IJESRT
INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

WATERMARK-BASED CORRELATION THROUGH STEPPING STONES IN ENCRYPTED ATTACK TRAFFIC

Mr. Sachin Ghadge, Prof. Praveen. D Sathya
MECSE, Shreeyash College of engineering and Technology, Dr. Babasaheb Ambedkar University
Aurangabad, Maharastra India

ABSTRACT

In global network intruders having attack their data directly from their own computer. Often, this attack is coming from directly “stepping stones” in order to conceal their identity of data and original data. To identify the source address and data of the attack behind the stepping stone(s), it is required to gather the incoming and outgoing flows or different connections of a stepping stone. To prevent attempts at destination of correlation, the attacker might encrypt or otherwise manipulate the way of connection traffic. Timing based correlation shows in quite effective in the correlating encrypted connections. However, timing based approach are subject to time factor perturbations that may be deliberately introduce by the attacker at traffic destination. In this project, our watermark approach is “active and passive” in that It embeds a unique and single watermark into the encrypted flows in networks by slightly defined the timing of selected packets. The unique and single watermark that is embedded in the encrypted flow gives us a number of benefits over active and passive timing based correlation in resisting timing perturbations by the attacker. A two-fold monotonically increasing compound mapping is created and proved to yield more distinctive and theoretical visible watermarks in the watermarked images. Security protection is defined by parameter and mapping data at destination have also been proposed to deter attackers from illicit image pattern.

.

KEYWORDS: Watermark Bit, Correlation, Tracing, Mapping, stepping stones.

INTRODUCTION

Intruders are attack their threads directly from their own computer. Often, they further their attacks through intermediate “stepping stones” in order to research their Identity and origin. To identify the source of the attack behind the stepping stone(s), it is necessary to correlate the incoming and outgoing flows or connections of a stepping stone. To prevent attempts at correlated; the attacker may encrypt data or otherwise manipulate the established connection traffic. Timing based correlation approaches have been shown to be quite affect in correlating connections. However, timing based correlation approaches are subject to timing perturbations that may be strongly introduced by the attacker at stepping stones. In recent years, unauthenticated accesses to the computer environment are increasing as various activities takes place on the network. The common way for network attacker to conceal their identity by connecting across particular hosts before attacking the final destination target. Intruders do not log in directly to their final destination targets from their self computers, but they firstly make access login data through various hosts and then to the another hosts address and continue this series several number of times which makes a "chain of intermediate hosts line" before breaking into their final targets. Therefore, it becomes necessary for the identification (attack target) to trace back the chain to find the origin of attacker. For this it is important to collect incoming packets and outgoing packets from source. So, correlation methods are needed to link connections between stepping stones in the network. The earlier recordation work on connection correlation was based on track the address and data of user activities or connection content (packet payload) was used.

Necessity

In existing connection establishment are based on three different characteristics,1)hosting; 2) component content and their comparisons; 3)time require for each characteristics. The hosting approach collects and tracks user data and his login activities operation at every stepping stones. And the components with comparisons are that contents between each stepping stones will be compared collect data. Time related approach makes use the incoming and outgoing times of packets to correlate connections. The hosting operation approach (DIDS and CIS). DIDS (Distributed Intrusion

Detection System) is a system where all TCP connections with users and logins within the global network are remotely monitored and the system keeps tracing of all the activities and the current states of users.
TCP connections and logins within the global network are remotely monitored and the network system keeps tracing of all the movement and the current activity of users. A hosting remotely monitors locate on each domain host in the network, gathering audit data about the host, which can be transmitted to the central DIDS director. The intruder attacker attacks the final destination target; this can be done by that the attacker will be compromising the sequence and number of nodes that is a system before attacking the final destination target. The compromised nodes are also known as ZOOBIES.

Objective

The objective of watermark is to make the correlation of encryption data connections probabilistically robust against sequence and random timing perturbations by the adversary. Unlike existing timing schemes, our watermark is active condition in that it encrypt a unique watermark into the encrypted pipe, by slightly providing delay timing of selected packets. If the watermark is both single and robust, the watermarked flows can be remotely identified and thus correlated at each stepping stone in the location.
1) While the attacker in network can add or update the secret key in watermarking, we can easily analysis and identify the intruder by using remotely observation.
2) All packets with watermark flow in the original flow are kept. No packets are dropped or damage from or added to the flow by the stepping stone.
3) While this scheme is public knowledge, the watermarking embedded and decoded parameters are secrets known only to the watermark embedded and the watermark detector(s).From above application examples, we summarize a list of properties and principles for designing flow watermarks. The challenge to build an efficient scheme lies in the difficulty to receive all desired properties simultaneously.

LITERATURE SURVEY

1. In June 2009 Mohd Nizam Omar and Rahmat Budiarto proposed Intelligent Network-Based Stepping Stone Detection Approach Shows the potential of SOM technique as to count the number of connection chains and shows protocol that include in network stepping stone detection. The effect of active and passive perturbation attack is not involved. They used SSD algorithm.
2. In May 2011 Xinyuan Wang, Douglas S. Reeves proposed Robust Correlation of Encrypted Attack Traffic through Stepping Stones by Flow Watermarking timing-based correlation techniques. Watermark-based correlation is provably effective against correlated random timing perturbation and Detection algorithm.
3. In May 2012 B. Bazeer Ahamedand and S. Hariharan proposed Implementation of Network Level Security Process through Stepping Stones by Watermarking Methodology they presented a novel active timing-based correlation approach to deal with random timing perturbations. makes no assumptions is consider about the original distribution of the inter-packet timing of the original packet flow. It also involve low complexity detection algorithm.
4. In March 2013 S. Jayanthi and G. Shamili proposed Detecting and Preventing Traffic Attack through Stepping Stones and securing the information using watermarking Does not require the adversary's timing perturbation shows the require substantially packets. They develop a suite of algorithms to infer watermark parameters, recover/duplicate watermarks.
5. In Jan 2014 Saptshree Dengle,Snehshree Dengle proposed Active Watermarking Approach in Detecting Encrypted Traffic Attack by Making Correlation Scheme Robust Analyze the active watermarking scheme for tracing through stepping stones. They identified the provable upper bounds on the number of packets
6. In May 2015 Shao-Da Huang, Tzong-Sun-Wu,Han-Yu Lin proposed A Provable Watermark-Based Copyright Protection Scheme developed signal processing techniques but focus on how to employ unpredictable signature-seeded pseudo random bit sequence negative watermark detection rate computationally negligible. Exhibited digital media under adversarial watermark removal attacks.

INDENTATIONS AND EQUATIONS

Detection: The probability real edge points detection should be maximized while the probability of falsely detection should be minimized non-edge points. This corresponds to maximizing the signal-to-noise ratio.Let $\pm i$ be the added delay to packet Pi, and $t0\ i$ be the time stamp of packet is distorted Pi, then $t0\ i = ti + \pm i$. The delays original and distorted inter-packet (IPD) between *Pi*+1 and *Pi* are *Ii* =*ti*+1 ¡ *ti* and *I0 i* = *t0 i*+1 ¡ *t0 i* respectively. Therefore,±*k* = *t0k* ¡ *tk* = ±1 +*k*X¡1*i* = 1(*I0i* ¡ *Ii*) The perturbed and original inter-packet timing characteristics of packet flow *P*1*; : : : ; Pn* can be represented with < *t*1*; I*1*; : : : ; In*¡1 > and < *t*01*; I*01 *; : : : ; I0n*¡1 > respectively. In, particular, < *I*01*; I0n*¡1 > the original inter-packet timing characteristics represents the distortion pattern. According to obtained results from section VI-A, in order to completely remove any hidden data from the original inter-packet timing characteristics, the purely needs to disturb < *t*1*; I*1*; In*¡1 > into an unique one. That means < *I*01 *; : : : ; I0n*¡1 > needs to be unique from < *I*1*; : : : ; In*¡1 >. Therefore, the distortion pattern < *I*01*; : : : ; I0n*¡1 > original inter-packet timing characteristics can be consider to be pre-determined .Perturbation theory comprises mathematical methods for finding an approximate solution to a problem, by starting from the determination solution of a related problem. A middle step has critical feature of the technique is a that breaks the problem into "solvable" and "perturbation" parts. Perturbation theory is obtained if the problem at hand cannot be solved exactly, but can be formulated by just adding a "small" term to the mathematical description of the exactly solvable problem.

Perturbation theory leads to an expression for the perfect solution in terms of a formal power series in some "small" parameter – known as a perturbation series – that quantifies the deviation from the exactly solvable problem. In this power series the leading term is the solution of the exactly solvable problem, while deviation describe further terms in the solution, because to the deviation from the initial problem. Formally, A having the approximation to the full solution , a series in the small quantity parameter (here called ε), like the following:

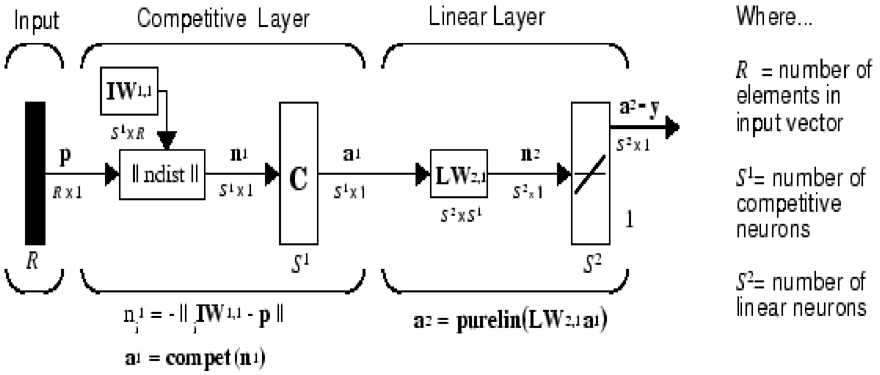
$$A = A_0 + \varepsilon^1 A_1 + \varepsilon^2 A_2 + \cdots$$

In this example, A_0 would be the known solution to the exactly initial problem solution and A_1, A_2,.. Represent the higher-order factors terms which may be found iteratively by some systematic procedure during execution. For small ε these higher-order terms in the series become successively smaller. An approximate "perturbation solution" is obtained by detachment the series, usually by keeping only the first successive two terms, the initial solution and the "first-order" perturbation correction:

$$A \approx A_0 + \varepsilon A_1$$

Architecture

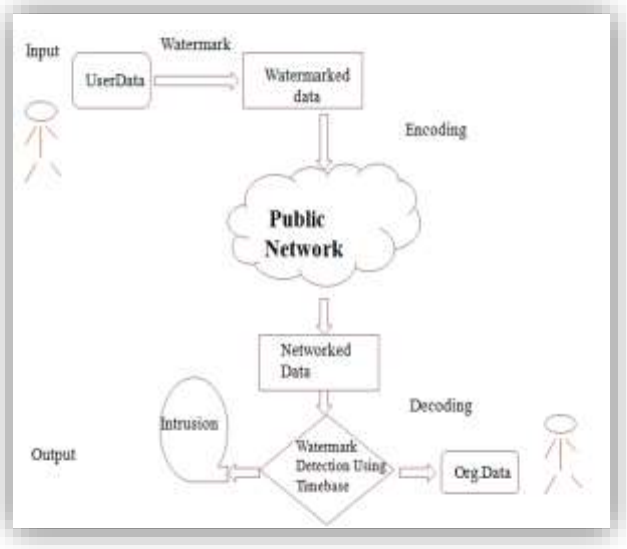
The LVQ network architecture is shown below.



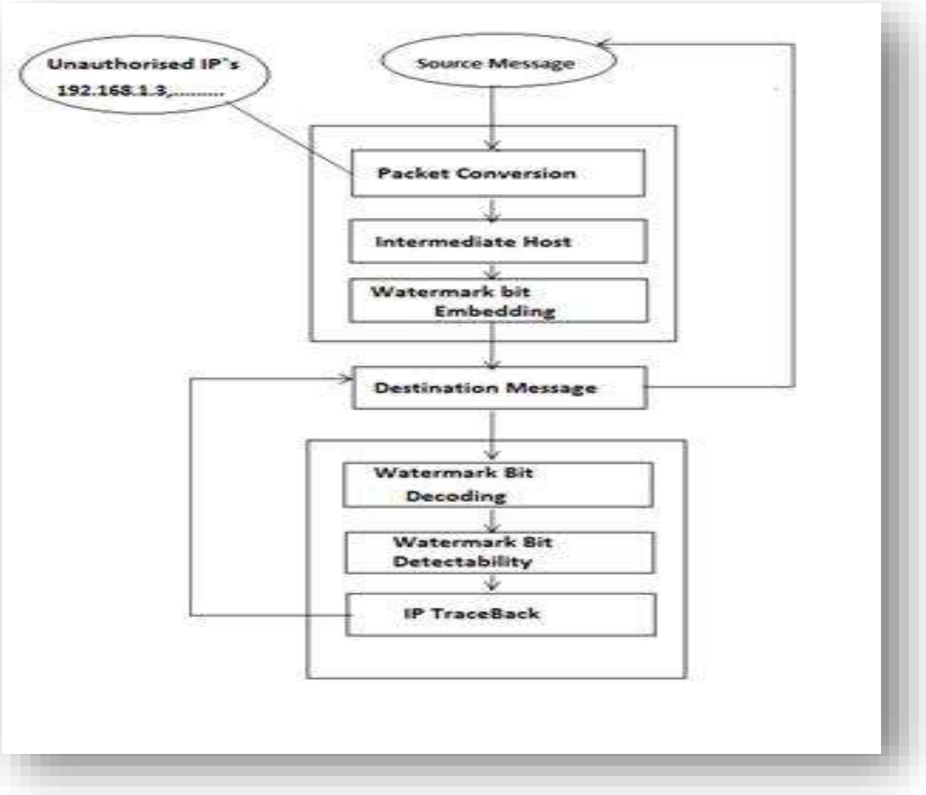
An LVQ network shows a first competitive layer and a second linear layer. The competitive layer describe classify input vectors in much the same way as the competitive layers of Cluster with Self-Organizing Map Neural Network described in this topic. The linear layer transforms the competitive layer's categories into target classifications defined by the user. The categories learned by the competitive layer are referred to as subclasses and the categories of the linear layer as target categories. Both the competitive and linear layers have one neuron per (sub or target) class. Thus, the competitive layer can describe up to S^1 subclasses. These, in turn, are combined obtained by the linear layer to form S^2 target classes. (S^1 is always larger than S^2.)For example, suppose neurons 1, 2, and 3 in the competitive layer all describe subclasses of the input space that belongs to the linear layer target class 2. Then competitive neurons 1, 2, and 3 get $LW^{2,1}$ weights of 1.0 to neuron n^2 in the linear layer, and weights of 0 to all other linear neurons. Thus, the linear neuron obtained a 1 if any of the three competitive neurons (1, 2, or 3) wins the competition and getting outputs a 1. This is how the categories of the competitive layer are combined into destination classes in the linear layer.

TABLES, FIGURES AND EQUATIONS

3.1 Architecture Diagram of Watermarking Methodology



Message Container and path sequence.



MODULE USED

1) Bit formation and Embedding- Decoding:

Generally, watermarking includes the selection of a watermark flow carrier, and the design of two complementary successive processes method: embedding and decoding. We collect the watermark signature in the registration. The watermark embedding process inserts the information in source by a slightly modified data of some property of the carrier. The watermark decoding process identified and extracts the watermark (equivalently, determines the original data of a given watermark). To correlate encrypted connections, we consider using the inter-packet timing as we have the watermark carrier property of interest. The encrypted watermark bit is guaranteed to be not corrupted by the timing perturbation. If the perturbation goes out of range, the embedded watermark bit may be altered by the attacker.

2) Limiting factor:

In practice, the number of packets with data available is the fundamental Limiting factor to the achievable effectiveness of our watermark based correlation. This set of operation to compare and evaluate the correlation effectiveness of our proposed active and passive watermark based correlation and previous passive timing-based correlation under various timing perturbations. By embedding a unique watermark into the inter-packet timing, we have sufficient redundancy; we can make the correlation of encrypted flows substantially more robust against random timing perturbations. We can correlate the signatures of watermark and identify it's the positive or negative factorization, if positive occurs it detect it is the authenticated user otherwise, if negative occurs it detect it is an Intruder.

3) Watermark Detection Model:

The watermark tracing approach exploits the observation that interactive connections are both direction. The idea is to watermark the backward traffic (from victim back to the attacker) of the both direction attack connections by slightly providing delay to the timing of selected packets. If the embedded watermark is both robust and unique, the watermarked back traffic can be effectively correlated and traced across stepping stones, from the victim all the way back to the attacker, assuming the attacker has not having full control on the attack target, the attack Target will initiate the tracing after it has identified the attack. Specifically, the attack target will watermark the backward traffic of the attack connection chain, and inform across the network about the watermark. The stepping stone in global the network will scan all traffic for the presents of the indicated watermark, and report to the target to specific if any occurrences of the watermark are detected.

4) Data Mapping:

One simple technique to achieve this is to use a secret key to get a pseudo-random sequence of numerical random values and add them to both of and for the pixels in the watermarking region. This technique is called as parameter randomization. This parameter exchange does not impact on lossless recoverability, because we can now recover and form the original pixel values by the compound mappings. We will refer to this technique in the sequel as mapping randomization. We also combine this technique with the parameter randomization technique to enhance the security level. Finally, the Authenticated user takes the file in zip format with proper password.

RESULT ANALYSIS

Compares with the Existing System

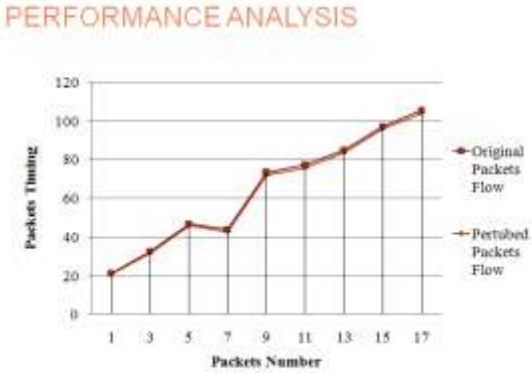


Fig 1. Difference between original and perturbed packet flow

PERFORMANCE ANALYSIS(Cont.)

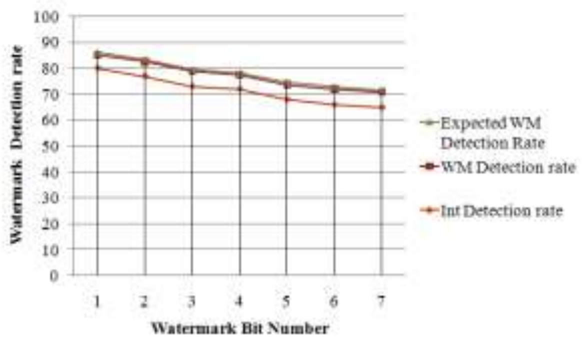
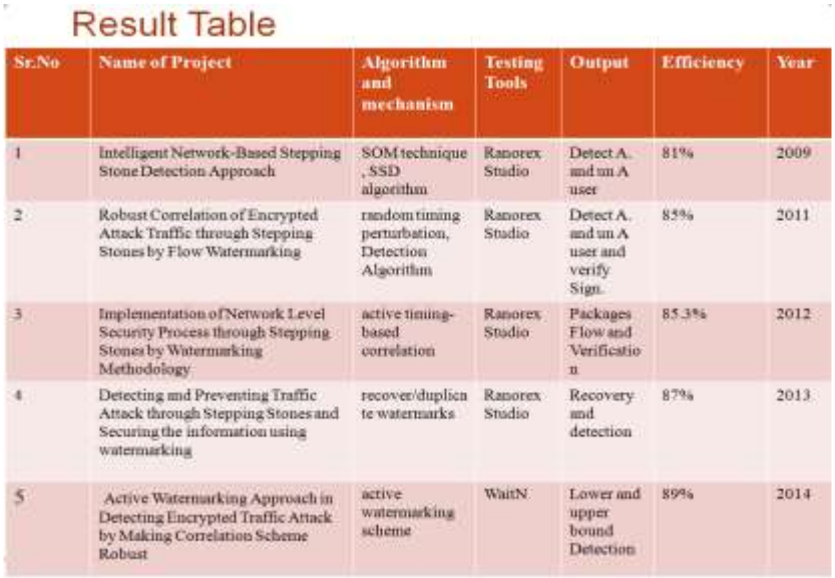


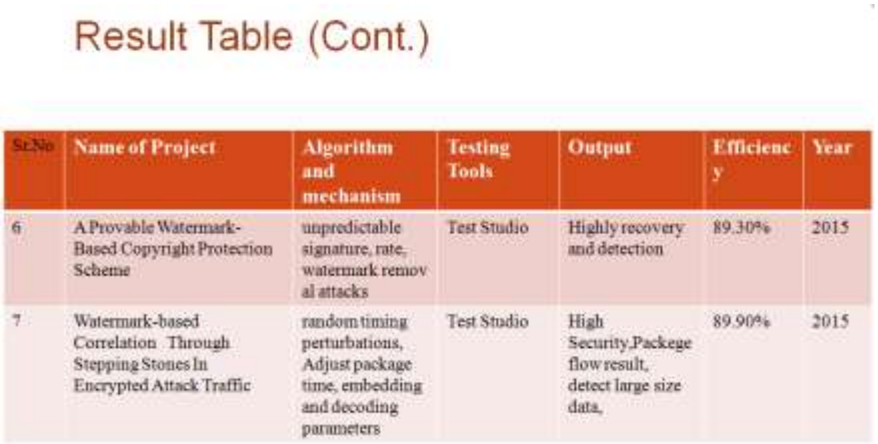
Fig2. Watermark detection rate with number of watermark bits



Result Table

Sr.No	Name of Project	Algorithm and mechanism	Testing Tools	Output	Efficiency	Year
1	Intelligent Network-Based Stepping Stone Detection Approach	SOM technique , SSD algorithm	Ranorex Studio	Detect A. and un A user	81%	2009
2	Robust Correlation of Encrypted Attack Traffic through Stepping Stones by Flow Watermarking	random timing perturbation, Detection Algorithm	Ranorex Studio	Detect A. and un A user and verify Sign.	85%	2011
3	Implementation of Network Level Security Process through Stepping Stones by Watermarking Methodology	active timing-based correlation	Ranorex Studio	Packages Flow and Verification	85.3%	2012
4	Detecting and Preventing Traffic Attack through Stepping Stones and Securing the information using watermarking	recover/duplicate watermarks	Ranorex Studio	Recovery and detection	87%	2013
5	Active Watermarking Approach in Detecting Encrypted Traffic Attack by Making Correlation Scheme Robust	active watermarking scheme	WaitN	Lower and upper bound Detection	89%	2014

Fig 3 Result table 1



Result Table (Cont.)

Sr.No	Name of Project	Algorithm and mechanism	Testing Tools	Output	Efficiency	Year
6	A Provable Watermark-Based Copyright Protection Scheme	unpredictable signature, rate, watermark removal attacks	Test Studio	Highly recovery and detection	89.30%	2015
7	Watermark-based Correlation Through Stepping Stones In Encrypted Attack Traffic	random timing perturbations, Adjust package time, embedding and decoding parameters	Test Studio	High Security,Packege flow result, detect large size data,	89.90%	2015

Fig 4 .Result Table 2

CONCLUSIONS

Tracing attackers in network traffic through stepping stones is a problem, when the attack traffic with encrypted data and its timing is manipulated (perturbed) to interfere with traffic determination. The uncertainty timing perturbation by the adversary can largely reduce the effectiveness of passive attacks, timing-based correlation techniques. We also presented an active timing-based correlation operation to deal with uncertainty timing perturbations. By embedding a unique watermark data into the inter-packet of message timing, with sufficient redundancy, we can make the correlation of encrypted flows substantially more robust and powerful against random timing perturbations. Watermark-based correlation is provably affected against correlated random timing considerations as long as the covariance of the timing perturbations on different packets is fixed.

REFERENCES

1. S. Snapp, et all. DIDS (Distributed Intrusion Detection System) – Motivation, Architecture and Early Prototype. In Proceedings of 14th National Computer Security Conference, 1991.
2. W. Stallings, Cryptography and Network Security—Principles and Practices, 3rd ed., Prentice Hall ~2002!.
3. H. C. Wong, M. Bern, and D. Goldberg, ‘‘An image signature for any kind of image,’’ in IEEE Int. Conf. on Image Processing, pp. I-409– I-412 ~2002!.
4. I. Cox, J. Bloom, and M. Miller, Digital Watermarking: Principles & Practice, Chaps. 1–2. Morgan Kauffman Publishers ~2002!.
5. J. Dittmann, A. Steinmetz, and R. Steinmetz, ‘‘Content-based digital signature for motion pictures authentication and content-fragile watermarking,’’ in IEEE Int. Conf. on Multimedia Computing and Systems, Vol. 2, pp. 209–213 ~2002!.
6. J. Fridrich and M. Goljan, ‘‘Images with self-correcting capabilities,’’ in Proc. Int. Conf. on Image Processing, Vol. 3, pp. 792–796 ~2003!.
7. C. Rey and J.-L. Dugelay, ‘‘Blind detection of malicious alterations on still images using robust watermarks,’’ presented at IEE Seminar on Secure Images and Image Authentication, Apr. 2000, pp. 7/1–7/6, IEE ~2003!.
8. A. C. Snoeren, C. Partridge, L. A. Sanchez and C. E. Jone et al. Hash-Based IP Traceback. Proceedings of the ACM SIGCOMM ’2001, November 2003.
9. D. Song and A. Perrig. Advanced and Authenticated Marking Scheme for IP Traceback. In Proceedings of IEEE INFOCOM’01, April 2004.
10. D. Song, D. Wagner and X. Tian. Timing Analysis of Keystrokes and Timing Attacks on SSH. In Proceedings of 10th USENIX Security Symposium, 2004.
11. S. Staniford-Chen, L. T. Heberlein. Holding Intruders Accountable on the Internet. In Proceedings of IEEE Symposium on Security and Privacy, 2005.
12. W. R. Stevens. <<TCP/IP Illustrated, Volume 1: The Protocol>>. Addison-Wesley Publishing Company 2005.
13. X. Y. Wang, D. S. Reeves, S. F. Wu and J. Yuill. Sleepy Watermark Tracing: An Active Network-Based Intrusion Response Framework. In Proceedings of 16th International Conference on Information Security (IFIP/Sec’01), June, 2007.
14. W. N. Lie, G. S. Lin, C. L. Wu and T. C. Wang, “Robust Image Watermarking on the DCT Domain,” in Proc. of IEEE Int. Sym. on Circuits and Systems, vol. 1, pp. 228-231, May 2007.
15. D.G. Luenberger, Optimization by Vector Space Method, John Wiley & Sons, Inc., 2007.
16. M. Wu, “Joint Security and Robustness Enhancement for Quantization Based Embedding,” IEEE Trans. on Circuits and Syst. for Video Technol., vol. 13, no. 8, pp. 831-841, Aug. 2008.
17. J. Li, M. Sung, J. Xu and L. Li. Large Scale IP Traceback in High-Speed Internet: Practical Techniques and Theoretical Foundation. In Proceedings of the 2009 IEEE Symposium on Security and Privacy, IEEE, 2009.
18. Active Watermarking Approach in Detecting Encrypted Traffic Attack by Making Correlation Scheme Robust(2011)
19. A. Blum, D. Song, and S. Venkataraman. Detection of InteractiveStepping Stones: Algorithms and Confidence Bounds. In Proceedings of the 7th International Symposium on Recent Advances in Intrusion Detection (RAID 2004). Springer, October 2012.
20. R. C. Chakinala, A. Kumarasubramanian, R. Manokaran, G. Noubir, C. Pandu Rangan and R. Sundaram. Steganographic Communication in Ordered Channels. In Proceedings of the 8th Information HidingInternational Conference (IH 2012), 2012 Mohd Nizam Omar and Rahmat Budiarto
21. Nedeljko Cvejic, Tapio Seppanen, “Digital Audio Watermarking Techniques and Technologies Applications and Benchmarks”, pages x-xi, IGI Global, Illustrated edition, August 7, 2013
22. Verma Harsh, Singh Abhishek, Kumar Raman, “Robustness of the Digital Image Watermarking Techniques against Brightness and Rotation Attack”, International Journal of Computer Science and Information Security, Vol. 5, No. 1, 2014]
23. Intelligent Network-Based Stepping Stone Detection Approach(2014) Xinyuan Wang, Douglas S. Reeves

24. Thanki Rohit, Kher Rahul, Vyas Divyang, “Comparative Analysis of Digital Watermarking Techniques”, LAP LAMBERT Academic Publishing, Germany, June 2015
25. J. Li, M. Sung ‘Probability and Statistics’.
26. D. Mitzel, and E. Estrin ‘Transactions on Signal Processing’.
27. ”Robust Correlation of Encrypted Attack Traffic throug Stepping Stones by Flow Watermarking” Xinyuan Wang, Member, IEEE, Douglas S. Reeves, Member, IEEE JUNE 2015.