An Internet Threat Evaluation Method based on Access Graph of Malicious Packets

June 16-23, 2007
FIRST 19th Annual Conference

Masaki Ishiguro[*1)] Hironobu Suzuki[*2)]
Ichiro Murase[*1)] Yoichi Shinoda[*3)] Shigeki Goto[*2)]

[*1)]Mitsubishi Research Institute, Inc [*2)]Waseda University
[*3)] National Institute of Information and Communications Technology, Japan

Agenda

- Introduction
  - Goal
  - Background history
  - System overview
- Threat evaluation method
  - Evaluation approach
  - Calculation method
- Experiments
  - MS SQL Incident
  - Windows File share Incident
- Conclusion and Future work

Goals

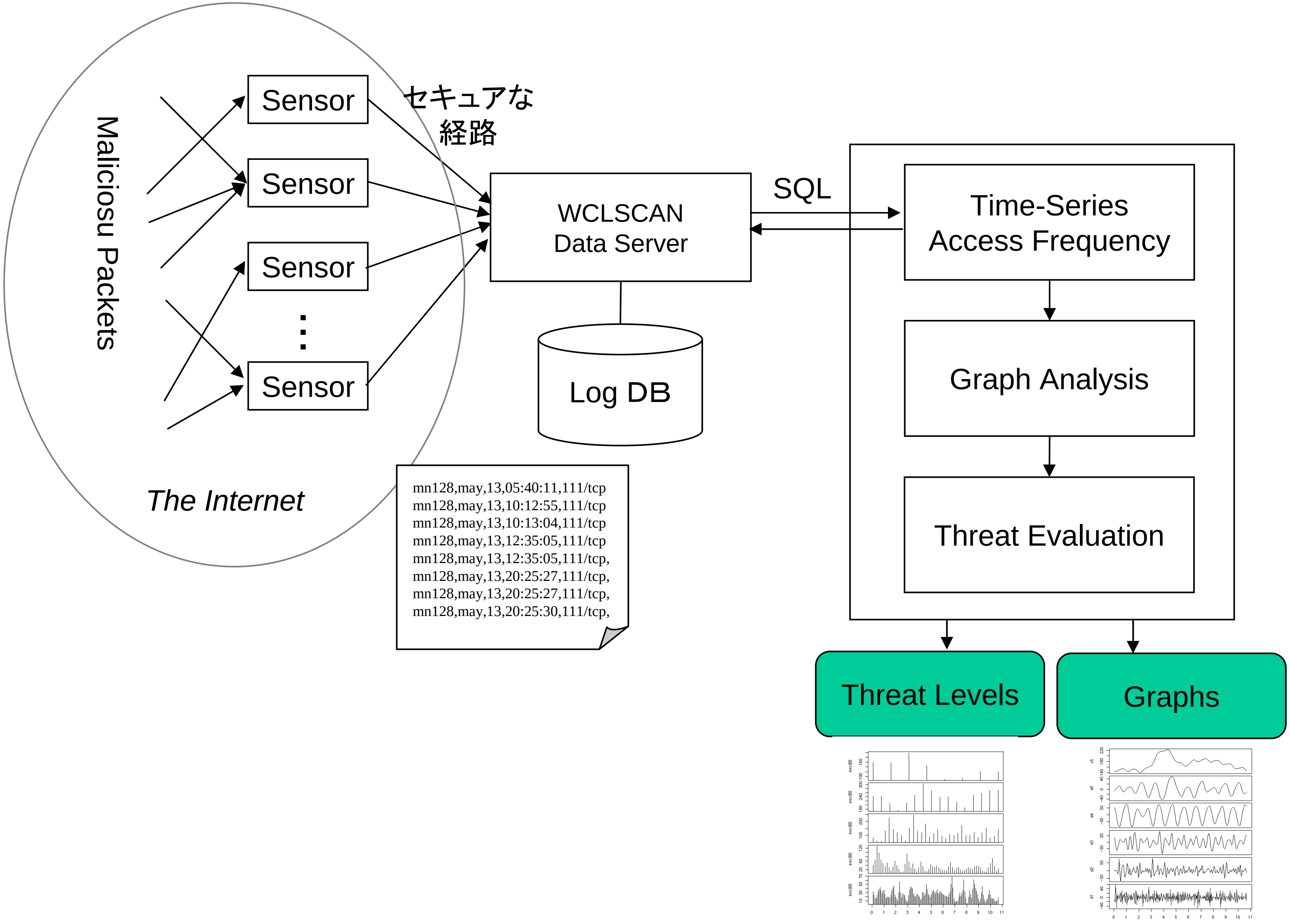
0-day attack

- Find “new” threats without human resources
- System never sleep, 24 hours/7 days
- Find threats from huge chaos data
- Show the simple conclusion
- Access the report in anytime from anywhere

Background History

- 1999 CLSCAN (common log scanner)
  - “pretty print” tool for syslog file of my Internet router
- 2001 Last 12 months log was analyzed
  - “*Internet security analysis using packet filter log*” , SEA software symposium 2001
- 2002 WCLSCAN project was started
  - Wide area version of clscan
- 2003 Internet Weather Report aka WCLSCAN
  - “threat calculation using Bayesian estimation” unit was added to WCLSCAN
- 2004, 4 sensor boxes have been running and provide information on www.clscan.org
- 2006, A Threat Evaluation Methods (Todays Topic)

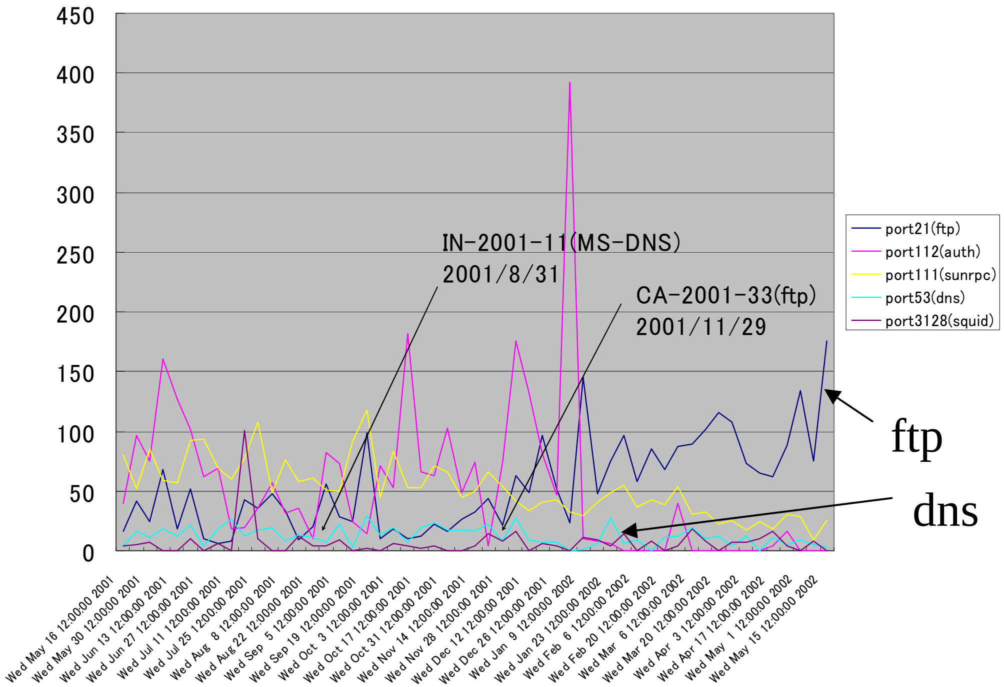
Our Internet Monitoring System



```
mn128,may,13,05:40:11,111/tcp
mn128,may,13,10:12:55,111/tcp
mn128,may,13,10:13:04,111/tcp
mn128,may,13,12:35:05,111/tcp
mn128,may,13,12:35:05,111/tcp,
mn128,may,13,20:25:27,111/tcp,
mn128,may,13,20:25:27,111/tcp,
mn128,may,13,20:25:30,111/tcp,
```

Monitored Data

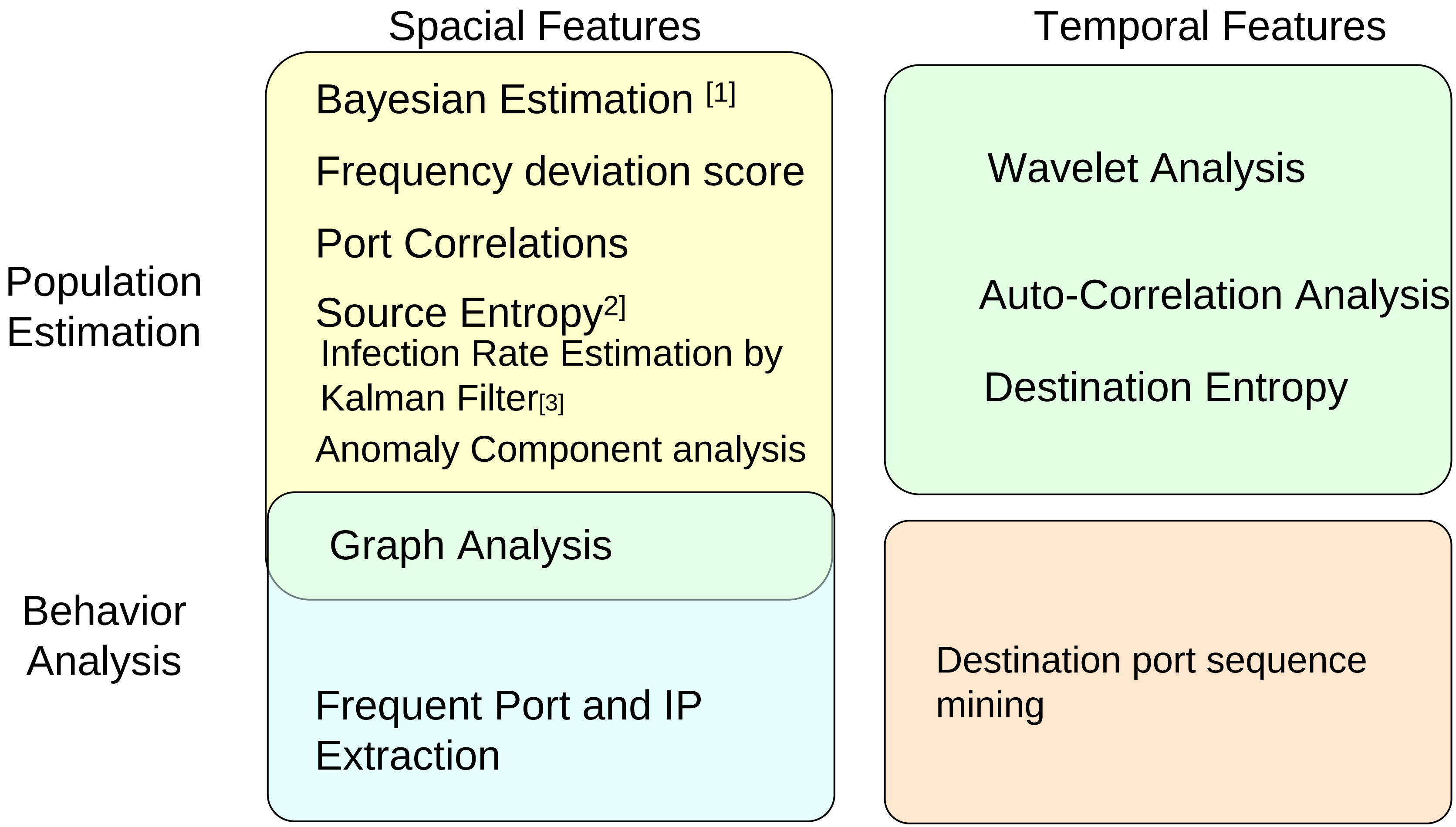
Date of Packet(Year, Month, Day, Time)
Protocol Type （TCP, UDP, ICMP）
Source IP Address
Source Port
Destination IP Address
Destination Port



IN-2001-11(MS-DNS)
2001/8/31

CA-2001-33(ftp)
2001/11/29

Related Work



	Spacial Features	Temporal Features
Population Estimation	Bayesian Estimation [1] Frequency deviation score Port Correlations Source Entropy[2] Infection Rate Estimation by Kalman Filter[3] Anomaly Component analysis Graph Analysis	Wavelet Analysis Auto-Correlation Analysis Destination Entropy
Behavior Analysis	Graph Analysis Frequent Port and IP Extraction	Destination port sequence mining

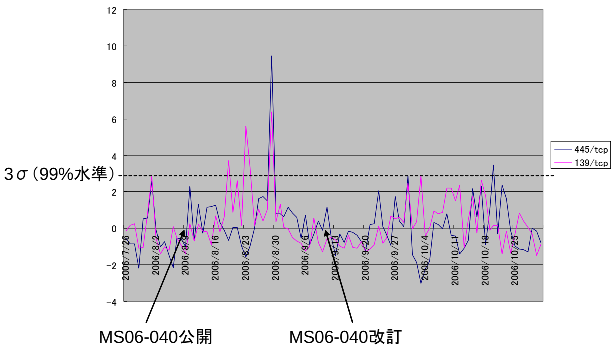
[1] Masaki Ishiguro et al, Internet Threat Detection System Using Bayesian Estimation, 16th Annual FIRST Conference on Computer Security Incident Handling, 2004

[2] C. Zou et al,"The monitoring and early detection of internet worms",IEEE/ACM Transaction on Networking,

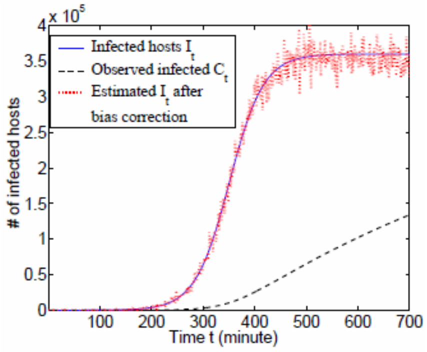
[3] Arno Wagner , Entropy Based Worm and Anomaly Detection in Fast IP Networks,14th IEEE International Workshop on Enabling Technologies

Evolution of Threat Evaluation Approach

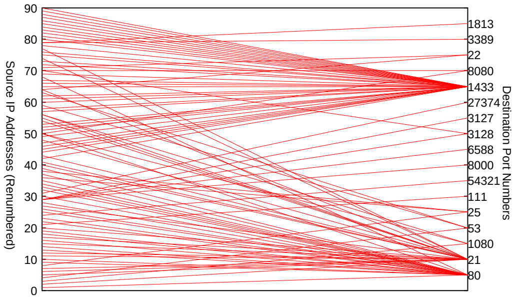
- Statistical analysis of Malicious Packet Counts



- Unique Source Address (Infected hosts)



- Analysis of Graph Structure
  - Consideration of vulnerability of destination ports as well as increase of unique source addresses



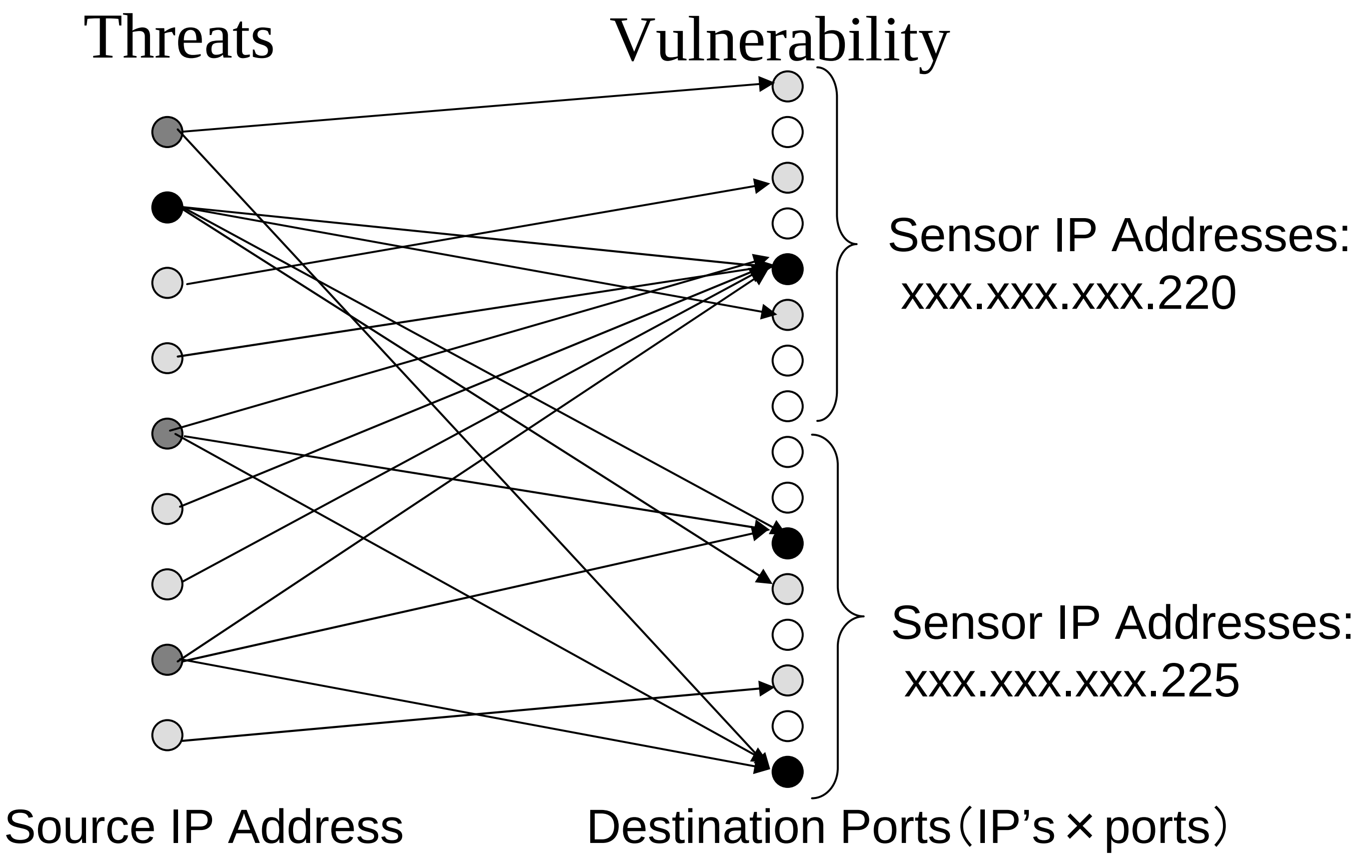
Relation between Threats and Vulnerability

Relationship 1

Vulnerability of a destination port is high if it gets access from many different source address with high threat level.

Relationship2

Threat level of a source address is high if it sends more packets to vulnerable destination ports.



Threat Calculation Method

Threat Vector $\mathbf{t} = (t_1, t_2, \cdots, t_n)$

Vulnerability Vector $\mathbf{v} = (v_1, v_2, \cdots, v_m)$

Relationship 1

$$\left\{\begin{array}{lcl} v_1 & = & c_1(w_{1,1}t_1 + w_{2,1}t_2 +, \cdots, w_{n,1}t_n) \\ & \cdots & \\ v_m & = & c_1(w_{1,m}t_1 + w_{2,m}t_2 +, \cdots, w_{n,m}t_n) \end{array}\right.$$

Relationship2

$$\left\{\begin{array}{lcl} t_1 & = & c_2(w_{1,1}v_1 + w_{1,2}v_2 +, \cdots, w_{1,m}v_m) \\ & \cdots & \\ t_n & = & c_2(w_{n,1}v_1 + w_{n,2}v_2 +, \cdots, w_{n,m}v_m) \end{array}\right.$$

$$\mathbf{v} = c_1 \underset{m\times n}{{}^tW} \mathbf{t}$$

$$\mathbf{t} = c_2 \underset{n\times m}{W} \mathbf{v}$$

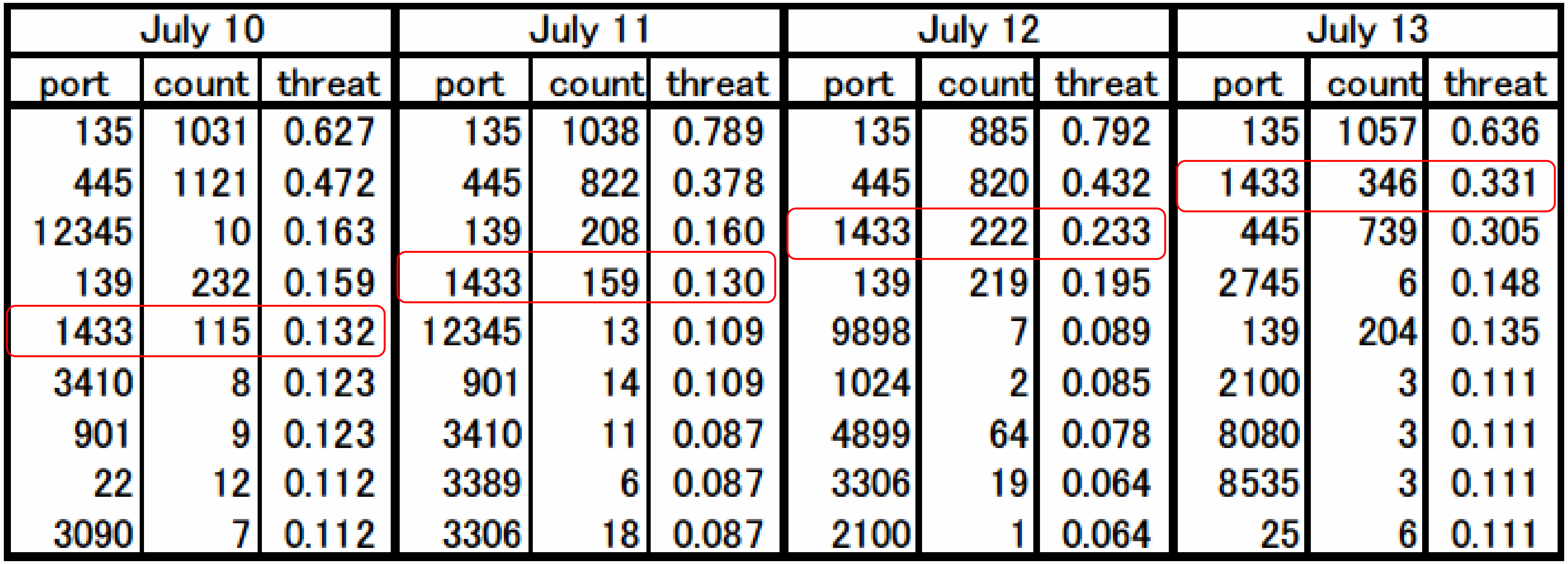
Eigenvalue Equations

$$\mathbf{v} = c_1 c_2 \underset{m\times m}{{}^tW W} \mathbf{v}$$

$$\mathbf{t} = c_1 c_2 \underset{n\times n}{W {}^tW} \mathbf{t}$$

Experiment1: Port1433 Incident (MS SQL)

- 2005/7



July 10			July 11			July 12			July 13		
port	count	threat	port	count	threat	port	count	threat	port	count	threat
135	1031	0.627	135	1038	0.789	135	885	0.792	135	1057	0.636
445	1121	0.472	445	822	0.378	445	820	0.432	1433	346	0.331
12345	10	0.163	139	208	0.160	1433	222	0.233	445	739	0.305
139	232	0.159	1433	159	0.130	139	219	0.195	2745	6	0.148
1433	115	0.132	12345	13	0.109	9898	7	0.089	139	204	0.135
3410	8	0.123	901	14	0.109	1024	2	0.085	2100	3	0.111
901	9	0.123	3410	11	0.087	4899	64	0.078	8080	3	0.111
22	12	0.112	3389	6	0.087	3306	19	0.064	8535	3	0.111
3090	7	0.112	3306	18	0.087	2100	1	0.064	25	6	0.111

Experiment2: Port 139 Incident (File Share)

- 2005/6

June 9			June 10			June 11			June 12		
port	count	threat	port	count	threat	port	count	threat	port	count	threat
135	2551	0.954	135	2174	0.883	135	2834	0.879	135	1906	0.846
445	751	0.209	445	1008	0.227	445	1308	0.244	445	989	0.249
1433	140	0.078	1080	4	0.104	12345	11	0.085	139	242	0.106
4899	43	0.052	44599	8	0.099	139	257	0.081	42857	2	0.102
1521	1	0.052	10589	4	0.099	21	4	0.077	4899	46	0.076
8535	1	0.052	8080	2	0.070	1433	142	0.065	143	1	0.076
8536	1	0.052	4899	47	0.070	44599	3	0.064	3306	9	0.076
2100	3	0.052	22	23	0.070	10589	3	0.064	1256	3	0.076
22	10	0.052	25	10	0.070	11524	2	0.064	2419	1	0.076
143	1	0.052	3306	4	0.070	42857	2	0.064	6346	3	0.076

Conclusion and Future Works

- We proposed a new threat evaluation method based on structure of access graph which are quite different from those based on the number of malicious packets.
- We demonstrated examples that our method responds better than the number of malicious packets

- Future Work:
- Optimization of edge weights of access graph
- Optimization of Unit time of our graph analysis
- Evaluation of Strength and weakness of our method depending on the types of incidents

ご静聴有りがたう御座ゐました