M. Deepika, G. Om Sai Prashant / International Journal of Engineering Research and Applications (IJERA) ISSN: 2248-9622 www.ijera.com
Vol. 2, Issue 6, November- December 2012, pp.925-930

Active Watermark-Based Correlation Scheme for Identifying Source of Attack in Presence of Timing Perturbations

M. Deepika, G. Om Sai Prashant
Associate Professor Aurora Technological and Research Institute Andhra Pradesh, India.
Aurora Technological and Research Institute Hyderabad Andhra Pradesh, India

ABSTRACT

Intruders have changed their mode of operandi in breaking security of IT systems. Of late they are using different strategies making attacks successfully. One of their strategies is to attack systems though some intermediary nodes in the network instead of making attacks from their own machine. This helps them in hiding their identity. Such attacks can be identified by verifying and correlating incoming and outgoing network flows that come through intermediary nodes used in routing the attacks. The problem with this approach lies in the fact that attackers may intentionally alter such flow by disguising it and fooling the detection systems. The existing timing based correlation approaches to solve this problem are inadequate when attackers intentionally introduce timing perturbations. This paper introduces a new correlation approach based on watermarking which is proved to be robust to address such problems. This is achieved by timing of some packets selectively and embedding watermark in to the encrypted flows. This approach is active and can resist timing perturbations done by attackers. The empirical results reveal that our approach is almost close to providing 100% true positives.

Index Terms – Network intrusion detection, timing correlation, network flows, intermediary nodes, and timing perturbations.

INTRODUCTION

The information systems in the real world have been victims of network based attacks. This is the cause of concern though there are many security mechanisms in place to prevent such attacks. As security mechanisms grow in robustness in addressing many possible attacks, the attackers are also changing their strategies in making attacks successfully. This has become every growing problem that needs continuous attention and need to have on going research efforts. When attack is made, it is very essential to have the ability to trace and identify the source of attack. When attackers conceal their identity by not making attacks from their machine directly, it is challenging to find the source of attack. Obviously attackers over network are using some intermediary nodes to execute their attacks. The intermediary nodes, as they believe, hide the identity of original attacker. This makes it difficult to identify the source of attack as the attack is made through intermediary nodes by even spoofing IP source address of attack traffic. IP traceback is the method to identify source of attack in such cases as described in [1] and [2]. As discussed earlier, the attackers take countermeasures to IP traceback by making network – based intrusions through intermediary nodes. This is achieved by attacker by using some remote login programs like SSH or Telnet and perform attacks from remote machines in the network. The IP traceback method being employed in the industry is not adequate to reach the actual source of attack as the intermediary nodes stand in between.

From the literature it is understood that the prior works in this area were based on the login activities of the tracking user at various hosts [3], [4]. This has limitations as it fails to reach the actual source of attack when there are manipulations in the middle. Later researchers focused on the process of comparing payloads or packets of all the connections that are to be correlated [5], [6]. These are effective but suffer from limitations such as inability to accurately finding the source of attack. To overcome these limitations, some researchers [7], [8], [9] have focused on the features or characteristics of connections for the purpose of comparing and correlating encrypted connections. Timing based correlations suffer from the drawback that the adversaries may be able to perturb the timing based correlations intentionally. Addressing this problem is challenging as the encrypted traffics are subjected to time based perturbations.

To overcome the drawback of the timing based correlations, this paper introduces an efficient correlation scheme that is provide to be robust to such attacks. The proposed scheme is watermark – based which is active in nature. This means that it dynamically embeds watermark into encrypted flows. This is performed by slightly adjusting the timing of selected packets. Out approach also needs significantly less number of packets to achieve this. This is in contrast to the existing passive timing correlation schemes. The experimental results reveal the fact that our approach is close to 100% true positives.

RELATED WORK

When attacks are made through intermediary nodes by intruders it is very challenging to establish the source of attack accurately. This section provides insights into the literature in which many existing works on similar lines are reviewed. The existing solutions pertaining to connection correlation such as CIS [3], SWT [6], Thumbprinting [5], and DIDS [4] were developed based on certain features or characteristics. They include inter-packet timing characteristics, host activity and connection content such as packet payload. The main drawback of these solutions is that the host activity related data collected from intermediary node which is intended to find the source of attack is not trustworthy. This is because the attack is expected to have full control over all intermediary nodesand his node it is possible for attacker to manipulate the traffic to conceal him from being traced. As the attacker has logged into remote intermediary machines through remote login programs such as SSH and Telnet he has gained access to the resources of the intermediary nodes. The drawback of content based correlation approaches is that they assume that the payload of packets is not changed across the intermediary nodes. As encryption of such content can be made by attacker, these approaches are suitable only for unencrypted connections. Other approaches such as timing based approaches passively monitor incoming and outgoing traffic and correlate the flows. The main drawback of these approaches is that the attacker can perform timing based perturbations to deceive the detection systems. Therefore these systems tend to fail in this case.

The first generation of correlation approaches that are timing – based were very effective. They were able to correlate encrypted connections and establish the actual source of attack successfully and accurately. However, in the later stages, the intruders changed the way they make attacks. They started using encrypted connections and also perform timing perturbations. Thus these first generation correlation approaches became ineffective hence they are vulnerable when attackers use active timing perturbations. In [8] Donoho et al. first of all identified the limits of the attackers on performing active timing perturbations and injection of bogus packets. They showed that correlation based long time behavior is possible in spite of timing perturbations from attackers. According to them this can be achieved by using multiple timescale analysis techniques. However, in [8] they could not provide information on tradeoffs between the scale of timing perturbation and the required level of correlation effectiveness and the packets needed. Other issue that could not be addressed by [8] is the jitter used by intruders.

Due to the drawbacks in first generation timing based correlation techniques, the coarse scale analysis causes false positives to be increased. However, the true positives are increased with timing – perturbed flows usage. The limitations of timing perturbations is studied in [8]. However, they did not address these problems in their paper. In [10] and [7] other positive timing based correlation methods came into existing and they consider false positives and true positives at a time. They tend to derive both lower and upper bounds on the number of packets required to achieve false positive rate and 100% rate of true positives. The work of those papers could not provide any experimental evidences. He et al. [11] and Zhang et al. [12] of late proposed many timing based correlation methods based on the assumptions used in [7] and their approach has been proved to be better. However, they are using passive timing based approach. Information theoretic game is explored in [13]. Their analysis is based on the packet reordering channels. Watermark based correlation is studied in [10] recently and provided a statistical method in order to detect the presence of watermark in packet flow. Their method has some assumptions such as having access to flows containing watermark and no watermark. Overall, the existing approaches are passive in measuring the possibility of timing perturbations done by intruders. The existing approaches fail to cope with when intruders use timing perturbations in encrypted connections. To overcome this problem, this paper proposed an active timing based correlation approach without any assumptions, without any usage of random process , robust and requires very less packets when compared to passive approaches for ensuring the same level of accuracy in finding the source of attack and showing close to 100% true positives.

OVERVIEW OF OUR APPROCH

The proposed watermarking-based correlation approach is aware of bidirectional communication nature of remote login programs such as SSH or Telnet. This is because when attacks are made by adversaries trough intermediary nodes and by making use of remote login programs, it is essential to trace it back from the victim node to the attacker's actual machine. Figure 1 shows outline of the proposed model. As can be seen in the figure, between attacker and intended victim or target machines, there lay a set of intermediary nodes named H1, H2, and H3. These are considered for illustrative purpose only. There might be *n* number of intermediary nodes between the source and destination. The attacker here does not make an attempt to execute attacks directly on the target. Instead, by using remote login programs such as Telnet, SSH, etc. he will execute the attacks through intermediary programs and machines. This makes the security personnel at target machine to establish the source of attack accurately.

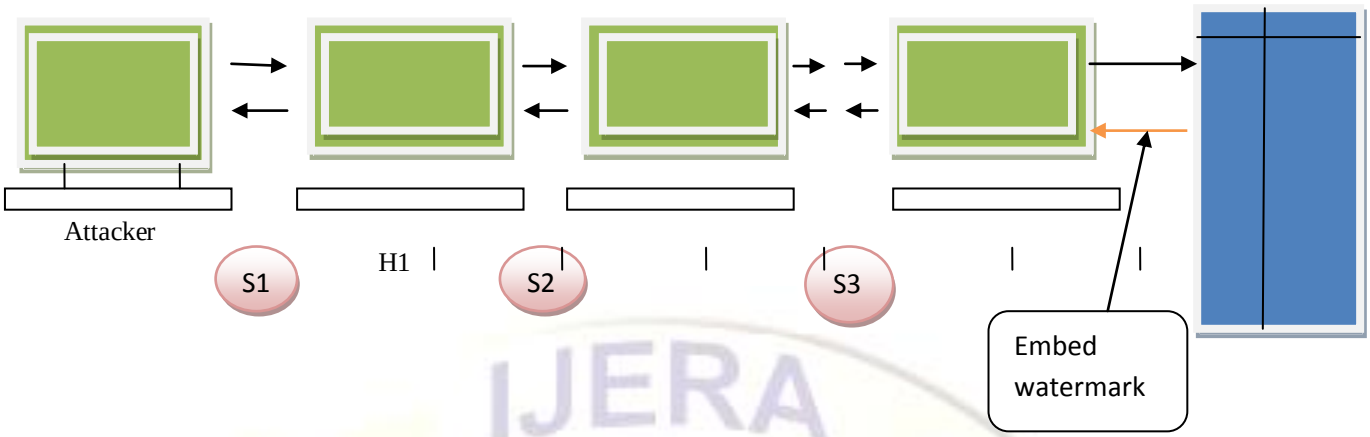

Fig. 1 – Outline of watermark tracing model

As can be seen in the proposed model in fig. 1, there are network sensors named S1, S2, and S3. These sensors are responsible to monitor the network flows and also involved in preventing disguised attacks from the adversaries. When attack is made by intruder, before it reaches final target machine, the proposed watermarking – based scheme will watermark the backward traffic and inform the fact to all sensors that are employed in the network. Afterwards, the sensors monitor the traffic and inform the target machines about any occurrence of watermark in the traffic flows. The sensors are deployed at strategic places such as edge router, firewall and gateway that are part of the network. The traffic that comes backward from the attack node back to actual source, the backward traffic which has been watermarked by the target's security framework, it can't be controlled by adversary. The attacker has no access to un-watermarked version of traffic. This very reason makes it difficult for the adversary to know the packets that are delayed. To follow any distribution mechanism to be effective the correlation method proposed here does not require the random timing perturbation provided by the attackers. Only one assumption made in this paper pertaining to timing perturbations.

PROPOSED WATERMARK BIT EMBEDDING AND DECODING

As intruders can perform timing based perturbations to encrypted flows, the watermark embedding is done at target machine. First of all IPD is quantized using the function.

$q(ipd,s)=round(ipd/s)$,

Then the embedding process is done using the function

$E(ipd,w,s) = [q(ipd + s/2,s) + \Delta] X s$,

In accordance with the above function, the watermark-bit-decoding is done as follows.

$d(ipdw,s) = d(ipdw,s) \bmod 2$.

EXPERIMENTAL RESULT

Analysis of Watermark Delectability

Watermark detection is a process of checking whether the given watermark is embedded in flows. The proposed watermark detector followed steps described here. Decode l-bit from given flow; compare the decoded l-bit (wf) with *w*; if the Hamming distance between *wf* and *w* indicate the decoded l-bits report that watermark is detected.

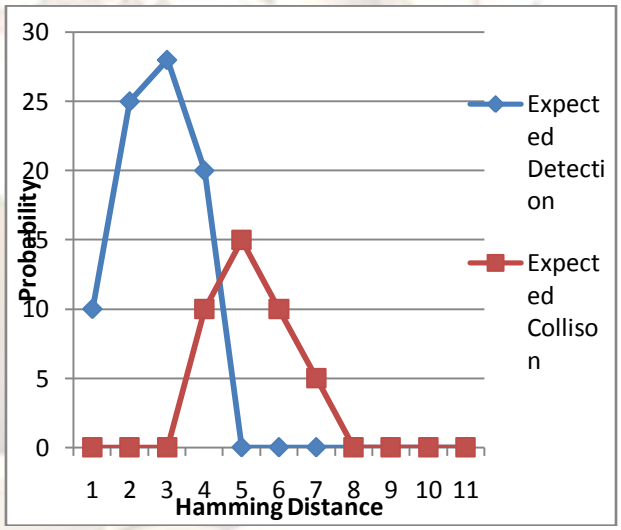

Fig. 2 – Effect of threshold on detection and collision rates of watermarking method

As can be seen in fig. 2, the derived probability distribution is plotted in Y axis while the Hamming distance in X axis for the expected detection and collision rates.

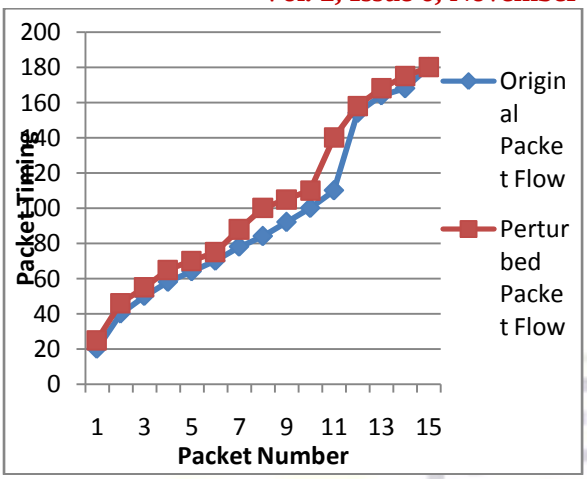

Fig. 3 – Difference between original and perturbed packet flow

As can be seen in fig. 3, it is evident that when timing based perturbation is employed, there is difference between the original packet flow and perturbed packet flow.

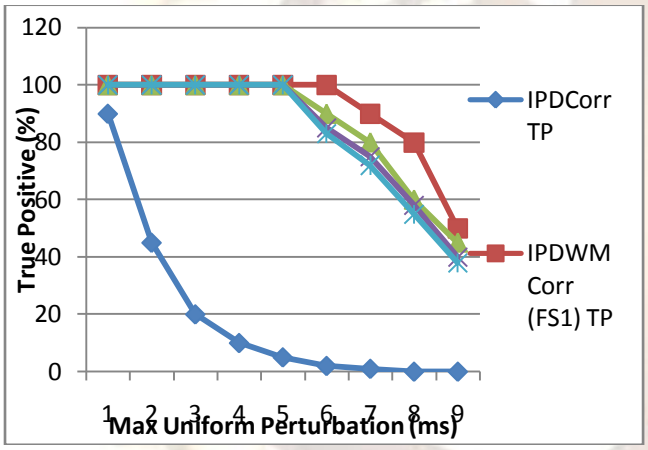

Fig. 4 – True positive rates of correlation

As can be seen in fig. 4, under uniformly distributed random timing perturbations, correlation true positives are visualized. It shows IPD based correlation and also watermark based correlation on FS1 and FS2 under various levels of uniformly distributed random timing perturbations.

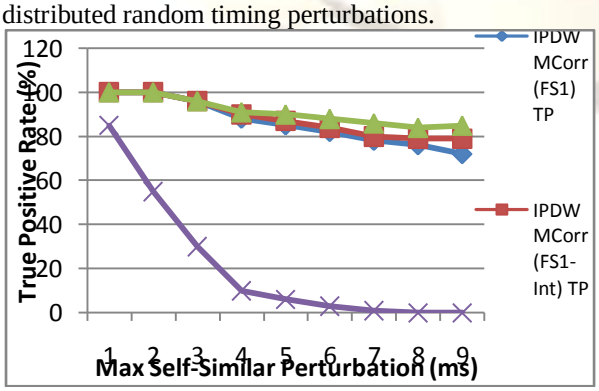

Fig. 5 – Correlation of true positive rates (Max self-similar perturbation)

As can be seen in fig. 5, the measured watermark correlation true positives under a variety of self similar perturbations are presented. It is evident that the bounded self – similar perturbation gives rise to much higher true positive rates.

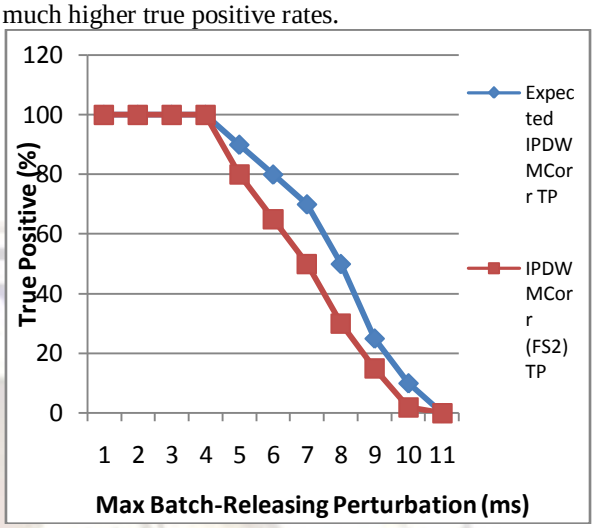

Fig. 6 – Correlation true positive rates (Batch releasing, random timing perturbations)

As can be seen in fig. 6, the measured watermark correlation true positives under batch releasing timing perturbations are presented. The measured true positive rates are close to the expected values. This indicates that our approach is effective.

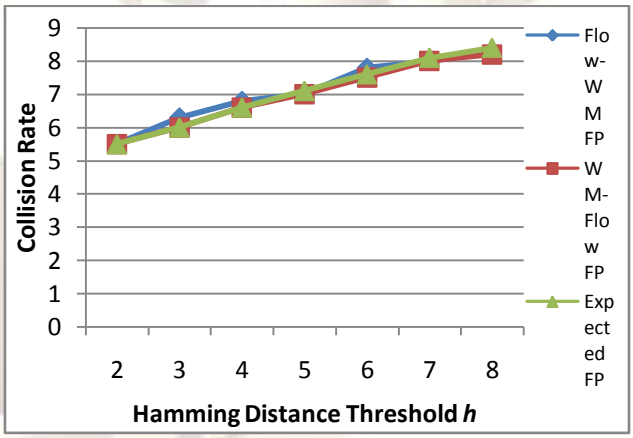

Fig. 7 – Correlation false positive rates vs. hamming distance

As seen in fig. 7, the false positive rates are shown for various hamming distance thresholds with fixed length 24 bit watermarks. Average of 100 separate experiments is presented in the figure.

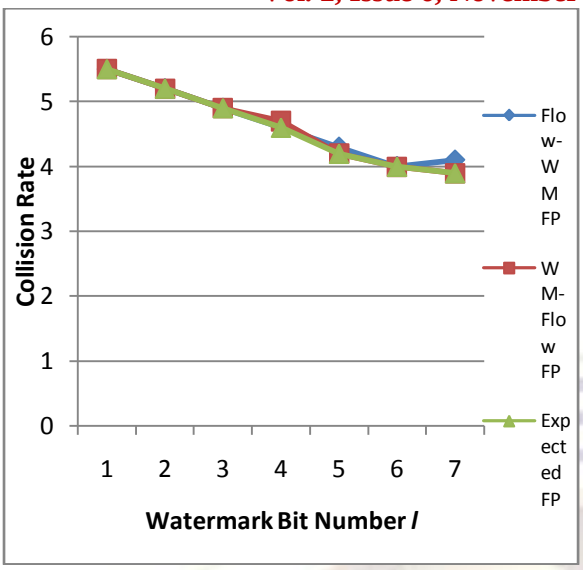

Fig. 8 - Correlation false positive rates vs. watermark bits

As seen in fig. 8, the false positive rates are shown for various watermark lengths with a fixed hamming distance length 5. Average of 100 separate experiments is presented in the figure. Figures 9, 10 and 11 are to present experimental results pertaining to watermark detection rate tradeoffs with redundancy number m, hamming distance threshold h and number of watermark bits l respectively. The results are average of 100 experiments for the measured watermark detection rates of FS1 and FS1-Int. In case of watermark detection rates of FS2 average of 10 experiments is used. They are part of the proposed quantitative tradeoff models.

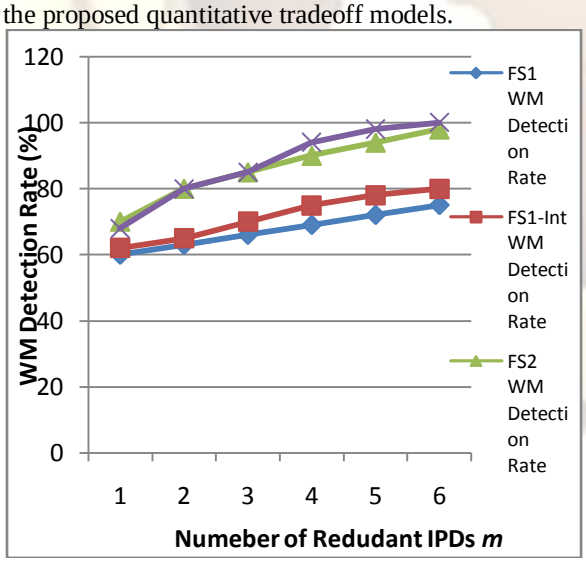

Fig. 9 – Watermark detection rate with redundancy number m

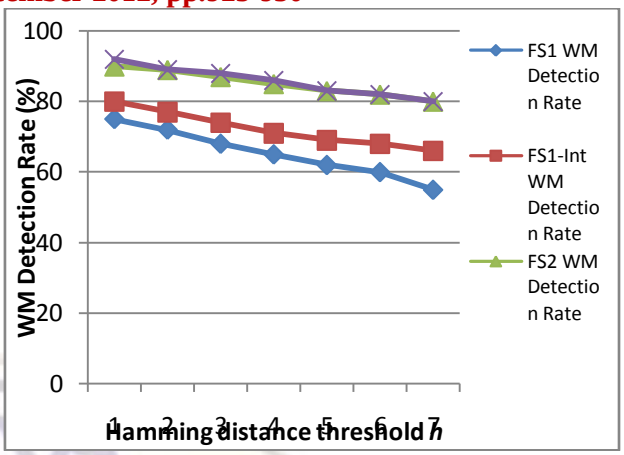

Fig. 10 - Watermark detection rate with hamming distance threshold *h*

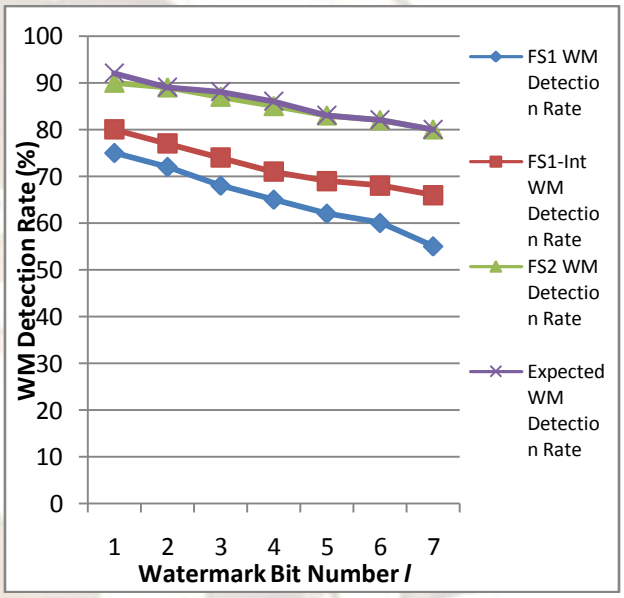

Fig. 11 - Watermark detection rate with number of watermark bits

CONCLUSION AND FUTURE WORK

Accurate identification of source of attack is a challenging problem when attackers make use of intermediary nodes to exercise their attacks. This is especially true when the traffic of attack is encrypted and the timing is altered by the attackers. The passive timing based correlation techniques are not effective for this reason. In this paper, we presented a new active timing based correlation approach that can effectively handle random timing perturbations. The proposed scheme embeds unique watermark into inter-packet timing in such a way that the encrypted flows are correlated and it is robust to random timing perturbations employed by intruders. The experimental results reveal the fact that the proposed approach results in close to 100% true positives and 0% false positives. When compared with passive correlation approaches our

approach has many advantages including no assumptions are required about original inter-packet timing and flow; very less packets are required. Further research can be made in the area of flow watermarking to make it more robust and works with even fewer packets.

REFERENCES

[1] M. T. Goodrich. Efficient packet marking for large-scale ip traceback.In *Proceedings of the 9th ACM Conference on Computer and CommunicationsSecurity (CCS 2002)*, pages 117–126. ACM, October 2002.

[2] 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 2004 IEEE Symposium on Security and Privacy*,IEEE, 2004.

[3] H. Jung. et al. Caller Identification System in the Internet Environment.In *Proceedings of the 4th USENIX Security Symposium*, USENIX, 1993.

[4] S. Snapp. et al. DIDS (Distributed Intrusion Detection System) -Motivation, Architecture, and Early Prototype. In *Proceedings of the14th National Computer Security Conference*, pages 167–176, 1991.

[5] S. Staniford-Chen and L. Heberlein.Holding Intruders Accountable onthe Internet. In *Proceedings of the 1995 IEEE Symposium on Securityand Privacy*, pages 39–49. IEEE, 1995.

[6] X. Wang, D. Reeves, S. F. Wu, and J. Yuill. Sleepy WatermarkTracing: An Active Network-Based Intrusion Response Framework. In*Proceedings of the 16th Internatinal Conference on Information Security(IFIP/Sec 2001)*, pages 369–384. Kluwer Academic Publishers, June2001.

[7] A. Blum, D. Song, and S. Venkataraman. Detection of Interactive Stepping Stones: Algorithms and Confidence Bounds. In *Proceedingsof the 7th International Symposium on Recent Advances in IntrusionDetection (RAID 2004)*. Springer, October 2004.

[8] D. Donoho. et al. Multiscale Stepping Stone Detection: Detecting Pairsof Jittered Interactive Streams by Exploiting Maximum Tolerable Delay.In *Proceedings of the 5th International Symposium on Recent Advancesin Intrusion Detection (RAID 2002): LNCS-2516*, pages 17–35. Springer,October 2002.

[9] K. Yoda and H. Etoh.Finding a Connection Chain for TracingIntruders. In *Proceedings of the 6th European Symposium on Researchin Computer Security (ESORICS 2000), LNCS-1895*, pages 191–205.Springer-Verlag, October 2002.

[10] X. Wang and D. Reeves. Robust Correlation of Encrypted Attack Traffic through Stepping Stones by Manipulation of Interpacket Delays. In*Proceedings of the 10th ACM Conference on Computer and CommunicationsSecurity (CCS 2003)*, pages 20–29. ACM, October 2003.

[11] T. He and L. Tong, Detecting Encrypted Stepping-Stone Connections.In *IEEE Transactions on Signal Processing*, 55(5), pages 1612-1623,2006.

[12] L. Zhang, A. G. Persaud, A. Johnson, and Y. Guan.Detectionof Stepping Stone Attack under Delay and Chaff Perturbations. In*Proceedings of the 25th IEEE International Performance Computingand Communications Conference (IPCCC 2006)*, April 2006.

[13] 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 2006)*, 2006.

AUTHORS

M.Deepika is Associate Professor in Aurora Technological amd Research Institute, Hyderabad, AP, INDIA. She has received B.Tech Degree Information Technology, M.Tech Degree in computer science and engineering. Her main research interest includes data mining. Cloud computing.

G. Om Sai Prashant is student of Aurora Technological amd Research Institute, Hyderabad, AP, INDIA. He has received B.Tech Degree in Information Technology, M.Tech Degree in Web Technologies. His main research interest includes data mining. Cloud computing.