

Big data-based data mining modeling for nuclear security analysis in the modified physical protection system

Tae Ho Woo*

Systemix Global Co. Ltd., 3 F, 494-48, Yonggang-dong, Mapo-Gu, Seoul, 121-876, Republic of Korea

**Corresponding author: thw_kor@hotmail.com*

1. Introduction

It is important to find out the terror plan before it really happens. Although there is a tendency to communicate for the plot of the terror among the terrorists which could be done by the telecommunication system like the mobile phone, the tapping is illegal in the normal situations. Therefore, it is necessary to make a system to detect the possible terror conspiracy without the tapping. In this study, the social networking-related strategies have been developed to prepare against the terror. Fig. 1 is the configuration for the normal communications between individuals by the mobile phones. Data mining is used as a major function for the algorithms which are combined with neural networking, systemic dynamics methods and big data. It is usual that terrorists make the terror incident after expressing and communicating with other companions. For the quantifications of the research, the exact tapping is replaced with the random number-generated quantity and its modifications. So, the result would be compared each other with the dimensionless numbers.

Usually an accident happens without any notifications, even though there are several causes. Otherwise, the terror incident happens by somebody, so to speak, terrorist. The terror case happens by any human intentional acts. So, the detection of the attempt of the terror attacks is a way for the data mining design. If the symptom of the terror incident is notified, the terror can be stopped by a relevant agency. Historically, the nuclear terror was not done in the massive scale previously, although the radioactive material poison attack happened using the polonium-210 previously [1]. Currently, the safety systems are equipped well in the commercial NPPs. But, the physical protection system (PPS) for the possible terror incident is not prepared comparatively much. There are some PPS like the barricade, security guard, and some more security procedures. The conventional PPS is composed of detection, delay, and response [2]. In the systems, each stage has particular characteristics where the detection state is to find out the intrusion and unauthorized behavior by some devices, the delay stage is to slow the malicious acts by some barriers, and finally the response stage is to interrupt the actions by several kinds of methods like the arms response [2].

It is said there are some data mining activities by the department of homeland security in the United States. The Automated Targeting System (ATS) compares several factors like traveler, cargo, and conveyance information against intelligence and other enforcement data, which is managed by U.S. Customs and Border Protection (CBP) and includes modules for inbound (ATS-N) and outbound (ATS-AT) cargo, land border crossings (ATS-L), and passengers (ATS-P) [3]. The Data Analysis and Research for Trade Transparency System (DARTTS) analyzes possible data like the trade and financial stuff, which is administered by U.S. Immigration and Customs Enforcement (ICE) [3].

2. Methods and Results

In the modeling of the study, the data mining is performed by the neural network incorporated with system dynamics (SD) method where the data feedback is applied to the neural network algorithm. The data mining technology is widely used in our lives for the commercialization, safety, security, and so on. Many applications in nuclear industry could give the reliability of the NPPs operations for the stable electricity supply.

2.1. Big data

This is used for the modeling in the basic data-bases which are shown in the modeling as input data. Since the topic should control many kinds of potential terrorism stuffs, the quantity is expressed huge as big data. Modern networking system in the society has been developed by the computer related systems like personal computer, mobile phone, and some other social networking systems. The processed data have increased in the human life following the system development. The big sized data processing is examined for the preventing nuclear terrorism in the NPPs. It is reasonable that highly developed networking system in current technology is utilized in this study. The social networking is used as a major role in simulating the nuclear accident. The big data mean the data more than terabytes in the computing memory [4]. In the previous work, Steed et al. studied for Exploratory Data analysis ENvironment (EDEN) in which the visual analytic system is performed for the analysis of complex earth system simulation data sets [5].

Regarding the methodological aspect, it is possible to use the big data in the case of the complex and complicated situations. For example, the traffic jam situations in the national migrations during the holidays of the 1st of year or Korean Thanksgiving Day seasons could be forecasted using the past experienced data. These data are really huge with many uncertainties like the weather, stock, and so on. However, the successfulness of the predictions are quite high comparing to the case without using the big data. In addition, it is also reasonable to use in the customer's trend in the department store for the deployment of the store. Usually, the 1st floor holds the luxury shop like the jewelry stuff, which could attract the rich person in the easy place to get there. Therefore, it is also reasonable to use the big data strategy in the terror related cases.

2.2. Cloud computing

The hidden layer is done as the cloud computing role in the modeling, which is in Fig. 2. The cloud computing is defined as a variety of computing concepts which are connected with many user's receiving systems like the mobile phone, personal computer, notebook, networking system and so on in the real-time communications [6]. So, this concept is used as the central control system in the modeling where the cloud behaviors as the management system. This looks like the satellite in the space which connects to many systems in the Earth.

In the aspect of applications, the cloud computing is used in the mobile phone which is attained by almost all persons. So, it is possible to connect with many people. That is to say, it is able to construct the social networking by the mobile phone-based systems where the cloud-like system is created. The cloud means the connections are accomplished by remote controlled system which looks like the cloud, because the cloud in the sky could see all things in the ground like the satellite in the space. So, it is reasonable to make the terror prevention system using the cloud computing.

2.3. Neural network

In modeling, this is shown as the networking, which is seen in the Fig. 2. There is one of important and famous algorithm in the networking quantification which is the neural networking system. This is a networking for an algorithm mimicking of biological neurons.

This has been used in the many kinds of technological systems where the connections of the objects are wisely expressed with some mathematical algorithm. Some famous controllers have been developed for the engineering systems. So, the terror-like social system could make use of the neural networking systems.

2.4. System dynamics (SD)

This is used in the data processing where the random sampling-based networking is used in the modeling as it is seen in the Fig. 2. The SD is applied for modeling in the study, which is one of non-linear algorithm. After developing by Jay Forrester in MIT, the algorithm has been used in the many fields of social humanities as well as science and technology. That is to say, SD is the application tools for the dynamical matters in the complex and non-linear systems [7].

This is one of complex algorithm with the feedback oriented information processing in which the data of information are manipulated by the operator. So, this logic can make the higher weighting in the suspicious data or systems which are the parts of terror incidents. There are many endeavors to develop in the industrial security as well as the safety using this method.

2.5. Data mining

This is expressed as the data characteristics, which is decided by the expert in the modeling. That is to say, the basic data would be decided as several kinds of random samplings. The data mining means the analyzing of collected information in the arbitrary data for the interested goal [8]. This has focuses on the financial, marketing, communication, and many more where the data make the operator decide the satisfactions of customers in marketing [8].

This is based on the big data where the traffic jam or department store cases are applied, which was explained in the section 2.1. Data mining is also used in the air traffic security for the potential terrorist could be detected using the face where the shape is changed by the emotional state of the passenger. The system of detecting a doubtful person in the airport has been commercialized for enhancing security in the airport.

2.6. Modeling

The modeling for study is constructed by the combinational algorithm of data mining based neural networking and SD methods. Fig. 2 shows the data flows from input layer to output layer via the hidden layer in the neural network system. The data numbers could exceed terabyte in the interested information processing. SD diagram is given in Fig. 3 in which there are 3 kinds of input data and hidden layer is shown as the processing of input and output data. Output data are written as 2 kinds of data. Output data mean as 4 kinds of faces which could happen as the nuclear terror incidents in [9]. Main characteristics of nuclear terrorism are the theft and releasing of radioactive nuclear material which are accomplished by intrusion or sabotage of matters in PPS. Table 1 shows the comparisons for raw score, banded system in Global Terrorism Index (GTI) 2011, and new index in which the raw data are given by weighting factor multiplied by

case numbers [10]. This means the tapping including the communications of possible terrorists is replaced with the random number generated quantity and its modifications. The quantified numbers are decided by the experts by subjective decisions. Table 2 is the classification of input and output data using a new index. In I1, the random number is generated between 0 and 1 for a normal distribution as mean of 0.3 and standard deviation of 0.3. Otherwise, the I2 gives as the conditional statement. If the random number between 0 and 1 is lower than 0.5, the value is 0. Otherwise it is 1. Following the mean and standard deviation values, the terror possibility could be changed. C12 is $C1 - C2$ which means that the input data of C1 are compared with the experts values of C2. If C1 is higher than C2, the output is positive. This expresses the tendency of terror incident would be higher than lower values or minus values. In O2, the values of C12 are integrated numerically with initial value of 0. Hence, this shows the accumulated terror possibilities. The meanings are explained in each case on Table 3. In hidden layer, the C1 is decided by the comparisons among the input data. C2 is made by the expert's judgment. Then, this is compared each other. So, the C2 could mean the current situation of terror incident. In output layer, the discrete value in O1 shows just the each result of the situation.

As a summary, the modeling is based on the combinational complex and artificial intelligence algorithms with mixings. This kinds of multi-systems are very common characteristics in modern science and technology, especially in 21st century. The highly and complicated technology combinations are used in the study. Non-exact solution of the method can be used by the best solutions using mixing of some methods.

3. Conclusions

The modeling has accomplished successfully the random number generation-based quantifications for nuclear security incorporated with several complex algorithms, which is one of popular strategies in the technological-social science areas. In the calculations, the data are processed in every 7.5 minutes in which this means the calculation is done as 8 times in one hour. So, the graph has 100 hours with 800 times data processing. This means that 35,673.516 graphs are available by the big data processing. That is,

$$1 \text{ terabyte} = 1,000,000,000,000 \text{ bytes} = 800 \times 2,500,000,000 \text{ bytes}$$

So, it is needed to process 2,500,000,000 bytes more for 1 terabyte. In the similar way,

2,500,000,000 bytes are done for 8 times in 1 hour. Then, 312,500,000 bytes are done for 24 hours in 1day. Then, 13,020,833.33 bytes are done for 365days in 1 year. Then, 35,673.516 years are done.

Therefore, about 35,673 years' data could be processed for the security. Input data are obtained by the random numbers in which Fig.4 shows one of input data. The higher values give the higher possibilities of terror incidents. Fig. 5 is the cloud computing data in which the values are $C1 - C2$ in each time. Fig.6 is the output data of the processing in the cloud computing. The value increases to -2.61741 in 63.125th hour. So, the terror possibilities would be highest in later time. In examination of the result, the graph of Fig. 6 shows the shape oscillates by time.

For real applications, the systematic research of the mobile phone based social networking service (SNS) is effective to prepare against the conspiracy of the terrorists. Fig. 7 is the procedure of nuclear security with data mining. In the last stage of this diagram, the international cooperation is suggested. The networking reaches to every country without interruption in which the effective performance can be accomplished by the multi-national information exchanges.

The realizations have been used in the airport for detecting the terrorist using the face shape by the big data-based data mining method. So, it is very reasonable to use in the nuclear terror preventions. In addition, the data control technique also could be applicable in the neutron population treatment in the reactor physics area where the non-exact solution of neutron behavior could be analyzed by the big data based algorithm like the data mining or neural network methods.

REFERENCES

- [1] Daily mail, Widow of poisoned Russian spy Alexander Litvinenko in tearful plea for cash to fund public inquiry into his death, The Mail on Sunday & Metro Media Group. 2013.
- [2] IAEA, Guidance for the Application of an Assessment Methodology for Innovative Nuclear Energy Systems INPRO Manual -Physical Protection, IAEA-TECDOC-1575 Rev. 1. IAEA, VIENNA, Austria, 2008.
- [3] U. S. Department of homeland security, Data Mining Report to Congress, Privacy Office, U.S. Department of Homeland Security, Washington, D.C., 2012, 2013.
- [4] P. Russom, Big Data Analytics, TDWI research, fourth quarter 6, 2011.
- [5] C.A. Steed, D.M. Ricciuto, G. Shipman, B. Smith, P.E. Thornton, D. Wang, X. Shi, D. N. Williams, Big data visual analytics for exploratory earth system simulation analysis. Computers & Geosciences 61 (2013) 71-82.
- [6] I. Ivanov, M. van Sinderen, B. Shishkov, Cloud Computing and Services Science, Springer, New York, NY, USA, 2012.
- [7] System dynamics society, The Field of System Dynamics. Albany, NY, USA, 2013.
- [8] Palace B, Data Mining, Technology Note prepared

for Management 274A Anderson Graduate School of Management at UCLA, Los Angeles, CA, USA, 1996.
 [9] C.D. Ferguson, W.C. Potter, A. Sands, L.S. Spector, F.L. Wehling, The Four Faces of Nuclear Terrorism. Center for nonproliferation studies, Monterey Institute of International Studies, Monterey, CA, USA, 2004.
 [10] Institute for Economics and Peace (IEP), 2012 Global Terrorism Index (GTI), Capturing the Impact of Terrorism from 2002 - 2011, Sydney, Australia, 2012.

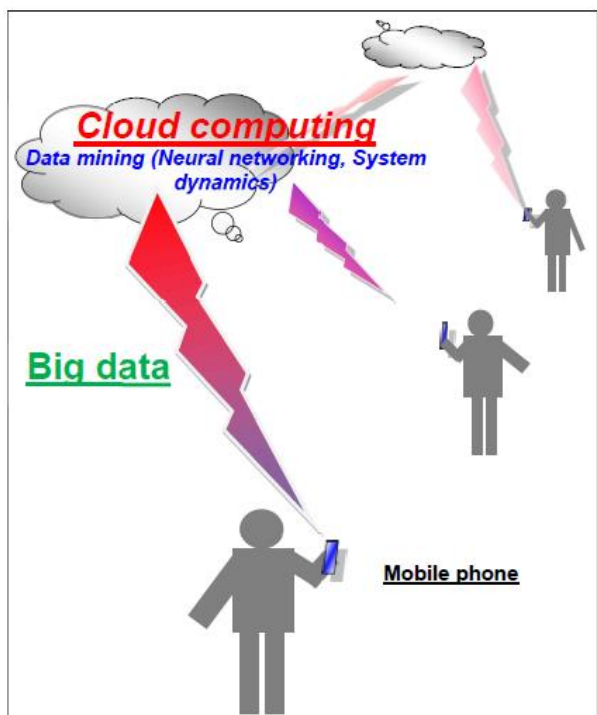


Fig. 1. Simplified configuration of cloud computing strategy.

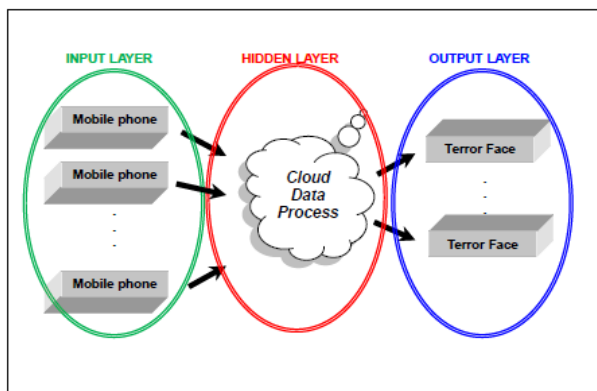


Fig. 2. Data mining systems by cloud computing strategy.

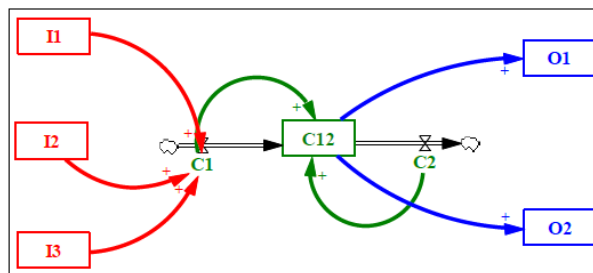


Fig. 3. Diagram by data mining of neural network incorporated with system dynamics.

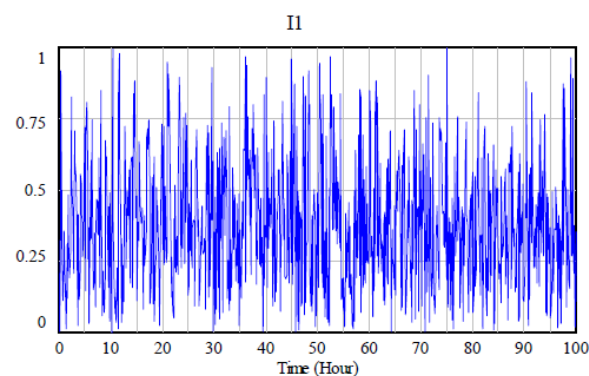


Fig. 4. Diagram of I1.

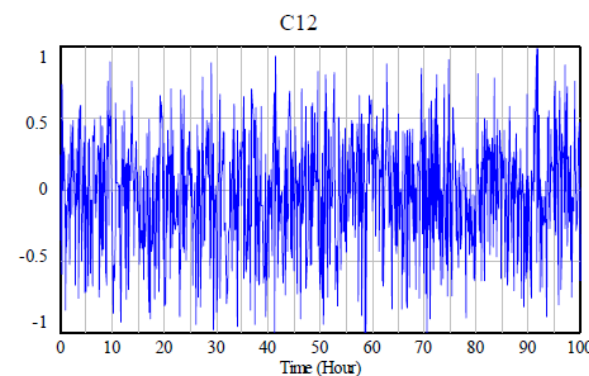


Fig. 5. Diagram of C12.

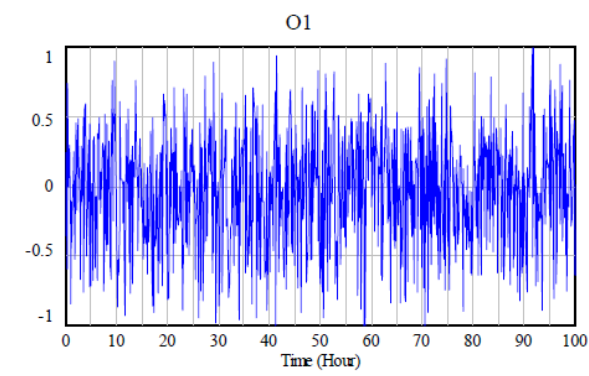


Fig. 6. Diagram of O1.

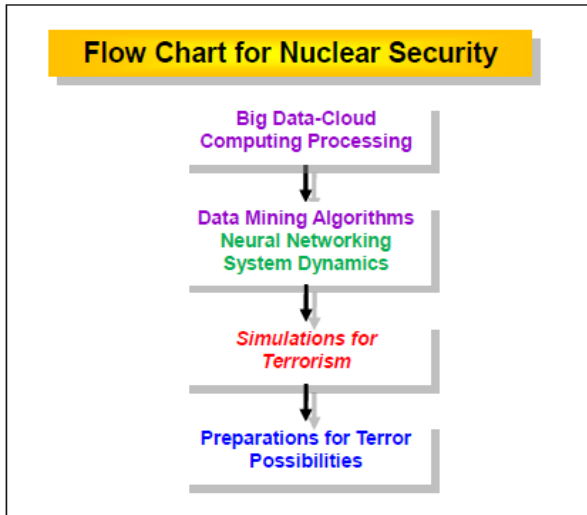


Fig. 7. Procedure of nuclear security with data mining.

Table I: Comparisons of raw score, banded system in Global Terrorism Index (GTI) 2011, and new index

Raw score	Global Terrorism Index (GTI) 2011	New index
0	0	0
.	.	.
6,000	5	0.5
.	.	.
12,000	10	1

Table II: Classification of input and output

Variable	Quantification
I1	Random # normal [0.3, 0.3]
I2	If random # < 0.5 [0, 1]
I3	Random # normal [0.6, 0.3]
C1	If random # < I1 [if random # < I2 [0, 1], I3]
C2	Random normal [0.6, 0.3]
C12	C1-C2
O1	C12
O2	INTEG [C12, 0]

Table III: Meaning of classification

Variable	Meaning
I1	Randomly happening event with specified mean and st. deviation
I2	Conditionally happening event
I3	Randomly happening event with specified mean and st. deviation
C1	Conditionally happening event (Input data are compared)
C2	Randomly happening event with specified mean and st. deviation (Expert's decided standard data are produced)
C12	Compared between C1 and C2
O1	Discrete output data
O2	Accumulated output data