2012 4th International Conference on Cyber Conflict
C. Czosseck, R. Ottis, K. Ziolkowski (Eds.)
2012 © NATO CCD COE Publications, Tallinn

Sensing for Suspicion at Scale: A Bayesian Approach for Cyber Conflict Attribution and Reasoning

Harsha K. Kalutarage, Siraj A. Shaikh, Qin Zhou and Anne E. James
Digital Security and Forensics (SaFe) Research Group
Department of Computing
Faculty of Engineering and Computing
Coventry University
Coventry, CV1 5FB, UK
{kalutarh, aa8135, cex371, csx118}@coventry.ac.uk

Abstract: Cyber conflict monitoring remains one of the biggest challenges today, amidst increasing scaling up of cyberspace in terms of size, bandwidth and volume. Added to this, the increased determination of cyber actors to operate beneath the threshold makes it ever more difficult to identify unauthorised activities with desired levels of certainty and demonstrability. We acknowledge a case for persistent and pervasive monitoring; detection of serious sabotage and espionage activities, however, is dependent, in part, upon the ability to maintain traffic history over extended periods of time, somewhat beyond current computational and operational constraints. This makes it crucial for research in cyber monitoring infrastructures, which are configured to handle cyberspace at live and modern scale and sense suspicious activity for further investigation. This paper explores Bayesian methods together with statistical normality to judge for effective activity attribution, particularly in high-volume high-scale environments, by combining both prior and posterior knowledge in the scenario. The set of experiments presented in this paper provides tactical and operational principles for systematic and efficient profiling and attribution of activity. Such principles serve a useful purpose for technologists and policy-makers who want to monitor cyberspace for suspicious and malicious behaviour, and narrow down to likely sources. The proposed approach is domain agnostic and hence of interest to a cross-disciplinary audience interested in technology, policy and legal aspects of cyber defence.

Keywords: *anomaly detection, Bayesian approach, reasoning, attribution, cyber attacks*

1. INTRODUCTION

Cyber conflicts are increasingly a part of mainstream warfare. Attribution of cyber activity — "knowing who is attacking you" or "determining the identity or location of an attacker or an attacker's intermediary" [1,2,3] – is naturally a vital ingredient in any cyber security strategy. Parker claims that 'a common problem with many analysis tools and techniques today is that they are simply not designed for purposes of attribution' [4]. According to [5,6], although current approaches are capable of alerting to suspicious activities, they are failing in this information age because when computers are under attack, the 'who' and 'why' are frequently unknown. Many researchers claim that completely depending on information derived from network traces will do little for cyber conflict attribution and detection, mainly due to the nature of Internet infrastructure, and therefore there is a need for approaches that combine technical solutions data with the information gathered from contextual analysis and intelligent services (combining prior belief with posterior knowledge) [7]. This paper aims to address this challenge and presents work ultimately contributing towards this goal.

Reconnaissance, the first phase of the anatomy of a cyber attack, can be further sub-divided into three incremental stages: casing, scanning, and enumeration. It is difficult to tackle suspicious activities at the casing stage, as everything seems to be legitimate. However in the second stage, scanning, the attacker attempts to send packets to the target IP address (range of IP addresses) with the goal of determining what machines are presented and reachable (ports) on the target network. Two most common examples of scans, among many others, are 'pings-ICMP' and 'SYN-TCP'. This offers a starting-point for detection of potentially suspicious activity. For enumeration, the attacker may follow up with various kinds of attempts to identify services. The detection of scan and enumeration attempts is made more difficult as attackers increasingly use slow scan rates to stay beneath the threshold. If an attacker is methodical enough to make only the slightest of changes at any one time and each step is spaced far enough apart, it will be difficult to detect by traditional signature matching algorithms. Often, network-based intrusions signatures are state-full and require several pieces of data to match an attack signature. If the length of *event horizon* (time amount from the initial data piece to the final data piece needed to complete the attack signature) is longer, intrusion detection systems (IDSs) cannot maintain state information indefinitely without eventually running out of resources. This helps slow attackers to hide behind noise and other traffic. Most current approaches do not track activity over an extended period of time, due to computational constraints and disc storage requirements. This paper develops an approach to serve as an early warning system for slow suspicious activities that warrant further investigation.

This work is inspired by Chivers *et al.*'s work [8,9] to adopt a Bayesian approach to combine both prior and posterior knowledge in the scenario and detect (with attribution) slow and suspicious activities in a cyber conflict. The series of experiments examines the effectiveness of such an approach under different parameters: multiple attackers, traffic volume, cluster size and event sampling.

The rest of this paper is organised as follows: Section 2 presents a brief overview of related work; Section 3 presents the underlying methodology and the theoretical account of the process; Section 4 overviews the experimental set up and Section 5 follows up with results and analysis.

Section 6 presents some early results on possible use of sampling. Section 7 concludes the paper.

2. RELATED WORK

Although a considerable number of anomalies-based intrusion detection approaches have been proposed during the last two decades, many of them are general in nature and quite simple [10-12]. They fail in attributing, in accumulating evidence, and also in scaling up. Since our approach accumulates evidence (both contextual and technical traits) over an extended period of time and uses that information to identify aberrant behaviours (see Sections 3 and 4) it differs from most of the above existing approaches, and can be known as an *incremental anomaly detection approach*. Based on an exhaustive survey of published incremental anomaly detection approaches, Bhuyan *et al*. conclude that most existing approaches have a high rate of false alarms, are non-scalable, and are not fit for deployment in high-speed networks [13]. On that perspective, the proposed approach differs from existing incremental approaches, since this is scalable in terms of storage and possible to incorporate with live analysis on high-speed networks. The proposed approach requires maintaining only a single value for a given node. Most of the current intrusion detection approaches do not accommodate integrating contextual information with attack detection and attribution and are heavily dependent on technical traits only [13-20]. Hence, our approach is significantly different from most of the existing approaches. However [8-12,22-24] can be identified as deviations from the current general and quite simple systems.

Kandias *et al*. propose a model to integrate the user's technological traits with data obtained from psychometric tests [24]. Although the authors focus on insider attacks, the core idea in their paper coincides (to some extent) with our work, since they do not depend completely on network traces. They combine users' (psychological) profiles with technical data. However, their model is highly subjective, organisationally dependent and does not accommodate any information gathered from contextual analysis. Most importantly, it cannot be applied to profile non-human actors. In contrast, ours can be used to profile human, non-human or even virtual actors and can be extended to accommodate a wide range of contextual information.

Chivers *et al*. provide a scalable solution to identify suspicious slow insider activities, combining evidence from multiple sources using the well-known Bayes' formula [8,9]. Although similarly motivated, our work mainly differs from the decision criteria used for the analysis as described in Section 3 and from the target domain. Also, we have discussed the possibility of extending the same formula to integrate contextual information on detection. Chivers *et al*. distinguish between anomaly and normal behaviours by setting a control (base line) and choosing the one most deviant from the control as an attacker. This is not practical, as it is very hard setting a predefined baseline for node behaviours and the authors have not discussed it. As we identified, when there are more than one attacker in a subnet with higher variations of nodes behaviours, this decision criterion does not work well. Comparison across subnets (i.e. using a common baseline for all subnets) is also problematic. Identifying anomaly nodes through visually inspected row score graphs is another issue in Chivers *et al*.'s work. Such a decision can be affected by even dimensions of the drawing canvas in a situation where there is a higher

variation in parameter values. In such a situation, *standardisation* of node scores should be performed, before any comparison, which has been ignored by their work. However Chivers *et al*. themselves identify a need for different decision criteria other than the *Maximum score function* method they used. We have incorporated the concept of statistical normality into our work when addressing these issues.

Basu *et al*. propose an approach which uses connection-based windows to detect low-profile attacks with a confidence measure while Streilein *et al*. use multiple neural network classifiers to detect stealthy probes [22,23]. [21,24,29] can be identified as much more similar studies to Chivers *et al*.'s work. In [23,24], users are profiled according to their behaviour and that information is used to identify users who require further investigations. Evidence accumulation as a means of detecting slow activities has been proposed by [21]. All the above approaches, except [8,9,21], require the storage of large volumes of event data for later analysis, and hence differ from our work. [21] differs from our work as it uses a counting algorithm instead of the Bayesian approach and also in its decision criteria. Importantly, all the above approaches, except [24], are profiling the suspected origins based on technical solution data only. Since our aim is not only to propose an efficient attribution methodology but also to conduct an investigation of its effectiveness under different conditions, certainly this work significantly differs from all the above works.

3. METHODOLOGY

We address the problem by dividing it into two separate smaller sub-problems: *Evidence fusion & aggregation (Accumulation)* and *Analysis (Anomaly definition)* assuming that exiting signature detection algorithms could be employed to detect the events (signature elements) of an attack pattern. The term node is used in this paper to denote anything in terms of identities, which can be a user, machine, account number or a location (physical or virtual), essentially the *visibility source* of a potential attack [2,3].

A. Evidence fusion & aggregation

According to Brackney *et al*., integrating information from many sources in a manageable and scalable fashion, in order to identify patient attackers, is still an important open question [18]. Chivers *et al*. claim that combining events from one or more sensors (possibly of various types) while reducing data without adversely impacting detection is a major challenge [8,9]. Both statements are talking about 'Evidence fusion & aggregation'. Chivers et al. use a Bayesian approach, while [21] uses a counting algorithm for this purpose. However [8,9] show that the Bayesian approach is superior to the counting algorithm. At this stage, we also used the simple Bayes' formula for evidence fusion, as described in the next sub-section. Jiang et al. show that probabilistic correlation works well in noisy environments [28]. However, investigating ways to apply other possible methods, instead of the simple Bayes' formula, such as Bayesian Belief network, Kernel Density Estimation (KDE), Dempster-Shafer theorem, Kalman Filter, Viterbi algorithm, Gi*, Evidential reasoning, Logic based fusion, Preference aggregation, Neural networks, Ontology & category theory for this task would be interesting and is left as future work in this ongoing work.

Bayesian approach

The posterior probability of the hypothesis H_k given that E is given by the well known formula:

$$P(H_k/E) = \frac{p(E/H_k) \bullet p(H_k)}{p(E)} \quad (1)$$

In order to fit this formula into our case, let H_k: hypothesise that k^{th} Node is an attacker and $E = \{e_1, e_2, e_3, \ldots, e_m\}$ the set of all suspicious evidence observed against node k during time t from m different independent observation spaces. Here $P(E)$ is the probability of producing suspicious events by node k, but on its own is difficult to calculate. This can be avoided by using the law of total probability and reformatted (1) as:

$$P(H_k/E) = \frac{p(E/H_k) \bullet p(H_k)}{\Sigma_i\, p(E/H_i) \bullet p(H_i)} \quad (2)$$

For independent observations, the joint posterior probability distribution:

$$P(H_k/E) = \frac{\Pi_j\, p(e_j/H_k) \bullet p(H_k)}{\Sigma_i\, \Pi_i\, p(e_j/H_i) \bullet p(H_i)} \quad (3)$$

Once we observed E from node k, to calculate the posterior probability of node k being an attacker $p(H_k/E)$, it is necessary to estimate:

1. $p(e_j/H_i)$ - likelihood of the event ej given the hypothesis H_i and,
2. $p(H_i)$ - prior probability

Assuming that we know the prior and likelihoods, it is obvious that (3) facilitates to combine evidence from multiple sources (contextual information) to a single value (posterior probability) which describes our belief, during a short observation period, that node k is an attacker given E. Aggregating short period estimations over time helps to accumulate relatively weak evidence for long periods. This accumulated probability term, $\Sigma_t\, p(H_k/E)$ (t is time) known as *profile value* hereafter, can be used as a measurement of the level of suspicion for node k at any given time. Schultz *et al*. claim that profiling suspected insiders provides one of the best ways of reverse engineering an attacker [25]. Although there are some significant differences between the characteristics of insiders and outsiders, profiling can still be used effectively in cyber conflict attribution, as shown in the rest of the paper.

B. Analysis

At any given time, given the profiles of all nodes, detecting suspicious profiles is the analysis stage as the attacker's activity pattern is now reflected by profiles. Bhuyan *et al*. claim that anomaly detection is usually flexible and sufficient to detect both unknown (novel) and known attacks [13]. When there is an attacker who violates legitimate users' activity patterns the probability that the attacker's activity is detected as anomalous should be high. We distinguish between anomalous and normal profiles using the concept of statistical normality.

Statistical Normality

The statistical approach to *normality* defines it in terms of a normal distribution curve. A normal curve is a statistical data distribution pattern occurring in many natural processes. As long as what is most common (average or most frequent) in the general population is considered as normal, any behaviour or characteristic that occurs only rarely can be regarded as abnormal. In a normal distribution, node profiles lying outside (around) three standard deviations from the mean can be considered as abnormal. This boundary may vary, so one may define abnormality beyond two standard deviations from the mean and hence select a wider selection of nodes for further investigation. One advantage of this is that confidence in attribution can also be expressed in probability terms. Calculating standardised node profiles (Z-scores) instead of node profiles themselves, will resolve the analysis problem better.

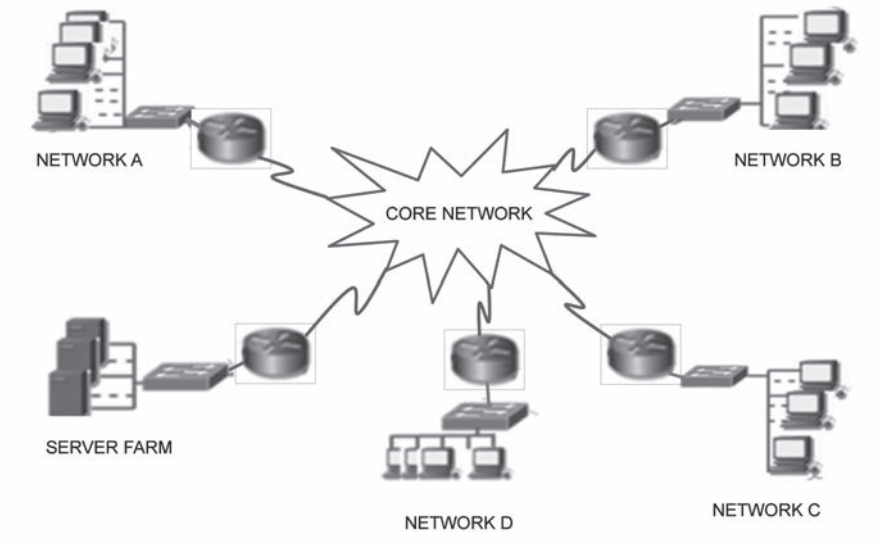
4. EXPERIMENTAL SETUP

To demonstrate the proposed approach, a series of experiments were conducted. Simulation was used to express network topology and traffic patterns of interest were generated using NS3 [26], assuming Poisson arrival model with inter-arrival time gap between two consecutive events as an exponential, to collect data on the characteristics and behaviour of several common network reconnaissance tools. Each simulation was run for a reasonable period of time to ensure that enough traffic was generated (over one million events).

A. Network Topology

Figure 1 shows the network topology used for our experiments. A total of 2,122 nodes were distributed among four networks labelled *A* (99 nodes), *B* (400 nodes), *C* (800) and *D* (800 nodes). In addition, a network dedicated to a server farm was simulated with 23 nodes.

FIGURE 1. THE NETWORK TOPOLOGY USED FOR EXPERIMENTS.
SOURCE FOR GRAPHIC SYMBOLS: FUNDAMENTALS OF NETWORK SECURITY GRAPHIC SYMBOLS, CISCO NETWORKING ACADEMY PROGRAM (FREELY AVAILABLE ON WWW).



B. Attacker Modelling

If λ_s . λ_1 are mean rates of generating suspicious events by suspicion and normal nodes respectively, we ensured maintaining $\lambda_s = (\lambda_1 . \lambda_1 + 3\sqrt{\lambda_1}]$ and $\lambda_1(=0.1)$ sufficiently smaller for all our experiments to characterise slow suspicious activities which aim atstaying beneath the threshold and hiding behind the background noise. $\sqrt{\lambda_1}$ is the standard deviation of rates of suspicious events generated by normal nodes.

C. Parameter Estimation

Prior probabilities and Likelihoods are assigned as follows.

$$P(H_m) = P(H_n) = \frac{1}{\text{Number of nodes in the scene}} \text{ , for all m, n and } m \neq n \quad (4)$$

$$p(e_j/H_m) = p(e_j/H_n) = k, \quad \text{for all j, m, n and } m \neq n \quad (5)$$

(4) assumes that all nodes in the scene have a same prior belief (equally likely) to be subverted. However, this is not the case in many situations. In cyber warfare, as many countries have a cold cyber war with other countries [6], one entity may have a higher prior belief of suspicion about the activities of another. In networks, an e-commerce server may have a higher chance to be subverted than a client node. In a company, an angry programmer attached to the IT department could be more dangerous than a loyal employee in the marketing department. Therefore if the analyst requires to distinguish between identities (or clusters of identities, for example, in case of identity is a geospatial location; a cluster can be a province, a country or even an alliance of countries), prior probability can be assigned separately. Since prior probabilities are based on previous experiences, $p(H_m)$ can be judged by the analyst, based on the information gathered from contextual analysis or intelligent services.

(5) explains the likelihood of producing event ej by any node if it is subverted. For the purpose of demonstration, we assigned arbitrary values (≤ 1) for k. However it can be estimated as follows. If ej is an event such as *UDP scan* or *land attack* which cannot be expected from a non-subverted node, then k can be assigned to one. However, k cannot always be one, for some suspicious events that appear as a part of attack signatures could also be originated from normal network activities. For example, a major router failure could generate many ICMP unreachable messages; an alert of multiple login failures could result from a forgotten password. An execution of *cmd.exe* could be part of a malicious attempt or a legitimate one, as it is frequently used by malicious programs to execute commands while it is also frequently used by legitimate users during their normal day-to-day operations. The question is how to estimate $p(e_j/H_m)$ if e_j becomes such an observation (true positives)? One possible answer would be using IDS evaluation datasets such as ISCX 2012 [32] or DARPA as corpuses and using similar techniques used in the natural language processing domain. Chivers *et al*. claim that, in some cases, the historical rate of occurrences of certain attacks is known and can be used to estimate the likelihood that certain events derive from such attacks or it may be sufficient to quantify

these frequencies, in a similar way to estimating risk likelihoods, to an accuracy of an order of magnitude [9].

5. RESULTS AND ANALYSIS

In this section, experimental results are presented along with the analysis.

A. Identifying Suspicious Nodes

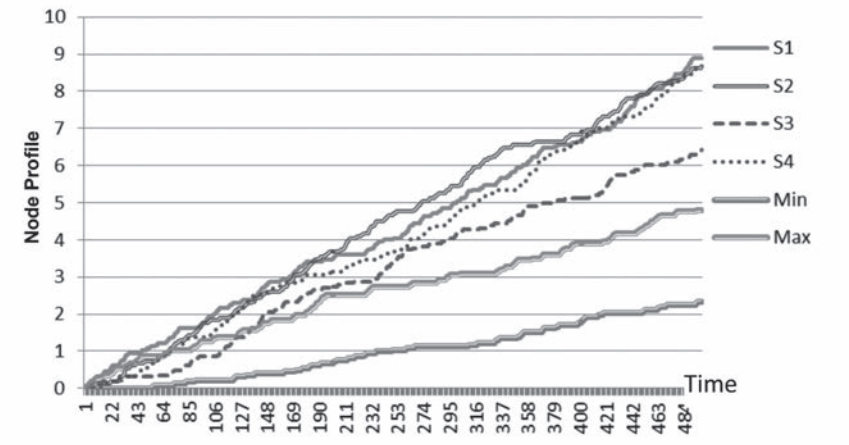
The proposed approach was tested against 25 (=5x5) test cases, varying subnet sizes {25,50,100,250,500} and number of attackers {1,2,4,7,10}, and it was observed that the proposed approach detected slow attackers well in all 25 cases. Due to space constraint only one test case, 100 size subnet with four attackers, is listed here.

Four low rate attackers were located in a 100 size subnet of network B. All clients generated innocent events (events such as forgotten password etc.) while four attackers generated low rate attack (reconnaissance) events. At each time point, node profiles were calculated for all 100 nodes in the subnet and converted to Z-scores. Node profiles and Z- scores were plotted as in Figures 2 and 3 respectively.

1) Maximum Score approach

As mentioned in Section 2, selecting suspicious nodes by looking at raw node profiles is problematic when there is more than one suspicious node. Although all suspicious nodes are above the *Max* line (after some time), setting this *Max* is problematic in real world implementations.

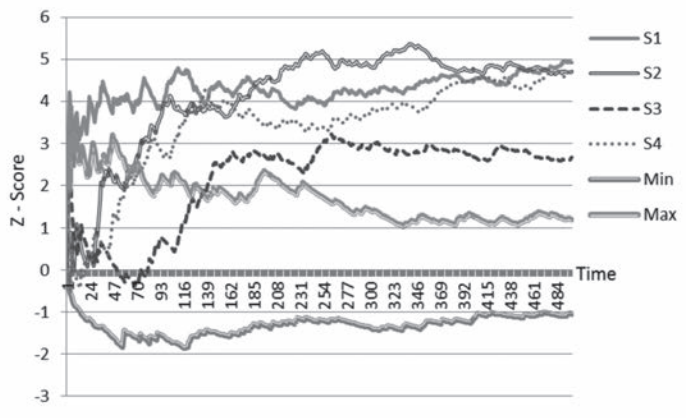
FIGURE 2. CUMULATIVE PROBABILITIES (NODE PROFILES), S1,S2,S3,S4 DENOTE ATTACKERS. MIN AND MAX REPRESENT THE MINIMUM AND MAXIMUM Z-SCORES OF NORMAL NODES AT EACH TIME POINT.



2) Z-Score approach

Attackers are always above or near around three standard deviations from the mean, and most importantly, there is a clear visual separation between a set of normal nodes and anomaly nodes. Graphs become more stable by the time (i.e. assuming stationary status), which means the proposed decision criteria are better for distinguishing anomalous profiles from normal profiles than the 'Maximum score approach'.

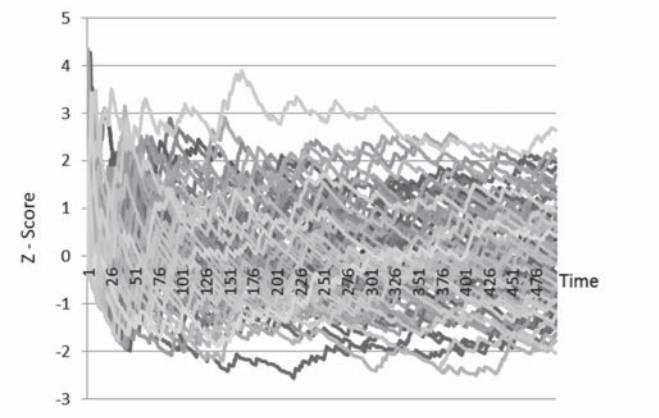
FIGURE 3. Z-SCORES OF NODE PROFILES. S1,S2,S3,S4 REPRESENT SUSPICIOUS NODES. MIN AND MAX REPRESENT THE MINIMUM AND MAXIMUM Z-SCORES OF NORMAL NODES AT EACH TIME POINT.



3) Best and worst cases

To investigate how the proposed approach works with best and worst cases, the above experiment was repeated twice, first without any attackers and then with all subverted nodes, and obtained the similar graphs as in Figure 4 in both cases. Most of the nodes are nearly between three standard deviations from the mean, and none of the nodes can be seen clearly separated from the majority. However this would not be a problem. If an analyst sees a similar graph, it would be safe to assume that all nodes are subverted (instead of assuming they are free of attackers) and to do further investigations on one or two nodes to verify. If investigated nodes are attackers it is reasonable to consider that all nodes are attackers or vice versa.

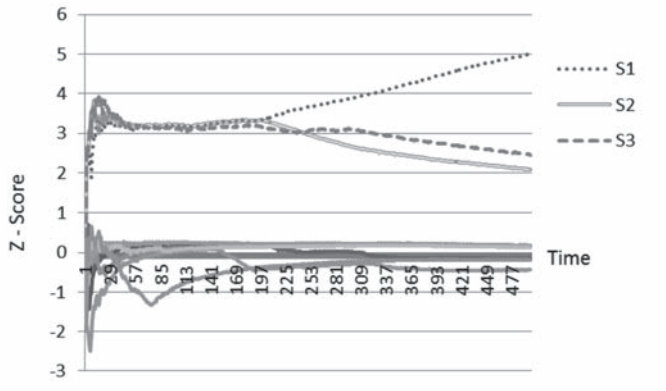
FIGURE 4. Z-SCORES OF NODE PROFILES, NO ATTACKERS, 100 SIZED SUBNET.



4) Node behaviour

To investigate how proposed Z-score graphs reflect the behaviour of nodes (identities), three attacker nodes were located in a 50 size subnet in network D. All others were innocents. Two out of three attackers stopped their reconnaissance attempts at 200 and 300 times respectively. As shown in Figure 5, when an attacker node changes its behaviour the relevant Z-score graph responds to that behaviour by changing its direction.

FIGURE 5. Z- SCORE GRAPHS ARE SENSITIVE TO NODE BEHAVIOUR. S1,S2,S3 ARE SUSPICIOUS NODES. ALL OTHERS ARE INNOCENTS.



B. Attacker Localisation

In a situation, there are multiple suspected sites to be investigated (e.g. different actors, subnets, LANs, locations etc) and determining the centres of attention would be problematic. Localisation of attackers' identities as much as possible, at least for an intermediary level, or choosing the smallest subset in which an attacker may be located, would greatly save the cost and time to be spent on investigations. To investigate the capability of the proposed approach herein: one attacker was placed in a subnet of network C. Scores were assigned (profiling) the Gateways of each subnet, using the formula:

$$\textit{Gateway score} = \frac{\textit{Cumulative Score}}{\textit{Number of nodes in the subnet}}$$

assuming each reconnaissance event can be reverse engineering only up to the gateways. They were converted to the Z-scores and Figure 6 was obtained. GA, GB, GC and GD are gateways of networks A, B, C and D respectively.

FIGURE 6. Z – SCORES OF GATEWAY SCORE OF EACH NETWOK.

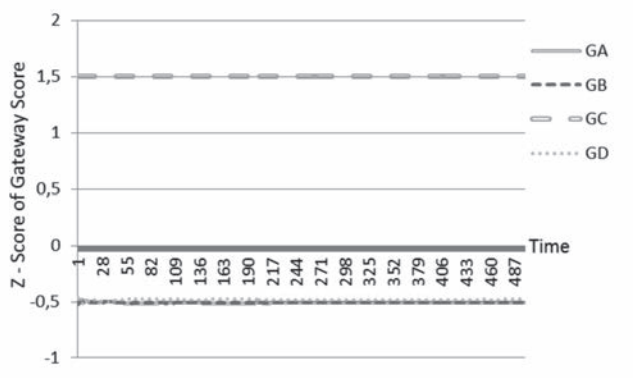


Figure 6 proves the proposed approach useful in attacker localisation.

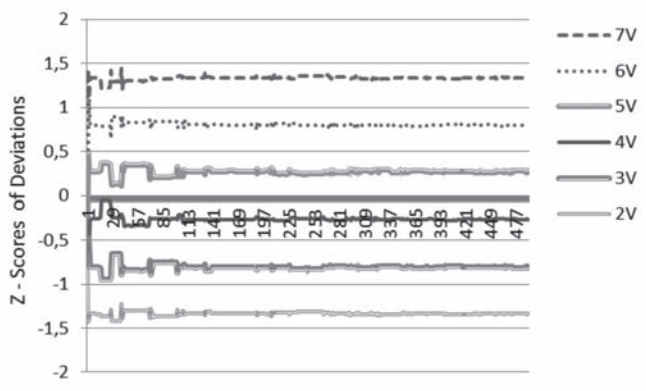
C. Network Parameters

In this section we investigate how different network parameters affect the attribution of slow activities.

1) Traffic Volume

An attacker was located in a 51 size subnet of Network C and generated events. The same experiment was repeated six times, keeping all parameters unchanged except the attacker's traffic volume. If the attacker's traffic volume is V the first time, then at each repetition the attacker's traffic volume was incremented by one time as *2V, 3V, ...,7V*. For each experimental run, the deviation of attacker's profile value from the average of normal (statistical norm) was calculated. Then the standardised deviations (z-scores of deviations) are plotted as in Figure 7. The graph tells us: 'the higher the traffic volume generated by attacker, the easier his detection will be.'

FIGURE 7. Z-SCORES OF DEVIATIONS OF CUMULATIVE NODE SCORES.



2) Cluster Size

To investigate how the identities' cluster size (here subnet size) affects detection, an attacker was located in a 500 size subnet and the same experiment was repeated six times by keeping all other parameters, except the subnet size, unchanged. Subnet size was changed to 400, 300, 200, 100, 50 and 25 at each experimental run and the graphs in Figures 8, 9 and 10 were obtained. Figure 8 and 9 say 'attackers have less chance to hide behind innocent events, when the cluster size decreases.' It is further reinforced by Figure 10 saying 'the smaller the cluster size, the better for detection of suspicious slow activities' in terms of security. But, in practice, it should be noted that partitioning a network into very small subnets would not be a feasible solution sometimes, as it depends on several other factors such as resources availability and user requirements. Figure 10 also suggests that 'going beyond 100 size cluster would not make any real sense in terms of detection.'

FIGURE 8. PERCENTAGES (%) OF SUSPICIOUS EVENTS GENERATED BY ALL INNOCENTS.

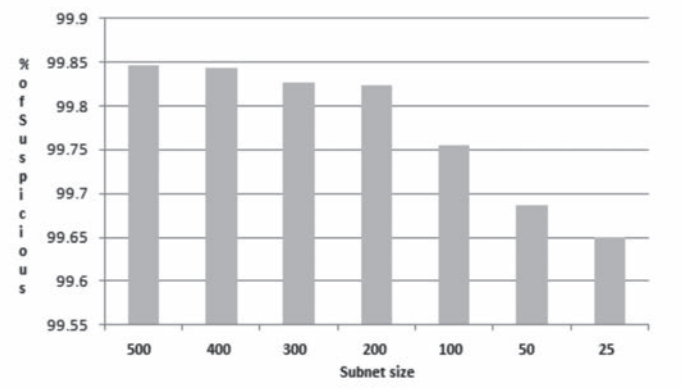
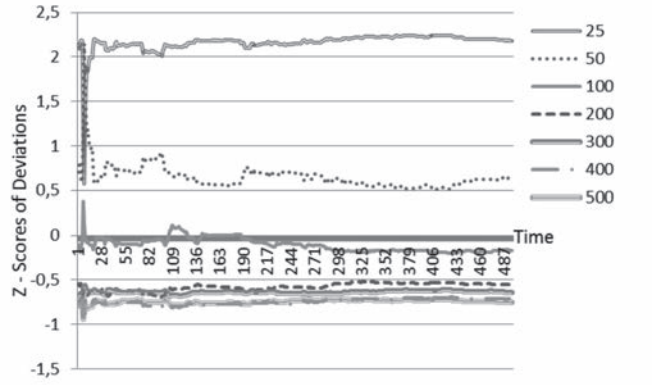
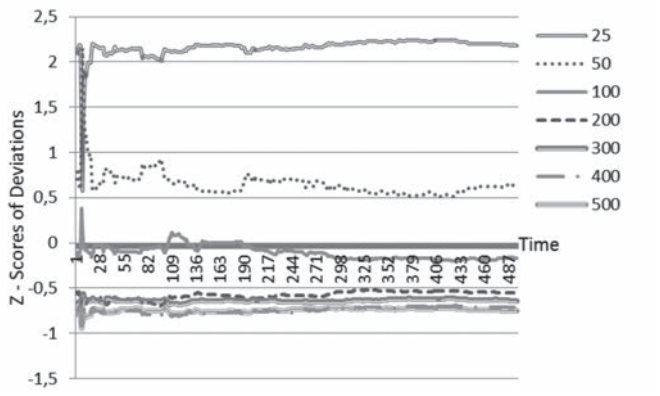


FIGURE 9. PERCENTAGES (%) OF SUSPICIOUS EVENTS GENERATED BY ATTACKER.



The authors would like to reiterate that a subnet equals a cluster of identities. For example, in a case of cold cyber war or in an attack like the well-known Georgia 2008 case, a cluster can be a country or a region of a suspected country and identity can be any physical or virtual location within that country or region.

FIGURE 10. Z – SCORES OF ATTACKER'S DEVIATIONS FROM THE AVERAGE.



3) Number of Attackers

Keeping all conditions unchanged, except number of attackers, the same experiment was repeated twice, first with two attackers and then with seven attackers. The attacker's node score (see Figures 11 and 12) is dependent on 'the number of attackers in his own subnet' (compare attackers' Z-scores). This rationalises the usage of 'Statistical normality' as the decision criteria and suggests defining 'one's abnormality' relative to his peers (i.e within the same domain, department, similar user group, region, country etc.) would give better results (in terms of lower false alarms) than defining it universally. Comparison of nodes profiles (as in Figure 2) regardless of their subnets would give higher false alarms.

FIGURE 11. Z-SCORE GRAPHS FOR SAME SIZE SUBNETS WITH DIFFERENT NUMBER OF ATTACKERS (250 SIZE SUBNET, TWO ATTACKERS)

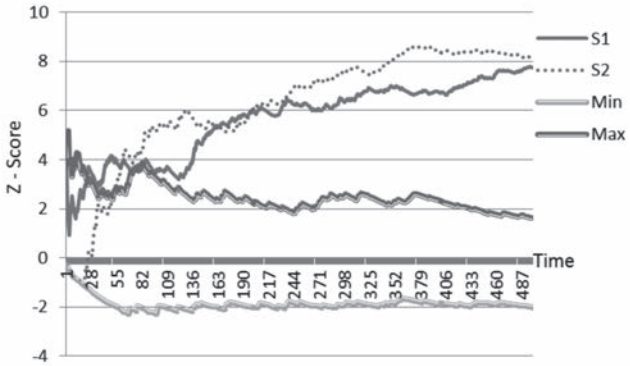
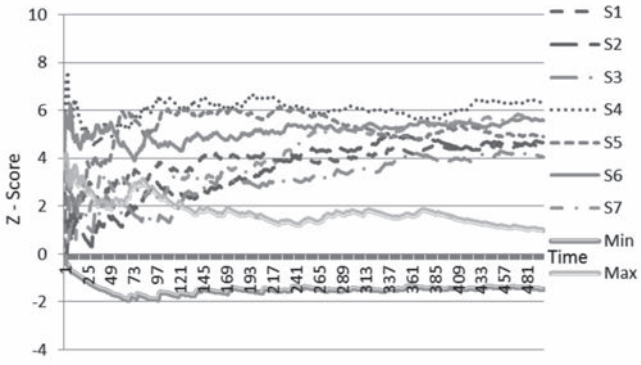


FIGURE 12. Z-SCORE GRAPHS FOR SAME SIZE SUBNETS WITH DIFFERENT NUMBER OF ATTACKERS (250 SIZE SUBNET, SEVEN ATTACKERS).



6. SAMPLING TECHNIQUES

Many IDSs such as Snort facilitate for logging data in a variety of ways for later analysis, as it is an essential part of any intrusion detection activity. If you are not looking at the logs and monitoring the alerts, then effort invested into an IDS can quickly become meaningless [27]. In a slow attack environment, logging is crucial as you cannot log everything during longer times. The large size/unmanageable nature of the target population is one of the main reasons for sampling instead of doing a census. As it is almost similar to the problem the analyst faces herein, the simple random sampling technique was used to investigate the usability of sampling for data logging in slow-attack environments.

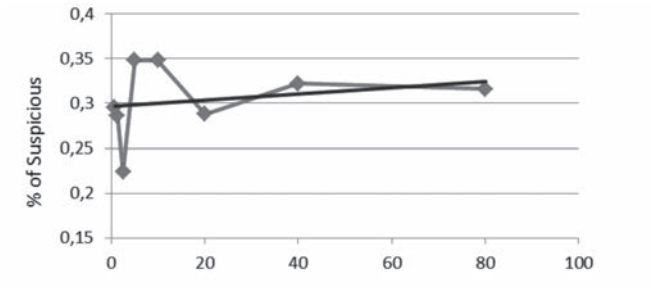
An attacker was located in a subnet of network C and 'stateless' attacks events were generated. The simulation was allowed to run 1440 time units. The whole period was divided into twelve blocks, and within each block, a sample was collected using an R [31] script. Finally, all twelve samples were combined together to make one final sample. The same experiment was repeated with different sample sizes in order to identify how sample sizes affect 'detection potential.' Table 1 and Graphs in Figures 13, 14, 15 and 16 show the experimental outcomes. We varied the sample sizes from 80% to 0.625% (see Table 1), always half of the previous size.

TABLE I. SAMPLING STATISTICS.

Sample Size as a % of population/whole observation)	80	40	20	10	5	2.5	1.25	0.625
Number of Attack Events selected	826	420	188	113	56	18	12	6
Number of Innocent Events selected	260244	130235	65200	32356	16043	8029	4188	2026
Percentage (%) of Attack Events	0.32	0.32	0.29	0.35	0.35	0.22	0.29	0.30

Although the fitted trend line in Figure 13 shows a very small positive trend between percentage of suspicious events and sample size, the real figures in the table explain that it would not be significant. Interestingly, 'in each sample, the percentage of suspicious events generated by the attacker is almost same as it is in the population (0.3)' is a good indicator that selected samples represent the intended population's characteristics, regardless of its size. Analyst may choose sampling techniques for long-term networking monitoring (it could not be for detection, but may be for other purpose of traffic analysis), deciding the sample size based on the resources availability and the intended purpose.

FIGURE 13. PERCENTAGE OF SUSPICIOUS EVENTS GENERATED BY ATTACKER.



Graphs in Figure 14, 15, 16 show that the analyst can enjoy the population characteristics (in terms of this analysis) even if the size of the sample is 5% of the entire data capture. This would be a good indicator, why?, if an analyst can reduce his focus by 95% it will reduce the time and cost too. However when the sample size is smaller than 2.5% of its population size, anomaly-based detection methods cannot be used. But the table explains that signature based detection methods can still be used, as it contains very few attackers' signatures. Generally using 10% size sample would be an ideal for detecting suspicious slow activities, whether it is based on anomaly or signature-based detection methods. However the authors do not generalise the optimal sample size as 10%. It could be highly subjective and varied according to the intended analysis. Further experiments are needed on this topic. At least at this stage, the authors have shown that some population characteristics remain unchanged in samples and, hence there is a possibility to use sampling techniques in this domain.

FIGURE 14. Z-SCORES, WHEN THE SAMPLE SIZE IS 10% OF WHOLE TRACE. S REPRESENTS THE SUSPICIOUS NODE. MIN AND MAX REPRESENT THE MINIMUM AND MAXIMUM Z-SCORES OF NORMAL NODES AT EACH TIME POINT.

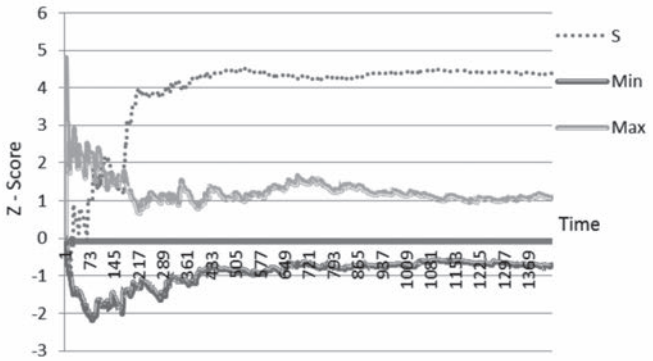


FIGURE 15. Z-SCORES, WHEN THE SAMPLE SIZE IS 5% OF WHOLE TRACE. S REPRESENTS THE SUSPICIOUS NODE. MIN AND MAX REPRESENT THE MINIMUM AND MAXIMUM Z-SCORES OF NORMAL NODES AT EACH TIME POINT.

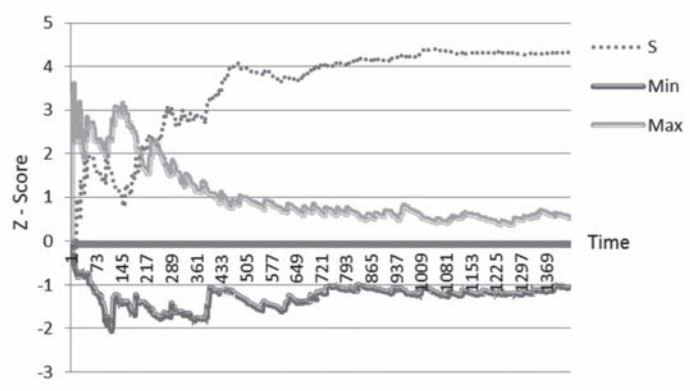
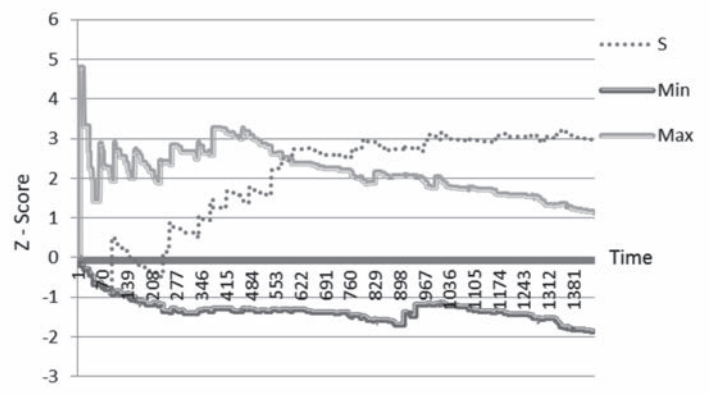


FIGURE 16. Z-SCORES, WHEN THE SAMPLE SIZE IS 2.5% OF WHOLE TRACE. S REPRESENTS THE SUSPICIOUS NODE. MIN AND MAX REPRESENT THE MINIMUM AND MAXIMUM Z-SCORES OF NORMAL NODES AT EACH TIME POINT.

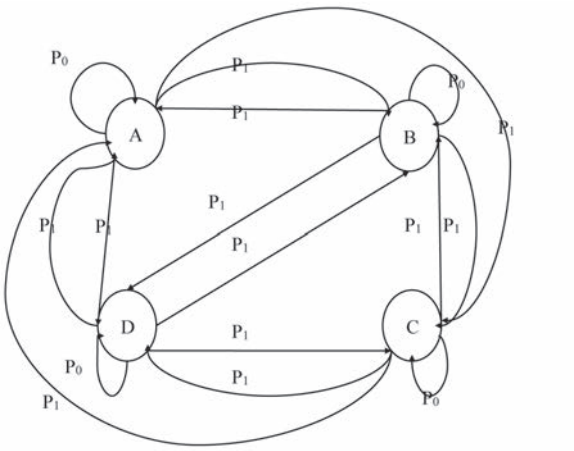


7. DISCUSSION

An efficient method for cyber conflict attribution (particularly slow activities) and an investigation of its effectiveness under different conditions have been provided. Breaking down the attribution problem into two sub-problems reduces the complexity of the problem, and explores ways to investigate alternative methods. The proposed approach is domain agnostic. It can be easily adjusted to use in many aspects of cyber warfare and help in actor intelligence: profiling adversarial technical capabilities; creating linkage between actor groups; tracking the supply chain; and differentiating between actors (e.g. state-sponsored or criminal) etc. It can be used for profiling any kind of actors, not only in the cyber domain but also in other domains such as crime and juridical sciences. Experimental outcomes and recommendations presented in Sections 5 and 6 provide tactical and operational principles for systematic and

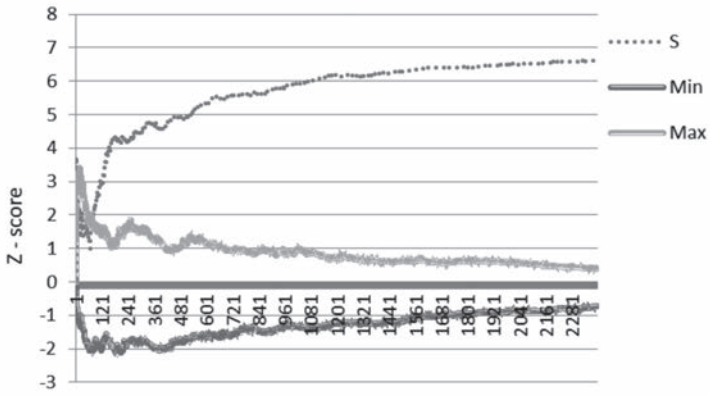
efficient profiling and attribution. They are particularly useful in the capacity planning stage of a network design process. Findings of how cluster size affects detection can be incorporated with existing clustering based analysis approaches [30,14]. In future, identifying the best performance method (among alternative methods such as using sensor fusion algorithms) and handling some miscellaneous issues, such as overcoming situations when the source of the event is unknown, will be addressed. Based on the idea derived from Section 6, an experiment was set up to investigate the possibility of using mobile sensors to slow activity detection. An attacker was located in network D. A Finite state automaton (see Figure 17) was used to control the sensor mobility (transitions). At any given state, the sensor spends a constant time interval for monitoring. Scores were updated only when the sensor had visibility to the target subnet.

FIGURE 17. FINITE STATE AUTOMATA USED FOR SENSOR MOBILITY, P_0=0 AND P_1=0.33.



As Figure 18 shows, it can identify the attacker, even using a mobile sensor. This could be mainly due to the cumulative nature of the proposed approach and the usage of automaton. It should be noted that the transition probabilities *(P_0, P_1)* of the above automaton can be estimated dynamically, based on evidence at the scene, in order to improve the quality of the detection, which is also left for future work.

FIGURE 18. Z-SCORE GRAPH FOR SENSOR MOBILITY. S REPRESENTS THE SUSPICIOUS NODE. MIN AND MAX REPRESENT THE MINIMUM AND MAXIMUM Z-SCORES OF NORMAL NODES AT EACH TIME POINT.



REFERENCES:

[1] S.W. Beidleman, 'Defining and Deterring Cyber War,' Msc Thesis, Dept. Military Strategy Planning and Operations, U.S. Army War College, Carlisle, Pa, 2009.

[2] D.Morrill. (2006, August 07). Cyber Conflict Attribution and the Law [Online]. Available: http://it.toolbox.com/blogs/managing-infosec/cyber-conflict-attribution-and-the-law-10949

[3] D.A.Wheeler and G.N.Larsen. (2003, October 30). Techniques for Cyber Attack Attribution. Inst. for Defense Analyses [Online]. Available: http://www.dtic.mil

[4] T.Parker. (2010). Finger Pointing for Fun, Profit and War? The importance of a technical attribution capability in an interconnected world [Online]. Available FTP: media.blackhat.com Directory: /bh-dc-11/Parker File: BlackHat_DC_2011_Parker_Finger_Pointing-wp.pdf

[5] S.Charney. (2009). Rethinking the Cyber Threat: A Framework and Path Forward [Online]. Available FTP: download.microsoft.com Directory: download File: rethinking-cyber-threat.pdf

[6] K.Saalbach. (2011). Cyberwar Methods and Practice [Online]. Available FTP: dirk-koentopp.com Directory: download File: saalbach-cyberwar-methods-and-practice.pdf

[7] N.Villeneuve and D.Sancho. (2011, September 26). The Lurid Downloader [Online] Available: http://www.trendmicro.com

[8] H.Chivers et al., 'Accumulating evidence of insider attacks,' in The 1st International Workshop on Managing Insider Security Threats 2009 (In conjunction with IFIPTM 2009) CEUR Workshop Proc., 2009, pp.34-51.

[9] H.Chivers et al., 'Knowing who to watch: Identifying attackers whose actions are hidden within false alarms and background noise,' Inform. Syst. Frontiers, Springer, 2010, doi: 10.1007/s10796-010-9268-7.

[10] A.Patcha and J.M.Park, 'An overview of anomaly detection techniques: Existing solutions and latest technological trends,' Elsevier Computer Networks, Vol. 51 Issue 12, pp. 3448–3470, August 2007.

[11] S.Kumar and E.H.Spafford, 'An application of pattern matching in intrusion detection,' The COAST Project, Dept. Comp. Sci, Purdue University, Tech. Rep, 1994.

[12] V.Chandola et al., 'Anomaly detection: A survey,' ACM Computing Surveys (CSUR), Vol. 41 Issue 3, July 2009, doi:10.1145/1541880.1541882.

[13] M.H.Bhuyan et al., 'Survey on Incremental Approaches for Network Anomaly Detection,' IJCNIS Vol. 3, December 2011, pp 226-239.

[14] C. Zhong and N. Li, 'Incremental Clustering Algorithm for Intrusion Detection Using Clonal Selection,' in Proc. PACIIA (1), 2008, pp.326-331.

[15] F.Ren et al., 'Using density-based incremental clustering for anomaly detection,' In: CSSE '08, DC, USA, IEEE Comput. Soc., 2008, pp. 986-989.

[16] R.Bejtlich, The Tao of Network Security Monitoring: Beyond Intrusion Detection, Addison-Wesley, 2005.
[17] A.J. Beecroft, 'Passive Fingerprinting of Comput. Network Reconnaissance Tools,' MSc Thesis, Naval Postgraduate School, Monterey, California, 2009.
[18] R.C.Brackney and R.H.Anderson, 'Understanding the insider threat,' Proc. March 2004 Workshop, RAND Nat. Security Research Division,Tech. Rep., 2004.
[19] W.Y.Yu and H.M.Lee, 'An incremental-learning method for supervised anomaly detection by cascading service classifier and its decision tree methods,' in PAISI '09 Proc. Pacific Asia Workshop on Intell. and Security Informatics, 2009@Springer Berlin /Verlag, doi:10.1007/978-3-642-01393-5_17.
[20] P.Laskov et al., 'Incremental support vector learning: Anal., implementation and applications,' J. of Machine Learning Research Vol.7, October 2006, pp. 1909-1936.
[21] T.Heberlein, 'Tactical operations and strategic intelligence: Sensor purpose and placement,' Net Squared Inc, Tech. Rep. TR-2002-04.02, 2002.
[22] W.W.Streilein et al., 'Improved detection of low-profile probe and novel denial-of-service attacks,' Int. Workshop on Statistical and Machine Learning Techniques in Comput. Intrusion Detection, 2002.
[23] R.Basu et al., 'Detecting low-profile probes and novel denial-of-service attacks,' IEEE SMC IAS Workshop, West Point, New York, USA, Tech. Rep., 2001.
[24] M.Kandias et al., 'Dimitris gritzalis: An insider threat prediction model,' in Trust, Privacy and Security in Digital Business, 2010@Springer Berlin/ Heidelberg, doi: 10.1007/978-3-642-15152-1_3.
[25] E.E.Schultz and R.Shumway, Incident response: A strategic guide for system and network security breaches, Indianapolis: New Riders, 2001.
[26] The NS3 discrete-event network simulator [Online]. Available: http://www.nsnam.org/
[27] J.Babbin et al., Snort Cookbook [Online]. Available: http://commons.oreilly.com/wiki/index.php/Snort_Cookbook
[28] G.Jiang and G.Cybenko, 'Temporal and spatial distributed event correlation for network security', Proc. Amer. Control Conference, Boston, MA, 2004, pp. 996-1001.
[29] P.G.Bradford et al., 'Towards proactive computer system forensics', Int. Conference on Information Technology: Coding and Computing, IEEE Comput. Soc., 2004, pp.648-653.
[30] M.C. Libicki (2009).Cyberdeterrence and Cyberwar. Santa Monica, CA: RAND Corporation [Online]. Available: http://www.rand.org/pubs/monographs/MG877
[31] The R project for statistical computing: R foundation for statistical computing [Online]. Available: http://www.r-project.org
[32] A. Shiravi et al.. (2012). ISCX Intrusion Detection Evaluation DataSet [Online]. Available: http://www.iscx.ca/datasets