International Journal of Pure and Applied Mathematics
Volume 119 No. 12 2018, 753-762
ISSN: 1314-3395 (on-line version)
url: http://www.ijpam.eu
Special Issue

Logo Based Pattern Matching Algorithm for Intrusion Detection System in Wireless Sensor Network

Dr.J.P.Ananth[1], Dr.S.Balakrishnan[2], S.P.Premnath[3]
[1]Department of Computer Science and Engineering
[2]Department of Information Technology,
[3] Department of Electronics and Communication Engineering,
[1,2,3]Sri Krishna College of Engineering and Technology, Coimbatore, India.

Abstract

Wireless Sensor Network (WSN) is a blend of detecting, calculation, and correspondence into a single little gadget. A sensor system comprises of a variety of various sensor systems of different sorts interconnected by a wireless communication network. Sensor information is shared between these sensor hubs and utilized as contribution to a distributed estimation system. The Intrusion Detection System (IDS) is one of the major defensive methods used against attack in WSN. There are three types of methodologies used for detecting intrusion on the network. They are "Signature based, Anomaly based and stateful protocol analysis." In this paper, we focus intrusion detection system methodology. The intrusion can happen either in the header part or payload part. Notwithstanding, intrusion detection system can be either "host-based or network-based" intrusion. Also, signature based detection strategies are utilized to identify anomalous bundles or assault in network- based intrusion detection systems. Thus, the difficulties are confronted by a large portion of the signature based detection systems. For instance, SNORT tool is to identify pernicious movement at higher activity system, which brought about a packet drooping and subjected the system where this signature based framework is arranged as a system edge security. In this paper, we propose Logo Pattern Matching (LPM) calculation that will diminish false alert tally. Intrusion can be conceivable on payload. So patterns are characterized to recognize the interruption however it is conceivable that example is typical information. IDS framework creates false alert to distinguish that pattern (example) in information.

Keywords: Wireless sensor network, intrusion detection system, snort, logo pattern matching.

1. Introduction

System (Network) security has as of late gotten a huge consideration because of the mounting security worries in today's systems (networks). An "Intrusion Detection System (IDS) investigates data from a PC or a system (network) to distinguish vindictive activities and practices that can trade off the security of a PC framework". At the point when a malevolent (malicious) conduct is distinguished, an alert (i.e. alarm) is dispatched. A "wide assortment of algorithms have been proposed which can recognize and battle with these security dangers". Among every one of these proposition, "signature based Network Intrusion Detection Systems (NIDS) have been a business achievement and have seen a broad selection". While, these frameworks as of now produce a few many million dollars in income, it is anticipated to ascend to more than 2 billion dollars by 2020.

A NIDS goes for identifying conceivable interruptions, for example, a vindictive action, PC assault and/or PC abuse, spread of an infection (virus), and so on, and alarming the best possible people upon identification. A NIDS screens and examines the information bundles that go over a system searching for such suspicious exercises. A "substantial NIDS server can be set up on the connections of a spine system, to screen all movement; or littler frameworks can be set up to screen activity coordinated to a specific server, switch, gateway, or router". Another class of NIDS can be setup at a brought together server, which will filter the framework records, searching for unapproved action and to keep up information integrity.

Generally, IDS's have been delegated either signature identification frameworks (likewise called negative methodology) or anomaly discovery frameworks (positive methodology). A hybrid intrusion recognition framework consolidates the systems of the two methodologies. The signature based methodology searches for the marks of known assaults (abuse of the framework assets), which misuse shortcomings in framework and application programming. It utilizes pattern matching strategies against an often upgraded database of assault marks. It is helpful to identify definitely known assaults or their slight varieties, however not the new ones or malevolent varieties that annihilation the pattern recognition engine. The anomaly-based methodology searches for conduct or utilization of PC assets going amiss from "ordinary" or "normal" conduct [1]. The hidden standard of this methodology is that "assault conduct" contrasts enough from "typical client conduct" accordingly it can be distinguished by indexing and recognizing the distinctions included. To start with, the "typical" conduct must be very much characterized, which is not a simple undertaking. When typical conduct is completely qualified, unpredictable conduct will be labeled as nosy.

The “pattern search problem in Intrusion Detection Systems is a specific issue in it's own particular right”. It requires thought of numerous issues connected with example seeks. The accompanying considerations ought to be obliged by any pattern search engine utilized for ongoing Intrusion Detection. "a) Multi-pattern search algorithms b) Pattern character case sensitivity c) Pattern sizes d) Pattern group size e) Alphabet size f) Algorithmic attacks g) Search text size h) Frequency of searches."

1.1 NIDS and Network Architecture

NIDS is deployed such that it screens the movement that crosses any given connection inside the system, in this manner giving an expanded security (appeared in Figure 1.1). Therefore the NIDS is conveyed close to the exchanging hubs inside the neighborhood organize, and close to the entrance switches at the system limit. In such designs, “the NIDS will no more screen the movement that has been obstructed by the firewall, which will prompt a quite decreased false alert rates”. A disadvantage however is that there will be numerous cases of NIDS, and it will get to be monotonous to stay up with the latest in say an expansive undertaking system. Such designs are prominent in “e-commerce back end systems, comprising of web and mail servers and database and capacity servers, as an expanded security is alluring there”. It likewise helps in keeping a tainted server to contaminate the others inside the system.

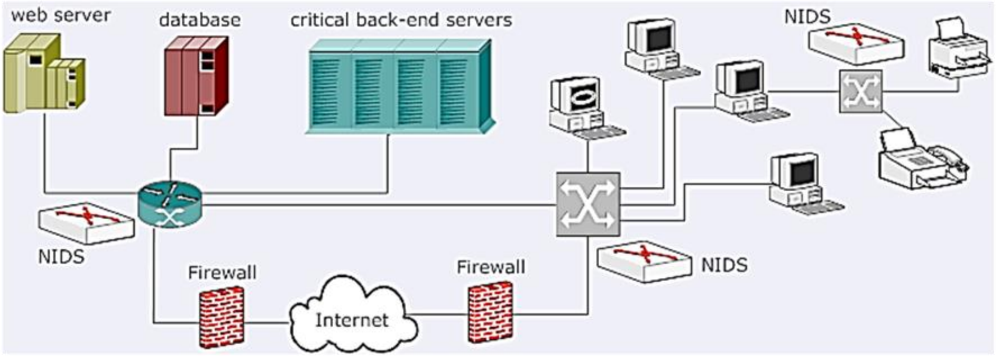

Figure 1. NIDS in complete deployment mode

2. Literature Survey

Pattern matching has been a subject of serious examination that has brought about a significant number of productions in the previous quite a few years. Two of the most famous calculations are those distributed by Aho and Corasick [2] and Boyer and Moore [3] just about 30 years back. The Aho-Corasick calculation develops a finite state machine (FSM) for recognizing all events of a given arrangement of examples by handling the contribution to a solitary pass, playing out a state move for every info character. The Boyer-Moore calculation, which is fundamentally a solitary example coordinating arrangement, abuses two heuristics (named "terrible character" and "great suffix") to skip bits of the information stream keeping in mind the end goal to enhance the normal execution. Ideas like Aho-Corasick and Boyer-Moore can be found in numerous other example coordinating calculations, for example, “the calculations by Commentz-Walter [4], Wu and Manber [5], and all the more as of late, the Aho-Corasick-Boyer-Moore (AC BM) calculation by Coit et al. [6] and the Setwise Boyer-Moore-Horspool (SBMH) plan by Fisk and Varghese [7]”. The heuristics connected by Boyer-Moore, Wu-Manber and different plans for skipping segments of the info empower a great normal case execution for an assortment of uses, yet the subsequent reliance of the preparing rate on information (and example) attributes is a vital disservice for interruption discovery, as it makes the NIDS defenseless against assaults that attempt to over-burden the example coordinating operation by creating most pessimistic scenario info situations [8] [9]. The Aho-Corasick calculation has a deterministic execution and, in this way, does not have this disadvantage. Rather, its principle inconvenience is the expansive stockpiling prerequisites expected to execute the FSM. Tuck et al. proposed two streamlined variants of Aho-Corasick to address this issue, and could accomplish a significant lessening away prerequisites [10].

Signature based IDS frameworks “require that their information bases should be redesigned routinely at various time interims in order to recognize the fast approaching strings created on the system”. This procedure is a “calm tedious and requires a snappy hidden framework to overhaul the database”. Two layer signature based model was proposed to address signature based recognitions with unequal databases. However, this model doesn't have any system for including, expelling or overhauling marks in the huge mark based database [11].

One of the soonest and productive calculations in accurate multi-design string coordinating is proposed by Aho-Corasick [12]; various late frameworks utilize this method. This calculation empowers string coordinating in a period direct in the span of the information. Sommer et.al, [13] proposed the idea that general expressions may end up being on a very basic level more proficient and adaptable when contrasted with careful match strings while indicating marks for NIDS. The most mainstream representation of customary expressions is limited state automata. This is proposed by Hopcroft et.al, [14]. There are two essential sorts: Deterministic Finite Automaton (DFA) and Non-deterministic Finite Automaton

(NFA). A DFA comprises of a letters in order, which is a limited arrangement of info images, a limited arrangement of states, an underlying state and a move capacity, which indicates the move from each state for each image in the letter set. In systems administration applications, the letter set by and large comprises of 256 ASCII characters. A key property of DFA is that in any given express, the move capacity gives back a solitary next state for any given information image; along these lines whenever, one and only state is dynamic in a DFA. The refinement between a NFA and a DFA lies in their move capacity: rather than giving back a solitary next state, the move capacity of a NFA gives back an arrangement of states, which might be a void set. In this manner, various states can be at the same time dynamic in a NFA.

The Karp-Rabin (KR) Algorithm was made by Michael Rabin and Richard Karp. They utilized a totally distinctive methodology than the single watchword techniques [15]. The Boyer-Moore calculation is one of the definite string coordinating calculations that utilized as a part of single example coordinating. The calculation utilizes two tables or capacities, which is utilized to move the sliding window to one side. The primary table is called "terrible character shift", while the second table called "great postfix shift". The calculation is quicker when it is working with little example size, yet it is slower when it is working with vast example size [16]. Balakrishnan [17] proposed the concept iTrust. iTrust is presenting an intermittently accessible Trusted Authority (TA) to judge the hub's conduct in view of the gathered steering proofs and probabilistically checking. We show iTrust as the Inspection Game and utilization diversion hypothetical investigation to exhibit that, by setting suitable examination likelihood. Balakrishnan et.al, [18] proposed an AODV (Ad-hoc on-demand distance vector) protocol for finding routes only as needed. Also, utilization of Sequence numbers to track precision of data.

3. System Architecture

False positive has turned into a basic component in deciding the accomplishment of IDS. False positive is an occasion that cautions IDS without assault being happened. So IDS improvement ought not just concentrate on its capacity in recognizing genuine assaults additionally on its capacity to stifle the false caution. For any framework, normally false cautions produced every day dwarf the genuine alerts. Around 96% of alarms created are affirmed as false positive.

We propose a calculation, "LPM" (Logo Pattern Matching) that will decrease false caution tally. Intrusion can be conceivable on payload. So examples (patterns) are characterized to distinguish the intrusion however it is conceivable that example is typical information. IDS framework creates false caution to distinguish that example in information. In this work, our speculation is that if example is rehashed more than two times, probability is higher that it will be intrusion rather than ordinary information. In such cases, the caution (alarm) will be activated. This calculation will make first table to decrease the quantity of correlation and afterward analyze.

The "two phases of the proposed algorithms are: (i) preprocessing phase, (ii) searching phase and two tables namely Prefix Table and Suffix Table are used".

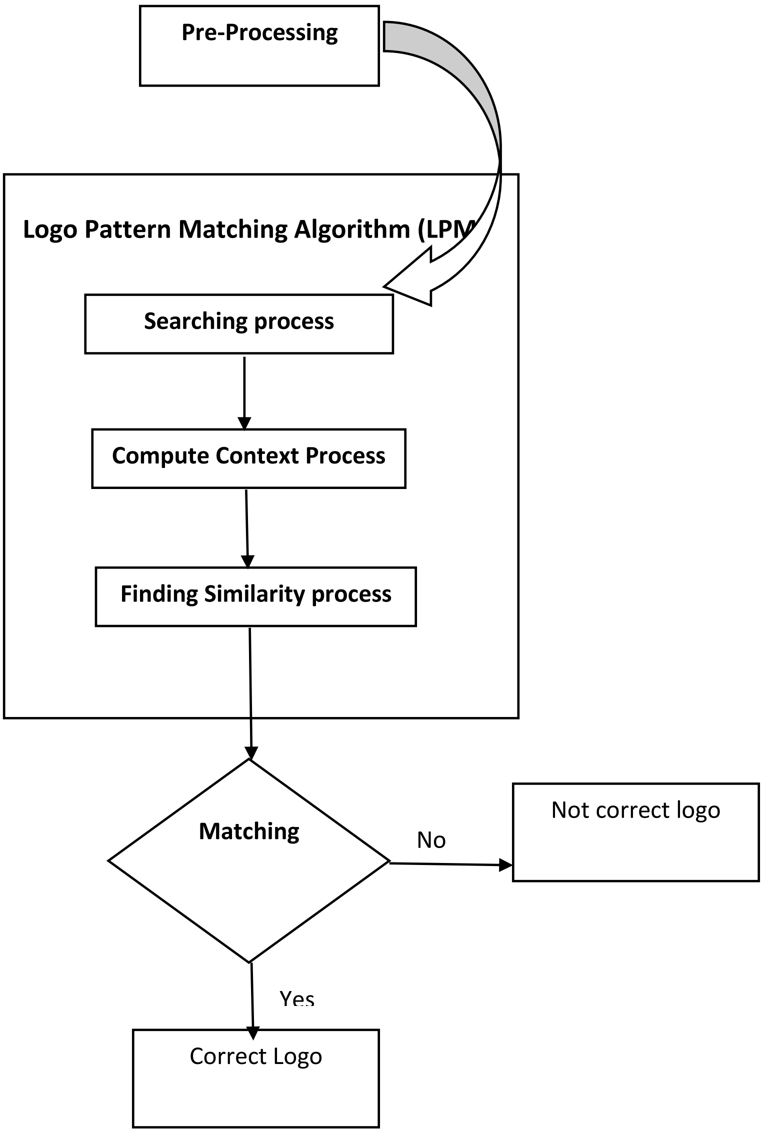

Figure 2. Process of Logo Pattern Matching Algorithm

Step1: (Pre-processing)

In the primary stage, the Logo Pattern Matching Algorithm (LPM) plays out the same preprocessing stage as in the current two calculations. It readies the hash capacity utilized as a part of KR calculation and the bmBc (Boyer-Moore bad character) table utilized as a part of HP calculation for the example.

The "bmBc table is the same as it was in the current HP calculation and the calculation preprocesses the example and makes tables, which are known as Boyer-Moore bad character (bmBc) and Boyer-Moore good-suffix (bmGs) tables". For every character in the letters in order set, an awful character table stores the movement esteem in light of the event of the character in the example.

Step 2: (Searching phase)

The "process of computing hash functions for the patterns and text window are exactly the same as the process of creating them in the existing Karp-Rabin Algorithm (KR) algorithm".

In the searching phase, the Logo Pattern Matching Algorithm (LPMA) performs "the comparison between the pattern and the text by utilizing the advantages of the KR and HP".

Step 3: (Compute Context)

"After the preprocessing phase has finished, the comparison start between the text and pattern by comparing the numerical value of pattern hash and window text hash".

Step 4: (Finding Similarity)

"Check for the two hash value similarities then the Logo Pattern Matching Algorithm performs the shifting and the logo is attached".

Step 5: (Matching phase)

Next Shift is to the right side based first character's values for the "window text in the bmBc table and this will speed up the algorithm during the comparison process and it reduces the number of character comparison by using the hash function".

6. Conclusion

Similarly as with most new advancements, wireless has a few vulnerabilities. Note that supreme security is a theoretical idea – it doesn't exist anyplace. All systems are helpless against "insider or outsider attacks, and listening stealthily (eavesdropping)". Nobody needs to hazard having the information presented to the easygoing spectator or open malevolent devilishness. Wireless IDS arrangements are accessible from both the open-source and business markets and both have their own preferences. In this paper, we presented a logo based pattern matching for recognizing intrusion. It is guaranteed that every one of the dangers to the system are distinguished without trading off the execution. It will be executed utilizing SNORT apparatus as a part of future.

References

1. García-Teodoro P., Díaz-Verdejo J., Maciá-Fernández G., Vázquez E. (2009), "Anomaly-based network intrusion detection: Techniques,

systems and challenges," Computers and Security, 28, 1-2, pp. 18-28 (2009)

2. A.V. Aho and M.J. Corasick, (1975), "Efficient string matching: An aid to bibliographic search," Communications of the ACM, vol. 18, no. 6, pp. 333-340, 1975.
3. R.S. Boyer and J.S. Moore, (1977), "A fast string searching algorithm," Communications of the ACM, vol. 20, no. 10, pp. 762-772, Oct. 1977.
4. B. Commentz-Walter, (1979) "A string matching algorithm fast on the average," Proceedings of the 6th Colloquium, on Automata, Languages and Programming, pp. 118-132, July 1979.
5. S. Wu and U. Manber, (1994), "A fast algorithm for multi-pattern searching," Technical report TR-94-17, Department of Computer Science, University of Arizona, May 1994.
6. C. Coit, S. Staniford, and J. McAlerney, (2002), "Towards faster string matching for intrusion detection," Proceedings of the DARPA Information Survivability Conference and Exhibition, pp. 367-373, 2002.
7. M. Fisk and G. Varghese, "Applying fast string matching to intrusion detection" [Online]. Available: http://woozle.org/~mfisk/ papers/.
8. V. Paxson, (1999), "Bro: A system for detecting network intruders in real-time," Computer Networks, vol. 31, no. 23-24, pp. 2435-2463, December 1999.
9. S. Antonatos, K.G. Anagnostakis, and E.P. Markatos, "Generating realistic workloads for intrusion detection systems," Proc. of the 4th ACM SIGSOFT/SIGMETRICS Workshop on Software and Performance, pp. 207-215, January 2004.
10. N. Tuck, T. Sherwood, B. Calder, and G. Varghese, "Deterministic memory-efficient string matching algorithms for intrusion detection," Proceedings IEEE Infocom, vol. 4, pp. 2628-2639, March 2004.
11. M. Salour and Xiao Su, (2007), "Dynamic Two-Layer Signature-Based IDS with Unequal Databases", presented at proceeding of the International Conference on Information Technology (ITNG'07), 2007.
12. A. V. Aho and M. J. Corasick, (1975), "Efficient string matching: An aid to bibliographic search," Communications of the ACM, Vol.18, issue 6, pp. 333-340, 1975. portal.acm.org/citation.cfm?id=360825.360855.
13. R. Sommer, V. Paxson, (2003), "Enhancing Byte-Level Network Intrusion Detection Signatures with Context," ACM conference on Computer and Communication Security, 2003, pp. 262-271.
14. J. E. Hopcroft and J. D. Ullman, (1979), "Introduction to Automata Theory, Languages, and Computation," Addison Wesley, 1979.

15. Ricardo A. Baeza-Yates, (1992), “String Searching Algorithms,” pp.1-18, 1992.
16. K. Prabha, S.Sukumaran, (2014), “Improved Single Keyword Pattern Matching Algorithm for Intrusion Detection System”, International Journal of Computer Applications, Vol. 90, No 9, March 2014.
17. S.Balakrishnan, “A Secure iTrust Scheme towards Trust Establishment in Delay Tolerant Networks Using Zone Based Routing Protocol”, Asian Journal of Information Technology, 2016, Vol. 15, Issue 22, pp. 4535-4540. DOI: 10.3923/ajit.2016.4535.4540
18. S.Balakrishnan, Vinod K, B. Shaji, “Secured and Energy Efficient AODV Routing Protocol For Wireless Sensor Network”, International Journal of Pure and Applied Mathematics, Vol. 119, No. 10, 2018, pp. 563-570.