IJCSE International Journal of Computer Sciences and Engineering Open Access

Review Paper Volume-2, Issue-12 E-ISSN: 2347-2693

Review on covert channel detection methods of TCP/IP header

Apurva N. Mahajan[1*] and I. R. shaikh[2]

[1*] *S. N. D. COE & RC Yeola, University Of Pune*
[2] *Assistant Professor, S. N. D. COE & RC Yeola*
www.ijcaonline.org

Received: 16 Dec 2014 Revised: 25 Dec 2014 Accepted: 26 Dec 2014 Published: 31 Dec 2014

Abstract— A covert channel is any methodology of communication that's acquainted illicitly transfer data, so breaking the security policy of a system. A network covert channel is a covert statement by hiding covert messages in to explicit network packets. Any shared resource will be probably used as a covert channel. In recent years with the growth of various hiding methods, network covert channel has become a new kind of intimidation for network security. A covert channel is an unplanned design within authentic communication whose axiom is to leak information as a part of undeveloped protocols. In fact, most detection systems can detect hidden data in the payload, but struggle to survive with data hidden in the IP and TCP packet headers. The huge number of protocols in internet seems ideal as a high-bandwidth vehicle for covert communication. Due to unwanted and malevolent nature of covert channel applications and as it poses a serious security threat to network, it is recommended to detect covert channels efficiently. This paper presents a criticism of TCP/IP covert channel design and their detection scheme and presents a proposed method based on Naive-Bayesian classifier to detect covert channels in TCP ISN and IP ID fields of TCP/IP packet.

Keywords— TCP/IP covert channel, TCP, IP, network security

I. Introduction

Covert channels are the variant of information hiding research area. Lampson proposed covert channel concept in 1973. Covert channel has been defined for the first time as a communication channel not designed for any kind of information transfer. Covert channels are system based. Covert channels are classified into covert storage channels and covert timing channels [1].

Covert storage channels can be described as writing of hidden information into storage location by transmitting party and subsequent retrieval of hidden information by receiving party. Stegnography can be a form of covert storage channel. Covert timing channel involves modifying the time characteristics to hide information [2].

In this modern time, internet becomes one of the most common ways of passing messages. Governments and military departments can be used covert channels to keep their communication secret. Network administrators can also use covert channels to keep communication related to network management secure [3]. TCP and IP are most widely used protocols in internet. There are several schemes developed to hide information in TCP/IP header. Information hidden in TCP ISN and IP ID fields are most difficult to detect. TCP/IP can be used to communicate across any set of interconnected network. TCP provides service of exchanging data directly between two network hosts. IP handles addressing and routing messages across one or more networks. TCP is used for reliable data transmission in transport layer whereas IP is network layer protocol. Both are carriers for covert channel [4],[5].

II. Related Work

Information and dissimulation is not new. Applications remain numerous and most recent techniques make such channels more difficult to detect. If covert channels are used to protect privacy or increase security of critical communication then it can be a good thing. And when applied to security policy bypassing, information leak or compromised system control the knowledge of such techniques become mandatory to enhance detection engine.
Stegnography is the first technique to write hidden messages to cover medium that suspects existence of message only from sender and intended recipient. Stegnography used in multimedia applications to distribute hidden data via audio, video or image files. Stegnography can be applied in digital watermarking for protecting copyrights [6]. Network stegnography is the general term that can classify information hiding techniques that may use to exchange stegnograms in telecommunication network [7].

An offline detection scheme proposed which used Support Vector Machine (SVM). SVM used to detect covert channels in TCP ISN. ISN, control flag, header checksum these three feature datasets evaluated. 5000 normal and 5000 abnormal data used to train SVM. SVM learning method has high correction data rate. This method is complicated and time consuming. Computational complexity of SVM is also more

[8]. Covert channel detection by using process query system (PQS) is developed which is a new type of information retrieval technology. In this PQS, queries are expressed as process descriptions. Queries allow PQS to solve large, complex information retrieval problems. System take input

from arbitrary sensors and then form hypotheses regarding the observed environments, based on given process queries by user [9].

Nushu covert channel method is based on neural network. ISN generation model is used for detection. Nushu is a new tool for covert channel implementation in linux. It is a proof-of-concept tool for linux. It uses existing ones data and doesn't generate any additional traffic i.e. it provides Passive Covert Channel(PCC). To detect covert channel using nushu , a model of ISN generation is constructed using experimental ISN data of original stack first. Prediction of successive ISN value is based on preceding ones. Neural network is most promising approach. Neural network creates model using experimental data without data generation algorithm information. Neural network can generalize and correct mistakes, obtain these properties during training. Training is iterative process with random initial parameters. Duration of process depends on training set. Now system must be trained to recognize presence of hidden channel. ISNs generated by standard stack are collected to form training set.

Neural network is trained to reproduce mapping Elman neural network is used. This network predicts successive ISN using all data received before. Calculate similarity threshold during training. Tested packets do not match normal stack model when threshold is exceeded. SYN-ACK-packets intercepted in controlled network, as training is completed. As soon as current ISN is intercepted, compare it with predicted value. If difference is higher than chosen threshold, then ISNs are not generated by original stack [10].

III. METHODOLOGY

A. *Phase Space Reconstruction*

Chaos can be defined as a random and non uniform phenomenon in the deterministic nonlinear system. Phase space reconstruction is the first step in nonlinear time series analysis of data from chaotic systems. It is a useful nonlinear/chaotic signal processing technique to characterize dynamic system, whether low-dimensional or high dimensional.

Phase space reconstruction consists of viewing a time series,

$X_k = X(K\tau)$, K=1, 2,….,N in a Euclidean space R^m, where m is the embedding dimension and τ is the sampling time. The points in form an attractor that preserves the topological properties of the original unknown attractor, is the expectation.

An attractor is a set towards which a dynamical system evolves over time. Geometrically, an attractor can be a point, a curve, a manifold, or even a complicated set with a fractal

structure known as a strange attractor. A dynamic system can be described by a phase space diagram, which must be a coordinate system. Coordinates are the variables that are necessary to describe the state of the system at any moment completely. Delayed coordinates method is used which based on the concept to reconstruct missing dimension using its previous and delayed function values as additional coordinates. A given time series, X_i, i =1,2,3,….,N can be reconstructed in a multidimensional phase space.

$Y_j = (X_j, X_j\text{-}\tau, \ldots, X_j\text{-}2\tau, \ldots., X_j\text{-}(m\text{-}1)\tau)$ – (1)

Where, j=1,2,…,N-(m-1), and m is the dimension of vector Y_j. & τ is delay time.

Further expanding, $Y=[Y_1, Y_2, \ldots, Y_j, \ldots, Y_M]$ –(2)

Where, Y_j is the vector of m dimension & M is the number of vectors in multidimensional phase space.

M can be given by, M=N-(m-1)

To reveal the hidden structure of ISNs, the phase space can be reconstructed by using "delayed coordinates".

Y_i=(ISN(i),ISN(i-1),…,ISN(i-(m-1)) –(3)

Where, i=1,2,…,N-m+1, N is number of ISNs. & m is the dimension. Vector Y_i is new phase space.

The first-order difference of input that is used in phase space reconstruction.By using first-order difference as coordinates, the phase space is constructed as follows:

x(n) = ISN(n)-ISN(n-1)
y(n) = ISN(n-1)- ISN(n-2)
z(n) = ISN(n-2)- ISN(n-3) - (4)
x(n), y(n), z(n) are called points coordinates.

B. *Proposed PRM Model*

There are various methods to estimate m including empirical methods. Different values of m such as 2, 3, 4, 5, 6 were tested in creating reconstruction phase space. & m=4, 5, 6 gave 100% detection accuracy rate. Larger the m, higher computational complexity is. Hence m=4 is selected. 4-dimensional vector coordinates are calculated as:

x(n)=ISN(n)-ISN(n-1)
y(n)=ISN(n-1)-ISN(n-2)
z(n)=ISN(n-2)-ISN(n-3)
w(n)=ISN(n-3)-ISN(n-4) –(5)

where, n= N,N-1,N-2,…,5 & N is the number of ISNs.

4-dimensional phase vector r_i is constructed as:

R_i =[x(i),y(i),z(i),w(i)], i=1,2,…,M,M=N-5 –(6)

R is used to represent phase space on dataset formed by (5) & (6).Number of phase vectors in phase space R is N-5, if number of ISNs is N.

R= $[r_1, r_2, .., r_m]$ –(7)

To extract features from dataset R, define distance between any two vectors r_i and r_j in the phase R as

$$d_{i,j} = \sqrt{(x(i)-x(j))^2 + (y(i)-y(j))^2} \ * \ \sqrt{+(z(i)-z(j))^2 + (w(i)-w(j))^2}$$

-(8)

Calculate distance between any 2 vectors in R with the help of 2D matrix.

j= 1,2,…,M & M is the number of vectors in the R.This is M*M matrix. There are M row vectors. Diagonal entries in D are all 0's, because each row vector represents distance between specified vector in R.

$d(j) = [d_j,1, d_j,2, \ldots, d_j,M], j=1,2,\ldots,M$ –(9)

where $d_{j,j} = 0$.

The variance for row vector d(j) is calculated as :

$\sigma_j^2 = 1$ –(10)

where, μ_j is mean value of row vector d(j).

σ_j^2 stands for variance of distance between specified vector r_j & all vectors in dataset R.

Variance vector σ^2 is obtained as

$\sigma^2 = [\sigma_1^2, \sigma_2^2, \ldots, \sigma_M^2]$ –(11)

Statistical feature is defined as

var_σ = -(12)

μ_σ is mean value of row vector σ^2. var_σ is variance of variance vector σ^2. 350 ISNs are used to obtain var_σ.
Then select threshold as :

T_ISN = var_σ /10 –(13)

Threshold is used in the classifier to test if coming ISN is a legal one or not. Only legal ISNs are used in constructing model.

C. Proposed classification algorithm

Legally generated ISNs are used to construct statistical model. Dataset of 745 ISNs are used & they are collected by winDump. Half ISNs are used to construct statistical model in 4-dimensional phase space. & then obtain third order feature var_σ of the proposed statistical model.
Other half is used for testing. Stego_ISNs are generated by covert_TCP. Stego algorithm is also used to test SVM based detection.
Proposed classifier is used to identify incoming ISN(k) as legal or illegal.
Coordinates of 4-dimensional phase vector p=(u,v,s,t) for incoming input ISN(k) are constructed as follows:
u = ISN(k)-ISN(k-1)
v = ISN(k-1)-ISN(k-2)
s = ISN(k-2)-ISN(k-3)
t = ISN(k-3)-ISN(k-4)

Then calculate distance between vector p=[u,v,s,t] & to obtain $d=[d_{k,1}, d_{k,2}, \ldots, d_{k,M}]$ i.e. distance vector, by vectors r_i in the training dataset R.

Second order feature: The variance of distance vector is calculated as σ_k^2, then third order feature var_σ is calculated & compared with T_ISN.

For legal ISN number, third order feature is larger than threshold, for stegnographic ISNs. It is smaller than threshold.
T_ISN is obtained based on statistical feature model which is constructed using legal ISNs. For any incoming ISN(k), var_σ is calculated by proposed algorithm.

IV. PROPOSED WORK

Naive bayes classifiers use to detect covert channels in TCP ISN field. These classifiers classify information into legal and illegal data. A novel detection scheme using PRM model is proposed in this study. Naive bayes classifiers use because of its high classification accuracy.
Below diagram can describe the overall predicted architecture of our system. So that one can easily find out the required modules for developing such type of systems. We introduce covert channel detection in TCP ISN by using PRM 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.

Source node sends information to the destination .Before sending information to destination, it requests for shortest path to reach the destination. For that we use AODV routing algorithm. Information sends through the node through which distance is minimum. These nodes are force covert nodes. After that information sends to ISP.

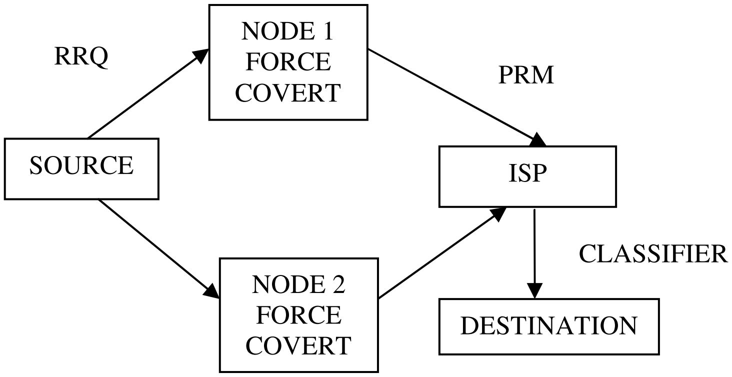

Fig.1 block diagram of proposed system

CONCLUSION

The huge amount of data transmitted over Internet by using TCP/IP protocols makes it ideal as a carrier in steganography. Attacks based on covert channels become a potential threat to the Internet. Covert channels based on the reserved fields, unused combinations of flag field of TCP/IP header, or modification of some header fields can be easily detected or removed. Detecting covert channels in TCP ISN field is known as one of the most difficult covert channels to be detected.

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 naïve 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] 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**, May**1998**.

[2] S. Attallah, "Trusted Computer System Evaluation Criteria", Tech. Rep. DOD 5200. 28-STD, **1985** [Online]. Available: http:// csrc.nist.gov/ publistications/history/dod85.pdf.

[3] V. Forte, C.Maruti, M. R. Vetturi, and M. Zambelli, "SecSyslog: An approach to secure logging based on covert channels," in Proc. First Int. Wksp. Systematic Approaches to Digital Forensic Engineering, pp. **248–263**, Nov. **2005**.

[4] Transmission Control Protocol (TCP), Information Sciences Institute, University of Southern California, RFC 793, Sep. **1981**.

[5] Internet Protocol (IP), Information Sciences Institute, University of Southern California, RFC 791, Sep. **1981**.

[6] M. Owens, "A Discussion of Covert Channels and Steganography", SANS (SysAdmin, Audit, Network, Security) Institute, 2002.

[7] K.Szczypiorski, "Steganography in TCP/IP Networks. State of the Art and a Proposal of a New System HICCUPS Institute of Telecommunications Seminar [Online], Retrieved Jun. 2010

[8] 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.