International Journal of Advanced Computer and Mathematical Sciences
ISSN 2230-9624. Vol 3, Issue 4, 2012, pp 530-534
http://bipublication.com

A NEW LAYERED APPROACH FOR INTRUSION DETECTION

[1]Ramana Rao H. C and [2]Janardhana Reddy Y. R
[1]PRRM Engineering College .Hyderabad
[2]AVR&SVR Engg College ,Nandyal,

Abstract:

In this paper, we have addressed the dual problem of Accuracy and Efficiency for building robust and efficient intrusion detection systems. Our experimental results show that CRFs are very effective in improving the attack detection rate and decreasing the FAR. Having a low FAR is very important for any intrusion detection system. Further, feature selection and implementing the Layered Approach significantly reduce the time required to train and test the model. Even though we used a relational data set for our experiments

INDEXING TERMS : Layered Approach, Random Fields

Existing System:-

The field of intrusion detection and network security has been around since late 1980s. Since then, a number of methods and frameworks have been proposed and many systems have been built to detect intrusions. Various techniques such as association rules, clustering, naive Bayes classifier, support vector machines, genetic algorithms, artificial neural networks, and others have been applied to detect intrusions. In this section, we briefly discuss these techniques and frameworks.

Experimental results on the benchmark KDD '99 intrusion data set show that our proposed system based on Layered Conditional Random Fields outperforms other well-known methods such as the decision trees and the naive Bayes.

The improvement in attack detection accuracy is very high, particularly, for the U2R attacks (34.8 percent improvement) and the R2L attacks (34.5 percent improvement). Statistical Tests also demonstrate higher confidence in detection accuracy for our method.

Proposed System:-

Other approaches for detecting intrusion include the use of autonomous and probabilistic agents for intrusion detection. These methods are generally aimed at developing a distributed intrusion detection system. To overcome the weakness of a single intrusion detection system, a number of frameworks have been proposed, which describe the collaborative use of network-based and hostbased systems . Systems that employ both signature based and behavior-based techniques are discussed in the authors describe a data mining framework for building adaptive intrusion detection models.

Two Approach

INTEGRATING LAYERED APPROACH WITH CONDITIONAL RANDOM FIELD

The LIDS represents a sequential Layered Approach and is based on ensuring availability, confidentiality, and integrity of data and (or) services over a network.

The goal of using a layered model is to reduce computation and the overall. This can be achieved by making the layers autonomous and self-sufficient to block an attack without the need of a central decision-maker.

Each layer is then separately trained with a small set of relevant features. Feature selection is significant for Layered Approach and discussed in the next section.

In order to make the layers independent, some features may be present in more than one layer. The layers essentially act as filters that block any anomalous connection, thereby eliminating the need of further processing at subsequent layers enabling quick response to intrusion. The effect of such a sequence of layers is that the anomalous events are identified and blocked as soon as they are detected.

Main Modules:-

CONDITIONAL RANDOM FIELD

The CRFs have proven to be very successful in such tasks, as they do not make any unwarranted assumptions about the data. Hence, we explore the suitability of CRFs for intrusion detection. system may consider features such as "logged in" and "number of file creations." When these features are analyzed individually, they do not provide any information that can aid in detecting attacks.

Probe layer

The probe attacks are expected at acquiring information about the target network from a source that is often external to the network

DoS layer,

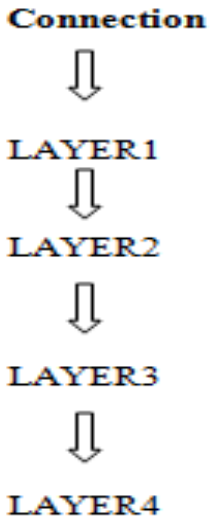
For the DoS layer, traffic features such as the "percentage of connections having same destination host and same service" and packet level features such as the "source bytes" and "percentage of packets with errors" are significant. To detect DoS attacks, it may not be important to know whether a user is "logged in or not."

R2L layer

The R2L attacks are one of the most difficult to detect as they involve the network level and the host level features. We therefore selected both the network level features such as the "duration of connection" and "service requested" and the host level features such as the "number of failed login attempts" among others for detecting R2L attack.

U2R layer (User to Root attacks)

The U2R attacks involve the semantic details that are very difficult to capture at an early stage. Such attacks are often content based and target an application. Hence, for U2R attacks, we selected features such as "number of file creations" and "number of shell prompts invoked," while we ignored features such as "protocol" and "source bytes."



Algorithm & Training:

Step 1: Select the number of layers, n, for the complete system.

Step 2: Separately perform features selection for each layer.

Step 3: Train a separate model with CRFs for each layer using the features selected from Step 2.

Step 4: Plug in the trained models sequentially such that only the connections labeled as normal are passed to the next layer. Testing

Step 5: For each (next) test instance perform Steps 6 through 9.

Step 6: Test the instance and label it either as attack or normal.

Step 7: If the instance is labeled as attack, block it and identify it as an attack represented by the layer name at which it is detected and go to Step 5. Else pass the sequence to the next layer.

Step 8: If the current layer is not the last layer in the system, test the instance and go to Step 7. Else go to Step 9.

Step 9: Test the instance and label it either as normal or as an attack. If the instance is labeled as an attack, block it and identify it as an attack corresponding to the layer name

CONCLUSIONS

In this paper, we have addressed the dual problem of Accuracy and Efficiency for building robust and efficient intrusion detection systems. Our experimental results show that CRFs are very effective in improving the attack detection rate and decreasing the FAR. Having a low FAR is very important for any intrusion detection system. Further, feature selection and implementing the Layered Approach significantly reduce the time required to train and test the model. Even though we used a relational data set for our experiments, we showed that the sequence labeling methods such as the CRFs can be very effective in detecting attacks and they outperform other methods that are known to work well with the relational data.This can further be extended to implement pipelining of layers in multicore processors,which is likely to result in very high performance.

REFERENCES

[1] Autonomous Agents for Intrusion Detection, http://www.cerias. purdue.edu/research/aafid/, 2010.

[2] CRF++: Yet Another CRF Toolkit, http://crfpp.sourceforge.net/, 2010.

[3] KDD Cup 1999 Intrusion Detection Data, http://kdd.ics.uci.edu/ databases/kddcup99/kddcup99.html, 2010.

[4] Overview of Attack Trends, http://www.cert.org/archive/pdf/ attack_trends.pdf, 2002.

[5] Probabilistic Agent Based Intrusion Detection, http://www.cse.sc. edu/research/isl/agentIDS.shtml, 2010. GUPTA ET AL.: LAYERED APPROACH USING CONDITIONAL RANDOM FIELDS FOR INTRUSION DETECTION 47 Authorized licensed use limited to: UNIVERSITY OF MELBOURNE. Downloaded on July 19,2010 at 00:12:39 UTC from IEEE Xplore. Restrictions apply.

[6] SANS Institute—Intrusion Detection FAQ, http://www.sans.org/ resources/idfaq/, 2010.

[7] T. Abraham, IDDM: Intrusion Detection Using Data Mining Techniques, http://www.dsto.defence./gov.au/publications/ 2345/DSTO-GD-0286.pdf, 2008.

[8] R. Agrawal, T. Imielinski, and A. Swami, "Mining Association Rules between Sets of Items in Large Databases," Proc. ACM SIGMOD, vol. 22, no. 2, pp. 207-216, 1993.

[9] N.B. Amor, S. Benferhat, and Z. Elouedi, "Naive Bayes vs. Decision Trees in Intrusion Detection Systems," Proc. ACM Symp. Applied Computing (SAC '04), pp. 420-424, 2004.

[10] J.P. Anderson, Computer Security Threat Monitoring and Surveillance, http://csrc.nist.gov/publications/history/ande80.pdf, 2010.

[11] R. Bace and P. Mell, Intrusion Detection Systems, Computer Security Division, Information Technology Laboratory, Nat'l Inst. of Standards and Technology, 2001.

[12] D. Boughaci, H. Drias, A. Bendib, Y. Bouznit, and B. Benhamou, "Distributed Intrusion Detection Framework Based on Mobile Agents," Proc. Int'l Conf. Dependability of Computer Systems (DepCoS-RELCOMEX '06), pp. 248-255, 2006.

[13] Y. Bouzida and S. Gombault, "Eigenconnections to Intrusion Detection," Security and Protection in Information Processing Systems, pp. 241-258, 2004.

[14] H. Debar, M. Becke, and D. Siboni, "A Neural Network Component for an Intrusion Detection System," Proc. IEEE Symp. Research in Security and Privacy (RSP '92), pp. 240-250, 1992.

[15] T.G. Dietterich, "Machine Learning for Sequential Data: A Review," Proc. Joint IAPR Int'l Workshop Structural, Syntactic, and Statistical Pattern Recognition (SSPR/SPR '02), LNCS 2396, pp. 15-30, 2002.

[16] P. Dokas, L. Ertoz, A. Lazarevic, J. Srivastava, and P.-N. Tan, "Data Mining for Network Intrusion Detection," Proc. NSF Workshop Next Generation Data Mining (NGDM '02), pp. 21-30, 2002.

[17] Y. Du, H. Wang, and Y. Pang, "A Hidden Markov Models-Based Anomaly Intrusion Detection Method," Proc. Fifth World Congress on Intelligent Control and Automation (WCICA '04), vol. 5, pp. 4348-4351, 2004.

[18] S. Dzeroski and B. Zenko, "Is Combining Classifiers Better than Selecting the Best One," Proc. 19th Int'l Conf. Machine Learning (ICML '02), pp. 123-129, 2002.

[19] L. Ertoz, A. Lazarevic, E. Eilertson, P.-N. Tan, P. Dokas, V. Kumar, and J. Srivastava, "Protecting against Cyber Threats in Networked Information Systems," Proc. SPIE Battlespace Digitization and Network Centric Systems III, pp. 51-56, 2003.
[20] S. Forrest, S.A. Hofmeyr, A. Somayaji, and T.A. Longstaff, "A Sense of Self for Unix Processes," Proc. IEEE Symp. Research in Security and Privacy (RSP '96), pp. 120-128, 1996.
[21] Y. Gu, A. McCallum, and D. Towsley, "Detecting Anomalies in Network Traffic Using Maximum Entropy Estimation," Proc. Internet Measurement Conf. (IMC '05), pp. 345-350, USENIX Assoc., 2005.
[22] K.K. Gupta, B. Nath, and R. Kotagiri, "Network Security Framework," Int'l J. Computer Science and Network Security, vol. 6, no. 7B, pp. 151-157, 2006.
[23] K.K. Gupta, B. Nath, and R. Kotagiri, "Conditional Random Fields for Intrusion Detection," Proc. 21st Int'l Conf. Advanced Information Networking and Applications Workshops (AINAW '07), pp. 203-208, 2007.
[24] K.K. Gupta, B. Nath, R. Kotagiri, and A. Kazi, "Attacking Confidentiality: An Agent Based Approach," Proc. IEEE Int'l Conf. Intelligence and Security Informatics (ISI '06), vol. 3975, pp. 285-296, 2006.
[25] C. Ji and S. Ma, "Combinations of Weak Classifiers," IEEE Trans. Neural Networks, vol. 8, no. 1, pp. 32-42, 1997.
[26] D.S. Kim and J.S. Park, "Network-Based Intrusion Detection with Support Vector Machines," Proc. Information Networking, Networking Technologies for Enhanced Internet Services Int'l Conf. (ICOIN '03), pp. 747-756, 2003.
[27] D. Klein and C.D. Manning, "Conditional Structure versus Conditional Estimation in NLP Models," Proc. ACL Conf. Empirical Methods in Natural Language Processing (EMNLP '02), vol. 10, pp. 9-16, Assoc. for Computational Linguistics, 2002.
[28] C. Kruegel, D. Mutz, W. Robertson, and F. Valeur, "Bayesian Event Classification for Intrusion Detection," Proc. 19th Ann. Computer Security Applications Conf. (ACSAC '03), pp. 14-23, 2003.
[29] J. Lafferty, A. McCallum, and F. Pereira, "Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data," Proc. 18th Int'l Conf. Machine Learning (ICML '01), pp. 282-289, 2001.
[30] W. Lee and S. Stolfo, "Data Mining Approaches for Intrusion Detection," Proc. Seventh USENIX Security Symp. (Security '98), pp. 79-94, 1998.
[31] W. Lee, S. Stolfo, and K. Mok, "Mining Audit Data to Build Intrusion Detection Models," Proc. Fourth Int'l Conf. Knowledge Discovery and Data Mining (KDD '98), pp. 66-72, 1998.
[32] W. Lee, S. Stolfo, and K. Mok, "A Data Mining Framework for Building Intrusion Detection Model," Proc. IEEE Symp. Security and Privacy (SP '99), pp. 120-132, 1999.
[33] A. McCallum, "Efficiently Inducing Features of Conditional Random Fields," Proc. 19th Ann. Conf. Uncertainty in Artificial Intelligence (UAI '03), pp. 403-410, 2003.
[34] A. McCallum, D. Freitag, and F. Pereira, "Maximum Entropy Markov Models for Information Extraction and Segmentation," Proc. 17th Int'l Conf. Machine Learning (ICML '00), pp. 591-598, 2000.
[35] A.K. McCallum, MALLET: A Machine Learning for Language Toolkit, http://mallet.cs.umass.edu, 2010.
[36] L. Portnoy, E. Eskin, and S. Stolfo, "Intrusion Detection with Unlabeled Data Using Clustering," Proc. ACM Workshop Data Mining Applied to Security (DMSA), 2001.
[37] A. Ratnaparkhi, "A Maximum Entropy Model for Part-of-Speech Tagging," Proc. Conf. Empirical Methods in Natural Language Processing (EMNLP '96), pp. 133-142, Assoc. for Computational Linguistics, 1996.
[38] M. Sabhnani and G. Serpen, "Application of Machine Learning Algorithms to KDD Intrusion Detection Dataset within Misuse Detection Context," Proc. Int'l Conf. Machine Learning, Models, Technologies and Applications (MLMTA '03), pp. 209-215, 2003.
[39] H. Shah, J. Undercoffer, and A. Joshi, "Fuzzy Clustering for Intrusion Detection," Proc. 12th IEEE Int'l Conf. Fuzzy Systems (FUZZ-IEEE '03), vol. 2, pp. 1274-1278, 2003.
[40] C. Sutton and A. McCallum, "An Introduction to Conditional Random Fields for Relational Learning," Introduction to Statistical Relational Learning, 2006.
[41] E. Tombini, H. Debar, L. Me, and M. Ducasse, "A Serial Combination of Anomaly and Misuse IDSes Applied to HTTP Traffic," Proc. 20th Ann. Computer Security Applications Conf. (ACSAC '04), pp. 428-437, 2004.
[42] W. Wang, X.H. Guan, and X.L. Zhang, "Modeling Program Behaviors by Hidden Markov Models for Intrusion Detection," Proc. Int'l Conf. Machine Learning and Cybernetics (ICMLC '04), vol. 5, pp. 2830-2835, 2004.
[43] C. Warrender, S. Forrest, and B. Pearlmutter, "Detecting Intrusions Using System Calls: Alternative Data Models," Proc. IEEE Symp. Security and Privacy (SP '99), pp. 133-145, 1999.
[44] I.H. Witten and E. Frank, Data Mining: Practical Machine Learning Tools and Techniques. Morgan Kaufmann, 2005.
[45] Y.-S. Wu, B. Foo, Y. Mei, and S. Bagchi, "Collaborative Intrusion Detection System (CIDS): A Framework for Accurate and Efficient IDS," Proc.

19th Ann. Computer Security Applications Conf. (ACSAC ’03), pp. 234-244, 2003.
[46] Z. Zhang, J. Li, C.N. Manikopoulos, J. Jorgenson, and J. Ucles, “HIDE: A Hierarchical Network Intrusion Detection System Using Statistical Preprocessing and Neural Network Classification,”
Proc. IEEE Workshop Information Assurance and Security (IAW ’01), pp. 85-90, 2001.