning
- Attribute inference attacks
- Model reuse attack

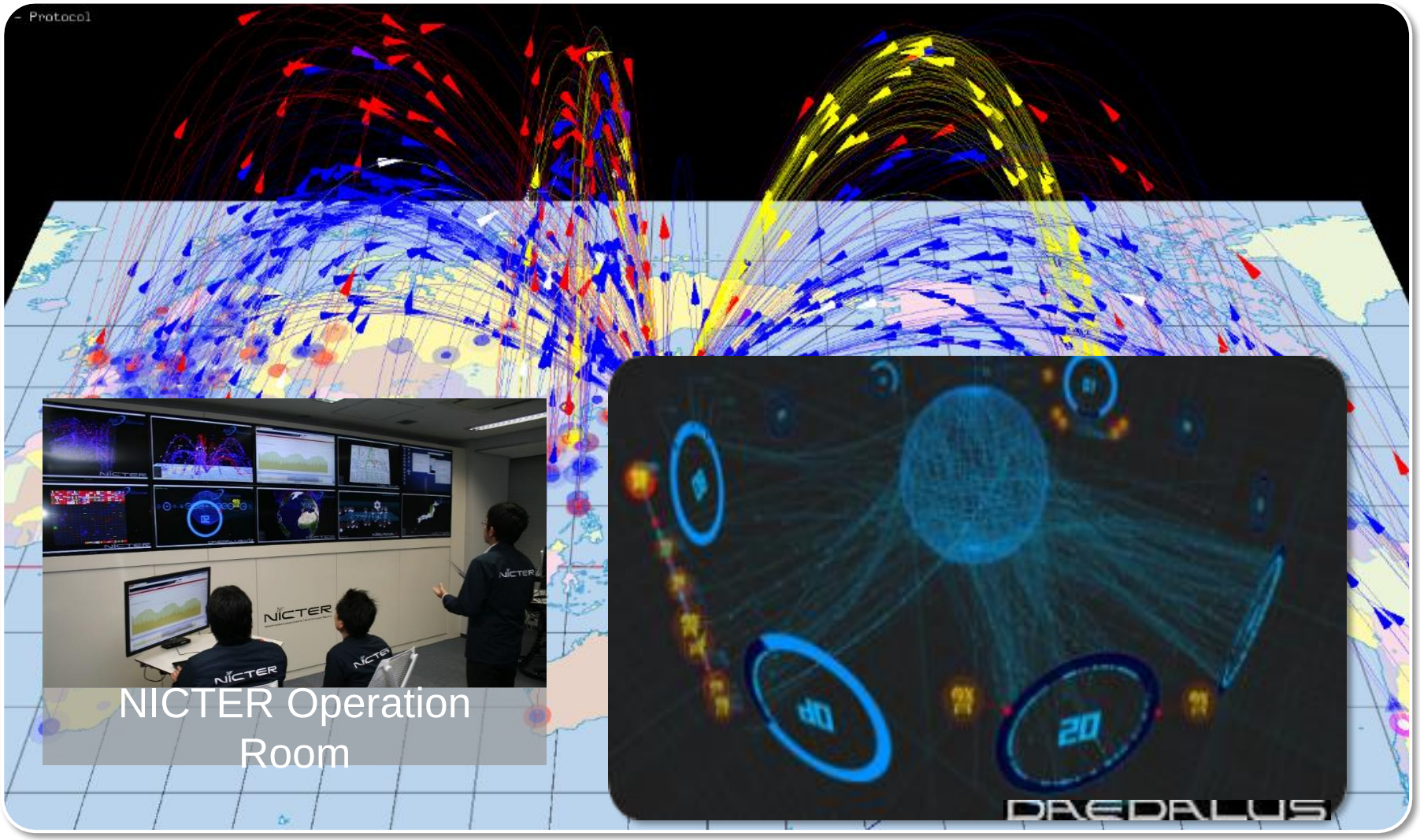
We worked on AI x cybersec. for more than a decade



- 11th International Data Mining and Cybersecurity Workshop (DMC), 2018
- 9th International Cybersecurity Data Mining Competition (CDMC), 2018

Our network monitoring systems accumulates data

- ✓ We monitor large-scale darknet spaces
- ✓ We built and have been operating systems, e.g., NICTER and DAEDALUS



NICTER Operation Room

Our dataset

Category	Examples of accumulated data
Darknet related data	Data on the traffic sent to unused IP address spaces. This includes pcap files, statistical information, and malicious host information.
Livenet related data	Traffic data within NICT. This includes pcap files, flow data, security alerts generated by security appliances.
Malware related data	Malware samples, static and dynamic analysis results, etc.
Spam related data	Spam (double bounce) mail data, statistical information, etc.
Android related data	APK files and applications' metadata, e.g., category and description of applications
Blogs and articles	Tweets, security vendor blogs, etc.
Web crawler	URL list, Web contents, their evaluation results, etc.
Honeypot data	Data from High-interaction/low-interaction honey pots and high-interaction/low-interaction client honey pots
Commercial Intelligence data	Information on the sites hosting malware, bot, C&C server list, domain history, malware samples, threat reports, etc. purchased from VirusTotal, SecureWorks, Anubis, DomainTools, Malnet, Team 5, etc.

With the dataset, we work on security automation

We conduct R&D on AI techniques that analyze and understand security situation and automate security operations within an organization.

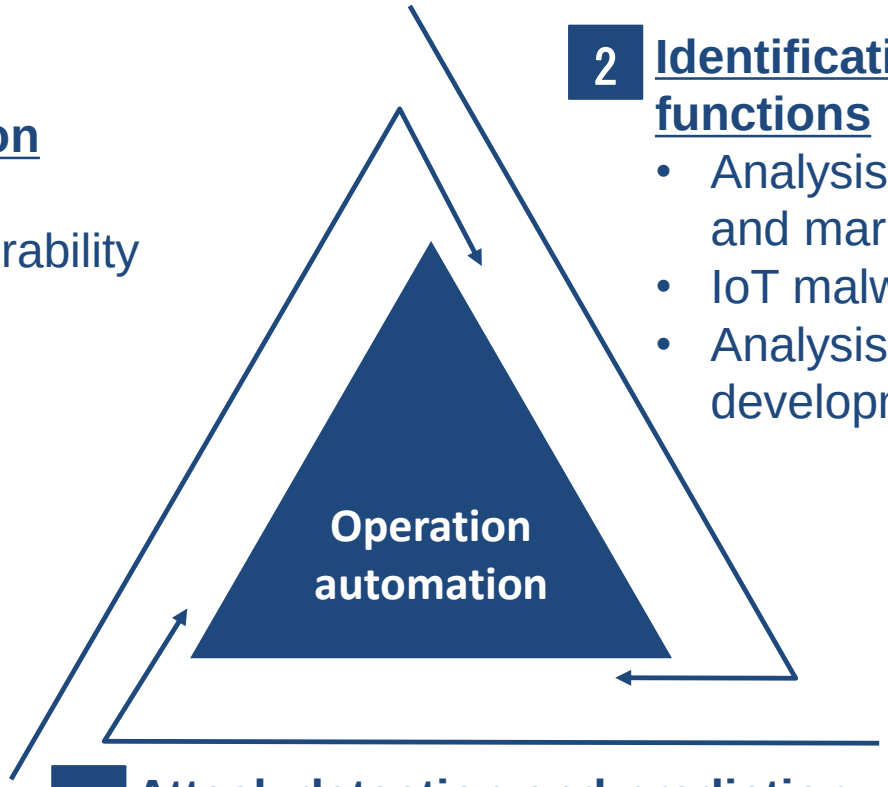
1

Priority determination

- Alert screening
- Evaluation of vulnerability severity

2 **Identification of malware functions**

- Analysis of Android apps and markets
- IoT malware analysis
- Analysis automation tool development



3 **Attack detection and prediction**

- Darknet analysis
- Threat estimation and prediction
- Encrypted traffic analysis

Agenda

1. The use of AI for cybersecurity in these days
2. Our research activities in a nutshell

Our research focus

We conduct R&D on AI techniques that analyze and understand security situation and automate security operations within an organization.

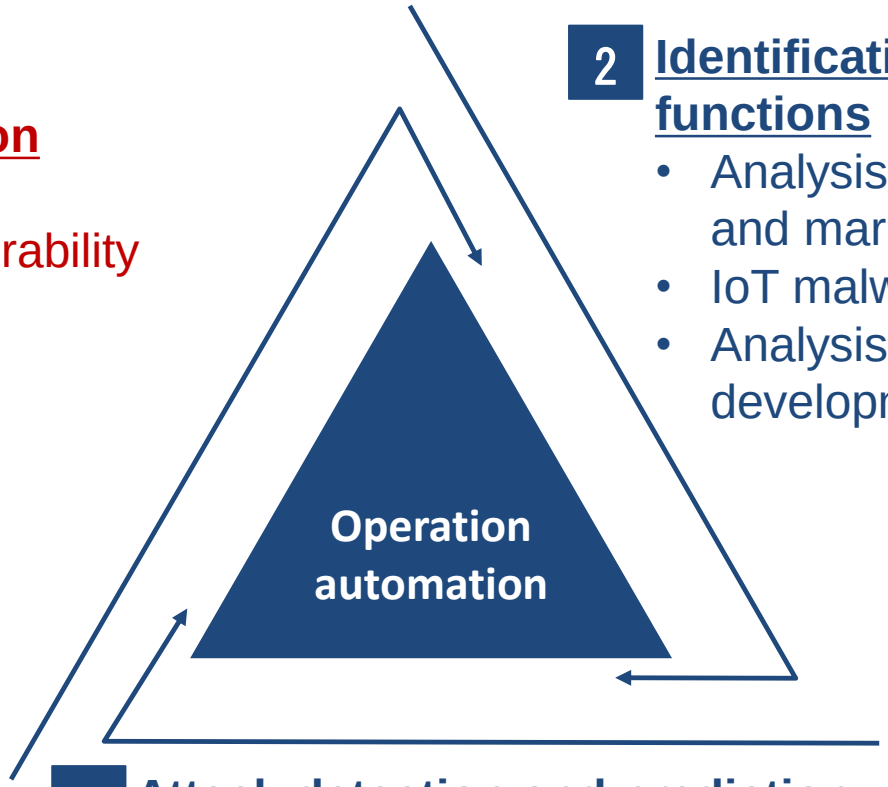
1

Priority determination

- Alert screening
- Evaluation of vulnerability severity

2 **Identification of malware functions**

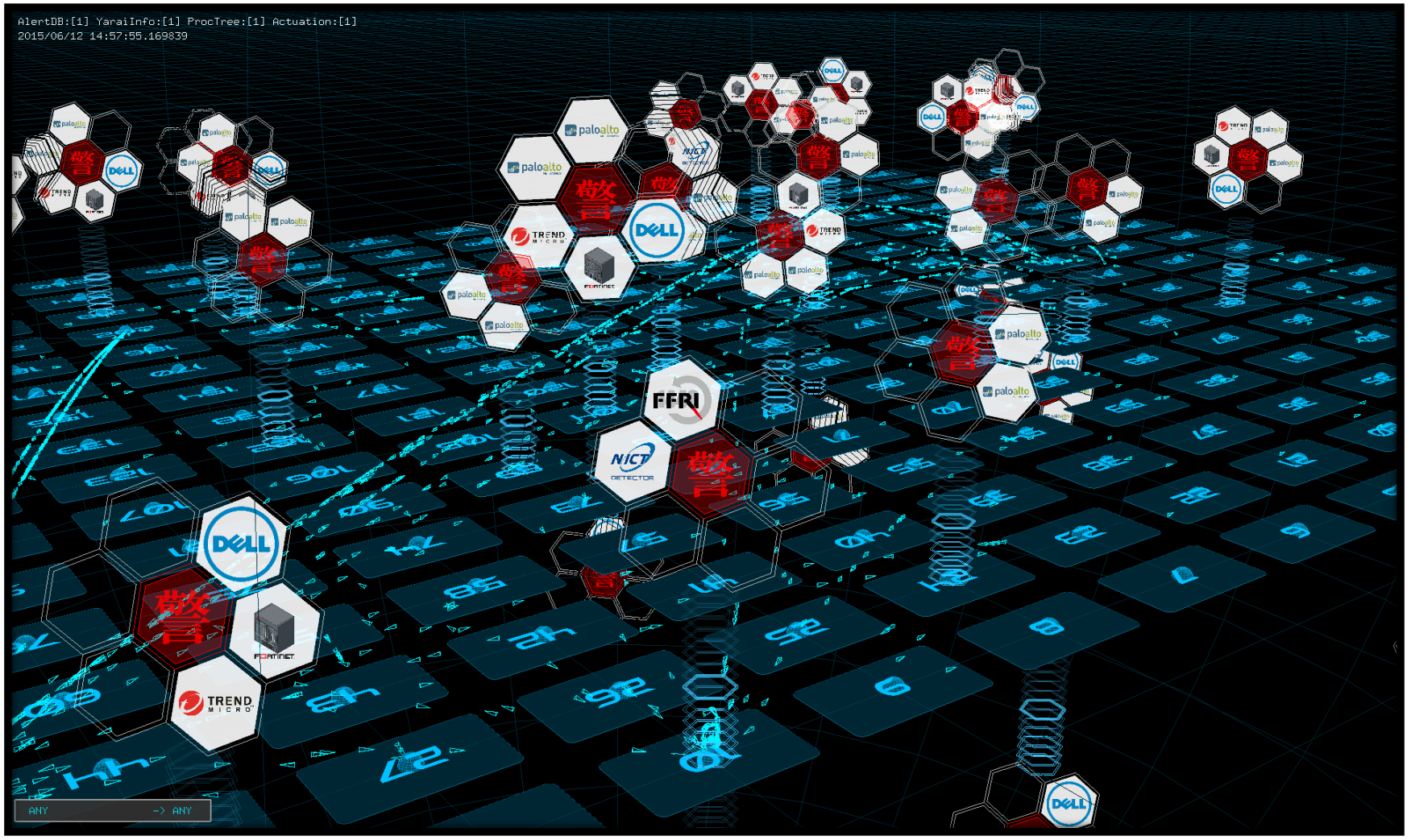
- Analysis of Android apps and markets
- IoT malware analysis
- Analysis automation tool development



3 **Attack detection and prediction**

- Darknet analysis
- Threat estimation and prediction
- Encrypted traffic analysis

NIRVANA Kai integrates security appliances



Alert screening and prioritization

Security appliances

Security alerts

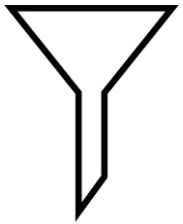
This filtering is currently conducted by static rule filtering and manual verification process

Important alerts

This filtering will be done by ML and automated verification process

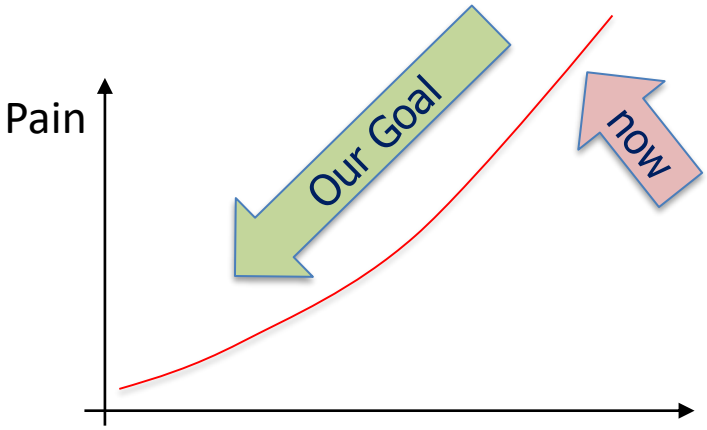
Difficulties in our current security operations

Alerts
~ 100,000/day



Suspicious
Alerts
~ 100/day

True
Alerts
~ 5 /day



One Expert
4 dedicated HOUR
Tedious Work

Our research focus

We conduct R&D on AI techniques that analyze and understand security situation and automate security operations within an organization.

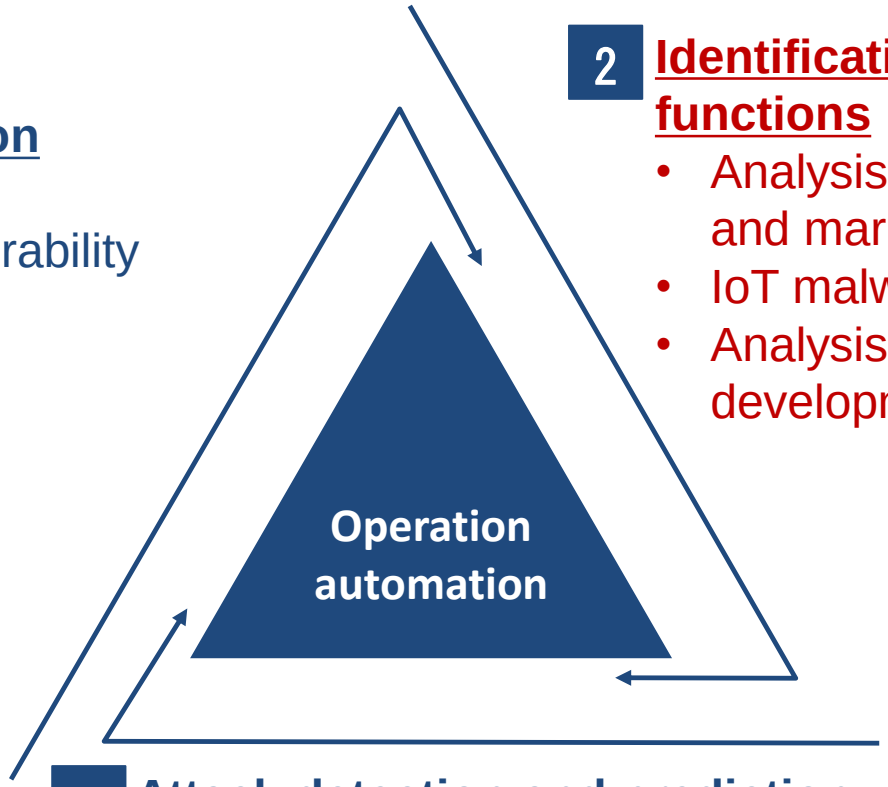
1

Priority determination

- Alert screening
- Evaluation of vulnerability severity

2 **Identification of malware functions**

- Analysis of Android apps and markets
- IoT malware analysis
- Analysis automation tool development

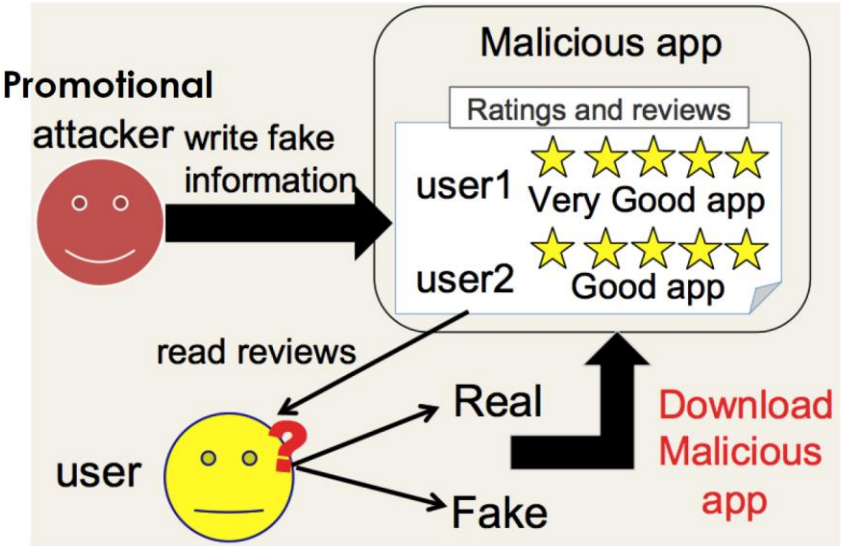


3 **Attack detection and prediction**

- Darknet analysis
- Threat estimation and prediction
- Encrypted traffic analysis

Current research issues on Android security

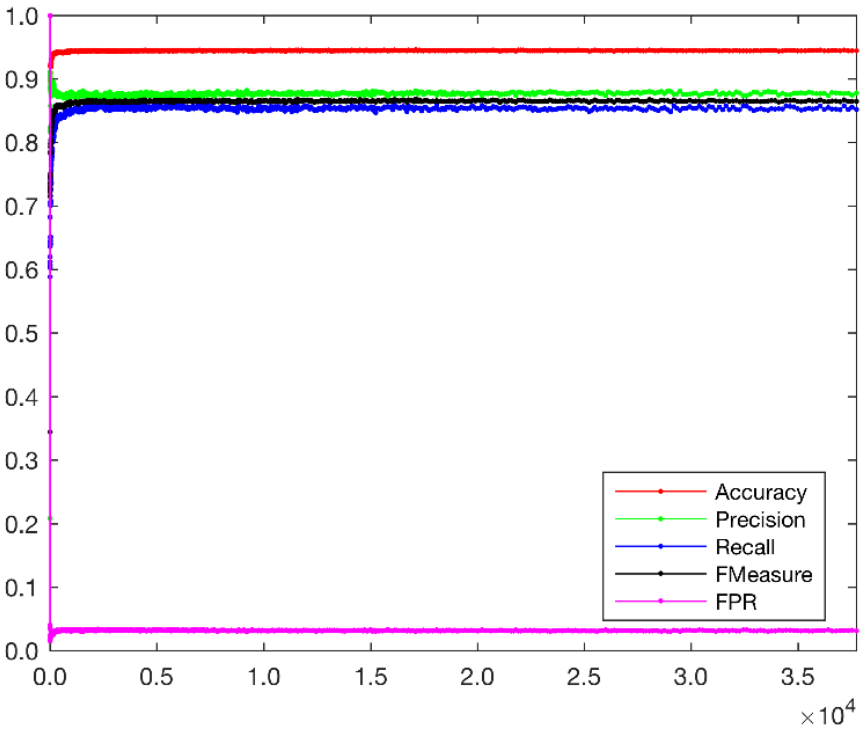
1. *Analyses on promotional attacks and demotional attacks* on Android app markets



2. *Android malware detection and classification*
   1. SVM/SVM-RFE approach using api calls, permission requests, category, and app descriptions.
   2. Deep learning approach
   3. Integration of static and dynamic analyses

Feature selection improves the performance further

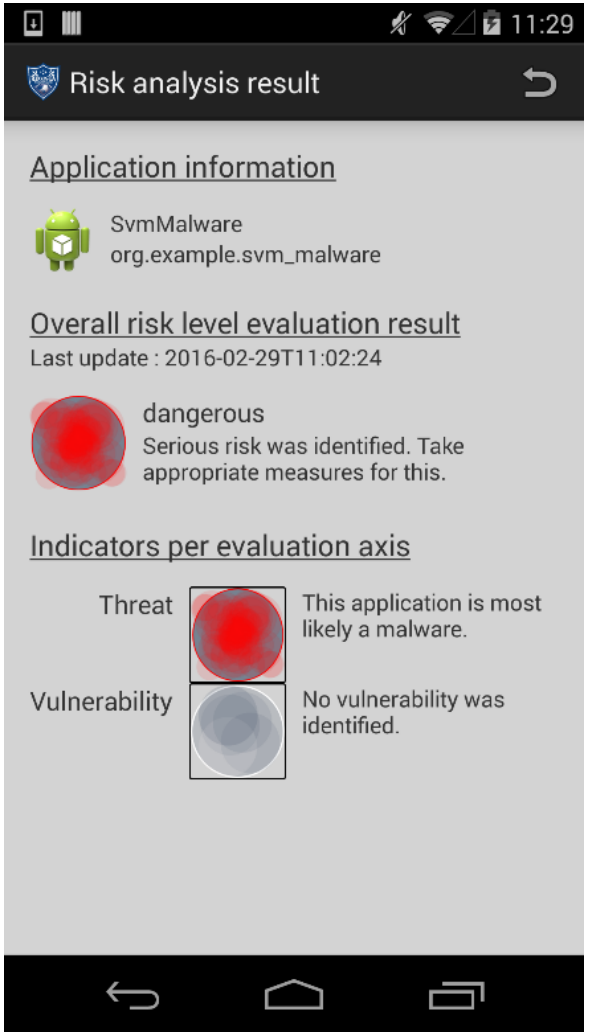
- *We detect malware using SVM-RFE (Accuracy = 94.59 %)*
  - *Features: permission requests, API calls, app categories, clusters(generated from app descriptions)*
  - *1,439 out of 30,000 features are used to produce the best performance*
  - *Influiential features:* API calls, some permission requests and application categories. (cluster feature barely contributes to the performance)

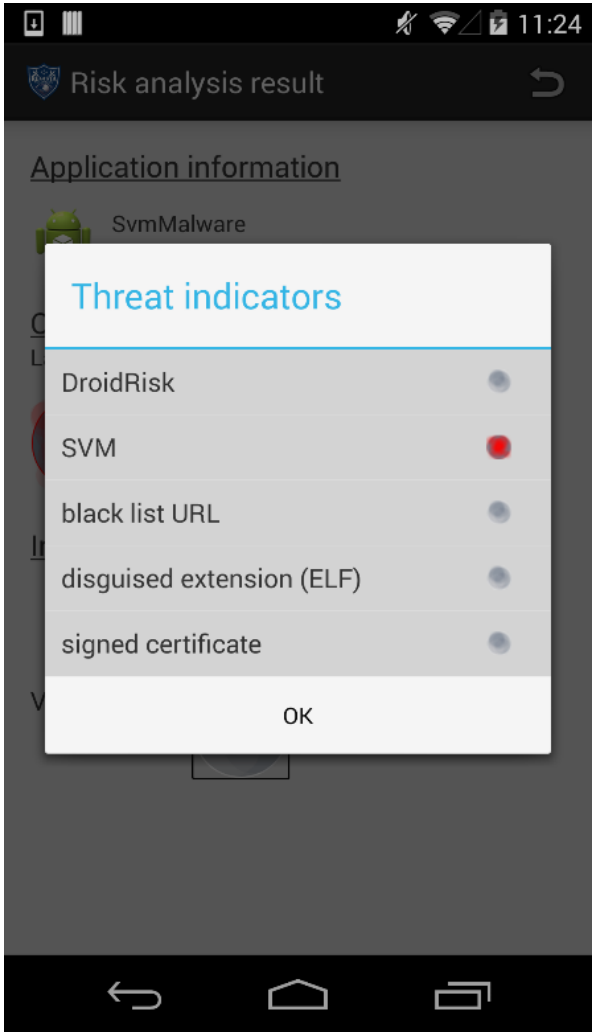


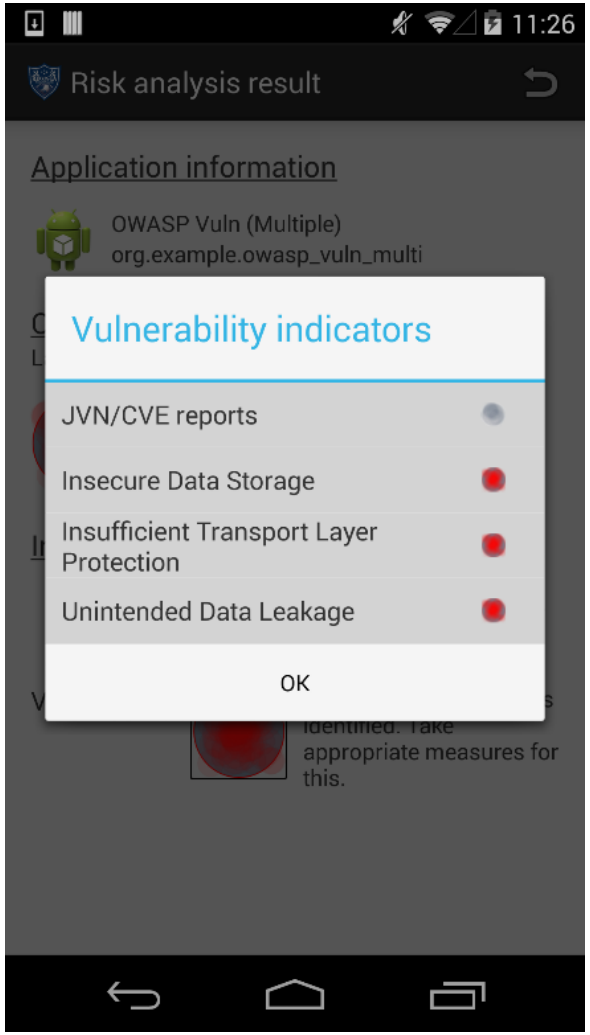
- We are currently evaluating the effectiveness of neural network techniques (Accuracy ≒99.79)

Source: T.Takahashi et al., “Android Application Analysis using Machine Learning Techniques,” AI for Cybersecurity, Springer, 2018

We identify Android malware and visualize that



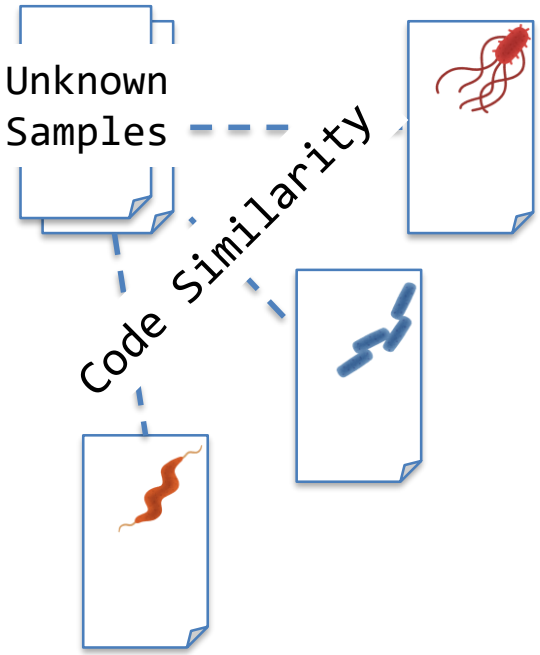


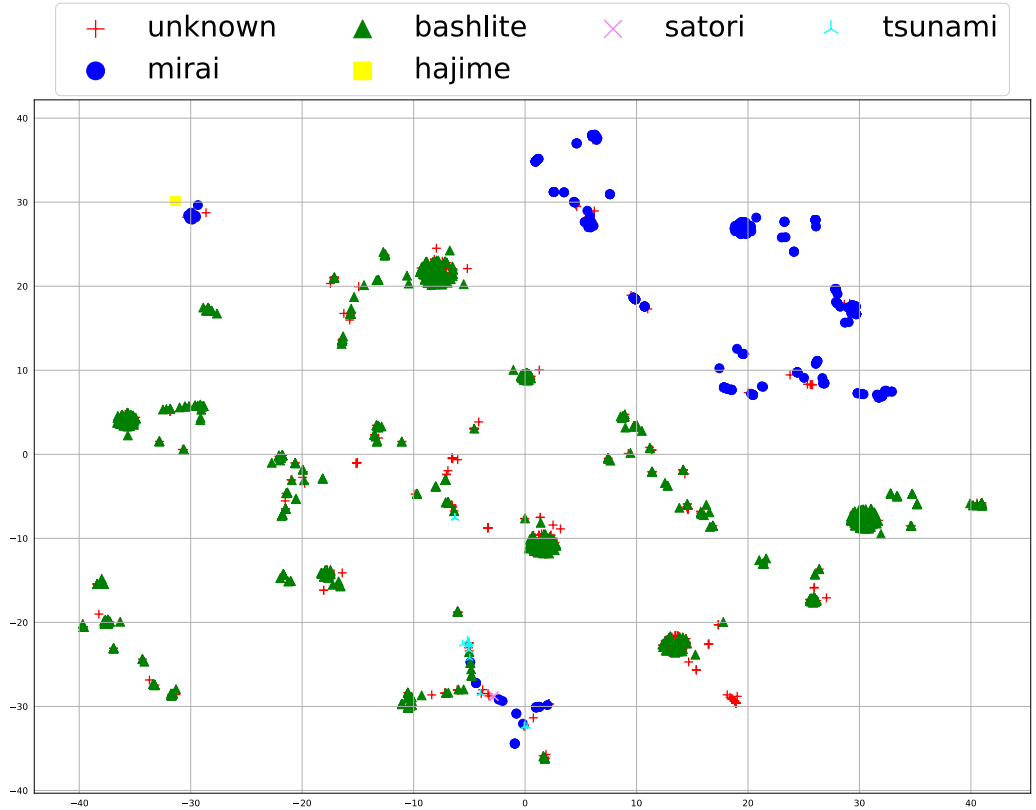


Dataset used in the paper will be available soon at http://mobilesec.nict.go.jp.

Clustering IoT malware in these days

We classified unknown samples into several malware families to catch a hint about efficient analysis of those samples.





Samples mapped on a two-dimensional plane with T-SNE

Our research focus

We conduct R&D on AI techniques that analyze and understand security situation and automate security operations within an organization.

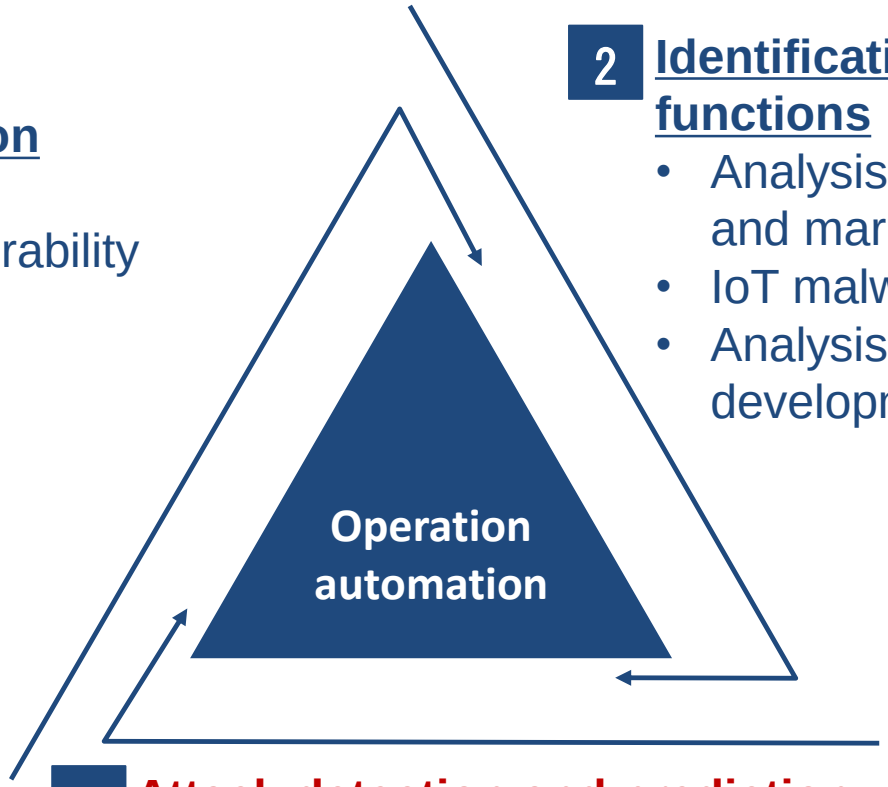
1

Priority determination

- Alert screening
- Evaluation of vulnerability severity

2 **Identification of malware functions**

- Analysis of Android apps and markets
- IoT malware analysis
- Analysis automation tool development

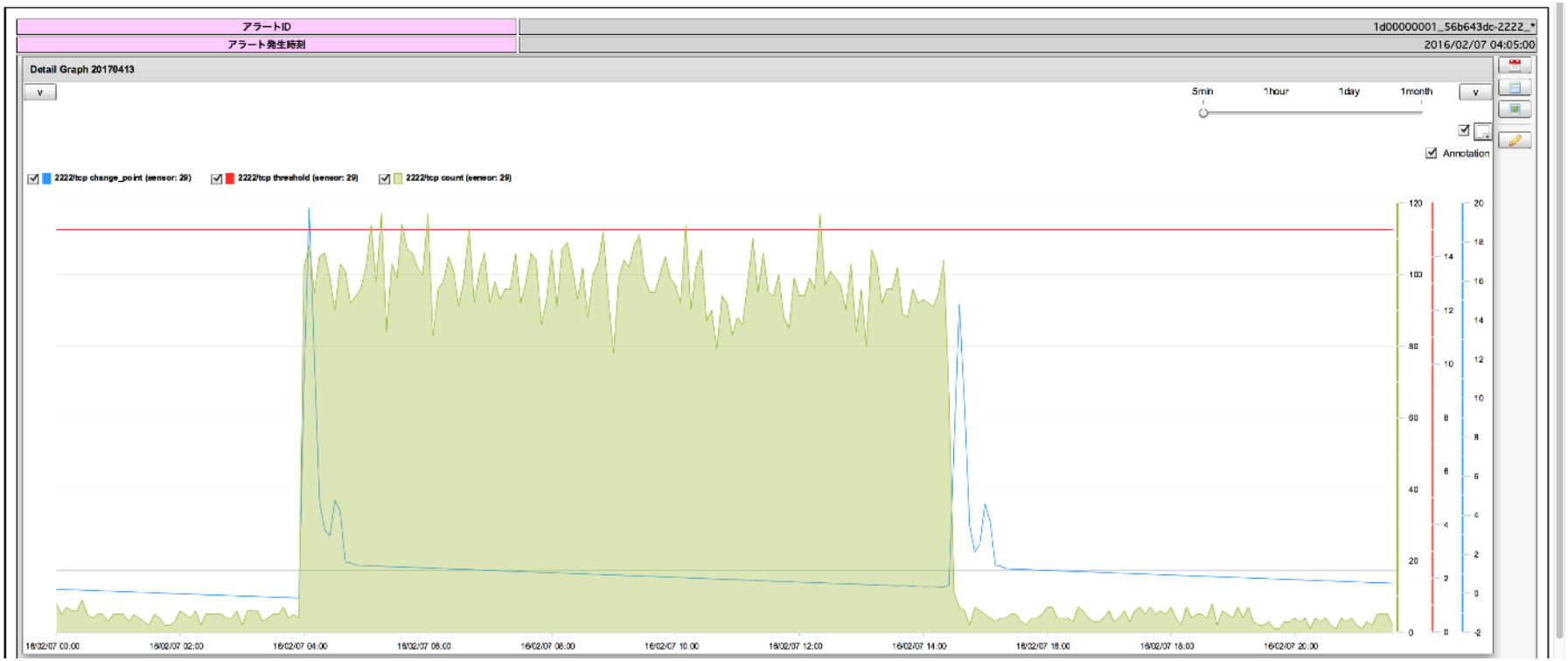


3 **Attack detection and prediction**

- Darknet analysis
- Threat estimation and prediction
- Encrypted traffic analysis

We detect the coordinated actions/trend changes

Activities of botnet computers are often coordinated (initiated by C&C server). Thus the change of the coordinated actions can be viewed as the change of traffic trend change



In this figure, horizontal axis represents time while the vertical axis represents the number of packets sent to our darknet address space. This figure represents an example case of a malware being activated and stopped at certain time.

Alerts are generated upon detecting such points

Filter Conditions

Alert Time | equal to | Add another condition

Reset all conditions

Search Refinement

prev:20 << 32227 32228 32229 32230 32231 32232 32233 32234 32235 32236 32237 32238 32239 32240 32241 >> next:20

Number of Display 20 Change Cur: 644661-644680 / All: 645089

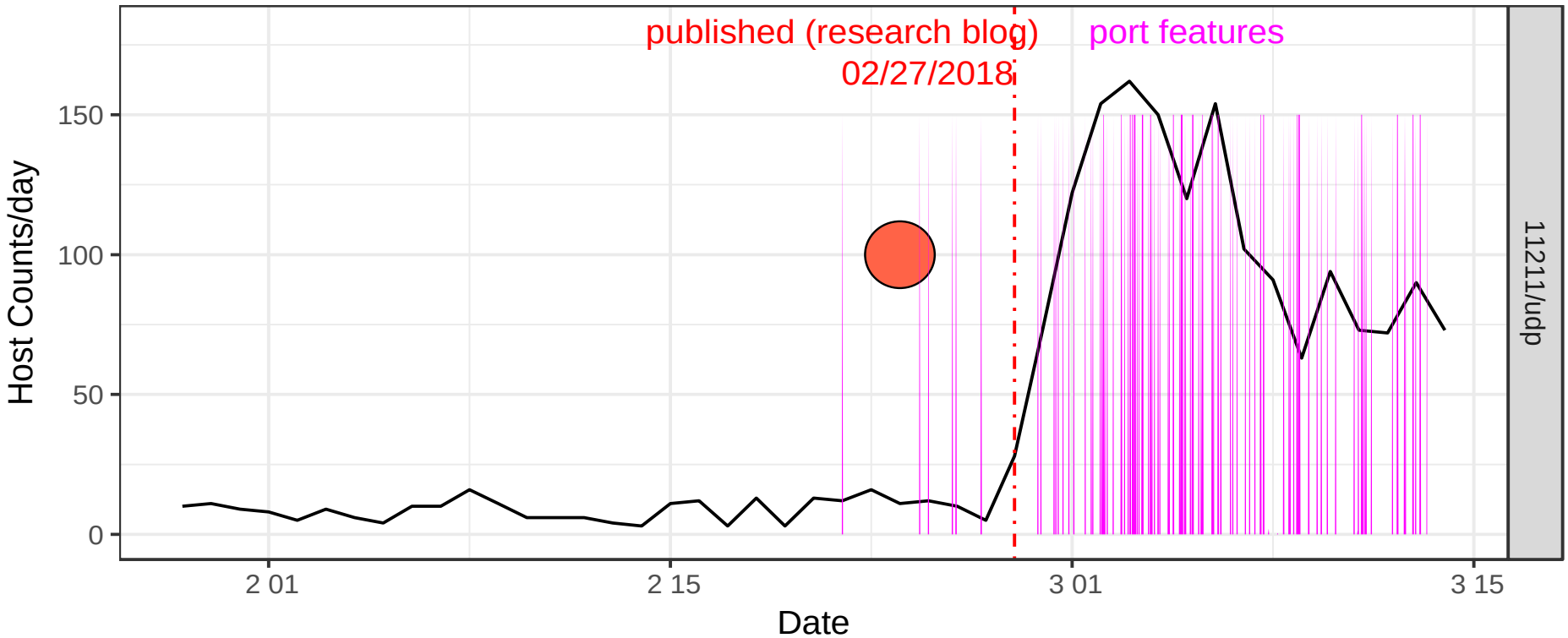
No.	Alert Time	Sensor ID	Type	Period	Target	Option	Change Point Score	Threshold	Cause Target	Cause Option	Cause Count	Detail Info
644661	2016/02/07 08:40:01	29	tpd00is	5m	0-1024	*	22.100443	15	139	*	1	detail
644662	2016/02/07 08:10:03	29	tpd00is	5m	1025-4999	*	19.837684	15	3306	*	2	detail
644663	2016/02/07 08:10:02	29	tpd00is	5m	0-1024	*	16.937063	15	22	*	22	detail
644664	2016/02/07 06:40:02	29	tpd00is	5m	0-1024	*	22.100847	15	102	*	3	detail
644665	2016/02/07 05:40:01	29	tpd00is	5m	0-1024	*	16.967217	15	143	*	5	detail
644666	2016/02/07 05:10:03	29	tpd00is	5m	10000-14999	*	19.401623	15	10000	*	10	detail
644667	2016/02/07 04:40:02	29	tpd00is	5m	1025-4999	*	19.725899	15	2222	*	108	detail
644668	2016/02/07 04:40:02	29	tpd00is	5m	5000-9999	*	19.676857	15	8000	*	99	detail
644669	2016/02/07 04:40:01	29	tpd00is	5m	0-1024	*	22.842586	15	502	*	22	detail
644670	2016/02/07 04:10:01	29	tpd00is	5m	0-1024	*	18.446741	15	80	*	26	detail
644671	2016/02/07 04:10:02	29	tpd00is	5m	5000-9999	*	22.09943	15	7071	*	7	detail
644672	2016/02/06 22:30:01	29	tpd00is	5m	0-1024	*	20.370078	15	82	*	5	detail
644673	2016/02/06 22:30:17	29	upd00is	5m	0-1024	*	16.89279	15	53	*	5	detail
644674	2016/02/06 22:30:18	29	upd00is	5m	53	*	16.892791	15	53	*	5	detail
644675	2016/02/06 22:00:02	29	tpd00is	5m	0-1024	*	16.683782	15	22	*	109	detail
644676	2016/02/06 22:00:17	29	tpd00is	5m	22	*	17.493689	15	22	*	109	detail
644677	2016/02/06 21:30:14	29	upd00is	5m	0-1024	*	22.100837	15	161	*	4	detail
644678	2016/02/06 21:30:16	29	upd00is	5m	161	*	22.100445	15	161	*	4	detail
644679	2016/02/06 21:00:02	29	tpd00is	5m	0-1024	*	16.178003	15	91	*	1	detail
644680	2016/02/06 19:10:02	29	upd00is	5m	5000-9999	*	22.100443	15	5006	*	6	detail

All Select/Unselect

prev:20 << 32227 32228 32229 32230 32231 32232 32233 32234 32235 32236 32237 32238 32239 32240 32241 >> next:20

1. Our prototype works, but it needs to work in real-time, and it needs to minimize false positives/negatives.
2. We are currently approaching this issue with glasso, NMF, and tensor decomposition techniques, respectively.

Real-time botnet detection using tensor decomposition



1. We were able to identify the coordinated action prior to the issue is published by a well-known research blog
2. We were able to identify the coordinated action before NICTER system identifies trend change

Major publications

1. H.Kanehara, Y.Murakami, J.Shimamura, T.Takahashi, D.Inoue, N.Murata, "Real-Time Botnet Detection Using Nonnegative Tucker Decomposition," ACM SAC, 2019.
2. B.Sun, T.Ban, S.Chang, Y.Sun, T.Takahashi, D.Inoue, "A Scalable and Accurate Feature Representation Method for Identifying Malicious Mobile Applications," ACM SAC, 2019.
3. T.Takahashi, T.Ban, "Android Application Analysis using Machine Learning Techniques," Intelligent Systems Reference Library, 181 - 205, 2019.
4. S.Chang, Y.Sun, W.Chuang, M.Chen, B.Sun, T.Takahashi, "ANTSdroid:Using RasMMA Algorithm to Generate Malware Behavior Characteristics of Android Malware Family," IEEE PRDC, 2018.
5. L.Zhu, T.Ban, T.Takahashi, D.Inoue, "Employ Decision Value for Binary Soft Classifier Evaluation with Crispy Reference," ICONIP, 2018.
6. R.Iijima, S.Minami, Z.Yunao, T.Takehisa, T.Takahashi, Y.Oikawa, T.Mori, "Poster: Audio Hotspot Attack: An Attack on Voice Assistance Systems Using Directional Sound Beams," ACM CCS, 2018.
7. T.Takahashi, B.Panta, Y.Kadobayashi, K.Nakao, "Web of cybersecurity: Linking, locating, and discovering structured cybersecurity information," Int J Commun Syst. 2017.

Our dataset, a vehicle for research collaborations

Category	Examples of accumulated data
Darknet related data	Data on the traffic sent to unused IP address spaces. This includes pcap files, statistical information, and malicious host information.
Livenet related data	Traffic data within NICT. This includes pcap files, flow data, security alerts generated by security appliances.
Malware related data	Malware samples, static and dynamic analysis results, etc.
Spam related data	Spam (double bounce) mail data, statistical information, etc.
Android related data	APK files and applications' metadata, e.g., category and description of applications
Blogs and articles	Tweets, security vendor blogs, etc.
Web crawler	URL list, Web contents, their evaluation results, etc.
Honeypot data	Data from High-interaction/low-interaction honey pots and high-interaction/low-interaction client honey pots
Commercial Intelligence data	Information on the sites hosting malware, bot, C&C server list, domain history, malware samples, threat reports, etc. purchased from VirusTotal, SecureWorks, Anubis, DomainTools, Malnet, Team 5, etc.