Improving Cyber Situational Awareness Through Data Mining and Predictive Analytic Techniques

Sina Pournouri(✉) and Babak Akhgar

The Cultural, Communication and Computing Research Institute,
Sheffield Hallam University, Sheffield, UK
{Sina.pournouri,babak.akhgar}@shu.ac.uk

Abstract. Due to the widespread usage of computer resources in everyday life, cyber security has been highlighted as one of the main concerns of governments and authorities. Data mining technology can be used for prevention of cyber breaches in different ways and Cyber Situational Awareness (CSA) can be improved based on analyzing past experiences in terms of cyber-attacks. This paper aims to investigate and review current state of CSA improvement through data mining techniques and predictive analytic and offers possible methodology based on data mining techniques which can be used by cyber firms in order to secure themselves against future cyber threats.

Keywords: Cyber security · Cyber threat · Cyber situational awareness · Data mining techniques

1 Introduction

Nowadays computers play crucial role in everyday life. Private and public organizations, banks, governments, Law enforcement agencies, intelligence services benefit from using computers in order to fulfil their business objectives. However this dependence to computer systems and digital resources also gives opportunity to cyber criminals in order to fulfil their aims too.

Cyber-attacks have huge implications on companies and governments. For instance unknown hackers targeted Sony Pictures and as a result some of their production including movies, contracts and market plans were leaked. Analysts predicted Sony financial loss was around 83 million dollars. (Savov 2014) Damages caused by cyber-attacks are not only always financial but also can lead to loss of reputation, customers and partners.

Prevention is always better than curing, so managers and authorities are always attempting to find an efficient way to be prepared and stay secured against current and future cyber-attacks. One of the common methods is applying security standards and policies including cyber security awareness programs. "How to improve cyber security awareness" is a significant challenge for security experts.

Understanding trends of cyber security can be divided into 2 different levels as follows:

© Springer International Publishing Switzerland 2015
H. Jahankhani et al. (eds): ICGS3 2015, CCIS 534, pp. 21–34, 2015.
DOI: 10.1007/978-3-319-23276-8_3

(1) Detection of weakness and bugs in the system: This step can be taken by security specialists by examining systems using penetration tests in order to find security bugs and weaknesses. By detection of weaknesses in the system, security managers can implement and design effective and solid security standards and procedures. In addition in order to fill technical bugs and gaps, security patches and equipment will be installed.

(2) Identifying cyber hackers and their methods: This level completes previous stage and it aids security managers to be aware of cyber-attacks recent methods. The concept of cyber-attack analysis will be highlighted in this stage in other words by analysing past historical cyber-attacks to cyber firms and finding relationship between different involved factors, a better landscape will be obtained and let managers to make effective decisions based on recent cyber threats.

2 Cyber Security

In order to understand the concept of cyber security, threats and possible breaches should be defined. Cybercrime is not a new phenomenon and all of cyber firms must be aware of that threat. Not only previous cyber-attacks should be taken into consideration but also there is always need for an efficient strategy against future cyber-attacks as they are getting more and more sophisticated.

FBI and Department of Justice in January 2012, Sony in May 2011 (Aspan and Soh 2011) and Citi Bank in June 2011 were victims of cyber-attacks and it proves that any organizations and businesses regardless of their size can be a potential target for cyber attackers (Das et al. 2013).

There are different categorization of cyber threats but mainly they are divided in following categories (Nikishin 2004):

1. Virus: a piece of malicious program which attacks to computer systems and spreads itself to different parts including disk without users' knowledge.
2. Worm: a malicious program which penetrates to the system and causes interruption in the process of computers.
3. Trojan horse: a program which has no sign of threat to computer systems in first place but can cause destruction and damage to systems.
4. Logic bomb: a piece of malicious program that penetrates to the systems and executes and damages at a specific date and time.
5. Key logger: key loggers are used to save key strokes and have multi usages. On one hand organization and companies use them for safety and monitoring employees and on the other hand cyber criminals install them on the victim's system to steal information.
6. DOS: Denial of Service attack is a malicious method targeting availability of the victim's server by sending too many request to it. After a while the server will be taken down by cyber attackers.
7. SQL injection: a piece of malicious codes that tries to compromise the database of a system and steal stored information.

8. Zero day threat: this threat refers to an unknown bug in a system which security experts are not aware of it or have not patched it yet and cyber attackers use it as a penetration gateway.
9. Phishing: refers to a malicious method tries to steal sensitive information such as credit card information and so on by tricking a victim through an electronic communication.

Above, cyber threats have been defined and cyber attackers and cyber role should be defined. According to Awan and Blakemore (2012) cyber attackers are divided into following groups:

1. White hat: This group also is defined as cyber security experts where they are hired by organizations and businesses to test their security standards and demands. The task of white hackers is identifying current and potential weaknesses of the computer and network systems through various approaches such as black box test or white box test. After identification of those weaknesses they present efficient solution in order to address them.
2. Black hat: This group refers to those hackers who use their abilities to attack systems and obtain unauthorized and sensitive information. Although Black hat hackers' motivations do not always focus on stealing information, sabotage and damaging to the systems are other motivations. (Jaishankar 2011). Cyber terrorists can be a subgroup of black hats whereas their motivations are illegitimate. According to Lewis (2002) those cyber attackers targeting critical infrastructure such as power, government operations in order to make public fear, are defined as cyber terrorists. There is some disagreements and dissimilarities between sociologists' definition of "cyber terrorism" term. Some authors like Aviksoo (2008) and Pollit (1998) define cyber terrorists as type of cyber attackers who follow political and social interests and carry cyber-attacks to achieve them. On the other hand some authors such as Cox (2015) the term of cyber terrorism makes sense when human casualties are the main risk as a result of cyber-attacks. However all of them agree on the type weapon for this act which is a computer and type of target which is critical infrastructures.
3. Grey hat: grey hat hackers can be a categorized in both previous groups. In other words they are judged based on the result of their performance whether is peaceful and leading to improvement of security standards or harmful and leading to security breaches.

3 Data Mining

According to Ledoltar (2013) data mining is increasingly used in everyday life when customers and data have become strategic goals. Data mining is a method of analysing, extracting and discovering useful information form huge amount of raw data. Ahlemeyer-Stubbe and Coleman (2014) suggest that managers try to use data mining methods for prediction of future behaviour business. Therefore cyber security is not an exception between those subjects which benefit from data mining techniques.

According to Dean (2014) and Odei Danso (2006) data mining methods are divided into two main categories; supervised and unsupervised.

Supervised method attempts to measure relationship and similarity between from known input and output. A threshold will be defined and based on first result and the comparison between threshold and error level, some changes will be applied to learning process to get better output. On the other hand unsupervised method tries to discover hidden pattern in input set and there is no need for modification. Figure 1 shows difference between unsupervised and supervised learning.

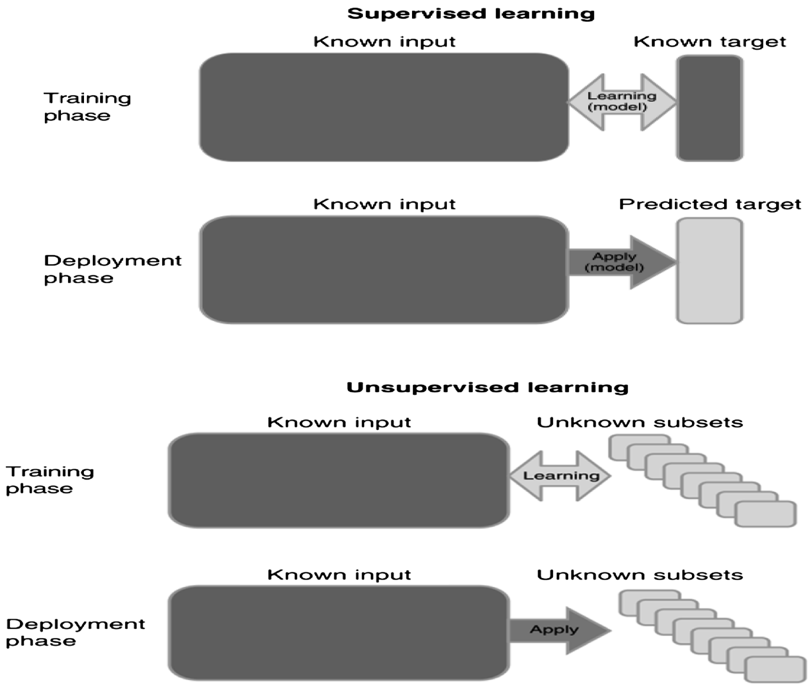

Fig. 1. (Ahlemeyer-Stubbe and Coleman 2014)

Data mining also includes different knowledge discovery techniques as follows (Dean 2014):

1. Regression analysis: This technique tries to establish a function leading to model the data.
2. Association rules: Refers to a technique used to discover interesting relationship among different variables in a data set.
3. Classification: Classification techniques are mainly used to classify data set into different subgroup and the result can be interpreted as a predictive model.
4. Clustering: This technique tries to arrange similar object in a specific groups based on their similarity factors.

4 Cyber Situational Awareness

According to Dua and Du (2011) in order to fill the present gaps in cyber security and deal with recent threats, an effective collaboration between cyber specialists and agencies is needed. These days cyber security researchers intend to design a solid and efficient framework maintaining confidentiality (the effort of keeping information secret between eligible and authorized parties and protecting it from unauthorized parties), Integrity (the ability of compatibility and accuracy of information) and Availability (Accessibility to cyber infrastructure and information) to protect computer and network systems (Dua and Du 2011). Cyber situational awareness is one of the frameworks designed by cyber security experts in order to preserve cyber security's interest and prevent from any sort of security breaches.

The definition of Situational Awareness should be taken into consideration. Situational Awareness is often described as an understating different factors in an environment which leads to predict and precept the near future events and trends for decision makers (Antonik 2007). In other definitions from other authors such as Tada and Salerno (2010) and Harrison et al. (2012) the factor of time plays crucial roles, in other words the time in situational awareness is coming with past information and learning from failures in order to analyse and extract any possible relations among them for a deeper and clearer understanding of future condition and situation. Situational Awareness (SA) is mainly divided to 2 different aspects as follows:

1. Cognitive aspect: from cognitive point of view, SA is mainly concerned with human perception. Endsley (1995) suggests that SA comes down to three main criteria: Basic perception of important data, Interpretation and conversion the data to knowledge and capability of using found knowledge for prediction of near future.
2. Technical aspect: In terms of technical according to Bryrielsson (2006) and Arnborg et al. (2000) SA is a combination of three main factors; arrange, analyse and integrate information as Arnborg et al. (2000) concentrates on arrange meaning collecting the data which suits the main demands and Bryrielsson (2006) reports that analyse and integration are two significant criteria in SA.

Franke and Bryielson (2014) and also Weick et al. (2005) suggest that Cyber Situational Awareness (CSA) can be a sub group of SA where the environment is cyber space and also in order to CSA, Data from IT equipment will be gathered and converted to suitable format for processing stage and that leads to better decision making. According to the study conducted by Weick et al. (2005) cyber sensors play prominent role gathering data for CSA improvement purposes in a deeper and detailed condition such as logs and data recorded by Intrusion Detection Systems.

Barford et al. (2010) categorizes existing methodologies for improving CSA into two main categories:

(1) Low level: Low level of improvement of CSA includes more technical factors rather than other factors including human factors. Vulnerabilities assessment, damage assessment and alert correlation are significant factors in the low level. For example security experts can correlate alert extracted from IDS and

vulnerabilities of their system in order to better understanding of current situation and predicting future issues occurring in the network.

(2) High level: High level of CSA is more general than low level in other words it is the combination of human elements and technical factors. Human elements includes human resources and human interference.

5 Existing Approaches

Ahn et al. (2014) suggest that big data analytic using machine learning, Artificial Intelligence and so on can benefit the improvement of CSA. Future and unknown attacks can be predicted by 3 main approaches:

(1) Classification: cyber experts use classification techniques to classify past cyber-attacks in order to define level of current and future threats in terms of cyber security.

(2) Regression analysis: regression analysis can help prediction of future cyber-attacks by probing similar behaviour among collected data from past. In other words regression analysis is a type of data mining technique which tries to find any possible patterns between different data based on their similarity and extends that pattern in order to predict the future.

(3) Relation rules: association techniques can find relationship between collected data and detect anomaly behaviour. For instance IDS alerts can be collected and by using association techniques anomaly behaviour can be detected from collected packets in the network.

Figure 2 shows the architecture of their approach. In the first step, event data from IDS, log files and network devices will be collected and then they will be formed into suitable and usable appliance for processing purpose. The third step includes applying various data analytic methods including regression analysis, association rules and classification techniques and so on. The result of third step feeds to fourth stage which is the process of interpretation for managers and make the valuable information meaningful for decision makers.

Although Ahn et al. (2013) do not purpose a real time algorithm, they introduce significant framework for obtaining valuable information raw collected data in order to improve CSA.

i-Hope framework is another approach presented by Das et al. (2013) trying to improve CSA by predicting cyber threats through formulation of goals as hypothesis and corroboration of them with mathematical and statistical analysis. Das et al. (2013) use CSI/FBI survey as their main resource of raw data and try to prove their hypothesis by applying Generalized Linear Model and it is a mathematical model used to predict uncertainty.

Das et al. (2013) formulate four main hypothesis as follows:

(H1) By increasing defence equipment within the system, the probability of cyber breach will be reduced.

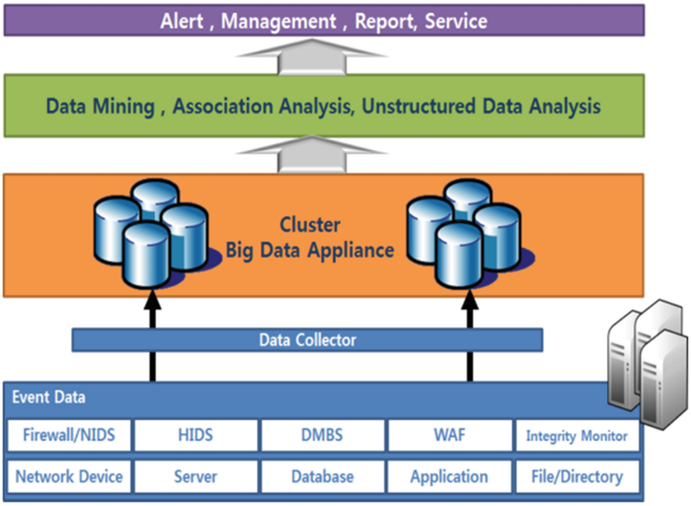

Fig. 2. Big Data analysis system architecture (Ahn et al. 2014)

(H2) By reporting the first attack to law enforcement agency, the probability of second or future attacks can be reduced.

(H3) Increasing IT budget leads to decreasing the chance of cyber-attacks.

(H4) By Increasing IT security outsourced, the probability of cyber-attacks will be decreased.

Das et al. (2013) apply chi-square and Deviance values to their data into a GLM and by interpretation of obtained models following answers are concluded to their hypothesizes:

Specify type of attack can be stopped only by installation of specify security equipment.

Reporting the first cyber-attack to Law enforcement agency does not have any influence on probability of future attacks.

The chance of security breach in a firm can be reduced by increasing ITO and IT budget.

Bayesian network has been suggested by Wu et al. (2013) in order to predict abd prevent cyber-attacks. Wu et al. (2013) take environmental criteria in target side into consideration and correlate them through Bayesian network analysis. These criteria are as follows:

Identify vulnerabilities and weaknesses: in order to this task in the system, powerful scanning tools such as Nessus are needed to identify security bugs.

The Usage Situation of Network: through this criterion, the usage condition of the network can be measured. The usage condition means the load of traffic which each node or device should deal with or number calls any of them received.

The Value of Asset in the Network: this can be done through an examination of each node in the network. Type of contained data and type of the task of each node will determine the value them.

Attack History: this attribute shows which nodes are more likely to be targeted by attackers based on previous observations in terms of cyber breaches.

Dut et al. (2012) propose an approach to defend against cyber-attacks based on Instance Based Learning Theory (IBLT). IBLT is a method providing accurate prediction of human behaviour. IBLT contains a storage called Instance including 3 main factors: (A) situation: the knowledge of features describing an attack. (B) Decision: an action taken against an attack. (C) Utility: the measurement of expected result of an action against an attack. Dut et al. (2013) reports that IBLT focuses on 3 main behaviours; Defender behaviour, Adversarial Behaviour and Tolerance level to threats. Therefore they try to improve CSA through behaviour of main key roles in a cyber-attack. Although this PhD project does not have accessibility to defenders behaviour, by analysing past cyber-attacks, cybercriminal behaviours will be analysed.

Morris et al. (2011) proposed an approach based on collecting intelligence and predicting future cyber-attacks through different levels of cyber missions. This approach concentrates on data gathering and giving solid result to decision makers level to improve CSA. Figure 5 shows their proposed method. This approach includes 8 different levels:

Mission tools: in this level regarding to type of attack, tools and general policies will be defined.
Mission need: based on outcome of level one, needs and different requests will be determined in order to combat against the cyber-attack.
Mission question: in order to identify weakness of the system, various questions will be brought up.
Mission area: it includes monitoring, Indication and warnings and counter intelligence. Based on questions which parts of cyber defence system should be activated.
Mission activity: in this level more specific activity against the cyber-attack will be agreed on.
Mission capability: in this level capabilities will be discussed based on taken actions.
Mission resources: type of resources needed by decision makers will be considered in this level.
Mission sources: the main concern of this level is where raw data comes from.

Musliner et al. (2011)'s approach is based on fuzzy logic which is an Artificial Intelligence technique. They divide the model into two various parts:

Proactive: Identification of weaknesses and vulnerabilities within the system in real time is the main task of this part.

Reactive: Identification of possible prevention techniques against cyber-attacks is the concern of reactive part.

The result of these two parts will be encoded into the fuzzy logic interference system as facts and rules. The outcome of the fuzzy system will build an automatic cyber shield against cyberattacks. Figure 3 shows Musliner et al. (2011)' s approach.

Schreiber-ehle and Koch (2012) suggest using JDL model of data fusion for improvement CSA in cyber defence. Data fusion is the process of combination different pieces of information in order to make them meaningful for better understating of different issues. Data fusion includes different sub-processing stages such as recording,

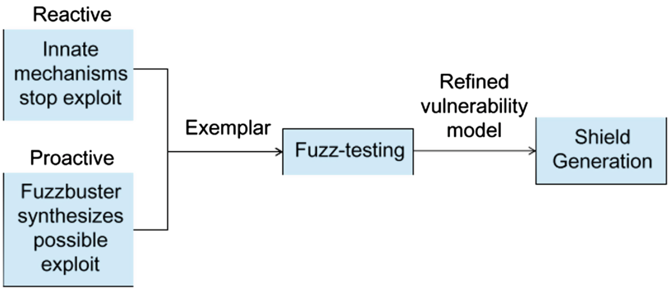
Reactive
Innate mechanisms stop exploit
Proactive
Fuzzbuster synthesizes possible exploit
Exemplar
Fuzz-testing
Refined vulnerability model
Shield Generation

Fig. 3. (Musliner et al. 2011)

storing, filtering, analysis and suitable projection of result of analysis. The JDL model of data fusion is a cohesive model showing each components of data fusion in a organised form helping manages to comprehend valuable information. Schreiber-ehle and Koch (2012) apply JDL model to cyber defence factors. Figure 4 Shows proposed JDL model of their methodology which has 6 levels as follows:

Level 0: is a component accepting input from monitoring sensors within the system. IDS is one of those sensors which monitors activities in the system and in case of suspicious one, it will notify the system administrator.
Level 1: level 1 aims to refine inputs from level 0 in other words those raw data extracted by level 0 needs to be pre-processed and allocated to specific objects on the system. For instance log files and IP addresses recorded by IDS need to map to relevant nodes or objects in the network.

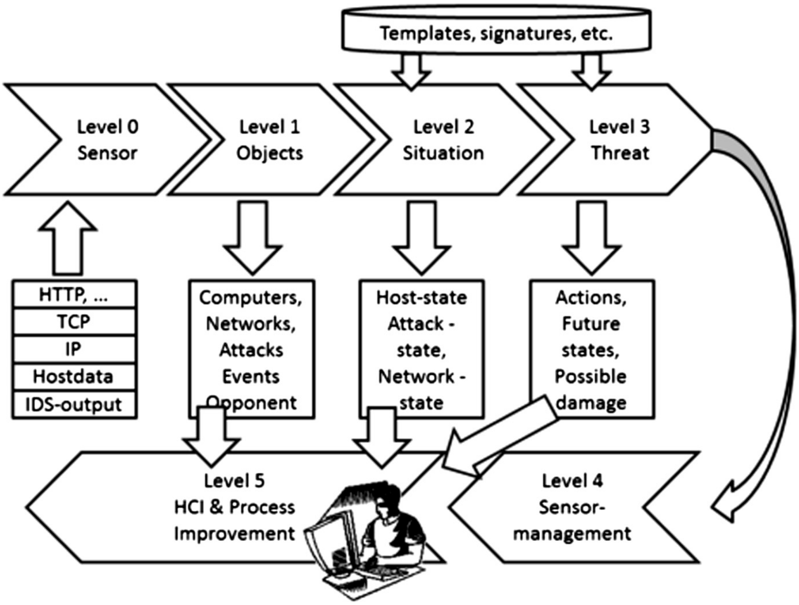
Templates, signatures, etc.
Level 0 Sensor
Level 1 Objects
Level 2 Situation
Level 3 Threat
HTTP, ...
TCP
IP
Hostdata
IDS-output
Computers, Networks, Attacks Events Opponent
Host-state Attack - state, Network - state
Actions, Future states, Possible damage
Level 5 HCI & Process Improvement
Level 4 Sensor-management

Fig. 4. (Schreiber-ehle and Koch 2012)

Level 2: level 2 investigates for current relationship between cyber entities through different types of analytic algorithms such as classification, clustering and so on.
Level 3: level 3 is the stage of prediction of future situation based on current and past condition in terms of enemies, threats, vulnerabilities, weaknesses and possible future operations to combat them. Information at this level is extracted from known attacks, signatures and templates and so on.
Level 4: level 4 operation includes observation of overall data fusion to maintain the system performance and try to improve it if it is feasible. Sensor and resource management is the main task at this level.
Level 5: level 5 provides Human-computer interaction where the process can be modified and refined by human experts.

Basically Schreiber-ehle and Koch (2012) draw general model of improvement of CSA and it can be suitable for projects in big scale.

Fayyad and Meinel (2013) design a methodology for prediction of new attack scenarios leading to improve CSA. Figure 5 shows their methodology. Fayyad and Meinel (2013) use three main resources of data; IDS data base, Attack graphs and vulnerability data base. IDS records all of the alerts and by applying clustering and aggregation algorithms, they will be formed into suitable format for correlation process. After correlating alerts, they will be stored in another built data set. Now it is time for processing attack graphs data. Fayyad and Meinel (2013) defines attack graph as “is a directed graph has two types of vertices, exploit and condition. An exploit is a triple (hs, hd, v), where hs and hd represent two connected hosts and v a vulnerability on the destination host”.

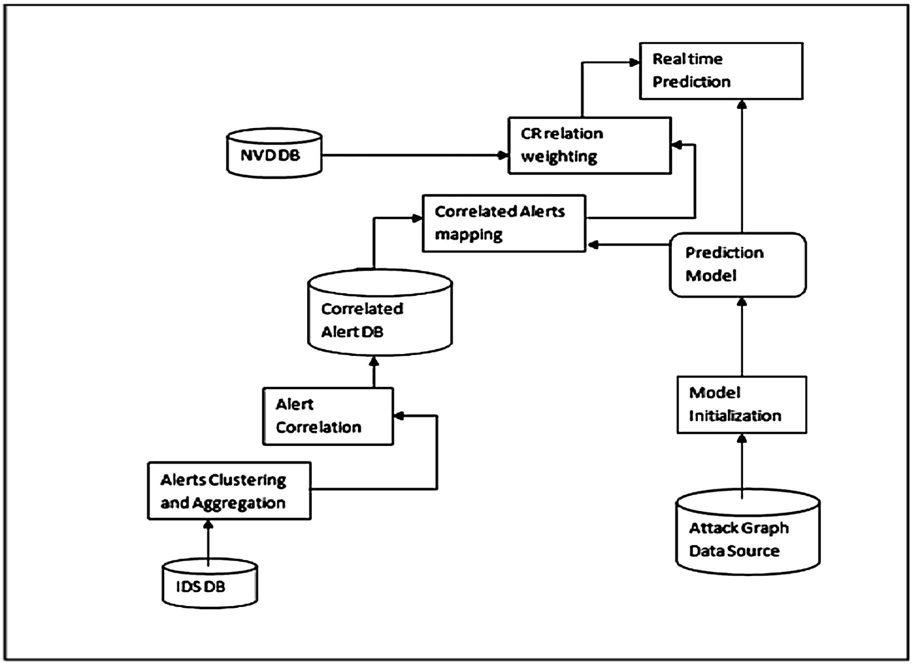

Fig. 5. (Fayyad and Meinel 2013)

By this definition attacks model will be initiated and then in another stage they will combined with correlated alerts. The next stage is processing vulnerability data base and trying to find relationship between them, attack graphs and correlated alerts. The result of this model provides a real time protection. The advantage of this methodology is staging each part of attack scenario which means defense indicators can have specific plan for each stage of ongoing attack when it happens.

5.1 Summary

To sum up, the following relations have been concluded between literatures and the possible methodology offered by this paper:

1. Framework: framework is highly important in any project and this research paper seeks to to make an effective framework which it not only fulfils the purpose of this paper but also delivers an operative method dealing with improvement of CSA. By reviewing above literatures, it can be understood that all of them, same principle in designing their framework where they have same blocks in their framework with different names and same operations, however, all of them do not look for same result and performance. Ahn et al. (2013) and Wu et al. (2013) and Schreiber-ehle and Koch (2012) do not propose a real time algorithm and their system mainly was designed for decision makes to improve CSA and protect their network through those solutions. Also the study by Das et al. (2013) follows same path but in different way. In other words they combine decision making process with financial issues in their approach of improvement of CSA. On the other hand Morris et al. (2011), Musliner (2011) and Fayyad and Meinel (2013) try to present real time framework of combating cyber breaches by probing technical elements in cyber security. This PhD can be a combination of both types of proposed models where by analysing past cyber breaches incidents in terms of not only type of attacks and methodology but also cyber attackers motivations and behaviours, it aims to present deeper understating of current situation of cyber environments and its players and predict future conditions based on current and past state.
2. Type of data and its collection: All of the literature mentioned in this paper focus on using Open Source Intelligence because they are publicly accessible and do not raise any ethical issues.
3. Type of analysis: All of proposed frameworks in this paper use statistical and data mining techniques.

6 Proposed Framework for Improving CSA

This paper aims to design a framework applying data mining and predictive analytic techniques to past historical data in terms of cyber-attacks to predict future cyber threats which will help to a deeper understanding of cyber situational awareness. This Paper is an ongoing research and it will be developed in near future.

The proposed framework includes following steps as it is showed in Fig. 6:

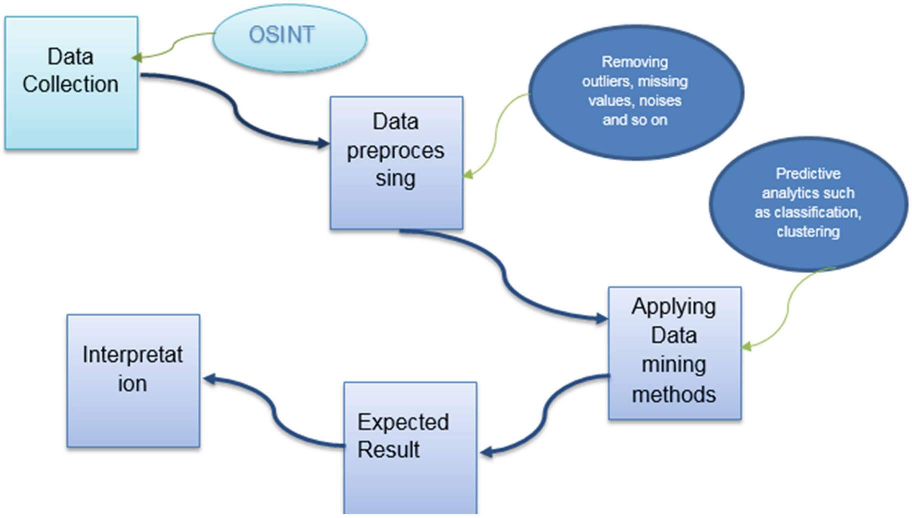
OSINT
Data Collection
Data preproces sing
Removing outliers, missing values, noises and so on
Predictive analytics such as classification, clustering
Applying Data mining methods
Interpretat ion
Expected Result

Fig. 6. (Proposed framework)

1. Data collection: as it mentioned this research paper will focus on cyber-attack historical data and the data will be collected from Open Source Intelligent (OSINT). Therefore it has been decided to use sources such as news and websites and any other sources which is publicly accessible and also their information does not raise any ethical issues.
2. Data type and attributes: the collected data initially will be nominal and categorical, however, based on requirements and type analysis they can be transformed into nominal and other formats. Data will include attributes such as type of attack, attacker, type of target, cyber motivation and so on.
3. Data pre-processing: sometimes because of operational errors and implementation of system, obtained data from real world has some errors, contradictions, incompatibility and missing values. This stage is based on Al-Janabi (2011) method which consists of following tasks to make the data ready for analysis purposes; dealing with missing values, removing noises, fixing incompatibilities and removing outliers.
4. Data mining techniques: Based on literature review, in order to find patterns among the data, data mining techniques will be applied. According to Ahn et al. (2014) using classification techniques can help cyber experts to find current patterns and based on findings try to predict the future patterns. Naïve Bayes and decision tree algorithms are two suitable techniques can extract valuable information from uncertain knowledge. (Bhardwaj and Pal 2011)
5. Expected result: after applying data mining techniques, the suitable result will be obtained including patterns relationships between different attributes and features of past cyber-attacks.
6. Interpretation of result: This stage is a crucial stage, Schreiber and Koch (2012) report that making the result meaningful to managers and decision makers is the

most significant stage of CSA improvement. For instance it should be determined which attributes or elements have more effect and cyber-attacks or which factors are weakest or strongest in the CSA.

7 Conclusion

Cyber Situational Awareness helps cyber defenders to adopt operative solutions dealing with cyber-attacks. This paper aims to review current frameworks improving CSA and propose an effective method to help cyber experts to understand their cyber situation deeply and combat cyber-attacks through implementation of security policies and countermeasures. The proposed framework is an ongoing and developing research which is based on data mining and predictive analytic techniques and it intends to highlight current issues in cyber firms through analysis of past cyber-attacks and predict future trends. This research will contribute to cyber managers to identify which attack methods are more favorable and help them to prioritize cyber security demands based on current and predicted future trends.

References

Aaviksoo, J.: Cyber-terrorism. Vital Speeches Day **74**(1), 28 (2008)

Ahlemeyer-Stubbe, A., Coleman, S.: A Practical Guide to Data Mining for Business and Industry. Wiley, New York (2014)

Ahn, S., Kim, N., Chung, T.: Big Data Analysis System Concept for Detecting Unknown Attacks (2014)

Al-janabi, K.B.S.: A proposed framework for analyzing crime data set using decision tree and simple k-means mining algorithms. J. Kufa Math. Comput. **1**(3), 8–24 (2011)

Antonik, J.: Decision management. In: Military Communications Conference (MILCOM 2007), Orlando, FL, USA, October 2007, pp. 1–5. IEEE (2007)

Aspan, M., Soh, K.: Citi says 360,000 accounts hacked in May cyber attack. Reuters (2011)

Awan, I., Blakemore, B.: Policing Cyber Hate, Cyber Threats and Cyber Terrorism. Ashgate, Farnham (2012). MyiLibrary

Barford, P., Dacier, M., Dietterich, T.G., Fredrikson, M., Giffin, J., Jajodia, S., et al.: Cyber SA: situational awareness for cyber defense. In: Jajodia, S., Liu, P., Swarup, V., Wang, C. (eds.) Cyber Situational Awareness, pp. 3–14. Springer, New York (2010)

Bhardwaj, B.K., Pal, S.: Data Mining: a prediction for performance improvement using classification. Int. J. Comput. Sci. Inf. Secur. **9**(4), 136–140 (2011)

Cox, C.: Cyber capabilities and intent of terrorist forces. Inf. Secur. J. Global Perspect. **24**, 1–8 (2015)

Das, S., Mukhopadhyay, A., Shukla, G.K.: i-HOPE framework for predicting cyber breaches: a logit approach. In: 2013 46th Hawaii International Conference on System Sciences, pp. 3008–3017 (2013)

Dean, J.: Big Data, Data Mining, and Machine Learning: Value Creation for Business Leaders and Practitioners. Wiley and SAS Business Series. Wiley, Hoboken (2014)

Dua, S., Du, X.: Data Mining and Machine Learning in Cybersecurity. CRC Press, Boca Raton (2011)

Dutt, V., Ahn, Y.-S., Gonzalez, C.: Cyber situation awareness: modeling detection of cyber attacks with instance-based learning theory. Hum. Factors J. Hum. Factors Ergon. Soc. **55**(3), 605–618 (2012). doi:10.1177/0018720812464045

Fayyad, S., Meinel, C.: Attack scenario prediction methodology. In: 2013 10th International Conference on Information Technology: New Generations, pp. 53–59 (2013). doi:10.1109/ITNG.2013.16

Franke, U., Brynielsson, J.: Cyber situational awareness – a systematic review of the literature. Comput. Secur. **46**, 18–31 (2014). doi:10.1016/j.cose.2014.06.008

Harrison, L., Laska, J., Spahn, R., Iannacone, M., Downing, E., Ferragut, E.M., Goodall, J.R.: situ: situational understanding and discovery for cyber attacks. In: 2012 IEEE Conference on Visual Analytics Science and Technology (VAST), pp. 307–308 (2012). doi:10.1109/VAST.2012.6400503

Jaishankar, K.: Cyber Criminology: Exploring Internet Crimes and Criminal Behavior. CRC, Boca Raton, London (2011). Dawsonera

Ledolter, J.: Data Mining and Business Analytics with R. Wiley and SAS Business Series. Wiley, Hoboken (2013)

Lewis, J.A.: Assessing the Risks of Cyber Terrorism, Cyber War and Other Cyber Threats, 1–12 December 2002

Morris, I., Mayron, L.M., Smith, W.B., Knepper, M.M., Ita, R., Fox, K.L., Corp, H.: A perceptually-relevant model-based cyber threat prediction method for enterprise mission assurance, pp. 60–65 (2011)

Musliner, D.J., Rye, J.M., Thomsen, D., McDonald, D.D., Burstein, M.H., Robertson, P.: FUZZBUSTER: towards adaptive immunity from cyber threats. In: 2011 Fifth IEEE Conference on Self-Adaptive and Self-Organizing Systems Workshops, pp. 137–140 (2011). doi:10.1109/SASOW.2011.26

Nikishin, A.: Malicious software–past, present and future. Inf. Secur. Tech. Rep. **9**(2), 6–18 (2004)

Odei Danso, S.: An exploration of classification prediction techniques in data mining: the insurance domain. Master Degree Thesis, Bournmouth University (2006)

Pollitt, M.M.: “Cyberterrorism — fact or fancy?”. Comput. Fraud Secur. **1998**(2), 8–10 (1998)

Savov, V.: Sony Pictures hacked: the full story (WWW Document). The Verge (2014). http://www.theverge.com/2014/12/8/7352581/sony-pictures-hacked-storystream. Accessed 6 April 15

Schreiber-Ehle, S., Koch, W.: The JDL model of data fusion applied to cyber-defence—a review paper. In: 2012 Workshop on Sensor Data Fusion: Trends, Solutions, Applications (SDF), 4–6 September 2012 (2012). doi:10.1109/SDF.2012.6327919

Wu, J., Yin, L., Guo, Y.: Cyber attacks prediction model based on Bayesian network. In: 2012 IEEE 18th International Conference on Parallel and Distributed Systems, pp. 730–731 (2012). doi:10.1109/ICPADS.2012.117