

International Journal of Computer Science and Mobile Computing



A Monthly Journal of Computer Science and Information Technology

ISSN 2320-088X

IJCSMC, Vol. 4, Issue. 6, June 2015, pg.1008 – 1012

RESEARCH ARTICLE

Detecting Covert Channels in TCP/IP Header with the Use of Naive Bayes Classifier

Ms. Apurva N. Mahajan¹, Prof. I. R. Shaikh²

¹Computer Engineering Department, S. N. D. College of Engg & Research Centre, Yeola, Maharashtra

²Computer Engineering Department, S. N. D. College of Engg & Research Centre, Yeola, Maharashtra
¹apurvambhavsar@gmail.com; ²imran.shaikh22@gmail.com

Abstract— *Covert channels via the widely used TCP/IP protocols have become a new challenging issue for network security. It is a methodology of communication which illicitly transfers data, it means by breaking security policy of system. Any shared resource will be used as a covert channel. It may be a good thing if covert channels are used to protect privacy or increase security of critical communication. Hidden data in the payload can detect by most of the detection systems in early days, but hidden data in IP and TCP packet headers survives a struggle. In this paper, proposed method is based on naive bayes classifier to detect covert channels in TCP ISN and IP ID fields of TCP/IP packets.*

Keywords— *TCP, IP, TCP ISN, IP ID, covert*

I. INTRODUCTION

Computer network is unpredictable due to information warfare and is prone to various attacks. Most of such attacks are devised using special communication channel called covert channel. Covert stands for hidden or non-transparent. Covert channels are scenario oriented. They do not have any concrete definition.

Covert channel concept has proposed by Lampson in 1973. There are different types of covert channels:

- Storage channel: It is based on used shared data storage area.
- Timing channel: Process needs time to perform operation that can be manipulated to provide information to other process.
- Termination channel: A process launches a task.
- Resource Exhaustion channel: Value is provided by availability of specific resource which may be filled up, overloaded.
- Power channel: Information is based on power consumption.

Covert channel has some characteristics like capacity, noise, transmission mode, etc. Capacity is the quantity of information which can transmit through channel. And noise is the amount of perturbations which can interfere with the information. Transmission mode can be synchronous when information is received and managed otherwise it can be asynchronous. Capacity is very important part of global quality of channel. A large capacity channel has possibility to leak more information. Noise will depend on nature of channel.

II. COVERT CHANNELS IN TCP/IP PROTOCOL

Now a days TCP and IP are most widely used internet protocols. Many different ways are used to hide information in TCP/IP header. TCP ISN and IP ID these are the two fields difficult to detect covert data. To communicate across any set of interconnected network, TCP/IP can be used.

Exchanging data directly between two network hosts provides by TCP. Addressing and routing messages across one or more networks handle by IP. IP is network layer protocol and TCP is used for reliable data transmission in network layer. Both are the carriers for covert channel [1].

TCP/IP header structure allows covert data within packet header fields which are unused, optional or required to hold random numbers. Grouping of covert channels which require to take random numbers are as follows:

A. Covert Channels by using Unused Header Bit Fields:

It consists of reserved fields which are used in future or unused such as 4 bit reserved field in TCP header, options and padding fields in TCP/IP, and unused bit of IP header's Type of Service (TOS) field which can be used to encode secret data.

B. Covert Channels by Modifying Some Header Fields of TCP/IP:

Some fields could be modified to encode secret data .TTL is a counter value which is decreased at each hop in IP. When TTL is zero then packet discarded. IP checksum field is modified to encode secret information and extension is added with the contents. TCP timestamp is used to improve TCP performance.

C. Covert Channels by using Some Header Fields which Require Random Number:

To coordinate which data has been transmitted and received providing reliable delivery , TCP sequence numbers are used. TCP sequence number depends on flag SYN. If SYN is set, then this is ISN. Otherwise it is accumulated sequence number.

Fig 1 and 2 show TCP and IP header structure having size 32 bit.

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
Source Port																Destination Port															
Sequence Number																															
Acknowledgment Number																															
Data Offset		Reserved		CWR	ECE	URG	ACK	PSH	RST	SYN	FIN	Window Size																			
Checksum																Urgent Pointer															
Options (Including timestamp)																								Padding							

Fig 1 TCP Header Structure

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31		
Version		Header length		Type of Service (TOS)				Total Length																									
Identification																X	DF	MF	Fragment Offset														
Time to Live (TTL)				Protocol				Header Checksum																									
Source Address																																	
Destination Address																																	
Options (if any)																												Padding					

Fig 2 IP header structure

III. LITERATURE SURVEY

A. Support Vector Machine

An offline detection scheme that used support vector machine (SVM) to detect covert channels in TCP ISN proposed previously. In this method, three feature dataset has been evaluated which includes ISN, control flag, header checksum fields of TCP header. This method uses 5000 normal and 5000 abnormal data to train SVM. Total numbers of packets are 10000. It has high correct detection rate. It is very complicated and time consuming [2].

B. Process Query System

In PQS, user queries are expressed as process descriptions. System can take input from arbitrary sensors and then forms hypotheses regarding observed environment, based on process queries given by user. Application programmer connects input event streams and then focuses on writing process models. Models can be constructed as state machine, formal language descriptions, hidden markov models or set of rules. Now track processes occurring in a dynamic environment and continuously present best possible explanation of observed events to the user [4].

C. Nushu

TCP ISN based covert channel is known as NUSHU for linux operating system it has been developed. As linux operating system is open to public, it's easy to implement nushu on it. To detect nushu covert channels, use of neural network has proposed. But it requires large number of training data than SVM. Computational complexity is very high and cannot be used online. For accuracy, it depends on quantity of ISNs [3].

IV. PROPOSED SYSTEM MODEL

Below diagram describes the overall predicted architecture of our system. So that anyone will easily find out the required modules for developing such type of systems. We introduce covert channel detection in TCP ISN in real instead of simulation, which is a new method for covert detection. It will work as follows, for all the coverts which are in TCP ISN field.

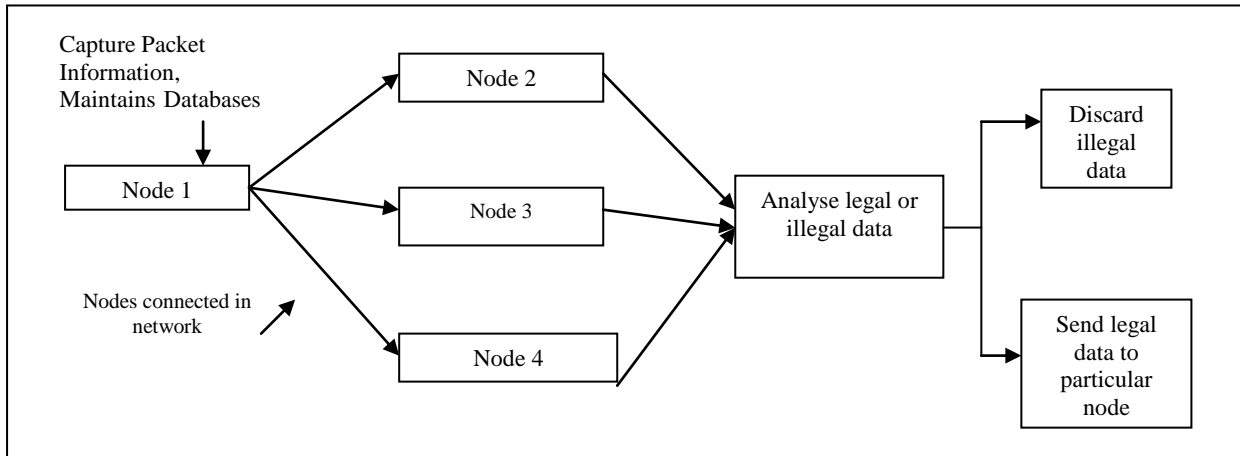


Fig 3 Block diagram of proposed system

Proposed system consists of 4 nodes. One node captures all packet header information of the systems which are connected in the network. Other nodes send information to each other. Information sends by typing IP address. If any system having that IP address is connected in the network then only information will send. Otherwise it shows error as no network connection.

After sending data, legal data will be easily stored at particular node. But in case of illegal data or covert data, it will give message as intruder detected and will block that covert with its IP address. So that from that blocked IP no one will send data next time. Proposed system maintains database of blocked IP addresses as well as covert data database with its parameters.

V. RESULTS

Table 1

Performance comparison between proposed naive bayes and the one using SVM

Case	No.of features	Total correctness
SVM	1	92%
Naive Bayes	1	94%

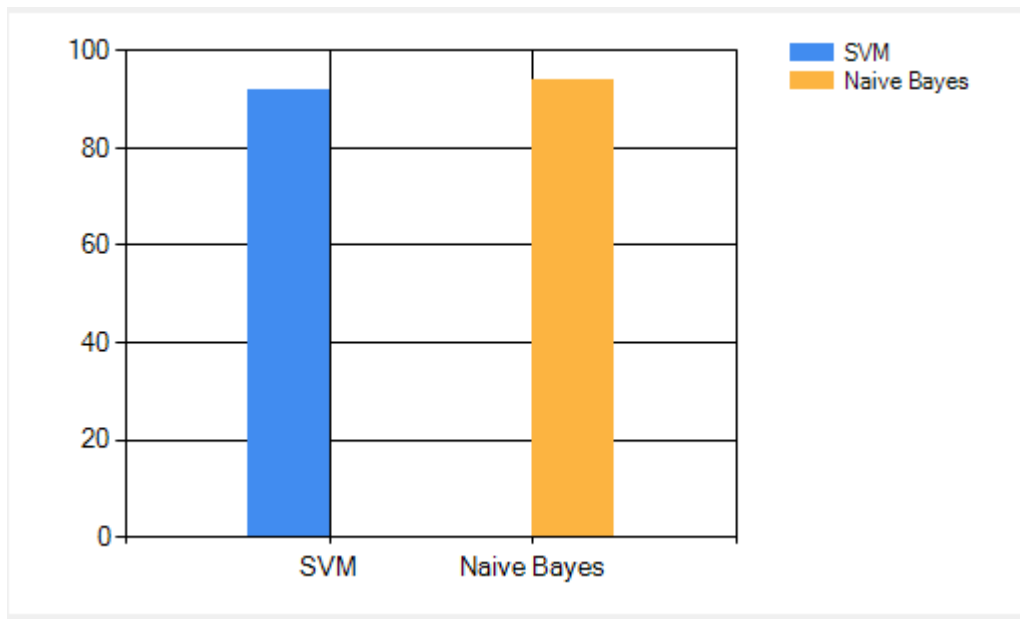


Fig 4. Graphical representation of Naive Bayes and SVM

Above table shows proposed naive bayes classifier outperforms Sohn’s SVM method in each field, and the fact considers that ISNs exhibit some structure to avoid receiving leftover packets from past connection. Thus computational complexity is reduced greatly and proposed classifier is much simpler than SVM. In Sohn’s SVM method, polynomial kernel is used, if linear kernel is used, total accuracy achieved will decrease to 87.9% . But this is not happened in case of naive bayes classifier. Computational complexity of SVM in training is evaluated as $O(n_samples^2 * m_features)$ which includes solving convex optimization problem. Computational complexity is evaluated as $O(3 * n_samples)$.

By using third order feature, it can clearly identify normal and stego_ISNs. It is known that network steganalysis is not like image steganalysis, which could be carried out offline. The covert channels need to be detected online because header information will be taken off as packet reached its final destination.

CONCLUSION

The huge amount of data transmitted over Internet by using TCP/IP protocols makes it ideal as a carrier in steganography. Occurrences based on covert channels become a potential threat to the Internet. Covert channels based on the reserved fields, vacant combinations of flag field of TCP/IP header, or change of some header fields can be easily detected or detached.

Detecting covert channels in TCP ISN field is known as one of the most difficult covert channels to be identified. The main intension of this proposed application is to detect covert channels in TCP ISN field in real instead of simulation method. The objective is to progress towards receiver fulfillment by returning the information that don’t have any stegoISNs.

To increase classification accuracy, we use naive bayes classifier. Therefore we need to focus on various aspects of detection methods and classifiers. We have analyzed possible covert channels and presented a practical method in detecting the covert channels in TCP ISN field, which can detect the covert channels using TCP ISN in an online fashion owing to its largely reduced computational complexity. Furthermore, the simulation results have shown that our proposed PRM outperforms the state-of-the-art method in detecting covert channels in TCP ISN in terms of accuracy and speed.

REFERENCES

- [1] H. Zhao, Yun-king Shi, “Detecting covert channels in computer network based on chaos theory,” IEEE, Vol 8, No. 2, Feb 2013.
- [2] T. Sohn, J. S. ,and J. Moon, “ A study on covert channel detection of TCP/IP header using support vector machine,” in proc. 5th Int . Conf. Information and communication security (ICICS 2003), pp. 313-324, Oct. 2003.
- [3] E. Tumoian and M. Anikeev, “Detecting NUSHU covert channels using neural networks”.
- [4] V. Berk, A. Giani, G. Cybenko, “Covert channel detection using process query system,” Flocon 2005.
- [5] David Muchene, K. Luli, C. Shue, “reporting insider threats via covert channels,” IEEE, 2013.
- [6] R. J. Anderson and F. A. P. Petitcolas, “On the limits of steganography,” IEEE J. Sel. Areas Commun., vol. 16, no. 4, pp. 474–481, May1998.

- [7] V. Berk, A. Giani, G. Cybenko, "Detection of covert channel encoding in network packet delays," Technical Report TR536, Revision 1, Aug 2005, revised Nov 2005.
- [8] Internet Protocol (IP), Information Sciences Institute, University of Southern California, RFC 791, Sep. 1981.
- [9] Transmission Control Protocol (TCP), Information Sciences Institute, University of Southern California, RFC 793, Sep. 1981.
- [10] Anjan K, Srinath N. K, Jibi Abraham, " Performance analysis of transport layer based hybrid covert channel detection engine," IJNSA, Vol.5, No.6, Nov. 2013.