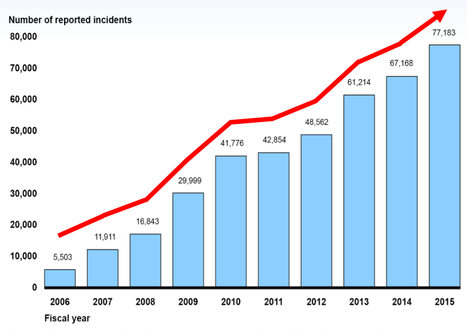
Analysis Between Cyberattack and Tweets

Junha Lee, Boyoung Han
IT Management, Seoul National University of Science and Technology

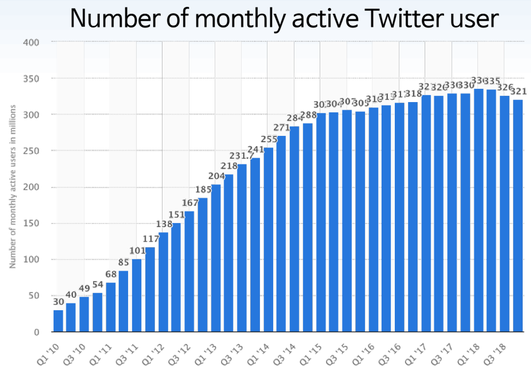
1. Background & Goal

Cyberattacks Reported by Federal Agencies



An increase in cyberattacks each year
→ Threats about cyberattacks are increasing

Number of monthly active Twitter user



The number of twitters sent every day
→ Lots of information in Twitter

➔ **Find Relations** between **Cyberattack** and **Hackers' Tweets**

2. Related Work & Differentiation

- Analyzing the Perceived Severity of Cybersecurity Threats Reported on Social Media
  - Analysis of text referring to the severity of cyberattacks online based on specific keywords
- Discovering Signals from Web Sources to Predict Cyber Attacks
  - Predict attack by analyzing cyber threats by analyzing articles posted on various websites through ML
- Detecting Denial-of-Service Attacks from Social Media Text : Applying NLP to Computer Security
  - Analysis of SNS users' responses to cyberattack(DDoS) using NLP models

➔ Common point : Predict cyberattack using SNS(Twitter) data

- **Different Point**
  - Focus on **Graph database**(user network)
  - **Time-series**(frequency of tweets)

3. Project Process ***For more details, feel free to ask us***

Data Collection

- User List
  - Provided by Recorded Future
- Tweepy
  - Twitter Developer API
  - Follow/ Following Relationship
- Scrapy
  - All historical Tweets Data
- News Data
  - Key-work : Hit by Cyberattack
  - Google News
  - Exact Attack date / Importance

Network Graph Analysis

- Follow / Following Relationship
- Neo4J
- Criteria
  - Random
  - Recorded Future
  - Betweenness Centrality
  - Closeness Centrality
- Clustering each Criteria

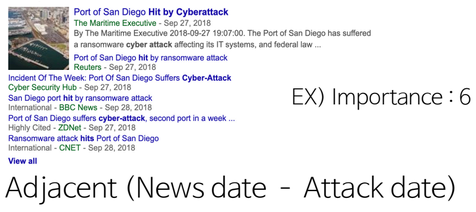
Criteria	Cluster 1	Cluster 2	Density
Random	65	1 …	65%
Recorded Future	95	1 …	95%
Betweenness	100	-	100%
Closeness	100	-	100%

Frequency Graph Analysis

- 2013 ~ 2018
- Criteria for Tweets Data
  - All
  - Keyword Filter

```
{
    "keyword": "take over Web server"
},
{
    "keyword": "telecom sit"
},
```

- Criteria for News
  - All
  - Importance



EX) Importance : 6

- Adjacent (News date - Attack date)

Relation Analysis

- Set the Period
  - Before the Attack Days
  - Attack Day
  - After the Attack Days
- Compare the Average
  - Year Average
  - Attack Average

4. Analysis & Conclusion

About Users : 4 Criteria (Described at the above part)

About Tweets : 2 Criteria (Described at the above part)

About News : 3 Criteria

	2013	2014	2015	2015	2017	2018	Total
All Attack	12	10	9	8	8	11	58
Important Attack	5	4	4	2	4	4	23
Adjacent Attack	7	7	7	3	3	3	30

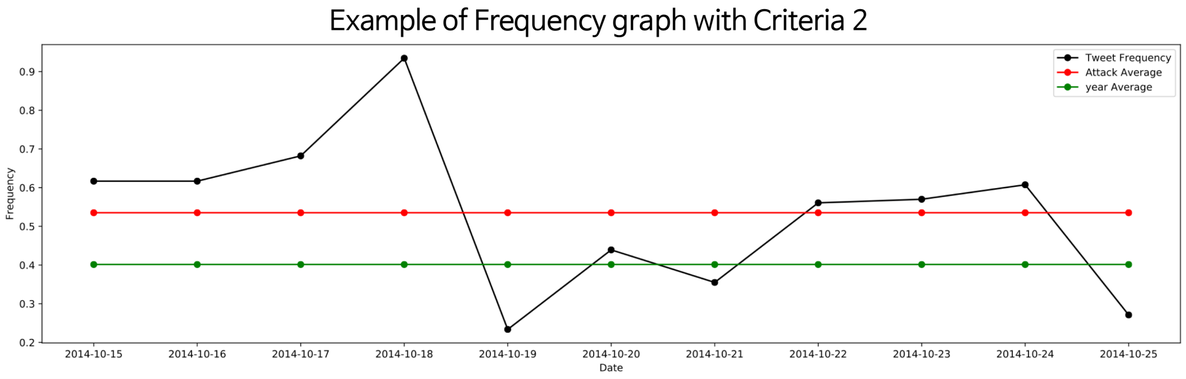
About Period : 3 Criteria

Criteria 1 : 15 Days (Before 7 days + Attack Day + After 7 days)
Criteria 2 : 11 Days (Before 5 days + Attack Day + After 5 days)
Criteria 3 : 7 Days (Before 3 days + Attack Day + After 3 days)

If an **Attack Average is higher** than Year Average, that attack <u>is assumed to</u> be **related to the tweet activity of users** involved in the Cyberattack.

(Attack Avg means the average of given period)

Example of Frequency graph with Criteria 2



The Result

R : Random / R_F : Random Filtered / RF : Recorded Future / RF_F : Recorded Future Filtered
BC : Betweenness & Closeness / BC_F : Betweenness & Closeness Filtered

		R	R_F	RF	RF_F	BC	BC_F
Criteria 1 : 15 Days	All Attack	38.89%	40.68%	42.37%	33.90%	44.07%	49.15%
	Important Attack	34.78%	47.83%	30.43%	34.78%	47.83%	43.48%
	Adjacent Attack	43.33%	50.00%	50.00%	43.33%	40.00%	53.33%
	Avg	39.03%	46.17%	40.93%	37.34%	43.97%	**48.65%**
Criteria 2 : 11 Days		R	R_F	RF	RF_F	BC	BC_F
	All Attack	30.51%	38.98%	40.68%	33.90%	42.37%	45.76%
	Important Attack	26.90%	43.48%	30.43%	34.78%	43.48%	39.13%
	Adjacent Attack	36.67%	46.67%	40.00%	40.00%	43.33%	56.67%
	Avg	31.36%	43.04%	37.04%	36.23%	43.06%	**47.19%**
Criteria 3 : 7 Days		R	R_F	RF	RF_F	BC	BC_F
	All Attack	37.29%	42.37%	44.07%	32.20%	40.68%	49.15%
	Important Attack	39.13%	52.17%	34.78%	30.43%	47.83%	47.83%
	Adjacent Attack	43.33%	50.00%	50.00%	40.00%	36.67%	60.00%
	Avg	39.92%	48.18%	42.95%	34.21%	42.25%	**52.33%**

1. Sort the Users involved in overall Cyberattacks by Betweenness Centrality
2. Filter the Tweets based on Keywords related to Cyberattacks

➔**More relevant to Cyberattacks**

5. Future Work

1. Subdivide the time series consider the time differences
2. Analyze Tweet text which was written BEFORE the attack by using NLP
3. Find deep relationship between users' Tweets and actual cyberattack

➔Possible to <u>detect</u> the potential Cyberattack

1. Subdivide the time series consider the time differences
2. Analyze Tweet text which was written AFTER the attack by using NLP
3. Find deep relationship between users' Tweets and actual cyberattack

➔Possible to <u>notify</u> the Cyberattack as soon as possible