Literature Survey on Signature Matching Schemes

[1] Balaji.S
Department of Computer Science and Engineering
Akshaya College of Engineering and Technology, Coimbatore

[2] Priyanka.N
Department of Computer Science and Engineering
Akshaya College of Engineering and Technology, Coimbatore

[3] Jeevanandham.S
Department of Computer Science and Engineering
Akshaya College of Engineering and Technology, Coimbatore

Abstract

An intrusion detection system is a hardware or software that is used to detect any intrusion that takes place inside a network or a computer system. All Intrusion Detection Systems use one of two detection techniques: Anomaly-based IDS, Signature-based IDS. The anomaly-based IDS identifies an attack by discovering significant deviations between the established behaviour and the current behaviour. The major drawback of anomaly detection is greater false alarm rate and if the malicious behaviour of the user falls under the accepted behaviour, then it goes unnoticed. So the Signature based IDS is used to increase the probability of intrusion detection. But in Signature based IDS, signature matching limits the performance. To improve the performance various techniques are used.

1. Introduction

An intrusion detection system (IDS) is a device or software application that monitors network or system activities for malicious activities or policy violations and produces reports to a management station. Intrusion detection system uses one of the two detection techniques: Anomaly-based IDS, Signature-based IDS.An Anomaly-Based Intrusion Detection System, is a system for detecting computer intrusions and misuse by monitoring system activity and classifying it as either normal or anomalous.In a signature based intrusion detection system a predetermined attack patterns in the form of signatures and these signatures are further used to determine the network attacks.

2. Adaptive Blacklist Based Packet Filter In Nids

To improve the performance of Signature based IDS adaptive blacklist based packet filter technique is used. The technique construct the packet filter using blacklist technique which in turn reduces the burden of signature based NIDS. The technique consists of: Blacklist packet filter and Monitor Engine.

2.1 BLACKLIST PACKET FILTER

Blacklist packet filter is used to filter the network packets based on blacklist. It is also used to compare the signature and the packet payload. The components are: Blacklist and Look-up table.

2.2 BLACKLIST

The blacklist contains all blacklisting IP addresses that can be updated periodically.

2.3 LOOK-UP TABLE

The look-up table consists of two sub- tables: Matched NIDS signatures and All NIDS signatures. The All NIDS signatures table stores all signatures that are currently active in the NIDS and the Matched NIDS signatures table records all the NIDS signatures that are not matched during the packet inspection.

2.4 MONITOR ENGINE

The monitor engine is used to collect the statistical data and calculate the confidences of IP addresses. The monitor engine periodically updates the blacklist in the blacklist packet filter and enables the packet filter to conduct packet filtration.

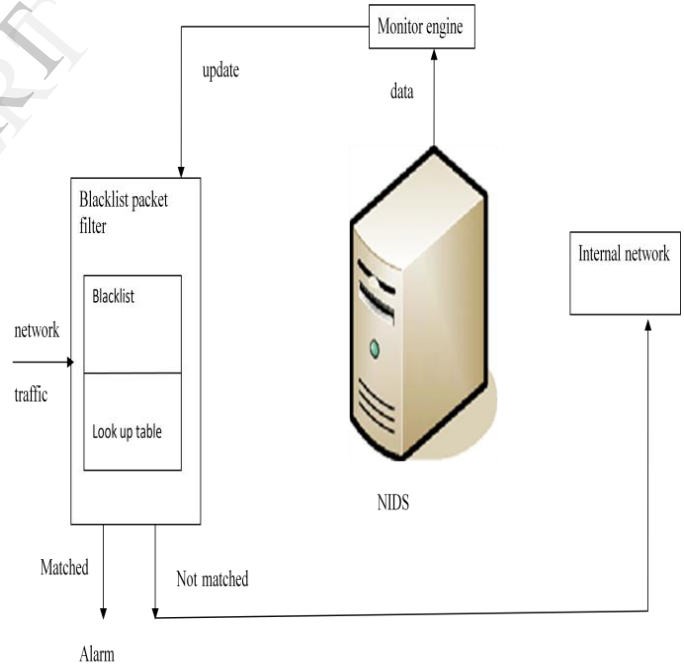

Fig 1. The interactions between the blacklist packet filter, the monitor engine and the network intrusion detection system[1]

When the network traffic reaches the blacklist packet filter and search for IP addresses in the blacklist.

- If the IP address of the packet is in the blacklist, then the blacklist packet filter will compare the payload of the packet with the signatures in its look

up table. If a match is identified, then the blacklist packet filter will block the packet and give an alert.

- If the payload of the packet does not match any signatures, then this packet will be sent to the internal network directly.
- If the IP address of the packet is not found in the blacklist, then the packet will be sent to the NIDS.

2.5 DRAWBACK

The technique uses the constant weight value to calculate the IP confidence which is used by monitor engine to filter the packets.The constant weight value is affected by the environmental changes.

3. Adaptive Character Frequency Based Exclusive Signature Matching Scheme

To improve performance , an adaptive character frequency-based exclusive signature matching scheme can be implemented in a signature-based NIDS .The scheme also reduces the time consumption to detect the intruder comparing with snort.The scheme consists of : SNS, SCQ1 SCQ2,MNS,Decisioncomponent,Communication component.

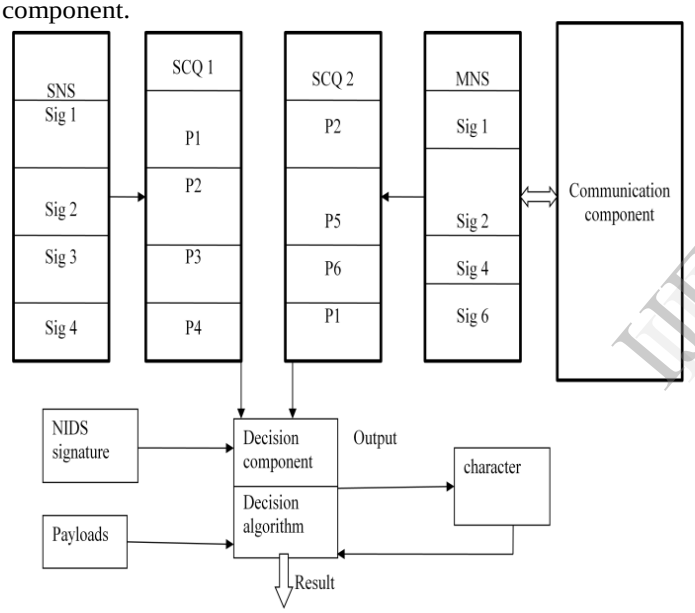

Fig 2. The architecture of our adaptive character frequency-based exclusive signature matching scheme [2]

3.1 SNS-Stored NIDS signatures

The table contains all active NIDS signatures that is currently used in signature database

3.2 SCQ1 table

The table computes character frequency based on NIDS signatures stored in SNS table

3.3 SCQ2 table

The table calculates and stores character frequency according to NIDS signatures in MNS table

3.4 MNS table

The table contains all NIDS signatures that are matched during the detection procedure

3.5 Decision component

The component is used to conduct exclusive signature matching and verifying the matching results. A decision algorithm is implemented in the component to perform the matching procedure.

3.6 Communication component

The component is responsible for communicating with central analysis server by transmitting status information.

- The input string is compared with set of substring if any of the character in substring does not exist in input string it is considered as mis-match.
- If mis-match does not occur the input string is again processed with Boyer and Moore algorithm.

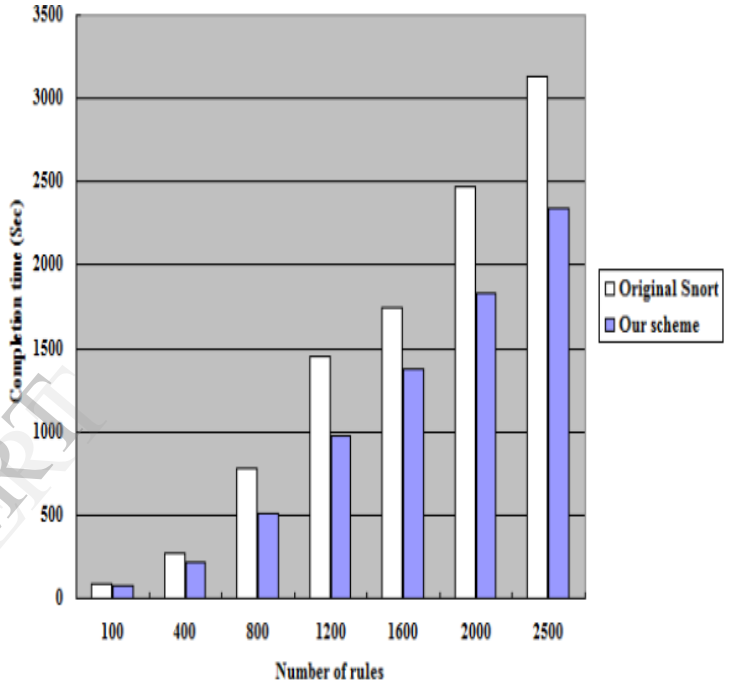

Fig 3. Performance of Snort and proposed scheme [2]

3.7 DRAWBACK

In adaptive character frequency-based exclusive signature matching scheme there is only one level of comparison so the chance for the intruder hacking the system is more.

4. Single Character Frequency Based Exclusive Signature Matching Scheme

To improve performance an single character frequency based exclusive signature matching scheme can be implemented in a signature-based NIDS. The scheme calculates the single character frequency from both stored and matched NIDS signatures. The scheme can adaptively choose the most appropriate character for conducting the exclusive signature matching.

4.1 DRAWBACK

Single Character Frequency Based Exclusive Signature Matching Scheme depends only on single character.So the chance for the intruder hacking the system is more.

5. A Fast Pattern Matching Algorithm With Multi-Byte Search Unit For High-Speed Network Security

Many pattern matching algorithms are used to improve the performance and reduces the time consumption in signature based Intrusion detection system. But most of the algorithms uses single-byte standard unit for search.The paper proposes the multi-byte search unit to improve the performance and reduces the time consumption. A new pattern matching algorithm called the L^{+1}-MWM algorithm for multi-pattern matching. Modified Wu–Manber (MWM) algorithm uses two-byte unit for search.The proposed algorithm enhance MWM algorithm by extending the length of the shortest pattern.The advantage of the proposed scheme are:

- L^{+1}-MWM algorithm improves the performance of the MWM algorithm by 20% when the pattern has shortest length and in normal traffic conditions.
- When the length of the shortest pattern in a rule set is less than 5, the L^{+1}-MWM algorithm provides 38.87% improvement.
- This algorithm implements in real campus network and provides 12.48% improvement.

6. Bit-Parallel Search Algorithms For Long Patterns

Signature Matching is a problem in Signature based IDS.The paper proposed an algorithm for searching long patterns.The algorithm includes three new bit-parallel algorithms BXS, BQL, and QF for searching longer patterns.

6.1 BXS

The first algorithm is BXS (BNDMq with eXtended Shift).This algorithm cut the pattern into m/w' where m-length of the pattern and w'-minimum of m and w, w-minimum shift , consecutive pieces of length w' except for the rightmost piece.The rightmost piece should be shorter.Then these pieces should be in length w'.Then BNDM algorithm is used to search the pattern. When searching this algorithm rotate the bits in D rather than just shifting them to the left as in the standard BNDM.
Where D-state vector

6.2 BQL

BQL (BNDMq Long) is second algorithm.The algorithm uses q-grams.The algorithm increases the effective alphabet size by using overlapping q-grams e.g. when using 3-grams the pattern "ACCTGGT" is processed as "ACC-CCT-CTG-TGG-GGT" where q=3.Then the searching process is similar to BXS.

6.3 QF

The third algorithm, QF (Q-gram Filtering).QF algorithm is used to reduce the space usage.

7. Compact Dfa: Generic State Machine Compression For Scalable Pattern Matching

Pattern matching is the core of all Intrusion Detection Systems. Pattern matching algorithms are based on Deterministic finite automata and this consists of hundreds of patterns, which requires high storage of memory, cost and power consumption. To overcome the problem generic DFA compression algorithm is proposed. The algorithm reduces the rule set to the minimum possible size, only one rule per state. The paper extends the generic DFA scheme, called Compact DFA for total memory minimization. The scheme aims at minimizing the product of the number of rules and the code width, rather than only the number of rules.It consists of three stages:State Grouping, Common Suffix Tree Construction , and State and Rule Encoding.

7.1 State Grouping

State Grouping is used to group the states, by calculating two parameters for each state: common suffix (CS) and longest common suffix (LCS). These parameters are used to encode each state with its label.

7.2 Common Suffix Tree

Common Suffix Tree is used to describe how to encode the rules with smaller number of bits and transform them from being defined on suffixes to being defined on prefixes.

7.3 State and Node Encoding

State and Node Encoding is used to encode the Common Suffix Tree and then the states and rules.

7.4 DRAWBACK

The scheme uses only one symbol at a time for state transitions in Deterministic finite automata.

8. A Memory-Efficient Bit-Split Parallel String Matching Using Pattern Dividing For Intrusion Detection Systems

The paper proposes a memory-efficient parallel string matching scheme to reduce the number of state transitions. The scheme divides the Long target patterns into subpatterns with a fixed length then the deterministic finite automata are built with the subpatterns. So that memory usage in homogeneous string matchers can be efficient.

8.1 *Architecture of FSM Tiles*

Multiple string matchers are adopted for parallel string matching.In a string matcher, several homogeneous FSM take n bits as an input at every cycle. In the state of each FSM tile, the pattern identification information is stored as a partial match vector, where the bit indicates whether the pattern is matched or not in the state.

8.2 Sequential Matching with Divided Patterns

The divided target pattern consists of successive matches is compared with its quotient vector and remnant pattern. If a target pattern is divided by a fixed length f, the sequential matches with the subpatterns in the quotient vector should be detected at f different points.

8.3 String Matching Engine Architecture

A character code of one byte from a payload is inputed in the quotient vector matcher. The quotient vector matcher consists of v string matchers. In the quotient vector matcher, only one bit in total temporary match vectors becomes true because only one subpattern can be matched in the quotient vector matcher per cycle.

In the proposed string matching engine, according to the fixed length of subpatterns in the quotient vector, f, the

numbers of states in the FSMs of the quotient vector matcher, the remnant pattern matcher, and the short pattern matcher were predetermined.

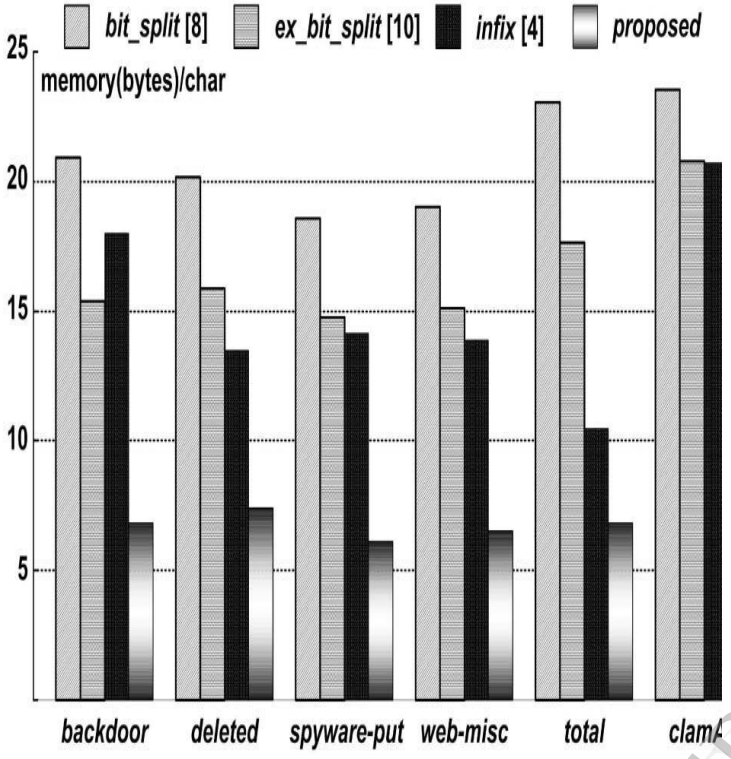

Fig 4. Summary of comparisons with existing bit-split string matching approaches in terms of normalized memory requirements [7]

The advantage of proposed system is: It reduces the total memory requirements of parallel string matching engines.

Table 1. Techniques Vs Parameters

S.No	Techniques	Parameters	Drawback
1	Adaptive Blacklist Based Packet Filter	The technique depends on IP confidence used by blacklist to filter the packets.	The technique uses the constant weight value which is affected by the environmental changes
2	Adaptive Character Frequency Based Exclusive Signature Matching Scheme	The technique depends on single and two length Character Frequency	The technique uses only one level of comparison so the chance for the intruder hacking the system is more.
3	Single Character Frequency Based Exclusive Signature Matching Scheme	The technique depends on single Character Frequency	The technique depends only on single character. the chance for the intruder hacking the system is more
4	A Fast Pattern Matching Algorithm With Multi-Byte Search Unit For High-Speed Network Security	The technique depends on two-byte unit for search	
5	Bit-Parallel Search Algorithms For Long Patterns	The technique depends on q-grams(q consecutive characters together)	

7. Proposed Solution

The performance analysis may include various parameters like ,bit-string length,packet size ,false match rate.Conditions from the information delivered between NIDS agents and central server.

8. Conclusion

An intrusion detection system (IDS) is important to monitor the network or system activities for malicious activities or policy violations and produces reports to a management.the system is more advantageous if the performance is more with less time consumption for intrusion detection.

9. References

[1] Yuxin Meng and Lam-For Kwok. "*Adaptive Blacklist-based Packet Filter with A Statistic-based Approach in Network Intrusion Detection"*. Journal of Network and Computer Applications. vol., no., In Press, Elsevier, 2013.

[2] Yuxin Meng and Wenjuan Li "Adaptive Character Frequency-based Exclusive Signature Matching Scheme in Distributed Intrusion Detection Environment" IEEE 11th International Conference on Trust, Security and Privacy in Computing and Communications,2012

[3] Y. Meng, W. Li, and L.F. Kwok, "Single Character Frequency-based Exclusive Signature Matching Scheme",*In: Proceedings of the 11th IEEE/ACIS International Conference on Computer and Information Science (ICIS 2012)*, In press,2012.

[4] Yoon-Ho Choi, Moon-Young Jung, Seung-Woo Seo "A fast pattern matching algorithm with multi-byte search unit for high-speed network security" Computer Communications 01/2011; 34:1750-1763. DOI:10.1016/j.comcom.2011.

[5] Hyun Jin Kim, Hong-Sik Kim, and Sungho Kang "A Memory-Efficient Bit-Split Parallel String Matching Using Pattern Dividing for Intrusion Detection Systems" *IEEE Transaction on Parallel and Distributed System,2011*

[6] Branislav Durian, Hannu Peltola, Leena Salmela and Jorma Tarhio" Bit-Parallel Search Algorithms for Long Patterns"

In:Proceedings of the 9th international conference on experimental algorithms(sea)2010.

[7] Bremler-Barr, A., Hay, D., Koral, Y.,CompactDFA: generic state machine compression for scalable pattern matching. In: Proceedings of the IEEE INFOCOM; 2010.

[8] H. Song, F. Hao, M. Kodialam, and T. V. Lakshman, "IPv6 lookups using distributed and load balanced bloom filters for 100 gbps core router line cards," in *Proc. IEEE INFOCOM*, 2009.

[9] D. Pao, W. Lin, and B. Liu, "Pipelined architecture for multi string matching," *IEEE Comput. Archit. Lett.*, vol. 7, pp. 33–36, 2008.

[10] W. Lin and B. Liu, "Pipelined parallel AC-based approach for multistring matching," in *Proc. IEEE International Conference on Parallel and Distributed Systems (ICPADS)*, 2008, pp. 665–672.