Effective network security monitoring: from attribution to target-centric monitoring

Shaikh, S.A. and Kalutarage, H.

Noname manuscript No.
(will be inserted by the editor)

Effective Network Security Monitoring: From Attribution to Target-centric Monitoring

Siraj Ahmed Shaikh · Harsha Kumara Kalutarage

Received: date / Accepted: date

Abstract Network security monitoring remains a challenge. As global networks scale up, in terms of traffic, volume and speed, effective attribution of cyber attacks is increasingly difficult. The problem is compounded by a combination of other factors, including the architecture of the Internet, multi-stage attacks and increasing volumes of nonproductive traffic. This paper proposes to shift the focus of security monitoring from the source to the target. Simply put, resources devoted to detection and attribution should be redeployed to efficiently monitor for targeting and prevention of attacks. The effort of detection should aim to determine whether a node is under attack, and if so, effectively prevent the attack. This paper contributes by systematically reviewing the structural, operational and legal reasons underlying this argument, and presents empirical evidence to support a shift away from attribution to favour of a target-centric monitoring approach. A carefully deployed set of experiments are presented and a detailed analysis of the results is achieved.

Keywords Communication Networks · Network Security · Attribution · Bayesian Statistics · Scalable Monitoring

Siraj Ahmed Shaikh
Faculty of Engineering and Computing, Coventry University, Coventry, CV1 5FB, UK.
Tel.: +44-7939-233995
E-mail: s.shaikh@coventry.ac.uk

Harsha Kumara Kalutarage
Faculty of Engineering and Computing, Coventry University, Coventry, CV1 5FB, UK.

1 Introduction

Network security monitoring poses a huge challenge. Cyber attacks are increasingly significant in terms of their disruptive and destructive impact; recent attacks [4, 53] serve as a reminder. As global networks scale up, both in terms of traffic volume and speed, the problem of detection of such activity is only going to get worse. What is it about modern networks that makes security monitoring such a challenge?

In particular, attribution is very difficult to achieve. This is demonstrated over and again, firstly, with the attack on Estonia in 2007, often cited as a first real instance of a cyber attack, with significant impact on the countrys critical infrastructure [34]. While this is attributed to non-state actors (with possible state support) from Russia [34], the fact remains that there was no official attribution to a state sponsor.

Stuxnet is another instance of a sophisticated cyber weapon purposefully-launched to target critical nuclear infrastructure in Iran (ultimately affecting other systems beyond just Iran). A careful examination of the malware only serves to prove that significant resources have been deployed to launch it successfully [16]; we have to acknowledge the current environment where several state and non-state actors have emerged to bear such capability [38]. Attribution to Israel and USA is of little comfort therefore [36]. Current policy discourse on cyber attribution is therefore heavily influenced by geopolitics. This fails to serve the agenda on effective network security monitoring as it ignores a complex threat from a variety of emerging state and non-state actors, amongst them insiders, who pose a real threat.

Finally, the recent attacks on Sony and the doubts whether North Korea was behind the attacks [47] makes a compelling case for the agenda of attribution to be in-

formed by technological challenges that exist. The purpose of this paper is to bring to light a variety of technological considerations that need to inform policy on attribution. We move beyond the current discourse and propose that we shift our focus on target-centric monitoring, which could offer a better potential for detection and could serve as an early warning system. This paper presents theoretical results to support this claim.

1.1 Rest of the paper

The rest of this paper is organised as follows. Section 2 motivates the reader and makes a constructive case for reconsidering attribution in network security monitoring. Section 3 reviews related work in this area and systematically points out other research that looks at network security monitoring but face similar challenges of achieving effective attribution. Section 4 describes the monitoring algorithm we use to demonstrate how shifting security monitoring from a focus on attribution to target-centric monitoring could be more effective. Section 5 presents the results of our empirical work and a detailed analysis for the benefit of the reader. Section 6 presents a discussion and further reflection, and concludes the paper.

2 Motivation

We motivate the problem on a variety of fronts including structural, operational and legal issues.

The technical architecture of cyberspace has its roots in connectivity and not accountability. Clark and Landau, experts in this area, have noted the problem of attribution and carefully examined solutions to tie together personally identifiable information with packet layer attribution, which they appropriately call the extreme of the accountable Internet [19]. The current architecture of the Internet does not permit such a provision, and any attempt to conceive of one, even at a technical level (by manipulating the IP layer), is not straightforward.

Moreover, multi-stage attacks, which most modern cyber attacks are, make it impossible to do any reliable attribution (for accountability and identification) [19]. Such attacks are realised when an attacker manages to use a different machine to launch an attack on the final target. There are multiple stages involved: the attacker would first compromise an intermediary machine and set it up to attack the final target; there may be several such machines, with each being used to compromise another. Once a complex web of anonymous mechanisms is set up, the attacker can then use these machines as a launchpad for the final attack; in some instances this activity carries on to allow for data exfiltration (illegal transfer out of data) to go on over a period of time, for example.

The ultimate aim here is to avoid any attribution back to the original attacker. There are secondary aims of trying to rally up as much computing power as possible to launch attacks using a higher communications bandwidth for maximum effect. Such stepping stones would ideally be found in foreign countries where foreign legal jurisdiction makes it even more difficult to carry out any post-incident response. Such is the appeal of this approach that several compromised machines are already controlled, commonly known as botnets, by botnet operators who lease out these machines in what has become an established trade in the cybercrime underground economy. Needless to say, effective attribution through such a clandestine infrastructure is near impossible.

From an operational perspective, an examination of Internet traffic characteristics reveals an increasing volume of "non-productive traffic" [39], which is essentially due to a variety of benign and malicious reasons, achieving no purpose. Non-productive traffic takes form in a variety of ways including

- continually growing scanning activity on public networks, partly due to search engines (like Google) collecting and indexing content for efficient search results (using technologies like Googlebot) [37]. The difficulty here arises from differentiating between legitimate scans to malicious attempts;
- backscatter traffic, which is essentially response traffic from other scanning and attack activity ongoing in cyberspace [39]; and
- a plethora of other network packet floating around due to misconfigured hosts and administrative errors.

Most of such Internet traffic is essentially defunct by the time it is visible on security sensors, but may still resemble genuine malicious attempts. Important to note here however is the difficulty this introduces for detecting purposeful malicious activity (targeted and deliberately designed attacks). This problem particularly manifests itself in the form of "high false positives", essentially mistakes made by security monitoring infrastructure failing to distinguish real attacks from benign activity. The ever increasing volume of such non-productive traffic, porportionally to rising Internet traffic, makes the problem of effective attribution only worse.

Beyond the technical, the increasing nature of state-sponsorship of large-scale cyber attacks means that nation states have to identify hostile states initiating the

attack. The only problem is that launch pads for most such attacks are found to be in non-hostile states [3], against which the victim state could only respond by applying the *unwilling or unable* test [22]. This is an underlying principle of international law which asserts that retaliation against an intermediary state used by an enemy to launch an attack is only permissible if the intermediary is either unwilling or unable to prevent the aggressor responsible from doing so.

Perhaps the greatest difficulty posed by any retaliatory cyber-attack is the geopolitics of the day. Political alliances, intelligence sharing, legal and ethical considerations, and potential sensitivity of offensive operations, all make it very difficult for nation states to launch such operations. The result is that the sort of public accusations of cyber attacks seen in the press and meant as a tool of deterrence are almost entirely useless. Essentially, the value of attribution to provide deterrence is increasingly futile.

On reflection one finds an opportunity to reconsider network security monitoring efforts. Perhaps then a shift of focus from security monitoring of the source to the target could provide for better network defence? Simply put, resources devoted to detection and attribution could be redeployed to efficiently monitor for targeting and prevention of attacks. Detection should aim to determine whether a node is under attack, and if so, effectively prevent the attack. This is a radical change whereby the cost of monitoring [43, 44] could be dramatically lowered and malicious activity is curbed at a much earlier stage in the attack cycle.

3 Related work

To demonstrate our idea of target centric monitoring we use a Bayesian-based traffic monitoring approach. Using Bayesian technique and its variants for intrusions detection can be found in [17, 18, 20, 45, 52]. The relevance of information fusion for network security monitoring can be found in [12, 15, 40, 48]. A scalable solution to identify insiders in a Bayesian framework is proposed in [17, 18]. A base line is defined and if a user profile is deviated from it an alarm is raised [8, 32]. Basu et al [2] uses connection based windows to detect low profile attacks with a confidence measure. Using multiple neural network classifiers to detect stealthy probes can be found in [46]. Evidence accumulation as a means of detection is proposed in [26]. Brynielsson et al [11] apply the same idea in a different domain (detecting lone wolf terrorists). Berk et al [5] combines traditional notion of Motive, Means, and Opportunity with behavioural analysis techniques to place each individual on a sliding scale of insider risk. Users' behaviour is compared both to their own baselines and to the behaviours of members in their peer groups, using the Euclidean distance. Eldardiry et al [23] proposes a method for detecting insiders with unusual changes in behaviour by providing a method to combine anomaly indicators from multiple sources of information for building a global model. Authors find outliers in that global model by comparing each user's activity changes to activity changes of his peer group. Axelrad et al [1] defines a Bayesian network model that incorporates psychological variables that indicate degree of interest in a potential malicious insider. Using a complex Bayesian networks to capture conditional dependencies between different attributes can be found in [20]. Chivers et al [18] demonstrate Bayesian approach is superior to the counting algorithm.

With respect to the Target-centric monitoring there is no established literature. Whyte et al [49] offer a different direction for security monitoring by proposing a class of scanning detection algorithms that focus on what is being scanned for instead of who is scanning. But such an approach is not completely independent from the source information either. It uses the source information of scan packets for victim detection. Our approach does not require any information about the source. It completely depends on destination information and allows for any suspicious event on the network to be accounted for. Most importantly, we acknowledge two types of uncertainties (*motivation* and *source*) of events as defined in [30, 31] in a Bayesian framework. Hence though we inspired by [49] our effort is different.

The motivation behind choosing a probabilistic approach for profiling is that a network event is not always easy to judge for malicious nature. Some suspicious events can appear as part of an attack signature as well as originate from benign network activity. For example, a major router failure could generate a flood of ICMP unreachable messages while some malicious program (viruses and worms) may generate the same for probing; such uncertainty needs to be acknowledged [27, 28, 29, 31]. We use a Bayesian technique to achieve this. Figure 1 is to demonstrate the need for acknowledging the event uncertainty in monitoring. In Figure 1, left graph obtained via the proposed approach in this paper, and using the same trace, right graph was obtained via the simple counting method. This type of counting method is used in heuristics developed in [49].

Proposed method is an anomaly detection based approach. In recent years numerous anomaly-based intrusion detection approaches have been proposed, but most of them are generic and simple [13, 33, 41]. Most of current incremental anomaly detection approaches have either high rate of false alarms, or suffer scalabil-

ity issues, or are not fit for deployment in high-speed networks (refer to survey paper [6]). Hence they are different from our approach which is minimal in terms of monitoring overhead and hence has the potential to scale up very well for large network topologies.

Importantly, all above approaches for network monitoring are source centric or use source information at some stage of the detection. But proposed approach in this paper is completely independent from the source information in monitoring. Hence this work is unique.

4 Bayesian-based traffic monitoring

The underlying idea of this paper is that network security monitoring should move from trying to attribute activity to suspicious sources to the targets of critical assets in a network system. To demonstrate this, we use an existing monitoring algorithm of same authors. In this section we provide substantial information about the monitoring algorithm, interesting readers are invited to refer [27, 28, 29, 31] for more information.

4.1 Monitoring Algorithm

Chivers et al [17, 18] use Bayes' formula for combining evidence from multiple sources to identify the source of suspicious activities. This is further developed by Kalutarage [27] to standardise profile scores using Z-Scores along with the concept of statistical normality. The problem of detection is broken down to two sub problems: *profiling* and *analysis*.

Profiling is the method for evidence fusion across space and time by updating node profiles dynamically based on changes in evidence. Simply put, we compute a suspicion score using the hypothesis given below for each node in the system during a smaller time window w, and that score is updated as time progresses to compute a node score for a lager observation window W.

4.2 Building the Hypothesis

Let $E = \{E_1 = \mathrm{e}_1,\ E_2 = \mathrm{e}_2,\ E_3 = \mathrm{e}_3, \ldots, E_m = \mathrm{e}_m\}$ be the set of all suspicious evidence observed against node k during time t from m different independent observation spaces. H_k is the hypothesis that k^{th} node being a victim of attacker(s). The node score is then calculated as follows.

$$p(H_k/E) = \frac{\prod_j p(e_j/H_k).p(H_k)}{\sum_i \prod_j p(e_j/H_i).p(H_i)} \quad (1)$$

Once the likelihood $p(e_j/H_i)$ and the prior $p(H_i)$ are known $p(H_k/E)$ can be calculated. Note that the profiling technique we use in this work can combines information gathered from different sources into a single score for a minimum computational cost. It reduces data into a single value which is important to maintain information about node activities for a W. The posterior probability terms ($p(H_k/E)$) can be accumulated by time and used as a metric to distinguish victims from other nodes. Node scores are updated at the end of each w by considering all the evidence observed during that period. Extending our approach to a very large scale attack surface is very simple as it is a matter of adding a new indicator (attack vector) in E. Existing domain knowledge will serve to enhance the performance of our monitoring algorithm since it takes advantage of prior knowledge about the parameters. Which is especially useful when technical data is scarce. However prior and likelihoods are the most critical parameters to our approach since Bayes factors are sensitive to them.

The analysis comprised of detecting anomalous profiles in a given set of node profiles. A statistical method is used to detect anomalies. Two techniques are used: *Peer analysis* and *Discord analysis*. Both techniques acknowledge the fact that baseline behaviour on networks is not necessarily stable. For example, operational or exercise deployments often mean the behaviour of nodes will potentially change dramatically.

4.3 Peer analysis

Aggregating short period (w) estimations over time helps to accumulate relatively weak evidence for long periods W. These accumulated probability terms $\sum_w p(H_k/E)$, known as node scores (χ), can be used as a measurement of the level of suspicion for a given node at any given time. For a given set of node profiles (e.g. profiles corresponding to a similar peer group), the univariate version of Grubb's test [25] is used to detect anomalous points. For each profile score χ, its z score is computed as:

$$z = \frac{\chi - \bar{\chi}}{s} \quad (2)$$

Where $\bar{\chi}$ and s are mean and standard deviation of data set. A test instance is declared to be anomalous at significance level α if z is greater than Grubbs' critical value (GC).

$$GC = \frac{N-1}{\sqrt{N}} \sqrt{\frac{t^2_{\alpha/N,N-2}}{N-2+t^2_{\alpha/N,N-2}}} \quad (3)$$

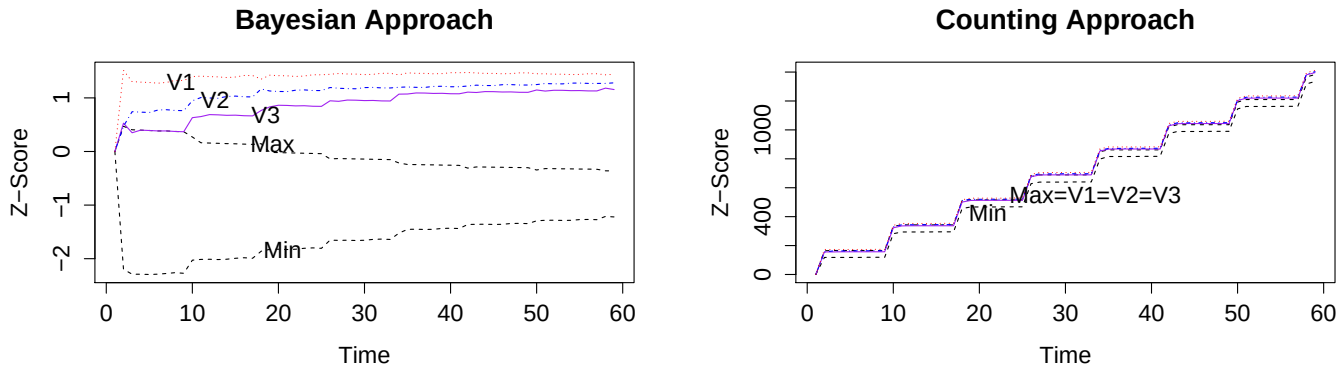

Fig. 1 A Comparison: Probabilistic vs Deterministic approaches. V_i denotes victim nodes while *min* and *max* denote minimum and maximum node scores of normal nodes in the same subnet.

where N is the number of profile points in the set, and $t_{\alpha/N,N-2}$ is the value taken by a t-distribution (one tailed test) at the significance level of $\frac{\alpha}{N}$ and degrees of freedom $(N-2)$. The α reflects the confidence associated with the threshold and indirectly controls the number of profiles declared as anomalous [12]. Note that the threshold adjusts itself according to current state of a network. This is a vertical analysis to detect one's aberrant behaviour with respect to her peers. In other words it compares each node's activity changes to activity changes of her peer group. Note that this analysis technique accounts for regular variations such as diurnal and familiarity. Looking at one's aberrant behaviour within a similar peer group gives better results in terms of false alarms than setting a universal baseline for the entire network [5, 23].

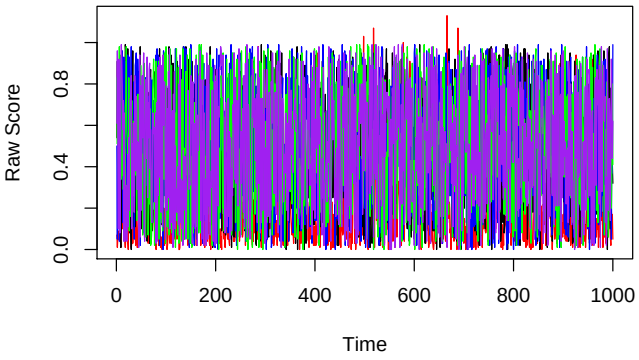

Fig. 2 Hiding behind innocent nodes (See magnified version in Figure 3) (quoted from [27]).

.

4.4 Discord analysis

When an attack is progressing malicious activities are occurring according to an on-off pattern in time. As a result, lack of agreement or harmony between points in the profile sequence of a given node can occur in a similar or different on-off fashion. This type of anomalies are known as discords [51]. These discords are random time context and peer analysis technique itself is not sufficient to detect them if the progression rate of malicious activities is far lower than the similar innocent activities. The Graph shown in Figure 2 presents such a situation. The objective of this analysis is to detect sub-sequences within a given sequence of profiles which is anomalous with respect to the rest of the sequence. Problem formulation occurs in time-series data sets where data is in the form of a long sequence and contains regions that are anomalous. The underlying assumption is that the normal behaviour of the time-series follows a defined random pattern, and a sub-sequence within the long sequence which does not conform to this pattern is an anomaly. In general, the purpose of this analysis is to detect one's aberrant behaviour with respect to her own behaviour regardless of her peers.

At the $(t-1)^{th}$ time point, using an auto-regressive integrated moving average model $ARIMA(p,d,q)$ [14] which describes the auto-correlations in the data, 95% Confidence Interval (CI) for the t^{th} profile score is predicted (see Figures 4 and 5). If the observed profile score at time t lies outside of the predicted CI then absolute deviation of the profile score from CI is calculated, for example the distance between points P1 and P2 in Figure 4. This deviation is used as a measure of non-conformity of a given profile score to the pattern of its own sequence (group norms). These deviations average out over the time to calculate the *anomaly score* for

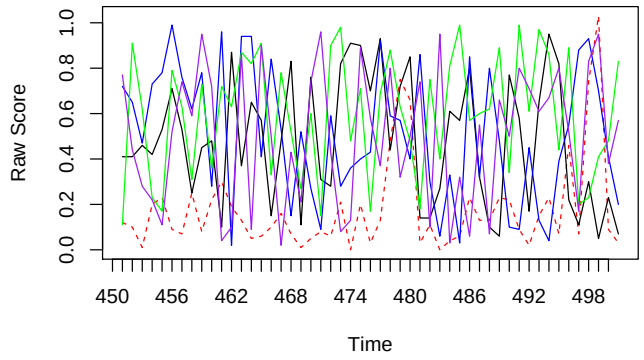

Fig. 3 Magnified version of Figure 2 - red dotted line denotes the attacker, all other lines denote innocent nodes (quoted from [27]).

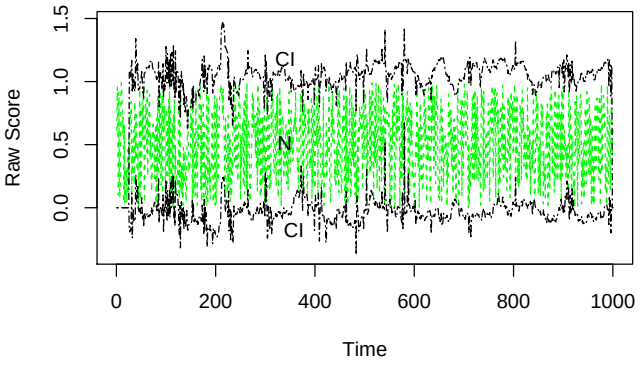

Fig. 5 Node scores and 95% CIs for a normal node. Black lines denote CIs while the green line denotes the normal node (N) (quoted from [27]).

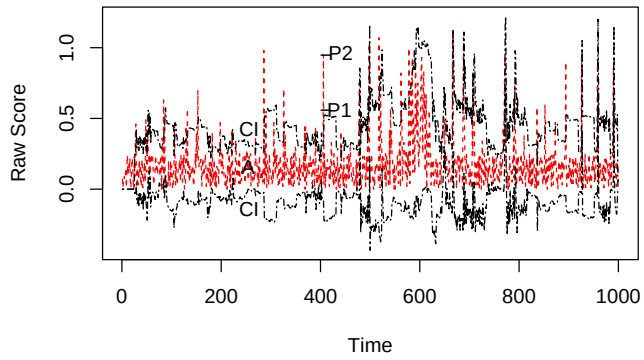

Fig. 4 Node scores and 95% CI intervals for the attacker node. Black lines denote CIs while the red line denotes the attacker (A) (quoted from [27]).

a given node. Note that this anomaly score is the average dissimilarity of profile scores with its own profile sequence of a node. This dissimilarity occurs randomly from time to time due to the deliberate intervention of the attacker. A node does exhibit sudden changes in behaviour when compared to its past behaviour is not necessarily suspicious as it could be a regular variation of the node behaviour [23]. Discord analysis technique uses in this work considers such variations as completely legitimate as it monitoring for changes to the changing pattern of node behaviour.

5 Experimental Set up and Results

We simulate a network with a set of attackers. We profile for both source and target of attackers using the above algorithm. We then compare the two analyses to offer insight into whether the source- or target-centric approach is more effective.

As it is very difficult to evaluate a novel algorithm based on live (or any raw) network traffic, very often simulation methods or some benchmark datasets (e.g. DARPA - KDD99 [35]) are used by researchers for evaluations of their algorithms [7]. However most of such datasets are also simulated traffic on real networks. Despite the significant contributions of benchmark datasets their accuracy and ability to reflect real world conditions has been extensively criticised [9, 10]. Therefore we set up a network, as shown in Figure 6, in a simulated environment using the network simulator ns-3. Poison arrival model was assumed to generate traffic patterns of interest. The simulation was run for a period of time to ensure that enough traffic was generated.

The network has ten subnets varying the size between 50 and 2, and any node is free to communicate with any other. Three attackers A_1, A_2, A_3 are planted in three subnets sizes 10, 25 and 50 respectively. All three attackers are launching attacks on two targets V_1 and V_2 in a given server farm. Anomalous traffic by means of unusual port numbers was generated in addition to generating usual traffic within and between subnets and to external networks. If λ_s, λ_n are mean rates of generating suspicious events by suspicious and normal nodes (i.e. the noise) respectively, we ensure maintaining $\lambda_s = \lambda_n \pm 3\sqrt{\lambda_n}$ and $\lambda_n (\leq 0.1)$ sufficiently smaller for our experiment to characterise suspicious activities. The idea to use the above relationship for generating attacker activities was to keep them within the normality range of innocent activities (i.e. back-

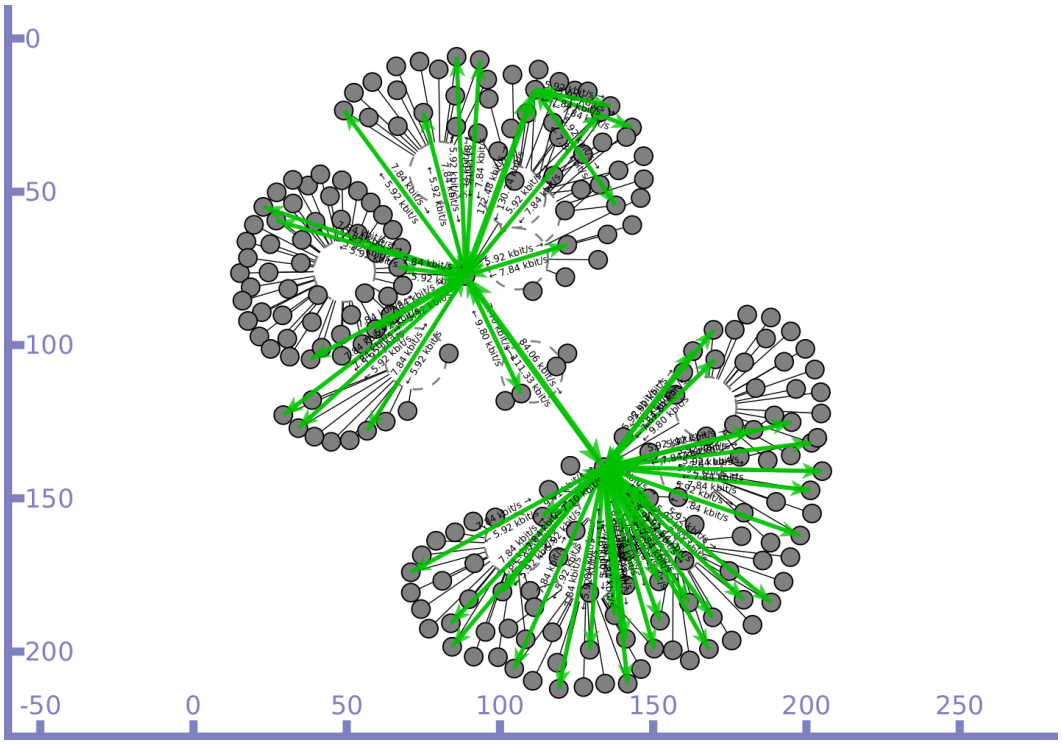

Fig. 6 Network topology: A run time instance.

ground noise). $\sqrt{\lambda_n}$ is the standard deviation of rates of suspicious events generated by normal nodes.

5.1 Prior and Likelihoods

The motivation behind the detection algorithm is that a network event is not always easy to judge for malicious nature. Some suspicious events can appear as part of an attack signature as well as originate from benign network activity. For example, a major router failure could generate a flood of ICMP unreachable messages while some malicious program (viruses and worms) may generate the same for probing; such uncertainty needs to be acknowledged [28, 29].

Prior probabilities and Likelihoods are assigned as follows.

$$p(H_1) = \frac{1}{2} = 0.5 \tag{4}$$

Equation 4 assumes that there is a 50% chance for a given node to be a victm. However, this is not the case in many situations. A node may have a higher prior belief of being a victim, such as a web server, than an ordinary client node. Since prior probabilities are based on previous experiences, $p(H_1)$ can be judged based on information gathered from contextual analysis. However if there is no basis to distinguish between nodes or groups of nodes equally likely (i.e.same probability of occurring) can be assumed. Alternatively the posterior probability of node k at time $t-1$ can be used as the prior of the same node at time t. This lets prior probabilities to adjust itself dynamically according to the suspicious evidences observed over time.

Likelihoods are assigned using

$$p(e_j/H_i) = k \tag{5}$$

where for all j,i. 5 denotes that the likelihoods of seeing the event e_j at a node when it is a victim of an attack. For the purposes of demonstration, arbitrary values ($\leq$1) for k are assigned to distinguish different types of events produced in the given scenario simulation. In actual implementation estimations of these types of likelihoods could be drawn from common classes of attacks and preconfigured normal traffic; solutions to empirically analyse day-to-day traffic and build statistical models of normal behaviour exist [21] for such purposes.

We present our experimental outcomes under four different cases. The cardinality of attacker to victim relationship in each case can be described as follows. The idea is to see whether target centric monitoring is sensitive to number of attackers and victims in the scene as our monitoring approach is aggregating profiles.

- One to one (1:1) - one attacker sends suspicious packets to only one target in the system
- One to Many (1:M) - three attackers send suspicious packets to only one target in the system

– Many to one (M:1) - one attacker sends suspicious packets to only two targets
– Many to many (N:M) - three attackers send suspicious packets to only two targets

5.2 Target-centric monitoring

Graphs in Figure 7 present monitoring outcomes by means of detection. Since we utilise the destination information of activities, our approach detects the targets of attackers (see Figure 7). *Min* and *Max* represent the minimum and maximum profile scores of normal nodes in each subnet where target node (denoted by V or V_i in graphs) is located. GC represents the Grubbs' critical value (threshold) for targets' subnet (i.e. the server farm). As is obvious from Figure 7, the proposed approach is capable of detecting targets of attack activities successfully in all four cases considered here. While the target is cut-off (or very close to) the threshold (GC), all other normal nodes in the target's subnet is significantly below from the threshold during the monitoring period.

However the M:1 case (single origin of activities, but many destinations) is significant. Only V_1 is above the threshold, and V_2 is among the normal nodes. However this should not be mistaken as the proposed approach is false negative on detecting V_2 in this case. In this case victim V_1 can be detected very early than detecting the corresponded attacker A (please compare with M:1 case in figure 8). Once V_1 is detected and prevented, then that case is turned into an 1:1 case which the victim can be detected very quickly than the attacker. Hence in M:1 case too both victims can be detected earlier using destination utilised monitoring approach than using the source utilised monitoring approach. Essentially, M:1 and 1:1 cases do not represent source collusion/distributed slow activities. We simulated these two (M:1 and 1:1) cases to see how proposed approach works on both. Specially, M:N and 1:M cases are very important here as both cases simulate the clouded and/or distributed type slow activities. As obvious from the experimental results destination utilised monitoring approach performs well in both cases.

5.3 A Comparison

This section compares target centric monitoring to traditional attribution-based approaches. The comparison is made under two different perspectives: first, comparing the actual detection on temporal aspects, and second, comparing the detection potentials.

5.3.1 Early detection

This perspective compares how early is abnormal activity detected using both approaches. To enable this comparison figure 8 was obtained using the same trace and the same Bayesian model used to obtain figure 7. The only change we made in obtaining graphs in figure 8 is utilising the source information of activities in profiling instead of destination information. *Min* and *Max* represent the minimum and maximum profile scores of normal nodes in each subnets where attack node (A_i) is located. GC represents the Grubbs' critical value (threshold) for attacker's subnet.

The results suggest that the victim-based approach is very quicker than the attribution-based approach in all cases (compare graphs in figures 7 and 8 temporally) in detection of the slow suspicious activities. Interestingly, in most cases presented in figure 8, source utilised monitoring approach failed to detect slow suspicious activities during the monitored period. In some cases activities are detected using the source utilised monitoring, however, after a considerable lag compared to the destination utilised monitoring approach.

5.3.2 Detection potential

A simple measure called detection potential is defined to explain how far an attacker node is deviated from the threshold. It helps to compare between different network conditions. The detection potential d is defined as:

$$d = z - GC \tag{6}$$

on the basis of the higher the detection potential the better for the detection. Figure 9 compares the *detection potential* across the two approaches in each cases.

As is obvious from Figure 9, the target-centric approach has a higher detection potential in all four cases. This means that there is a higher chance of detection of suspicious activities using our approach than most traditional source-centric approaches. Most importantly, detection potential of the source-centric approach has higher variations (fluctuations) while detection potential of target oriented approach is more stable. This is a good indication to imply that source oriented monitoring may have more false negatives (and positives) than target-centric monitoring approach. Future work will build on this further to establish this principle for actual real network examples where a diverse set of attack cases are prevalent.

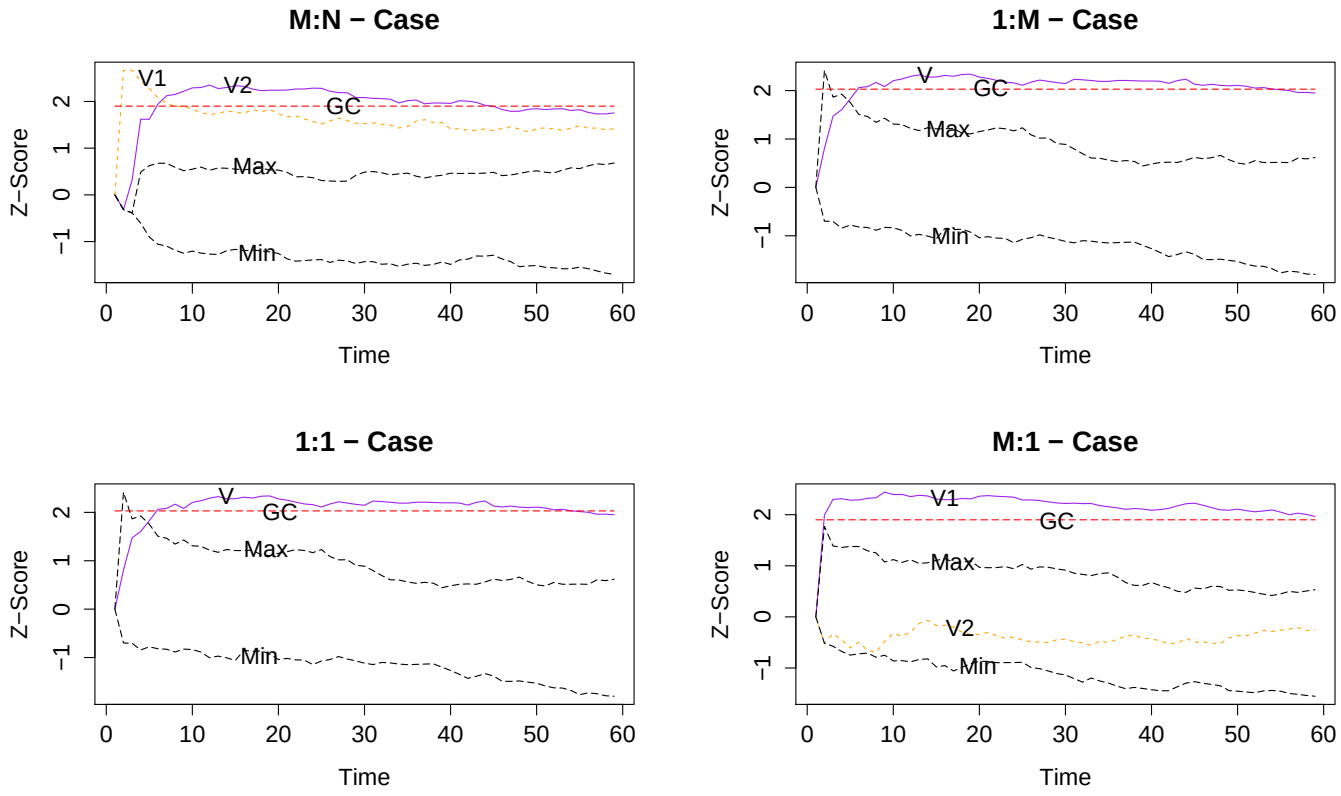

Fig. 7 Target centric monitoring. The cardinality of attacker to victim relationship in each case is described in section 5

6 Discussion

The purpose of this paper is to convey the core principle. No doubt this has to manifest itself in a variety of design principles, mechanisms and strategies that we hope will follow: recent work suggests that this is increasingly acknowledged within the wider security research community. Exposure maps are one such example [49, 50] where nodes are profiled for potential attack vectors available to an attacker and traffic activity is then assessed for port scanning attempts. An adaptive approach to network traffic analysis would also be welcome to address selective monitoring and collection of packets. Little research has considered this problem. One hardware-based approach to characterise unlikely uninteresting traffic more cheaply that can be devoid of further more expensive software-based analysis exists [24]; this is demonstrated to be effective for potential Gigabit Ethernet operations. However, further work is needed to allow for traffic monitoring to be sensitive to the type of services a given node may be vulnerable against. This will help avoid undue attention to suspicious traffic that will not prove harmful.

One difficulty with attribution discussed earlier is that attacks are carried out in multiple stages using compromised machines as stepping stones (or in the form of botnets). The focus on targeted nodes takes into account the importance of preventing such compromise, which in itself should help to undermine attacks. Make no mistake that attribution remains important but this is best left to be carried out by dedicated cybercrime units, perhaps operating at regional or national level providing for a coordinated response for potential attribution. Only then are the complexities involved in responding to large-scale organised attacks could be overcome both technically and otherwise.

This paper proposed a method to combine the output of several information sources to a single score. It acts as a data reduction method and enables to propose a lightweight monitoring scheme for the problem which is essential in near-real-time analysis of slow, sophisticated targeted attacks. It promises scalable means for detection. Experimental results offer a promise for the feasibility of target detection in network security monitoring.

Target-centric monitoring provides for effective detection of slow and suspicious activities as one does not have to rely on possible source aggregation. Note there is no guarantee that a publicly visible source of an event is the authentic source. Source-centric mon-

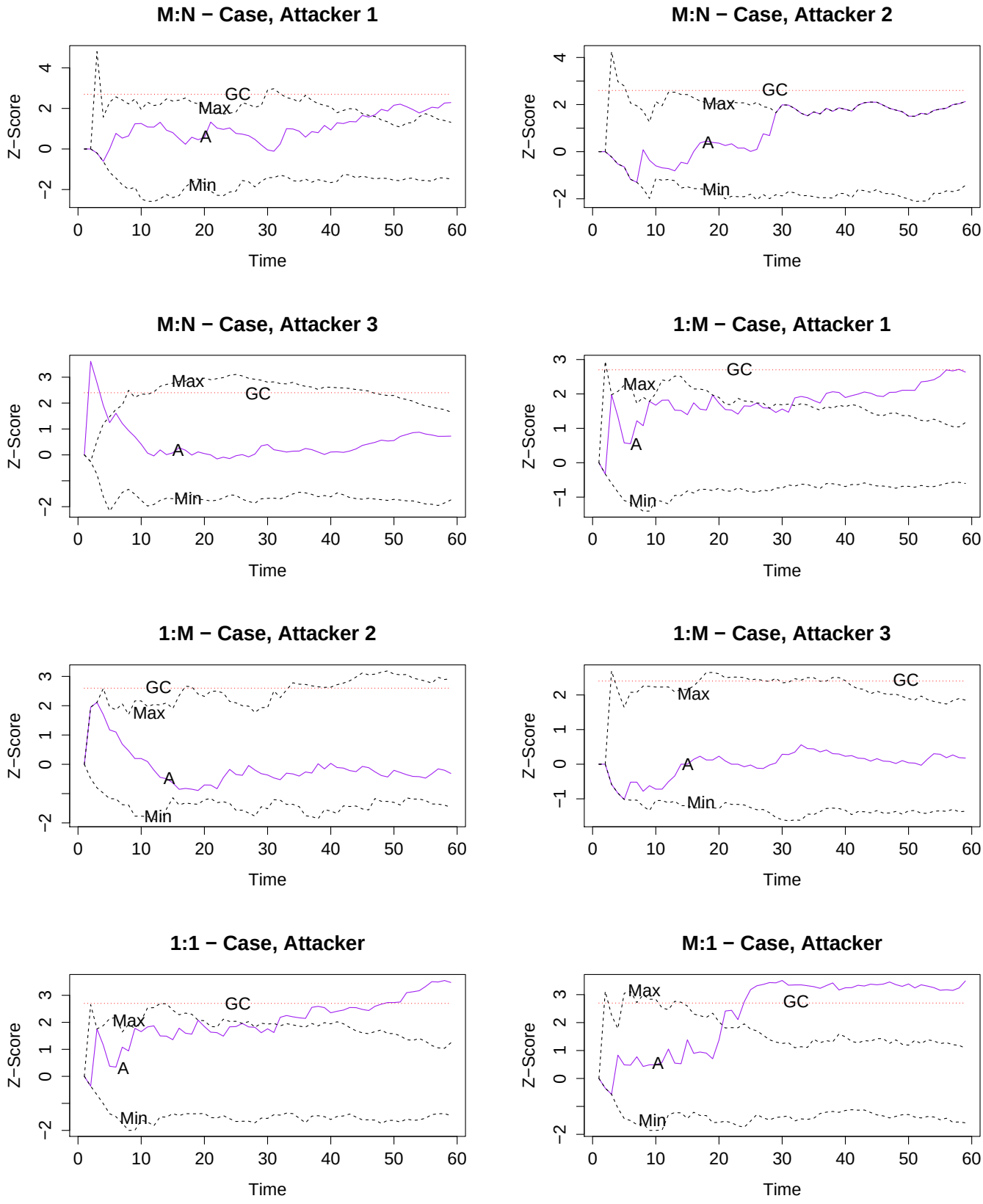

Fig. 8 Source centric monitoring. The cardinality of attacker to victim relationship in each case is described in section 5

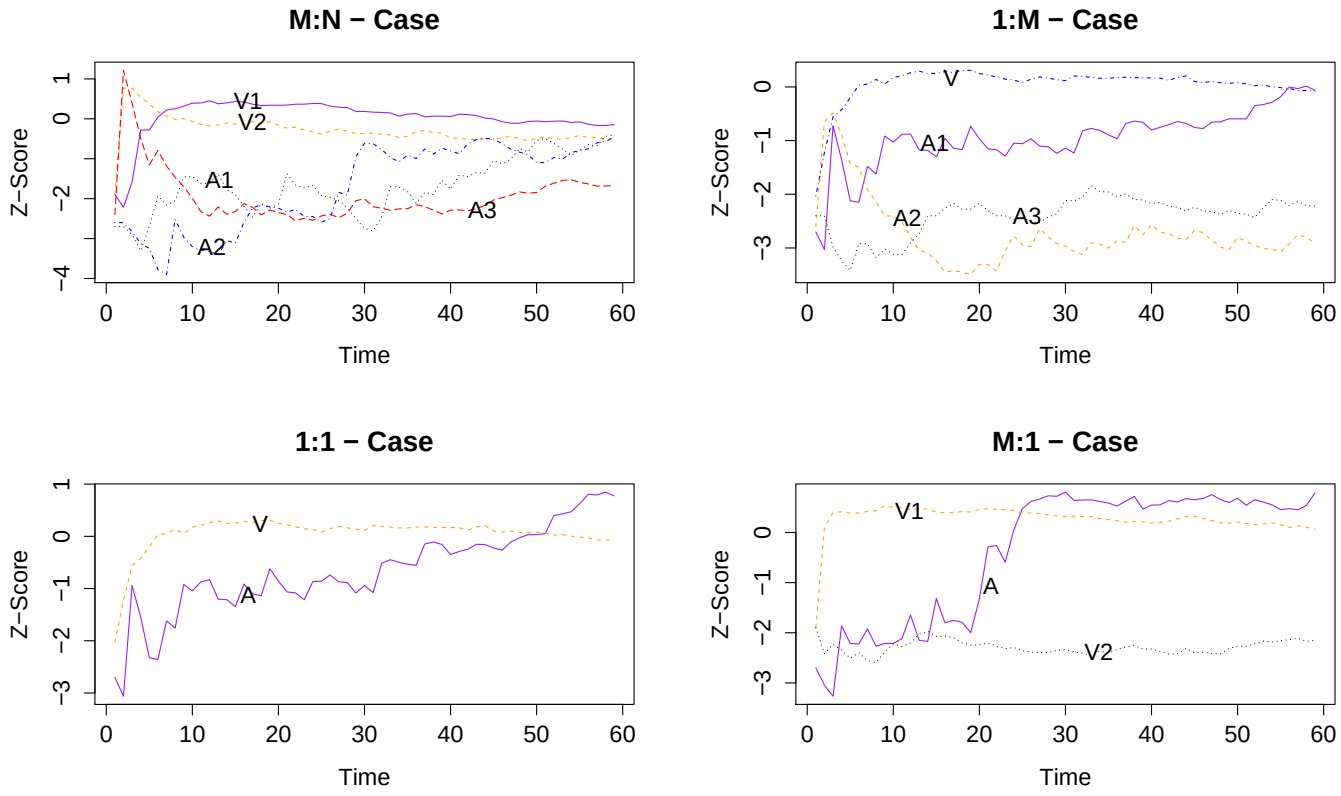

Fig. 9 A comparison of detection potential for each case.

itoring is vulnerable to this. Attacker tactics such as source collusion and source address spoofing are common and therefore make such attacker detection very hard. It becomes worse if the attack is stealthy as current computational constraints do not permit to maintain attack state over extended periods of times to correlate between suspicious events. This could be particularly useful for monitoring of high-profile nodes that are at particular risk from sophisticated insider attacks or purposefully designed cyber weapons. The focus on targeted nodes takes into account the importance of preventing such compromise, which in itself should help to undermine attacks. The main contribution of this paper is a shift to the focus of analysis.

Our approach should not be mistaken as a host based monitoring scheme. It focuses the event analysis stage of a monitoring system. Most of existing monitoring schemes share a common feature which we called source-centric analysis. They perform analysis based on source information (in fact perceived last hop) of activity either it is a host based or a network based monitoring system, and utilise that information at some stage of detection assuming that suspicious activity can be attributed to a meaningful specific source or an intermediate [49]. In a modern network such an assumption is not valid anymore. As mentioned in [42], most of the existing solutions become less effective when the attack is launched from distributed sources. We consider the actual reason behind such a deficiency is that their dependency on the Source IP addresses (or perceived last hop) of activities for attack detection, i.e. the source-centric analysis. What we propose in this paper is to move away from source-centric analysis to destination-centric analysis. Note that this should not be mistaken as a host based monitoring scheme, as it is not a problem of location of IDS deployment. It is a matter of whether monitoring system depends on source information of activities for attack detection or not.

Our method completely depends only on information within the control and ignores depending on any source information to protect networks. This work demonstrated the core principle taking into account the importance of preventing such compromise. Moreover, our approaches to tracing the source of such activity and a target-centric method to monitoring offer means to significantly improve network security monitoring against increasing volumes of traffic, spoofing attempts and collusion.

Acknowledgements Both authors have carried out this work partially under a grant (EP/L022656/1) received by the Engineering and Physical Sciences Research Council (EPSRC) of the UK.

References

1. Axelrad ET, Sticha PJ, Brdiczka O, Shen J (2013) A bayesian network model for predicting insider threats. In: 2013 IEEE Security and Privacy Workshops, URL https://www.ieee-security.org/TC/SPW2013/papers/data/5017a082.pdf
2. Basu R, Cunningham RK, Webster SE, Lippmann RP (2001) Detecting low-profile probes and novel denial-of-service attacks. Tech. rep., IEEE SMC IA&S Workshop 2001, West Point, New York, USA
3. BBC (2013) China IP address link to South Korea cyber-attack. http://www.bbc.co.uk/news/world-asia-21873017
4. BBC (2014) Hack attack causes 'massive damage' at steel works. http://www.bbc.co.uk/news/technology-30575104
5. Berk VH, Cybenko G, Souza IGd, Murphy JP (2012) Managing malicious insider risk through bandit. In: System Science (HICSS), 2012 45th Hawaii International Conference on, IEEE, pp 2422–2430
6. Bhuyan MH, Bhattacharyya DK, Kalita JK (2011) Survey on incremental approaches for network anomaly detection. In: International Journal of Communication Networks and Information Security (IJCNIS)
7. Bhuyan MH, Bhattacharyya D, Kalita JK (2012) Survey on incremental approaches for network anomaly detection. arXiv preprint arXiv:12114493
8. Bradford PG, Brown M, Self B, Perdue J (2004) Towards proactive computer system forensics. In: International conference on information technology: Coding and computing,IEEE Computer Society
9. Brown C, Cowperthwaite A, Hijazi A, Somayaji A (2009) Analysis of the 1999 darpa/lincoln laboratory ids evaluation data with netadhict. In: Computational Intelligence for Security and Defense Applications, 2009. CISDA 2009. IEEE Symposium on, IEEE, pp 1–7
10. Brugger ST, Chow J (2007) An assessment of the darpa ids evaluation dataset using snort. UCDAVIS department of Computer Science 1(2007):22
11. Brynielsson J, Horndahl A, Johansson F, Kaati L, Mårtenson C, Svenson P (2013) Harvesting and analysis of weak signals for detecting lone wolf terrorists. Security Informatics 2(1):11
12. Chandola V, Banerjee A, Kumar V (2009) Anomaly detection: A survey. ACM Comput Surv 41(3):15:1–15:58, DOI 10.1145/1541880.1541882, URL http://doi.acm.org/10.1145/1541880.1541882
13. Chandola V, Banerjee A, Kumar V (2009) Anomaly detection: A survey. In: ACM Computing Surveys 41
14. Chatfield C (2003) The analysis of time series: an introduction. CRC press
15. Chatzigiannakis V, Androulidakis G, Pelechrinis K, Papavassiliou S, Maglaris V (2007) Data fusion algorithms for network anomaly detection: classification and evaluation. In: Networking and Services, 2007. ICNS. Third International Conference on, IEEE, pp 50–50
16. Chen T, Abu-Nimeh S (2011) Lessons from stuxnet. Computer 44(4):91–93, DOI 10.1109/MC.2011.115
17. Chivers H, Nobles P, Shaikh SA, Clark JA, Chen H (2009) Accumulating Evidence of Insider Attacks. In: Proceedings of the 1st International Workshop on Managing Insider Security Threats (MIST-2009)
18. Chivers H, Clark JA, Nobles P, Shaikh SA, Chen H (2013) Knowing who to watch: Identifying attackers whose actions are hidden within false alarms and background noise. Information Systems Frontiers 15(1):17–34
19. Clark DD, Landau S (2010) The problem isn't attribution: it's multi-stage attacks. In: Proceedings of the Re-Architecting the Internet Workshop, ReARCH '10, pp 11:1–11:6
20. Das K, Schneider J (2007) Detecting anomalous records in categorical datasets. In: Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining, ACM, pp 220–229
21. Davidoff S, Ham J (2012) Network Forensics: Tracking Hackers Through Cyberspace. Prentice Hall
22. Deeks A (2013) The geography of cyber conflict: Through a glass darkly. Journal of International Law Studies 89:1–20
23. Eldardiry H, Bart E, Liu J, Hanley J, Price B, Brdiczka O (2013) Multi-domain information fusion for insider threat detection. In: 2013 IEEE Security and Privacy Workshops, URL https://www.ieee-security.org/TC/SPW2013/papers/data/5017a045.pdf
24. Gonzalez JM, Paxson V, Weaver N (2007) Shunting: a hardware/software architecture for flexible, high-performance network intrusion prevention. In: Proceedings of the 14th ACM conference on Computer and communications security, pp 139–149

25. GRUBBS RE (1969) Procedures for Detecting Outlying Observations in Samples. Technometrics 11(1):1–21
26. Heberlein T (2002) Tactical operations and strategic intelligence: Sensor purpose and placement. Net Squared Inc, Tech Rep TR-2002-0402
27. Kalutarage H (2013) Effective monitoring of slow suspicious activites on computer networks. PhD thesis, Coventry: Coventry University, URL "https://curve.coventry.ac.uk/open/file/afdbba5c-2c93-41a7-90c3-2f0f3261b794/1/Kalutarage2013.pdf"
28. Kalutarage HK, Shaikh SA, Zhou Q, James AE (2012) Sensing for suspicion at scale: A bayesian approach for cyber conflict attribution and reasoning. In: 4th International Conference on Cyber Conflict (CYCON) 2012, NATO CCDCOE, pp 1–19
29. Kalutarage HK, Shaikh SA, Zhou Q, James AE (2013) How do we effectively monitor for slow suspicious activities? In: Proceedings of the International Symposium on Engineering Secure Software and Systems (ESSoS-DS 2013) CEUR Workshop Proceedings
30. Kalutarage HK, Shaikh SA, Zhou Q, James AE (2013) Monitoring for slow suspicious activities using a target centric approach. In: Information Systems Security, Springer Berlin Heidelberg, pp 163–168
31. Kalutarage HK, Shaikh SA, Zhou Q, James AE (2013) Tracing sources of anonymous slow suspicious activities. In: Network and System Security, Springer Berlin Heidelberg, pp 122–134
32. Kandias M, Mylonas A, Virvilis N, Theoharidou M, Gritzalis D (2010) An insider threat prediction model. In: Trust, Privacy and Security in Digital Business, Springer, pp 26–37
33. Kumar S, Spafford EH (1994) An application of pattern matching in intrusion detection. In: Technical Report CSDTR-94-013 The COAST Project, Department of Computer Sciences, Purdue University, West Lafayette, IN
34. Lesk M (2007) The new front line: Estonia under cyberassault. IEEE Security & Privacy 5(4):76–79, DOI http://doi.ieeecomputersociety.org/10.1109/MSP.2007.98
35. Lippmann RP, Fried DJ, Graf I, Haines JW, Kendall KR, McClung D, Weber D, Webster SE, Wyschogrod D, Cunningham RK, et al (2000) Evaluating intrusion detection systems: The 1998 darpa off-line intrusion detection evaluation. In: DARPA Information Survivability Conference and Exposition, 2000. DISCEX'00. Proceedings, IEEE, vol 2, pp 12–26
36. Nakashima E, Warrick J (2012) Stuxnet was work of U.S. and Israeli experts, officials say. http://www.washingtonpost.com/world/national-security/stuxnet-was-work-of-us-and-israeli-experts-officials-say/2012/06/01/gJQAlnEy6U_story.html
37. Notknown (2006) Googlebot. http://www.lightbluetouchpaper.org/2006/02/15/complexities-in-criminalising-denial-of-service-attacks/
38. Paget F (2013) Hacking Summit Names Nations With Cyberwarfare Capabilities. http://blogs.mcafee.com/mcafee-labs/hacking-summit-names-nations-with-cyberwarfare-capabilities
39. Pang R, Yegneswaran V, Barford P, Paxson V, Peterson L (2004) Characteristics of internet background radiation. In: Proceedings of the 4th ACM SIGCOMM conference on Internet measurement, pp 27–40
40. Parikh D, Chen T (2008) Data fusion and cost minimization for intrusion detection. Information Forensics and Security, IEEE Transactions on 3(3):381–389
41. Patcha A, Park JM (2007) An overview of anomaly detection techniques: Existing solutions and latest technological trends. In: Computer Networks (Elsevier)
42. Peng T, Leckie C, Ramamohanarao K (2004) Proactively detecting distributed denial of service attacks using source ip address monitoring. In: NETWORKING 2004. Networking Technologies, Services, and Protocols; Performance of Computer and Communication Networks; Mobile and Wireless Communications, Springer, pp 771–782
43. Shaikh SA, Chivers H, Nobles P, Clark JA, Chen H (2008) Characterising intrusion detection sensors. Network Security 2008(9):10–12
44. Shaikh SA, Chivers H, Nobles P, Clark JA, Chen H (2008) Characterising intrusion detection sensors, part 2. Network Security 2008(10):8–11
45. Siaterlis C, Maglaris B (2004) Towards multisensor data fusion for dos detection. In: Proceedings of the 2004 ACM symposium on Applied computing, ACM, pp 439–446
46. Streilein WW, Cunningham RK, Webster SE (2002) Improved detection of low profile probe and novel denial of service attacks. In: Workshop on Statistical and Machine Learning Techniques in Computer Intrusion Detection
47. Time LA (2014) Sony insider – not North Korea – likely involved in hack, experts say. http://www.latimes.com/entertainment/envelope/

cotown/la-et-ct-sony-hack-inside-job-not-north-korea-20141231-story.html
48. Vokorokos L, Chovanec M, Látka O, Kleinova A (2008) Security of distributed intrusion detection system based on multisensor fusion. In: Applied Machine Intelligence and Informatics, 2008. SAMI 2008. 6th International Symposium on, IEEE, pp 19–24
49. Whyte D, van Oorschot PC, Kranakis E (2006) Exposure maps: removing reliance on attribution during scan detection. In: Proceedings of the 1st USENIX Workshop on Hot Topics in Security (HOTSEC'06)
50. Whyte D, Oorschot PC, Kranakis E (2007) Tracking darkports for network defense. In: 23rd Computer Security Applications Conference, pp 161–171
51. Yankov D, Keogh E, Rebbapragada U (2008) Disk aware discord discovery: finding unusual time series in terabyte sized datasets. Knowledge and Information Systems 17(2):241–262
52. Ye N, Xu M, Emran S (2000) Probabilistic networks with undirected links for anomaly detection. In: IEEE Systems, Man, and Cybernetics Information Assurance and Security Workshop, pp 175–179
53. ZDNet (2015) Sony takes $15M hit after North Korea cyberattack. http://www.zdnet.com/article/sony-hack-cost-it-15-million-so-far/