Chapter 8

CONCLUSIONS & FUTURE WORK

Computer network traffic is dynamically changing. These changes are measured either across spatial perspectives or across temporal perspectives. The traffic features change with respect to time as well as space. Network traffic anomalies also go through changes. The anomalies too change along time and space dimensions. Network anomaly detection systems are therefore, constantly upgraded to meet the changing needs of network traffic and anomalies. Further, network administrators do round the clock job of monitoring and maintaining local networks and protecting them from internal and external threats. They rely heavily on network monitoring tools and visualizers. In this thesis, our aim, therefore, was to develop a methodology for detection of network based anomalies and also provide a visualization map.

8.1 CONCLUSIONS

This thesis proposed the use of self-similarity as a measure for detection of anomalies in computer network traffic. We did extensive survey of the literature available in the area of distributed denial of service attacks and various detection techniques being used for detection and mitigation of these attacks. While doing the survey
we realized that there was need for development of anomaly detection techniques that could be applied to aggregated network traffic and could provide both spatial and temporal dimensions of the traffic. And in doing so be able to differentiate network anomalies from regular traffic flows.

The main contributions of this thesis are:

1. To provide a collaborative solution for detection of rate-based network anomalies we have provided unified detection methodology, Multi Scale Network Anomaly Detection (MS-NAD). MS-NAD is based on the concept of measuring the variances in the self-similar behaviour of network traffic and using it for detection of different kinds of anomalies in the network traffic.

2. We have shown that by using the variance in the self-similar parameter $H$, traditional flooding based as well as new breed of Distributed PDoS attacks could be detected. It was measured and analyzed (Chapter 5), that presence of flooding-based DDoS attacks in the regular traffic decreases the self similar nature of the traffic. This was because of sudden surge in DDoS flooding packets that led to ‘0 byte per window’ syndrome where the impact of small sized and slow rate flows diminishes. The Hurst values in case of flooding-based DDoS attacks decrease by 0.045 or 4.5% and based on this threshold, the attacks were successfully detected.

3. In Chapter 5 it was also realized that, there is significant difference in the impact of ‘flash crowds’ on the regular traffic in comparison to the flooding-based DDoS attacks. Flash Crowds increase the expected energy and thus self-similarity in the upper scales of the logscale plots. This was due to gradual increase in the flash crowds and sudden decline in the build down phase in contrast to DDoS attacks. Moreover flash crowds are a resultant of high rise in traffic to websites and are TCP supported HTTP packets only. The regular TCP flows have no impact on self-similar nature of the traffic in smaller scales because of their long-lived large time interval flows. Therefore in the presence of flash crowds the
self-similar nature increases.

4. For distinguishing between flooding-based attacks and flash crowds MRO maps proved useful. The flooding-based attacks show sudden increase in the MRO maps spread across all scales whereas for flash crowds the build up phase grows gradually and spread only to a range of scales. Flash crowds could therefore be distinguished from flooding-based DDoS attacks.

5. On the one hand we were not only able to successfully detect distributed flooding-based denial of service attacks but also distinguish them from their look alikes—‘flash crowds’, we were also able to successfully detect the new breed of DDoS attacks (Chapter 6). Distributed PDoS attacks are a distributed form of PDoS or LDoS attacks. The PDDoS attacks are a new category of attacks and therefore needed detailed study. These attacks send stealthy pulses to the TCP servers resulting in degradation in its QoS. The attacks are difficult to detect using the traditional statistical techniques. Large number of traces were generated ranging from low traffic strength to heavy traffic strength and impact of low strength attacks to heavy duty attacks was measured. The results were computed with reference to Logscale plots, $H$ values and $MRO$ maps.

6. In Chapter 6 we first measured self-similarity for attack free TCP traffic. The Hurst values were 0.75 approximately for all TCP traffic strengths. We also realized that TCP traffic has presence of self-similar behaviour only in selected range of scales, i.e. the scales that map to the RTTs of TCP flows. The dip in the logscale plots around scales 3, 4 and 5 was seen for low, medium and heavy TCP traffic strengths, respectively. The dip was around 80ms to 320ms time scales and matched the RTTs of the TCP flows.

7. The detection of DPDoS attacks from regular traffic was done by measuring variances in Hurst values and analyzing logscale plots. Logscale plots for heavily utilized bottleneck links were same because of no variance in expected energy
values for heavy traffic and high attack strengths. For low strength attacks logscale plots had dip seen at scales 5 and 6 for shorter pulses and at scales 3 and 4 for longer pulses. Similarly, for medium strength attacks dips were recorded between scales 4 to 7 and for heavy strength attacks dips were measured between scales 3 to 8. **Hurst values for low, medium and heavy attack traffic flows** lied in the range of 0.8 to 1. The variances in the Hurst values were in the range of 0.04 to 0.115 which were above threshold value and henceforth, presence of attacks could be detected.

8. **MRO maps** helped in pointing the PoP points as well as duration and strength of the attack for DPDoS attacks with different pulse lengths and burst rates. For low strength attacks the visibility could be measured in scales 2 to 9, for medium strength attacks visibility could be measured in scales 2 to 7 and for heavy strength attacks visibility could be measured in scales 2 to 6. The variations in color bars in upper scales proved helpfull in distinguishing attacks with different pulse lengths and time durations as well. We, therefore were able to detect more than 90% of the PDDoS attacks using Hurst values, logscale plotting and MRO maps. We therefore, successfully detected distributed PDoS attacks from normal traffic behaviours.

9. We **tested our wavelets based estimation of Hurst values for DoS attacks in the KDD Dataset as well**, to detect changes and henceforth trace an attack in the network traffic. The **overall accuracy of the proposed model was 99.79% with true positive rate of 99.95%, true negative rate of 99.13%, false positive rate of 0.87% and false negative rate of 0.05%**. We therefore could conclude that variation in self similar nature of the network traffic could be considered for detection of DoS attacks particularly long-duration attacks.

10. We therefore concluded that the variability in self-similar behaviours of regular network traffic and anomalous traffic can be measured using the Hurst parameter and could be used for detection of malicious
traffic flows from legitimate traffic flows. We used wavelets based fast pyramidal algorithm for computing coefficients of aggregated traffic flows at different scales and measured variances in expected energies at different scales independently. The variances in computed values were measured and used to compute Hurst values, to plot logscale plots and MRO maps. Hurst values helped in distinguishing attack traffic from normal traffic. MRO maps proved useful in pointing out point of presence of an attack as well as strength of an attack. Logscale plots proved usefulness in detection of DPDoS attacks that rely on hitting the RTTs of the TCP flows.

8.2 FUTURE WORK

1. We have used single input parameter (aggregated source bytes count) for our detection technique. This single parameter, although effective in identifying the types of attack presented in this work, may not hold for a wider range of anomalies where the variations in other features may be more subtle. Next step would therefore be to explore the possibility of other input parameters.

2. We observed that the detection of DDoS attacks based on self-similarity was significantly affected by network congestion. An in-depth study of the effects of rate limiting and queue length on self-similar traffic would be useful to establish a relation between level of congestion and reliability of DDoS detection.

3. Look out for applicability of property of self-similarity for detection of other type of anomalies in the network traffic like Worms, Outages, Port Scans, DNS amplification attacks as well.