3-5 Internet Alive Monitoring around Disaster Areas using Large-scale Darknet

Mio SUZUKI

During a large scale disaster, power outages and physical damage to items such as network devices and optical fiber can result in the Internet malfunctioning. In this report, we discuss the potential for Internet alive monitoring using large-scale darknet monitoring during such disaster and in reconstruction efforts.

1 Introduction

Since the Great East Japan Earthquake and Tsunami on March 11, 2011, there has been a heightened awareness in all areas regarding disaster preparedness. We the current authors have conducted monitoring of the darknet for many years in the field of cyber-security [1]–[3], but at present we are working on ideas for dealing with alive monitoring that estimates network dead-or-alive status during disasters by way of darknet monitoring, expecting that such monitoring results can be used during a disaster and in reconstruction efforts after such.

During a large scale disaster, power outages and physical damage to items such as network devices and optical fiber can result in the Internet malfunctioning. One can see that accurately and broadly surveying and identifying these malfunctions will be significant in reconstruction. However, because the Internet is an autonomously distributed network, information on those malfunctions is usually contained within the organization in which they were generated and it is difficult to gain cross-functional information from outside.

At the laboratory of the current authors, experimental operation of NICTER (Network Incident analysis Center for Tactical Emergency Response), which employs a large scale darknet monitoring network, is being carried out regularly. In the present report, there is a discussion of the potential for Internet alive monitoring through combining the information obtained through said experimental operation with other publicly-available information [4][5]. There is also a discussion of the potential for swiftly executing alive monitoring by cross-functionally working in units of prefecture, city, and AS to perform passive monitoring that does not put a strain on the network, followed by a discussion of setup and implementation for a system that automatizes monitoring.

2 Darknet monitoring

“Darknet” refers to the space on the Internet that can be reached and has IP addresses that are not yet used. Sending packets to unused IP addresses is something that is unlikely to happen within the sphere of normal Internet use, but in actuality there is a considerable number of packets that arrive at the darknet. Most of these packets originate from unauthorized activities on the Internet and represent, for example, backscatter from scans or exploit code sent out by remotely-infecting malware and responses to SYN flood attacks that spoof source IP addresses. Due to this, by monitoring packets that arrive at the darknet it is possible to ascertain trends in unauthorized practices occurring on the Internet. The greatest benefit of monitoring the darknet is the lack of need to differentiate between authorized and unauthorized traffic, as the majority of packets can be considered unauthorized.

When monitoring the darknet, a server machine for aggregating and responding to packets, called a “sensor,” is located inside the network that is to be monitored. Sensors are classified variously depending on the level of response to the packet source. From among those classifications, a representative sensor is the “black hole sensor.”

A black hole sensor is a sensor that does not respond at all to the packet source. This sensor is easy to maintain and is useful in large scale darknet monitoring. Another of its benefits is that, because the sensor does not respond to packets, it is difficult to detect its existence from the outside. However, while it is possible to monitor scans, which represent the early stages of malware infection activities, it cannot monitor behavior thereafter.

With NICTER, on which research and development is progressing at the laboratory of the current authors, black hole sensors have been established for multiple darknets that are sporadically located around Japan and the world (a total of over 300,000 addresses) and monitoring is regularly performed [1]–[3].

3 Network alive monitoring, using a large -scale darknet monitoring network, based on source address geographic information

This section will present a discussion on the potential for alive monitoring during a disaster by combining the aforementioned darknet monitoring technology with source address geographic information.

3.1 Basic idea

In a large scale darknet monitoring network, there are always massive amounts of packets coming in from all around the world (as mentioned previously, most of them are unauthorized). Among these, there are, of course, also those that are sent from hosts in Japan toward the darknet. It can be confirmed that hosts which have sent packets to the darknet are unauthorized and that they are active on the Internet. Thus, before and after the occurrence of a large-scale disaster, by mapping host groupings monitored on the darknet with physical geographic information there is potential for estimating dead-or-alive status of the Internet in the disaster-affected area and surrounding areas. In other words, the concept is that unauthorized traffic directed toward the darknet can be used against itself and, via passive monitoring that doesn't put a strain on the network, alive monitoring can be realized. The below will work to verify this basic idea through the use of actual data from the time of the Great East Japan Earthquake and Tsunami.

3.2 Number of unique hosts in the six prefectures of the Tohoku region during the Great East Japan Earthquake and Tsunami

Presented in this section are the results from aggregating fluctuations in the numbers of unique hosts in the six prefectures of the Tohoku region which were obtained by darknet monitoring before and after the Great East Japan Earthquake and Tsunami on March 11, 2011. The packets that arrived at NICTER's Class B darknet (approx. 65,000 IPv4 addresses) in the one month between March 1 and March 31, 2011 were used as the data for the basis of aggregation. Further, the GeoIP City Database (April 2011 version) of the company MaxMind was used to perform matching with geographic information from source IP addresses.

Figure 1 shows the number of unique hosts in one day for packets sent to the darknet from the six prefectures of Tohoku. As for the graphs from here on, unless otherwise specifically noted, the vertical axis represents the number of unique hosts that were monitored in units of one hour, the horizontal axis represents the date and time, and the red vertical line in the Figure represents the date and time of the disaster. From the graph it can clearly be seen that after the disaster there is a trend toward reduction, and after that there is a smooth transition back toward recovery.

Next, when narrowing it down to the unit of the prefecture it can be seen that, of the six prefectures of Tokoku with the exception of Aomori, there is a clear trend toward a reduction in the number of unique hosts in each prefecture. As one example here, Fig. 2 shows the number of unique hosts when focusing on one day in Miyagi prefecture. In this case, too, it can be seen that Miyagi generally follows the same trend as all six Tohoku prefectures.

Until here, darknet monitoring results in one-day units had been used, but in order to perform swift Internet alive monitoring, it is preferable to shorten the intervals between aggregations of unique host numbers. Thus, the intervals between the aggregations of unique hosts were shortened to ten minutes, and the application of the simple moving average (144 segments) is shown in Fig. 3. From this graph the reduction in unique host numbers immediately after the disaster and the recovery in numbers after it can be confirmed.

This is to say, by measuring the number of unique hosts that send packets to the darknet, there is the potential for swiftly estimating in around ten minutes the dead-or-alive status of the Internet in each prefecture.

3.3 Number of unique hosts at the unit of city level during the Great East Japan Earthquake and Tsunami

Until now at the laboratory of the current authors, the numbers of unique hosts in darknet monitoring had been aggregated down to the unit of prefecture but in this report such data was combined with the city-unit data of the GeoIP City Database and those results were aggregated to plot out the graphs.

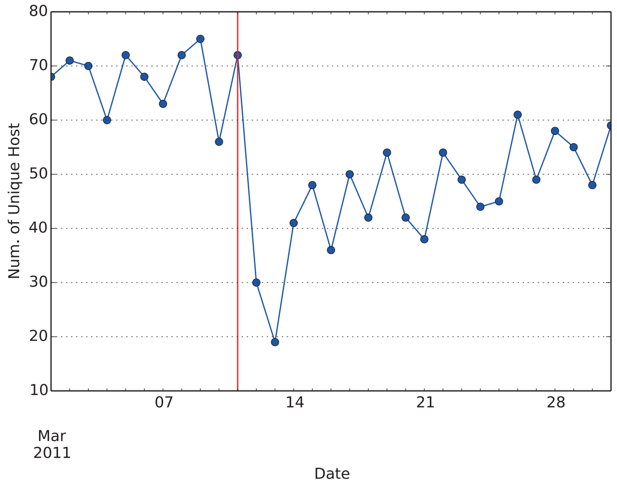

Fig. 1 Number of unique hosts in the six Tohoku prefectures in one day

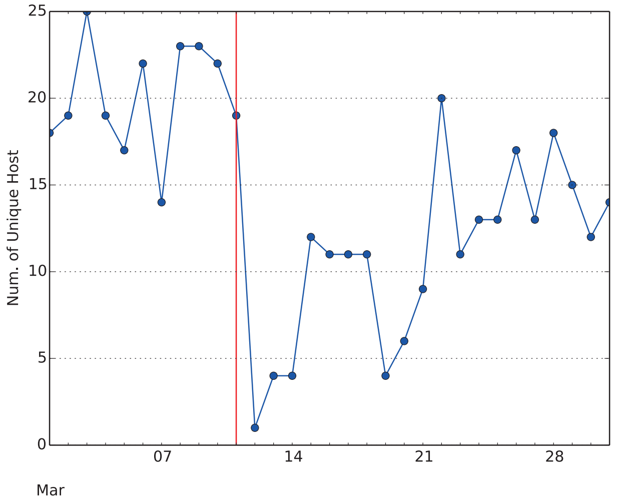

Fig. 2 Number of unique hosts in Miyagi prefecture in one day

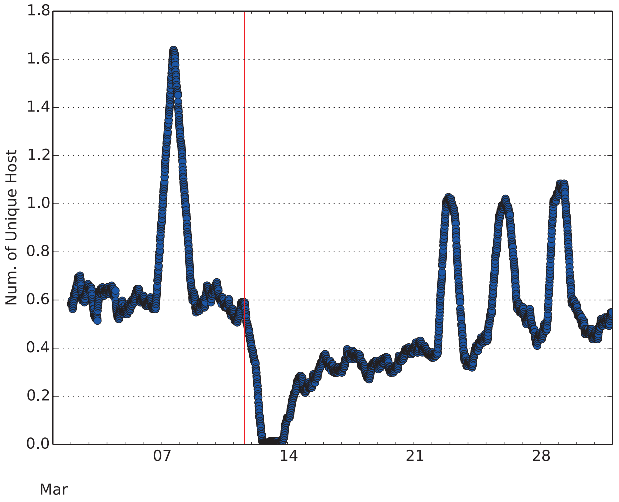

Fig. 3 Number of unique hosts in Miyagi prefecture in ten minutes (144 segment moving average)

First, because, as NICTER had shown in the past, there is somewhat of a correlation between the number of unique hosts being monitored and general population numbers, a focus was placed on Sendai city, Miyagi prefecture, which is the city with the largest population out of the six Tohoku prefectures. As a point of reference, the estimated population of Sendai city as of December 2014 was 1,074,125 people. Figure 4 shows the number of unique hosts in one day sending packets to the darknet from Sendai. Further, Fig. 5 shows a representation of a 24 segment moving average for the number of unique hosts in one hour, while Fig. 6 shows a representation of a 144 segment moving average for the number of unique hosts in ten minutes.

For all Figures, the vertical red line therein indicates the time of the disaster.

The graphs share the same tendencies as the graphs showing prefectural trends shown in the previous section.

However, in the graphs for unique host numbers for one hour and for ten minute units, the numbers for unique hosts being monitored are low at a maximum of, for example, 2.8 or 1.6, and due to random increase/decrease of multiple unique hosts, it can be expected that there unfortunately will be easy fluctuations in trends. In fact, when graphs for multiple other cities were consulted, trends could often not be read in the case of graphs with low maximal unique host numbers.

From these results, it can be said that there is potential for swiftly estimating the dead-or-alive status of the Internet in around ten minutes for aggregations on the city-unit level, too, in the case where there is a certain number of unique hosts.

4 Alive monitoring system setup and implementation

Based on the potential discussed in Section **3**, system setup and implementation was carried out in order to automatize alive monitoring. Figure 7 shows the flow of network malfunction estimation on the alive monitoring system.

In regards to the flow of network malfunction estimation on the alive monitoring system, first the malfunction estimation unit for aggregation is defined, and, after setting the system, the number of unique IP addresses is monitored in each of the aggregation cycles that were decided on for the malfunction estimation unit. In regard to the number of unique IP addresses for each of the aggregation cycles, prediction of future values is carried out using the time

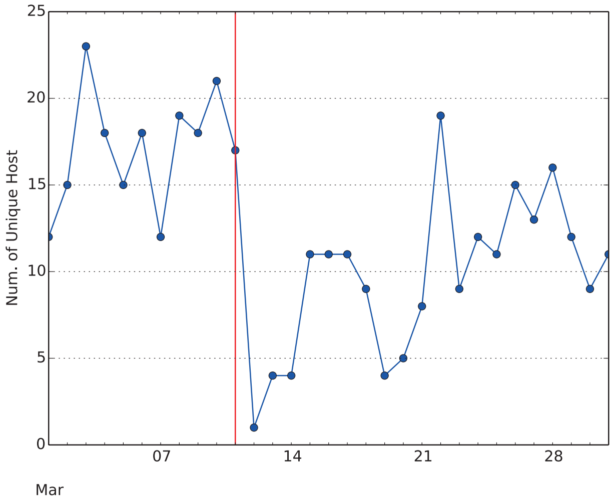

Fig. 4 Number of unique hosts in Sendai city, Miyagi prefecture in one day

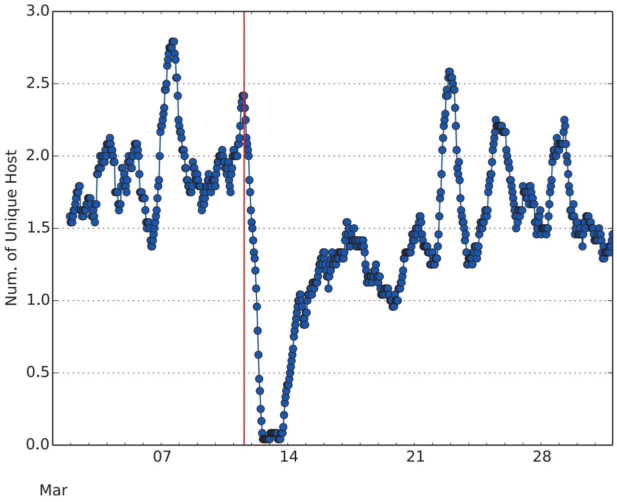

Fig. 5 Number of unique hosts in Sendai city, Miyagi prefecture in one hour (24 segment moving average)

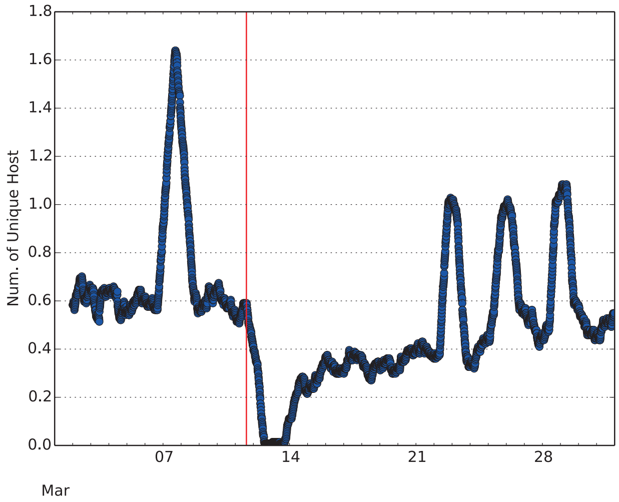

Fig. 6 Number of unique hosts in Sendai city, Miyagi prefecture in ten minutes (144 segment moving average)

series analysis method. Monitoring is continued in this way and malfunction is estimated when a deviation between the predicted value and the actual value is detected. The important elements in this flow of malfunction estimation—such as methods used for malfunction estimation unit, aggregation of number of unique IP addresses, and estimation of malfunction—are outlined from Subsection **4.1** onward.

4.1 Unit of malfunction estimation

As for the unit of malfunction estimation, in addition to the geographical regions mentioned in Section **2**, there could also be those of, for example, AS, IP network address, and domain name. By using these units with the additional information that is represented in the GeoIP City Database, there is the possibility of eventual change in the range of the address. Further, address ranges take on the shape of groupings. For example, the system is composed such that the United States, Canada, and Mexico are included in the "North America" group, and each country has multiple address ranges. These groups are registered in the system as malfunction estimation units, and monitoring and aggregation is carried out for each unit.

4.2 Monitoring and aggregation of number of unique IP addresses

In the alive monitoring system, in regard to the registered malfunction estimation unit, the number of unique IP addresses is monitored and recorded for each of the set time units. Based on the results discussed in Section **3** concerning unit times, the units are set at intervals of ten minutes, one hour, and one day.

4.3 Malfunction estimation method

In regard to the number of time sequence unique IP addresses for each time unit monitored and aggregated in Subsection **4.2**, model estimation is carried out using the ARIMA (Auto Regressive Integrated Moving Average) model, and prediction of future values is carried out based on that model. Further, updating of this model is carried out at fixed intervals. Monitoring is continued in this manner, and malfunctioning is estimated by detecting deviation between the predicted values and the actual values.

The ARIMA model is one of the time sequence methods that use previous data to predict the future. The ARIMA model is adapted for the data difference of the AR (Auto Regressive) model and the MA (Moving Average) model. This ARIMA model is used for short period estimation for

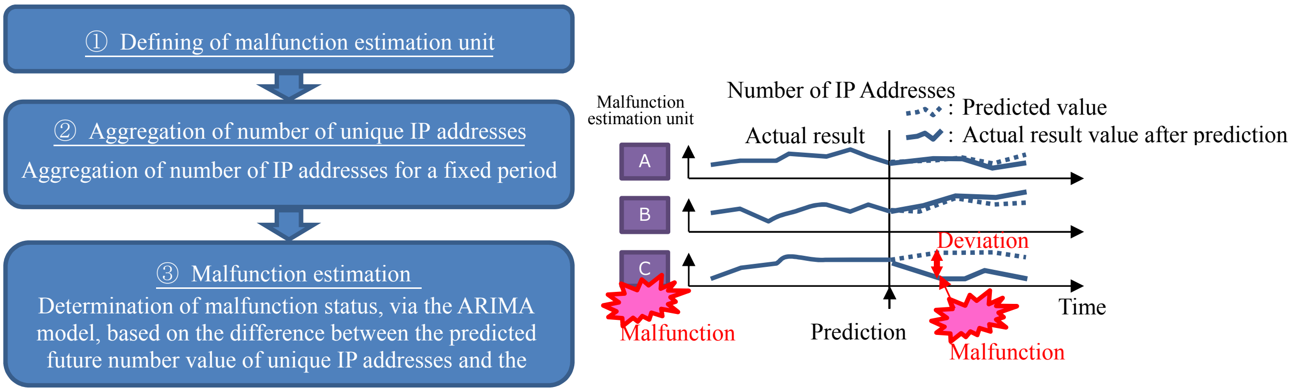
① Defining of malfunction estimation unit
② Aggregation of number of unique IP addresses
Aggregation of number of IP addresses for a fixed period
③ Malfunction estimation
Determination of malfunction status, via the ARIMA model, based on the difference between the predicted future number value of unique IP addresses and the
Malfunction estimation unit
Number of IP Addresses
Actual result
: Predicted value
: Actual result value after prediction
A
B
C
Deviation
Malfunction
Prediction
Time
Malfunction

Fig. 7 Flow of network malfunction estimation for alive monitoring system

cycles of ten minutes, one hour, and one day. Moreover, the numerical variation used in prediction is only for one previous piece of variable data.

Further, threshold values are set at the four stages of narrow and wide confidence segment upper and lower bounds in order to estimate malfunction. Each of the confidence segments have the likelihood of entrance into that segment defined. In that way, the monitored value of the number of unique IP addresses is estimated as a malfunction when it goes below the lower bound of the wide confidence segment.

5 Summary

Discussed in this report was the potential for swiftly estimating prefecture-, city-, and AS-unit cross-functional Internet alive monitoring with passive monitoring that does not place any strain on the network—by using the NICTER large scale darknet monitoring system that is in continued regular actual operation at the laboratory of the current authors—and by combining that data with other data. And, further discussed was the setup and implementation of systems for automatizing monitoring. As a result of considerations, if the number of unique IP addresses being monitored in the range of a given malfunction estimation unit is sufficient, it can be said that malfunction estimation is possible.

Moving forward, while continuing to discover issues and perform improvements through system operation, the current authors hope to build up a system that will be useful in the event of an actual disaster.

References

1 K. Nakao, K. Yoshioka, D. Inoue, and M. Eto, “A Novel Concept of Network Incident Analysis based on Multi-layer Observations of Malware Activities,” The 2nd Joint Workshop on Information Security (JWIS07), pp.267–279, 2007.

2 D. Inoue, M. Eto, K. Yoshioka, S. Baba, K. Suzuki, J.Nakazato, K. Ohtaka, and K. Nakao, “nicter: An Incident Analysis System toward Binding Network Monitoring with Malware Analysis,”WOMBATWorkshop on Information Security Threats Data Collection and Sharing (WISTDCS 2008), pp.58–66, 2008.

3 Koji Nakao, Daisuke Inoue, Masashi Eto, and Katsunari Yoshioka, “Practical Correlation Analysis between Scan and Malware Profiles against Zero-Day Attacks based on Darknet Monitoring,” IEICE Trans. Information and Systems, vol.E92-D, no.5, pp.787–798, 2009.

4 D. Inoue, J. Nakazato, J. Shimamura, M. Eto, K. NAkao,” A Practical Usage of Large-scale Darknet Monitoring for Disaster Recovery,” ICSS2011, Mar. 2011.

5 M. Suzuki, J. Shimamura, J. Nakazato, D. Inoue, M. Eto, K. Nakao, ”Internet Alive Monitoring Method around Disaster Areas Using Large-scale Darknet, Autonomous System Information, and Geographical Information,” ICSS2015, Feb. 2015.

Mio SUZUKI, Ph.D.
Researcher, Cybersecurity Laboratory,
Cybersecurity Research Institute
Network Security, Network Operation