An Efficient Classification Model for Detecting Advanced Persistent Threat

Saranya Chandran[1], Hrudya P[2], Prabaharan Poornachandran[3]
Amrita Center for Cybersecurity

Amrita Vishwa Vidyapeetham

Amritapuri Campus, Kollam

[1] *saran.cr4@gmail.com*

[2] *hrudyap@am.amrita.edu*

[3] *praba@am.amrita.edu*

***Abstract*—Among most of the cyber attacks that occured, the most drastic are advanced persistent threats. APTs are differ from other attacks as they have multiple phases,often silent for long period of time and launched by adamant, wellfunded opponents. These targeted attacks mainly concentrated on government agencies and organizations in industries, as are those involved in international trade and having sensitive data. APTs escape from detection by antivirus solutions, intrusion detection and intrusion prevention systems and firewalls. In this paper we proposes a classification model having 99.8% accuracy, for the detection of APT.**

***Keywords*—Advanced Persistent Threat, model, targeted attack.**

I. Introduction

Mathematical model is a conceptual model which uses mathematical languages to describe a system i.e, it help us to study the real world problem and reach mathematical conclusions about it. The steps in the process of mathematical modelling is shown in figure 1. First we studied the real world problem and identified the dependent and independent variables and also the changes in the variables. Then we applied our mathematical skills to obtain equations to describe the problem. By this method we can formulate a mathematical model. Second task is to solve the formulated mathematical model and there by deriving mathematical conclusions. The third task is to interpret this mathematical conclusions as information about the real word data by giving explanations or making predictions. The final task is to test our predictions by checking against new data. If it is not work well with reality ,we need to revise our model or formulate a new model and start the cycle again[43].

We use mathematics to model the problems in cyber security. The major challenges includes how we model networks, its dynamics, threat discovery etc[1]. Graph theory is a widely used approach to model networks[3],[4],[5],[6],[7], to find the distance between the attackers and the compromised hosts[4]. Machine learning techniques are used to find the probability for finding threats[1]. By using the concepts of game theory defenders can predict the properties of future threats[1],[5]. The models in cybersecurity should have the ability to describe the problem well and analyse the system[8].

978-1-4799-8792-4/15/$31.00 ©2015 IEEE

Advanced Persistent Threat is a targeted attack mainly concentrated on organizations for the purpose of stealing sensitive data or causing specific damage. As the name specifies it is advanced i.e, APT utilizes different types of vulnerabilities identified within the organization. It works on multiple phases over a long period of time i.e, they may take months or even years for the fulfilment of their objective. Due to the advanced nature of APT, large number of features may be required for accurate detection. Since it is persistent, modelling must follow changes of features over time.

Operators behind APT uses different techniques and technologies for the intrusion phase. They often merge multiple methods, tools, and techniques in order to reach and compromise their target and maintain access to it. They give priority to a particular task, not for immediate financial gain. The attackers are directed by external body and their operation is lead through continuous monitoring and interaction for achieving the defined objectives. Since their prime goal is to maintain longterm access to the target, they use low and slow approach. APTs are a threat because they have both the potential and purpose. APT attacks are executed by coordinated human actions, rather than by code itself. The operators have a specific objective and are talented, prompted, systematic and well funded.

Stuxnet is an APT attack which is detected on June 2010, but the Kaspersky Lab experts reported that the first form of the worm appeared in June 2009[32]. Its speculted purpose was to deliberately damage the Iranian nuclear program,more explicitly Natanz uranium enrichment plant. Another one Duqu was detected in September 2011. It is belived that it has been active since February 2010. These have similarities with stuxnet, but its main aim is espionage rather than sabotage. After the initial infection it is remained active for 30 days before its gets self destructed. The next threat Flame which is believed that it had been already active for five to eight years and is detected in May 2012. There is no direct connection between flame and the above described threats. Like Duqu, Flame was also a targeted data stealing malware but it is more widespread as it had infected thousands of windows systems. Another information stealing threat is Red October which was discovered in October 2012. It is believed that it has been active since May 2007 and its targeted organisations include political , governmental and scientific institutions. A recent poll by the isaca[9] revealed that many organizations are unprepared to deal with these targeted and persistent threats. APTs targeted organizations include google, Adobe Systems, RSA, Sony, Juniper Networks and Rackspace i.e, domains located in every region of the world. In the case of Sony , the attackers exploited a number of vulnerabilities and is carried over several months. They compromised some 100 million customer records and affecting company's web operations in several countries. Another reported APT attack is Operation Aurora in January 2010 , a wide scale attack and is targeted thirty four different organisations.

The cyber security company FireEye Labs[10] recently analysed some new samples of terminator RAT that was sent through spear fishing emails to targets in Taiwan. This terimnator RAT now become more sophisticated in recent APT attacks. For example , in RSA APT attack they used Remote Access Toolkit to communicate with the remote C&C centre. FireEye found that Government was the mostly targeted in the world in 2013, with 84 of the 159 malware families documented by FireEye. FireEye analyzed nearly 40,000 unique, advanced attacks , more than 100 per day on average. From these, they categorized over 4,000 unique attacks as APT-directed. And they discovered nearly 18,000 unique malware infections due to APT activity. In 2013, FireEye tracked 159 malware families associated with APT activity. And they discovered initial C&C infrastructure within 206 national top-level.

APT is differ from other attacks mainly based on its sophisticated nature and the use of multiple techniques for its operation. Since it is targeted it causes serious damage to the organisation. Also APT is not controlled by our traditional anti virus and firewall systems. In this paper we are giving an overview of APT, use of mathematical model for the detection of these threats. Finally proposed an accurate classification model for the detection of APT.

II. RELATED WORK

Over two thirds of data leakage now occur by the sophisticated APT attacks[11]. They have accessibility map which shows the mapping between actors and data i.e, it specifies which actor is permitted to have access to which data. In[12] they proposed an attack pyramid with the goal of APT on the top of it. This attack is carried out by multiple phases of operations. The system developed in [13] mapped this phases into a markov chain i.e the future state X_{n+1} of a discrete time markov process is dependent only on the current state X_n and not on any of the previous states. After compromised a system in the network, most of the malwares starts to communicate with the C&C sever controlled by the attacker using http protocol. Also they are trying to connect to the C&C servers by quering domain names which are algorithmically generated. The abnormal domain name queries are detected from a large set of DNS logs[14]. The model proposed in [15] separates malicious and geniune http requests and classifies them.

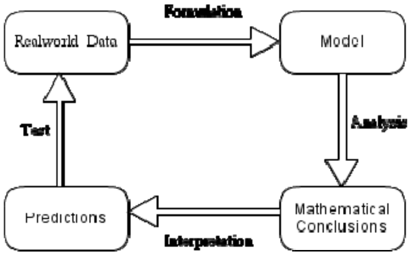


Fig. 1. Process of mathematical modelling

Advanced Persistent Threat concentrated on organizations intellectual property, financial assets and status. APT operations may have multiple phases running in parallel, each consisting of one or more operations. Preparation for the attack and gaining the initial entry point are essential for APT [16]. Most APTs start their exploitation techniques with social engineering and spear phishing. Also most of the APT type malware includes a dropper , a droppee and some more trojans [16].

A model is proposed for the detection of APT that uses a Finite Angular State Velocity Machine. APT may be detected by evaluation of large numbers of features that changed based on the systems operation, eg: CPU usage. This model allows the vector to detect change in a feature if an APT has infected a system. In this model the probability of the malware can be calculated using naïve bayes[17].

III. PROPOSED SYSTEM

Figure 2 describes system architecture. Initially gather some malwares for execution. Then comes the analysis and model blocks. The detailed description of these two blocks is given in the figure 3. In the analysis phase we are extracting the features CPU usage , memory usage of the systems , open ports and also the number of files in system32 folder from a normal system. We are selected known APT malwares from the open malware site. Next we are executing the malwares on the system which is used by a normal user. Here malwares are analysed dynamically. After this, repeat the process of feature extraction. In the third phase we are doing analysis. Here the training of the features is done, and creating the model. Try training set with different machine learning algorithms and find out the accurate one. Next is testing phase. Extract features from a targeted system,input the features to the model. If an APT attack had happened to the system then the model produces an alert signal.

A. Random Forest Model

Mathematical models have different forms. Here we use random forest model. It is used ensemble learning method, i.e using multiple learning algorithms for getting good performance[1]. Bagging is a type of ensembles in which it trains each model in the ensemble using a randomly drawn subset of the training set[2]. The random forest algorithm combines random decision trees with bagging to achieve very high classification accuracy[2]. A learning set of L consists of data $\{(y_n,x_n),n=1,.....,N\}$ where y is the class label[40]. There is some procedure for using this learning set to form a predictor $\varphi(x,L)$ I.e, if the input is x then predict y by $\varphi(x,L)$. Given a sequence of learning sets $\{L_k\}$ each consisting of N independant observations from the same underlying distribution as L. By using this {Lk} we can develop a better predictor than the single learning set predictor $\varphi(x,L)$ i.e working with a sequence of predictors $\{ \varphi(x,L_k)\}$[40].

Random Forest is a classifier that consists a number of decision trees and the resulting class is predicted based on the votes from individual trees. Here there are two classes normal(n) and compromised(c). Also extracting four random features and a total of 958 different values for each feature. Constructed 10 different decision trees from these values.

Each tree is constructed using the following algorithm:

1. Let the number of training values be *Y*, and the number of attributes in the classifier be *X*.
2. The number *x* of input values to be used to determine the decision at a node of the tree; *x* should be much less than *X*.
3. Choose a training set for this tree by choosing *y* times with replacement from all *Y* available training cases.
4. For each node of the tree, randomly choose *x* variables on which to base the decision at that node. Calculate the best split based on these *x* variables in the training set.
5. Each tree is fully grown and not pruned [18].

After constructing decision trees, again extract new features from the system and is pushed down the trees. Each tree predict one class i.e, that class got one vote and finally random forest predicts the class which have the maximum votes.

B. Class Prediction

The data set was randomly divided into a test set T and a learning set L. The sizes of test set are adhoc, mostly chosen so that L would be reasonably large. A classification tree was constructed from L, with selection done by 10-fold cross validation . Running the test set T down this tree gives the misclassification rate $e_s(L,T)$. A bootsrap sample L_B is selected from L, and a tree grown using L_B and 10-fold cross validation. This is repeated 10 times giving tree classifiers $\varphi_1(x),.....\varphi_{10}(x)$. If$(j_n,x_n)$Є T,then the estimated class of x_n is that class having more votes in $\varphi_1(x),.....\varphi_{10}(x)$. The proportion of times the estimated class differs from the true class is the bagging misclassification rate $e_B(L,T)$.

C. Random forest's theoretical background

Given an ensemble of classifiers φ_1 (x),φ_2 (x), ... ,φ_K (x) , and with the training set drawn at random from the distribution of the random vector Y,X, the margin function is defined as

$$mg(X,Y\ ) = av_k\, I(\ \varphi_k(X)=Y\ ) - \max_{j\neq Y} av_k\, I(\varphi_k(X)= j) \quad (1)$$

where I(•) is the indicator function.

The margin measures the extent to which the average number of votes at X,Y for the right class exceeds the average vote for any other class. The larger the margin, the more confidence in the classification[41]. The generalization error is given by

$$PE^* = P_{X,Y}\,(mg(X,Y\ ) < 0) \quad (2)$$

where the subscripts X,Y indicate that the probability is over the X,Y space.

In random forests, $\varphi_k(X) = \varphi(X,L_k)$.
As the number of trees increases PE* converges to Eq.3.

$$P_{X,Y}\,(P_L\,(h(X,L)=Y)-\max_{j\neq Y} P_L(h(X,L)= j) < 0) \quad (3)$$

The margin function for a random forest is

$$mr(X,Y\ ) = P_L\,(h(X,L)=Y\ )- \max_{j\neq Y} P_L(h(X,L)= j) \quad (4)$$

Strength of the set of classifiers {h(x,L)} is

$$s= E_{X,Y}\, mr(X,Y\ ) \quad (5)$$

IV. EXPERIMENTAL RESULTS

Detection of APT was modeled in a number of desktops having Intel dual core processor in virtual local area network, that are running Windows XP operating system. Here we are monitoring the systems for three months and extracted features for these period of time. The training set was given to different machine learning algorithms and the result is shown on the following table.

TABLE I. Comparison of models

Model	Accuracy
Random Forest	99.8%
Naïve Bayes	72.65%
Part	53.85%
J-48	50%

The comparison suggest that random forest model is more accurate for APT detection.

The following graphs shows the values of the features extracted from a normal system and a compromised system for a period of ten hours.

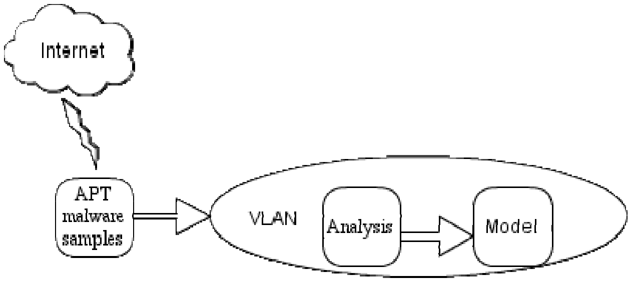


Fig. 2. System Architecture

Figure 4 gives the details of the system's usage of cpu. The amount of memory used by the system before and after executing the malware is shown in figure 5. While executing the malware it uses more memory. The binary here adds some dll files in system32 folder for performing some malicious operations. Figure 6 gives the number of files in system32 folder of a normal system and also added by the malware. The number of open ports of both the normal and compromised systems are shown in figure7 and figure8.

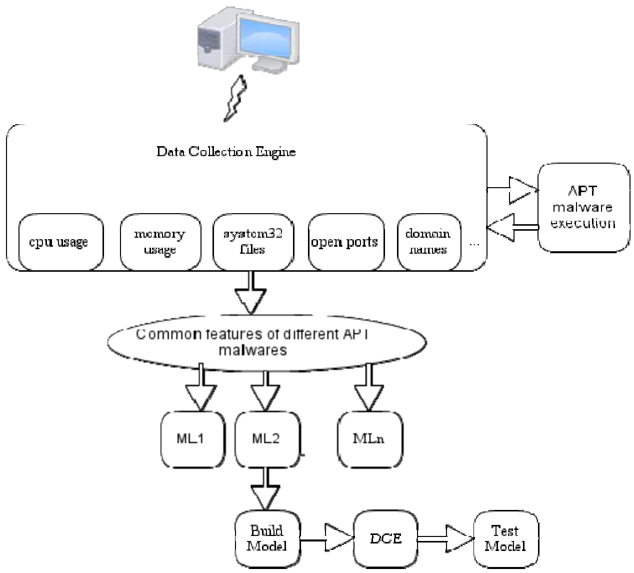


Fig.3. Detailed Architecture

In this model we also extracted DNS queries from both the normal and compromised system. The figure 9 shows the graphical representation.

Normal system queries for benign domain names gmail, yahoo, facebook, twitter, gstatic, google. In the graph these are represented by numbers 1,2,3,4,5,6. Again extracted DNS queries by the infected system. It queries for usual domain names as well as names that are algorithmically generated which is used in APT attacks. The identified names are a.najwahaifamelema1.com, a.najwahaifamelema2.com etc and are represented by number 7. The other set of names are lh4.google.com, lh5.google.com, lh6.google.com etc and are represented by number 10 and the other set are clients2.google.com, clients5.google.com etc are represented by number 11. The next set of names identified are lh4.googleusercontent.com, lh5.googleusercontent.com and are represented by the number 12. These domain names are frequently queried for connecting to C&C servers.

We executed different binaries for analysis and one of them is a bot. Here we also captured the network traffic from the infected system and analysed it. From the figure 10 we observed some dns queries in the traffic.

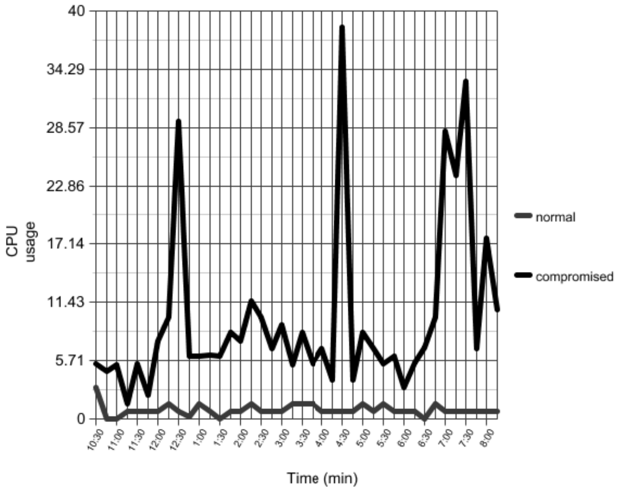


Fig. 4. CPU usage of normal and compromised system

Fig. 5. Memory usage of normal and compromised system

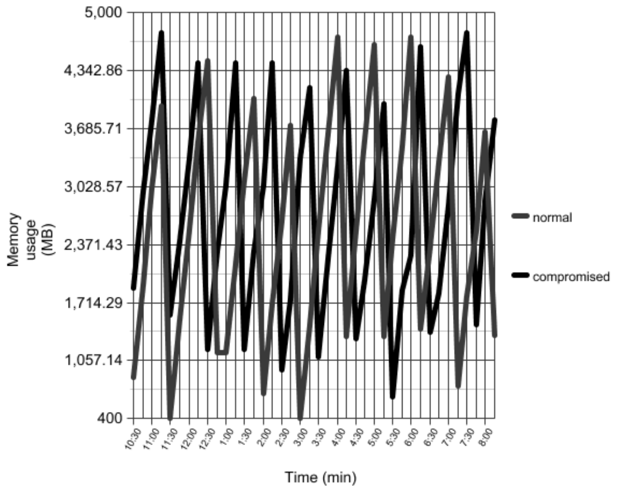


Fig. 5. Memory usage of normal and compromised system

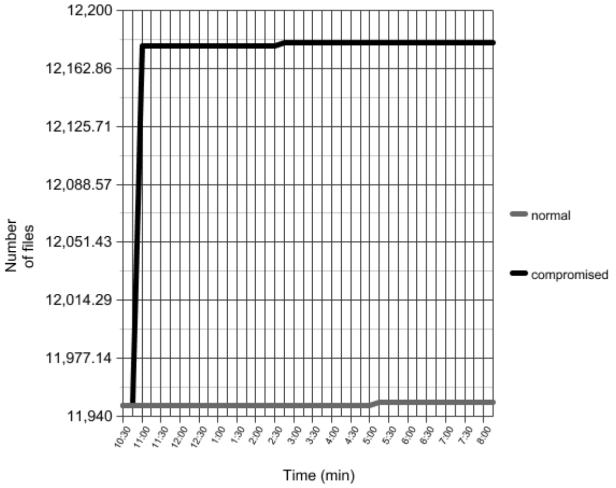


Fig. 6. Number of files in system32 folder of normal and compromised system

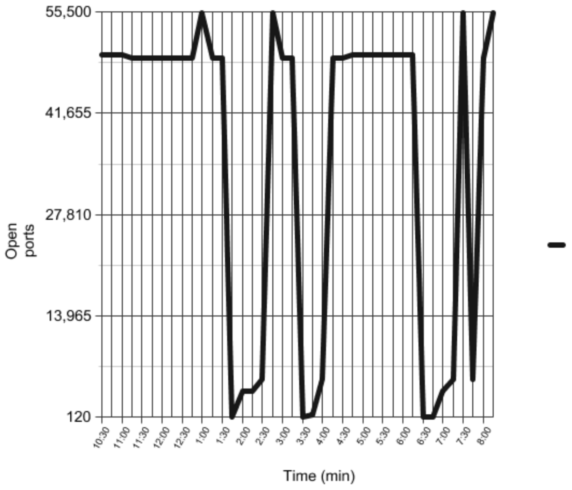


Fig. 7. Open ports in normal system

In defining our classification model we use random forest algorithm. Out of total values ten trees are constructed by considering four random features. The time took to build the model is 0.24 seconds.Here 99.8% of total training instances are classified correctly. the root mean squared error is 0.0058. The area under the ROC curve for the normal class is 0.999 and the area under the ROC curve for the compromised class is 1.

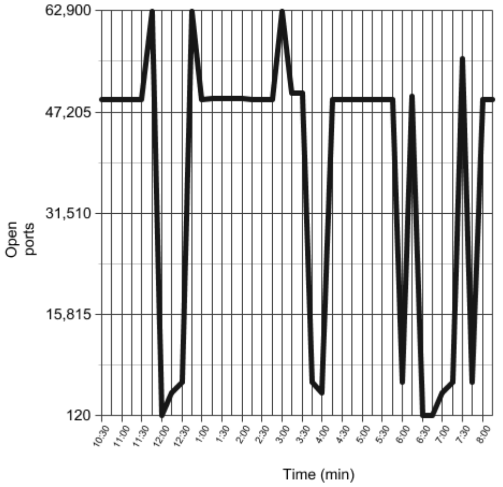


Fig. 8. Open ports in compromised system

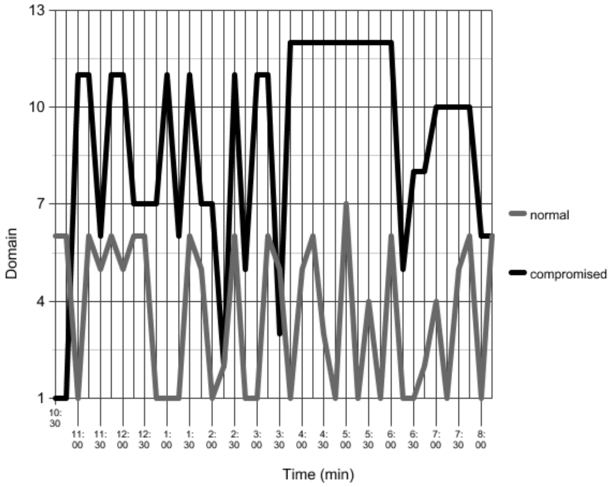
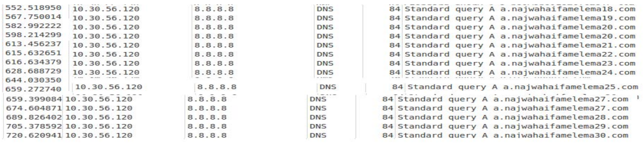


Fig. 9. DNS queries by normal and compromised systems



```
552.518950  10.30.56.120   8.8.8.8   DNS   84 Standard query A a.najwahaifamelema18.com
567.750014  10.30.56.120   8.8.8.8   DNS   84 Standard query A a.najwahaifamelema19.com
582.992222  10.30.56.120   8.8.8.8   DNS   84 Standard query A a.najwahaifamelema20.com
598.214299  10.30.56.120   8.8.8.8   DNS   84 Standard query A a.najwahaifamelema20.com
613.456237  10.30.56.120   8.8.8.8   DNS   84 Standard query A a.najwahaifamelema21.com
615.632651  10.30.56.120   8.8.8.8   DNS   84 Standard query A a.najwahaifamelema22.com
616.634379  10.30.56.120   8.8.8.8   DNS   84 Standard query A a.najwahaifamelema23.com
628.688729  10.30.56.120   8.8.8.8   DNS   84 Standard query A a.najwahaifamelema24.com
644.030350
659.272740  10.30.56.120   8.8.8.8   DNS   84 Standard query A a.najwahaifamelema25.com
659.399084 10.30.56.120    8.8.8.8   DNS   84 Standard query A a.najwahaifamelema27.com
674.604871 10.30.56.120    8.8.8.8   DNS   84 Standard query A a.najwahaifamelema27.com
689.826402 10.30.56.120    8.8.8.8   DNS   84 Standard query A a.najwahaifamelema28.com
705.378592 10.30.56.120    8.8.8.8   DNS   84 Standard query A a.najwahaifamelema29.com
720.620941 10.30.56.120    8.8.8.8   DNS   84 Standard query A a.najwahaifamelema30.com
```

Fig. 10. Network traffic from infected system

Kappa statistic is a value that compares an observed accuracy with an expected accuracy.

$$K= (oa - ea)/(1 – ea) \qquad (6)$$

From our experiment we got Kappa statistic as 0.9979. The root mean square error (RMSE) and mean absolute error are used as a standard statistical metric to measure model performance[44]. MAE is suitable to describe uniformly distributed errors and RMSE is better for random distribution. RMSE and MAE are calculated for the dataset as

$$MAE = (1/n) \sum_{i=1}^{n} \left| e_i \right| \qquad (7)$$

$$RMSE=\sqrt{}(1/n)\sum_{i=1}^{n} e_i^2 \qquad (8)$$

where n-total number of samples and ei -error for each sample. Result shows MAE as 0.001 and RMSE as 0.0194.

Relative absolute error and Root relative squared error is calculated by Equation 9 and Equation 10 respectively. The resultant values are 0.2091 % and 3.877 %.

$$RAE= (\left| p_1\text{-}a_1 \right| +..+ \left| p_n\text{-}a_n \right|)/(\left| \mathring{a}\text{-}a_1 \right| +..+ \left| \mathring{a}\text{-}a_n \right|) \qquad (9)$$

$$RRSE = \sqrt{} ((p_1\text{-}a_1)^2+..+(p_n\text{-}a_n)^2)/((\mathring{a}\text{-}a_1)^2+..+(\mathring{a}\text{-}a_n)^2) \qquad (10)$$

p_1, p_2, .., p_n – predicted values
a_1, a_2,.., a_n - actual values
å - average

V. CONCLUSION

Detection of targeted and slow moving threats such as APT requires the ability to model large amounts of data over a long period of time. The proposed mathematical model describes accurate method for APT detection. Here we proposed random forest model which is more accurate and is used for the modelling of APT with less false positives and false negatives. Research needs to be further conducted to extract more features that differentiate APT from other attacks.

References

[1] Daniel M. Dunlavy, Bruce Hendrickson, and Tamara G. Kolda , “Mathematical Challenges in Cybersecurity”, SANDIA Report,Printed February 2009.

[2] Daniel Lawson An Introduction to Mathematical Modelling. Retrievedfrom http://www.maths.bris.ac.uk/~madjl/course_text.pdf.

[3] Mark Burgess, Geoffrey Canright and Kenth Engo Monsen, “A graph-theoretical model of computer security”. International Journal of Information Security, November 2004, Volume 3, Issue 2, pp 70-85.

[4] Hogan E et.al, "Towards a multiscale approach to cybersecurity modeling," *Technologies for Homeland Security (HST), 2013 IEEE International Conference on* , vol., no., pp.80,85, 12-14 Nov. 2013.

[5] Hird J., Koelle R. and Kolev D., "Towards mathematical modelling in security risk management in system engineering," *Integrated Communications, Navigation and Surveillance Conference (ICNS), 2013* , vol., no., pp.1,13, 22-25 April 2013.

[6] Bimal Kumar Mishra and Dinesh Saini, “Mathematical models on computer viruses”. Birla Institute of Technology and Science, Mathematics Group, Pilani 333031, India,Birla Institute of Technology and Science, Computer Science & Information System Group, Pilani 333031, India, Applied Mathematics and Computation (2006).

[7] Caiyi Zhu and Xiangkun Dai, "Model of trust management based on finite state machine," *Cyber Security, Cyber Warfare and Digital Forensic (CyberSec), 2012 International Conference on* , vol., no., pp.161,164, 26-28 June 2012.

[8] D.Elliott Bell , Concerning “Modeling”of Computer Security Concerning Trusted Information Systems,Inc.IEEE, 1988.

[9] Advanced Persistent Threat Awareness 2014 - white paper,ISACA.

[10] Evasive Tactics: Terminator RAT, FireEye, October 24, 2013.Retrived from https://www.fireeye.com/blog/threat-research/2013/10/evasive-tactics-terminator-rat.html.

[11] Mustafa T, "Malicious Data Leak Prevention and Purposeful Evasion Attacks: An approach to Advanced Persistent Threat (APT) management," *Electronics, Communications and Photonics Conference (SIECPC), 2013 Saudi International* , vol., no., pp.1,5, 27-30April2013.

[12] Giura P and Wei Wang, "A Context-Based Detection Framework for Advanced Persistent Threats," *Cyber Security (CyberSecurity), 2012 International Conference on* , vol., no., pp.69,74, 14-16 Dec. 2012.

[13] Ioannou G et.al , "A Markov multi-phase transferable belief model: An application for predicting data exfiltration APTs," *Information Fusion (FUSION), 2013 16th International Conference on* , vol., no., pp.842,849, 9-12 July 2013.

[14] Begleiter et.al , "A fast and scalable method for threat detection in large-scale DNS logs," *Big Data, 2013 IEEE International Conference on* , vol., no., pp.738,741, 6-9 Oct. 2013.

[15] Zarras A et.al , "Automated generation of models for fast and precise detection of HTTP-based malware," *Privacy, Security and Trust (PST), 2014 Twelfth Annual International Conference on* , vol., no., pp.249,256, 23-24 July 2014.

[16] SANS Institute InfoSec Reading Room A Detailed Analysis of an Advanced Persistent Threat Malware33814.pdf. Retrieved 2014-06-06.

[17] Gregory Vert1, Bilal Gonen2 and Jayson Brown3 ,1,3 Department of Computer Information Systems, Texas A&M University – Central Texas, Killeen, TX, U.S.A. 2 Department of Computer Science, University of West Florida, Pensacola, FL, U.S.A. , “A Theoretical Model for Detection of Advanced Persistent Threat in Networks and Systems Using a Finite Angular State Velocity Machine (FAST□VM)”,International Journal of Computer Science and Application (IJCSA) Volume 3 Issue 2,May 2014.

[18] Random Trees, OpenCV 2.4.9.0 documentation. Retrieved from http://docs.opencv.org/modules/ml/doc/random_trees.html. December 5, 2014.

[19] “Life cycle_of_an_APT”. DELL Secure Works, Counter Threat Unit research2012. Retrieved 2014-07-07.

[20] Shui Yu1, Guojun Wang2 and Wanlei Zhou3 ,1,3 Deakin University, 2 Central South University, “Modeling Malicious Activities in Cyber Space”.

[21] A websense white paper on Advanced Persistent Threat and other Advanced Attacks. Retrieved from https://www.whitepaper-websense-advanced-persistent-threats-and-other-advanced-attacks-en.pdf.

[22] Frank R.Giordano, William P. Fox and Steven B.Horton,”A First Course in Mathematical Modelling Fifth edition”.

[23] D. Elliott Bell and Leonard J.LaPadula, “Secure Computer Systems: Mathematical Foundations”November, 1996 An electronic reconstruction by Len LaPadula of the original MITRE Technical Report 2547, Volume I dated 1 March 1973.

[24] Cyber flow Analytics, A nextgeneration approach to fighting Advanced Persistent Threats in cyber espionage. Retrieved from http://www.cyberflowanalytics.com/pdf/CyberFlow%20DataSheet.pdf. October 15,2014.

[25] “Advanced Persistent Threat: A brief description”. Damballa . Retrieved August 13, 2014.

[26] Introduction to decision trees and random forests, Ned Horning American Museum of Natural History's Center for Biodiversity and Conservation.

[27] “Persistent threats and how to monitor and deter them ”, Colin Tankard, Digital Pathways . Retrieved December 5, 2014.

[28] Ibrahim Ghafir and Vaclav Prenosil , “Advanced Persistent Threat Attack Detection: An Overview ”, International Journal of Advancements in Computer Networks and Its Se curity– IJCNS , Volume 4 : Issue 4 , [ISSN 2250-3757] ,27 December,2014.

[29] Frankie Li , Anthony Lai and Ddl Ddl , "Evidence of Advanced Persistent Threat: A case study of malware for political espionage," *Malicious and Unwanted Software (MALWARE), 2011 6th International Conference on* , vol., no., pp.102,109, 18-19 Oct. 2011.

[30] Mirza N.A.S et.al , "Anticipating Advanced Persistent Threat (APT) countermeasures using collaborative security mechanisms," *Biometrics and Security Technologies (ISBAST), 2014 International Symposium on* , vol., no., pp.129,132, 26-27 Aug. 2014.

[31] Yuan Wang et.al , "A Network Gene-Based Framework for Detecting Advanced Persistent Threats," *P2P, Parallel, Grid, Cloud and Internet Computing (3PGCIC), 2014 Ninth International Conference on* , vol., no., pp.97,102, 8-10 Nov. 2014.

[32] Virvilis N and Gritzalis D, "The Big Four - What We Did Wrong in Advanced Persistent Threat Detection?," *Availability, Reliability and Security (ARES), 2013 Eighth International Conference on* , vol., no.,pp.248,254,2-6Sept.2013.

[33] Almasizadeh J and Azgomi M.A., "A New Method for Modeling and Evaluation of the Probability of Attacker Success," *Security Technology, 2008. SECTECH '08. International Conference on* , vol., no.,pp.49,53,13-15Dec.2008.

[34] Murakami T, Kumano S and Koide H , "An implementation of tracing attacks on advanced persistent threats by using actors model," *Soft Computing and Intelligent Systems (SCIS), 2014 Joint 7th International Conference on and Advanced Intelligent Systems (ISIS), 15th International Symposium on* , vol., no., pp.1316,1320, 3-6Dec.2014.

[35] Adebayo O.S and AbdulAziz N, "An intelligence based model for the prevention of advanced cyber-attacks," *Information and Communication Technology for The Muslim World (ICT4M), 2014 The 5th International Conference on* , vol., no., pp.1,5, 17-18 Nov. 2014.

[36] Xupeng Fang et.al, "A Game Model for Predicting the Attack Path of APT," *Dependable, Autonomic and Secure Computing (DASC), 2014 IEEE 12th International Conference on* , vol., no., pp.491,495, 24-27 Aug. 2014.

[37] Vance A, "Flow based analysis of Advanced Persistent Threats detecting targeted attacks in cloud computing," *Infocommunications Science and Technology, 2014 First International Scientific-Practical Conference Problems of* , vol., no., pp.173,176, 14-17 Oct. 2014.

[38] Bhatt P, Toshiro Yano E and Gustavsson P.M, "Towards a Framework to Detect Multi-stage Advanced Persistent Threats Attacks," *Service Oriented System Engineering (SOSE), 2014 IEEE 8th International Symposium on* , vol., no., pp.390,395, 7-11 April 2014.

[39] Frank Giordano et.al, *A first course in mathematical modeling*,5th ed., Cengage Learning,14-Feb-2013, Mathematics.

[40] Breiman, L. [1996a] Bagging Predictors, Machine Learning, 26,No. 2, 123-140.

[41] Breiman,L.Randomforests.MachineLearning,45(1):5 32,2001.18

[42] Amit,Y. and Geman, D. [1997] Shape quantization and recognition with randomized trees, Neural Computation 9, 1545-1588.

[43] James Stewart, *Metric International version Calculus,* Cengage Learning EMEA, 6th ed., 2008.

[44] Chai T and Draxler R. R. *Root MeanSquare Error (RMSE) or Mean Absolute Error (MAE)? - Arguments Against Avoiding RMSE in the Literature*. Geosci. Model Dev.2014, 7, 1247-1250.

[45] Sundus Juma , Zaiton Muda and Warusia Muda , “Reducing False Alarm Using Hybrid Intrusion Detection Based on X-means Clustering and Random Forest” , J*ournal of Theoretical and Applied Information Technology ,* October 20, 2014. Vol. 68 No.2 ISSN: 1992-8645.