Real-time behaviour profiling for network monitoring

Kuai Xu* and Feng Wang

Arizona State University,
4701 W. Thunderbird Road,
Glendale, AZ 85306, USA
E-mail: kuai.xu@asu.edu E-mail: fwang25@asu.edu
*Corresponding author

Supratik Bhattacharyya

SnapTell Inc.,
Palo Alto, CA 94306, USA
E-mail: supratik@gmail.com

Zhi-Li Zhang

Department of Computer Science and Engineering,
University of Minnesota,
4-192 EE/CS Building, 200 Union Street SE,
Minneapolis, MN 55416, USA
E-mail: zhzhang@cs.umn.edu

Abstract: This paper presents the design and implementation of a real-time behaviour profiling system for internet links. The system uses flow-level information, and applies data mining and information-theoretic techniques to automatically discover significant events based on communication patterns. We demonstrate the operational feasibility of the system by implementing it and performing benchmarking of CPU and memory costs using packet traces from backbone links. To improve the robustness of this system against sudden traffic surges, we propose a novel filtering algorithm. The proposed algorithm successfully reduces the CPU and memory cost while maintaining high profiling accuracy. Finally, we devise and evaluate simple yet effective blocking strategies to reduce prevalent exploit traffic, and build a simple event analysis engine to generate ACL rules for filtering unwanted traffic.

Keywords: real-time traffic monitoring; behaviour profiling; profiling-aware filtering algorithms.

Reference to this paper should be made as follows: Xu, K., Wang, F., Bhattacharyya, S. and Zhang, Z-L. (2010) 'Real-time behaviour profiling for network monitoring', *Int. J. Internet Protocol Technology*, Vol. 5, Nos. 1/2, pp.65–80.

Biographical notes: Kuai Xu is currently an Assistant Professor at Arizona State University. He received his PhD in Computer Science from the University of Minnesota in 2006, and his BS and MS from Peking University, China, in 1998 and 2001. His research interests include network security and cloud computing. He is a member of ACM and IEEE.

Feng Wang received her BS from Wuhan University in 1996, MS from Beijing University in 1999, and PhD from the University of Minnesota in 2005, all in Computer Science. She is currently an Assistant Professor at ASU. Her research interests include wireless networks and combinatorial optimisation. She is a member of the IEEE.

Supratik Bhattacharyya received his MS and PhD in Computer Science from the University of Massachusetts, Amherst. He is currently with SnapTell Inc, Palo Alto, CA. He was a Distinguished Member of Technical Staff at Sprint Labs in Burlingame CA. His research interests are in mobile communication and services and in mining network traffic data.

Zhi-Li Zhang received his BS in Computer Science from Nanjing University, China, in 1986 and his MS and PhD in Computer Science from the University of Massachusetts in 1992 and 1997. He is currently a Professor of Computer Science and Engineering at the University of Minnesota. His research interests include computer communication and networks. He is a member of IEEE, ACM and INFORMS Telecommunication Section.

1 Introduction

Recent years have seen significant progress in real-time, continuous traffic monitoring and measurement systems in IP backbone networks (Iannaccone, 2005; Iannaccone et al., 2001). However, *real-time* traffic summaries reported by many such systems focus mostly on volume-based heavy hitters or aggregated metrics of interest (Keys et al., 2005), which are not sufficient for finding interesting or anomalous behaviour patterns. This paper[1] explores the feasibility of building a real-time traffic *behaviour profiling* system that analyses vast amount of traffic data in IP backbone networks and reports *comprehensive behaviour patterns* of significant end hosts and network applications.

Towards this end, we answer a specific question in this paper: is it feasible to build a *robust* real-time traffic behaviour profiling system that is capable of continuously extracting and analysing 'interesting' and 'significant' traffic patterns on high-speed internet links, even in the face of sudden surge in traffic. We address this question in the context of a traffic behaviour profiling methodology developed for IP backbone networks (Xu et al., 2005a). The behaviour profiling methodology employs a combination of data-mining and information-theoretic techniques to build comprehensive behaviour profiles of internet backbone traffic in terms of communication patterns of end hosts and applications. It consists of three key steps: significant cluster extraction, automatic behaviour classification and structural modelling for in-depth interpretive analysis. This three-step profiling methodology extracts hosts or services that generate significant traffic, classifies them into different *behaviour classes* that provide a general separation of various normal and abnormal traffic as well as rare and anomalous traffic behaviour patterns (see Section 2 for more details). The profiling methodology has been extensively validated *offline* using packet traces collected from a variety of backbone links in an IP backbone network (Xu et al., 2005a, 2005b).

To demonstrate the operational feasibility of performing *online* traffic behaviour profiling on high-speed internet backbone links, we build a prototype system of the aforementioned profiling methodology using general-purpose commodity PCs and integrate it with an existing real-time traffic monitoring system operating in an internet backbone network (Xu et al., 2007). The real-time traffic monitoring system captures packets on a high-speed link (from OC12 to OC192) and converts them into 5-tuple flows (based on source IP, destination IP, source port, destination port, protocol fields), which are then continuously fed to the real-time traffic profiling system we build. The large volume of traffic flows observed from these links creates great challenges for the profiling system to process them *quickly* on commodity PCs with *limited memory* capacity. We incorporate several optimisation features in our implementation such as efficient data structures for storing and processing cluster information to address these challenges.

After designing and implementing this real-time traffic profiling system, we perform extensive benchmarking of CPU and memory costs using packet-level traces from internet backbone links to identify the potential challenges and resource bottlenecks. We find that CPU and memory costs linearly increase with flow arrival rate. Nevertheless, resources on a commodity PC are sufficient to continuously process flow records and build behaviour profiles for high-speed links in operational networks. For example, on a dual 1.5 GHz PC with 2048 MB of memory, building behaviour profiles once every 5 minutes for a 2.5 Gbps link loaded at an average of 209 Mbps *typically* takes 45 seconds of CPU time and 96 MB of memory.

However, resource requirements are much higher under anomalous traffic patterns such as sudden traffic surges caused by denial of service (DoS) attacks, when the flow arrival rate can increase by several orders of magnitude. We study this phenomenon by superposing 'synthetic' packet traces containing a mix of known DoS attacks (Hussain et al., 2003) on real backbone packet traces. To enhance the robustness of our profiling system under these stress conditions, we propose and develop sampling-based *flow filtering* algorithms and show that these algorithms are able to curb steep increase in CPU and memory costs while maintaining high profiling accuracy.

Given the prevalent exploit traffic, we further consider blocking strategies the real-time profiling system can deploy to reduce such unwanted traffic. Based on the characteristics of exploit traffic, we devise several heuristic rules that the profiling system can employ to reducing unwanted traffic, and evaluate their cost and performance. By replaying packet traces collected from backbone links to the real-time profiling system, we find that simple blocking strategies could potentially reduce substantial exploit traffic in backbone networks.

The contributions of this paper are as follows:

- We present the design and implementation of a real-time traffic profiling system for link-level internet traffic, and demonstrate its operational feasibility by benchmarking CPU and memory costs using packet traces from an operational backbone.
- We propose a new filtering algorithm to improve the robustness of the profiling system against traffic surges and anomalous traffic patterns, and show that the proposed algorithm successfully reduces CPU and memory costs while maintaining high profiling accuracy.
- We devise, evaluate and deploy simple yet effective the access control list (ACL) rules in the real-time profiling system to reduce prevalent exploit behaviour in backbone networks.

The remainder of this paper is organised as follow. Section 2 describes a behaviour profiling methodology that automatically discovers significant behaviour patterns from massive traffic data. Section 3 introduces the real-time profiling system and discusses its functional modules as well as the interfaces with continuous monitoring systems and an event analysis engine. Section 4 is devoted to

performance benchmarking and stress test of the profiling system using a variety of packet-level traces from OC-48 backbone links, and synthetic traces that mix various attacks into real backbone packet traces. In Section 5, we propose and evaluate sampling-based filtering algorithms to enhance the robustness of the profiling system against sudden traffic surges. Section 6 devises and evaluates several simple blocking strategies in the real-time profiling system to reduce unwanted exploit traffic. Finally, Section 7 concludes this paper.

2 Behaviour profiling methodology

In light of wide spread cyber attacks and frequent emergence of disruptive applications, we developed a general traffic profiling methodology that automatically discovers significant behaviours with plausible interpretations from vast amount of traffic data. The profiling methodology uses 5-tuple flows, i.e., source IP address (`srcIP`), destination IP address (`dstIP`), source port number (`srcPrt`), destination port number (`dstPrt`), and protocol, collected in a time interval (e.g., 5 minutes) from internet backbone links. We focus on the first four feature dimensions in 5-tuples, and extract clusters along each dimension. Each cluster consists of flows with the same feature value in a given dimension. The value and its dimension are denoted as *cluster key* and *cluster dimension*. This leads to four groups of clusters, i.e., `srcIP`, `dstIP`, `srcPrt` and `dstPrt` clusters. The first two represent a collection of host behaviour, while the last two yield a collection of application behaviours that aggregate flows on the corresponding ports.

2.1 Extracting significant clusters

Due to massive traffic data and wide diversity of end hosts and applications observed in backbone links, it is impractical to examine all end hosts and applications. Thus, we attempt to extract *significant* clusters of interest, in which the number of flows exceeds a threshold. In extracting such clusters, we introduced an entropy-based algorithm that finds adaptive thresholds along each dimension based on traffic mix and cluster size distributions.

The intuitive idea of this algorithm is:

1 to extract clusters in each dimension whose cluster keys are distinct in terms of size distributions

2 to repeat this process until the size distribution of the remaining clusters in the dimension is (nearly) random.

To quantify the 'randomness' of cluster size distribution, we use an information-theoretic measure, *relative uncertainty* (UR) (also known as standardised entropy), which provides a measure of randomness between 0 and 1. If RU on a given variable X is close to 1, it indicates that the observed values of X are closer to being uniformly distributed and thus look less distinguishable from each other.

By applying this algorithm on a variety of backbone links, we see that the number of significant clusters extracted along each dimension is far less than the total number of values. This observation suggests that this step is very useful and necessary in reducing traffic data for analysis while retaining most interesting behaviours.

2.2 Behaviour classification

Given the extracted significant clusters, the second step of the profiling methodology is to classify their behaviours based on *communication patterns*. The flows in each significant cluster, e.g., a `srcIP` cluster, share the same feature value in `srcIP` dimension, thus most behaviour information is contained in the other 'free' features including `dstIP`, `srcPrt`, `dstPrt`, which might take any possible values.

Traditional approaches mostly focused on volume-based information, e.g., unique number of `dstIP`'s or `dstPrt`'s in examining the patterns of such clusters. However, the traffic volume often is unable to uncover comprehensive communication patterns. For example, if two hosts communicate with 100 unique `dstIP`'s, we cannot safely conclude that their communication patterns from `dstIP` feature are the same without further investigation. A simple example is that one host could be a web server talking to 100 clients, while another is an infected host randomly scanning 100 targets. More importantly, the number of flows associated with each `dstIP` is very likely to be different. For the case of the web server, the numbers of flows between clients and the server tend to be diverse. On the other hand, the number of probing flows between the scanner and each target is often uniform, e.g., one in most cases. This insight motivates us to use RU again to measure the feature distribution of free dimensions for all significant clusters.

Suppose the size of a cluster is m and a free dimension X may take N_X discrete values. Moreover, let $P(X)$ denote a probability distribution, and $p(x_i) = m_i/m, x_i \in X,$ where mi is the frequency or number of times we observe the feature X taking the value x_i. Then, the RU in the feature X for the cluster is defined as

$$RU(X) \coloneqq \frac{H(X)}{H_{\max}(X)} = \frac{H(X)}{\log \min\{N_X, m\}}, \tag{1}$$

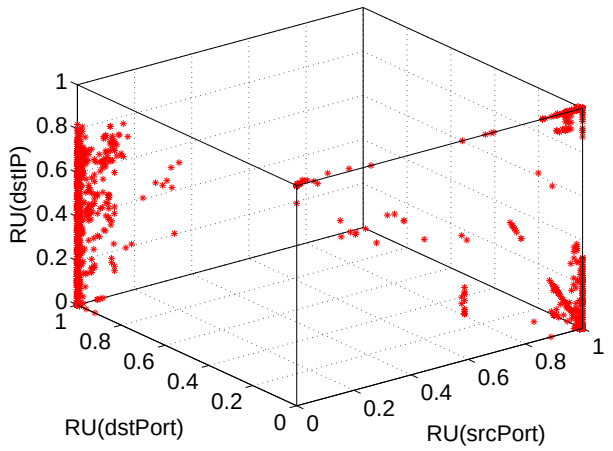
where $H(X)$ is the (empirical) entropy of X defined as

$$H(X) \coloneqq -\sum_{x_i \in X} p(x_i) \log p(x_i). \tag{2}$$

We use RU to measure feature distributions of three free dimensions. As a result, we obtain a RU vector for each cluster, e.g., $[RU_{srcPrt}, RU_{dstPrt}$ and $RU_{dstIP}]$ for `srcIP` clusters. Recall that RU is in the range of [0, 1], so we could represent the RU vector of each `srcIP` cluster as a single point in a three-dimensional space. Figure 1 represents each

srcIP cluster extracted in each 5-minute time slot over a 1-hour period from an OC-48 backbone link as a point in a unit cube. We see that the points are 'clustered', suggesting that there are few underlying common patterns among them. Such observation holds for other dimensions as well. This leads to a behaviour classification scheme which classifies all srcIP's into *behaviour classes* based on their similarity/dissimilarity in the RU vector space.

Figure 1 The distribution of RU on free dimensions for srcIP's from an OC-48 backbone link during a 1-hour period (see online version for colours)



By applying the behaviour classification on backbone links and analysing their temporal properties, we find this scheme is robust and consistent in capturing behaviour similarities among significant clusters. Such similarities are measured on the feature distribution of free dimensions of these clusters, hence provide useful insight in communication patterns of end hosts and applications (Karagiannis et al., 2005; Xu et al., 2005a).

2.3 Structural modelling

To provide a plausible interpretation for behaviour patterns, we adopt *dominant state analysis* technique for modelling and characterising the interaction of various feature dimensions in a cluster. The idea of dominant state analysis comes from structural modelling or reconstructability analysis in system theory (Krippendorff, 1986; Cavallo and Klir, 1979; Zwick, 2004) as well as more recent graphical models in statistical learning theory (Jordan, 2004).

The intuition behind dominant state analysis is described below. Given a srcIP associated with significant traffic flows, all flows can be represented as a 4-tuple (ignoring the protocol field) $\langle u, x_i, y_i, z_i \rangle$, where the srcIP has a fixed value u, while the srcPrt (X dimension), dstPrt (Y dimension) and dstIP (Z dimension) may take any legitimate value. Hence each flow in the cluster imposes a 'constraint' on the three 'free' dimensions X, Y and Z. Treating each dimension as a random variable, the flows in the cluster constrain how the random variables X, Y and Z 'interact' or 'depend' on each other, via the (induced) *joint* probability distribution $\mathcal{P}(X, Y, Z)$.

The objective of dominant state analysis is to explore the interaction or dependence among the free dimensions by identifying 'simpler' subsets of values or constraints [called *structural models* in the literature (Krippendorff, 1986)] to represent the original data in their probability distribution. Given the probability information, we can not only *approximately* reproduce the original flow patterns, but also explain the *dominant* activities of end hosts or applications.

2.4 Properties of behaviour profiles

We have applied the profiling methodology on traffic data collected from a variety of links at the core of the internet through *offline* analysis. We find that a large fraction of clusters fall into three typical behaviour profiles: server/service behaviour profile, heavy hitter host behaviour, and scan/exploit behaviour profile. These behaviour profiles are built based on various aspects, including behaviour classes, dominant states, and additional attributes such as average packets and bytes per flow. These behaviour profiles are recorded in a database for further event analysis, such as temporal properties of behaviour classes and individual clusters, or behaviour change detection based on RU vectors.

The profiling methodology is able to find various interesting and anomalous events. First, it automatically detects novel or unknown exploit behaviours that match typical exploit profiles, but exhibit unusual dominant states (e.g., dstPrt's). Second, any atypical behaviour is worth close examination, since they represent as 'outliers' or 'anomaly' among behaviour profiles. Third, the methodology could point out deviant behaviours of end hosts or applications that deviate from previous patterns.

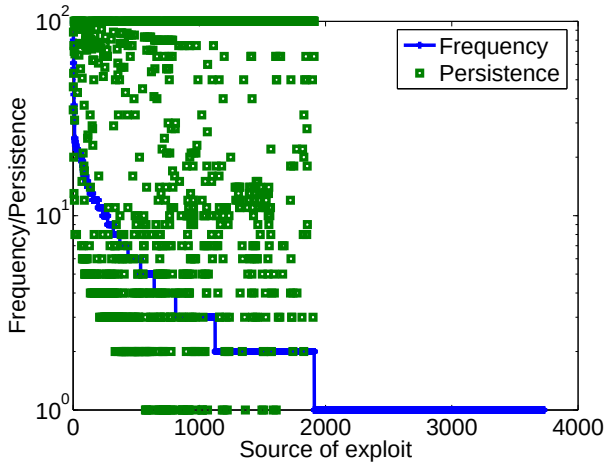
2.5 Characteristics of exploit traffic

Given the prevalent exploit activities, we further introduce several metrics to study the characteristics of exploit traffic. The *frequency*, T_f, measures the number of 5-minute time periods (over the course of 24 hours) in which a source is profiled by our methodology as having an exploit profile. The *persistence*, T_p, measures (in *percentage*) the number of *consecutive* 5-minute periods over the total number of periods that a source sends significant amount of exploit traffic. It is only defined for sources with $T_f \geq 2$. Hence $T_p = 100(\%)$ means that the source continuously sends significant amount of exploit traffic in all the time slots it is observed. Finally, we use the *intensity*, I, to relate both the temporal and spatial aspects of exploit traffic: it measures the (average) number of distinct target IP addresses per minute that a source touches in each 5-minute period. Thus it is an indicator how fast or aggressive a source attempts to spread the exploit.

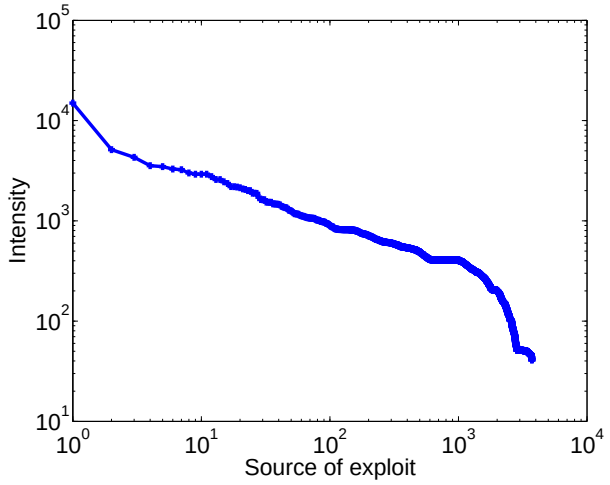
Figures 2(a) and 2(b) show the distributions of the frequency vs. persistence and the distribution of intensity

for the exploit sources from an OC-48 backbone link during 24 hours. From Figure 2(a) we observe that frequency follows a Zipf like distribution: only 17.2% sources have a frequency of 5 or more, while 82.8% sources have a frequency of less than 5. In particular, over 70% of them have frequency of 1 or 2. Furthermore, those 17.2% frequent ($T_f \geq 5$) sources account for 64.7%, 61.1% and 65.5% of the total flows, packets, and bytes of exploit traffic. The persistence varies for sources with similar frequency, but nearly 60% of the sources ($T_f \geq 2$) have a persistence of 100 (%): these sources continuously send exploit traffic over time and then disappear.

Figure 2 Characteristics of exploit traffic for the sources with exploit profiles in the backbone link during a 24-hour period (a) frequency (T_f) and persistence (T_p) (b) intensity (I) (see online version for colours)



(a)



(b)

The exploit intensity illustrated in Figure 2(b) also follows a Zipf like distribution. The maximum intensity is 21 K targets per minute, while the minimum is 40 targets per minute. There are only 12.9% sources with an intensity of over 500 targets per minute, while nearly 81.1% sources have an intensity of less than 500 targets per minute. Those 12.9% aggressive ($I \geq 500$) sources account for 50.5%, 53.3% and 45.2% of the total flows, packets and bytes of exploit traffic. The relatively small numbers of sources, which frequently, persistently or aggressively generate exploit traffic, are candidates for blocking actions. In the rest of this paper, we will demonstrate the feasibility of designing and implementing a real-time traffic profiling system that uses flow-level information generated from 'always-on' packet monitors and reports significant online events based on communication patterns of end hosts and applications even faced with anomalous traffic patterns, e.g., DoS attacks or worm outbreaks.

3 Real-time profiling system

In this section, we first describe the design guidelines for the profiling system and then present the overall architecture, functional modules and key implementation details.

3.1 Design guidelines

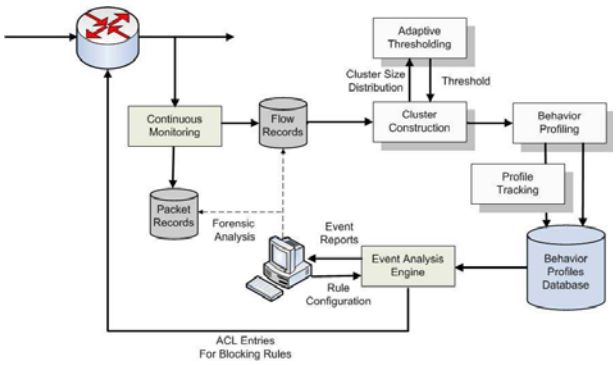
Four key considerations guide the design of our profiling system:

- *Scalability:* The profiling system is targeted at high-speed (1 Gbps or more) backbone links and hence must scale to the traffic load offered by such links. Specifically, if the system has to continuously build behaviour profiles of significant clusters once every time interval T (e.g., $T = 5$ minutes), then it has to take less than time T to process all the flow records aggregated in every time interval T. And this has to be accomplished on a commodity PC platform.
- *Robustness:* The profiling system should be robust to anomalous traffic patterns such as those caused by DoS attacks, flash crowds, worm outbreaks, etc. These traffic patterns can place a heavy demand on system resources. At the same time, it is vital for the profiling system to be functioning during such events since it will generate data for effective response and forensic analysis. Therefore the system must adapt gracefully to these situations and achieve a suitable balance between profiling accuracy and resource utilisation.
- *Modularity:* The profiling system should be designed in a modular fashion with each module encapsulating a specific function or step in the profiling methodology. Information exchange between modules should be clearly specified. In addition, the system should be designed to accept input from any packet or flow monitoring system that exports a continuous stream of flow records. However, the flow record export format has to be known to the system.
- *Usability:* The profiling system should be easy to configure and customise so that a network operator can focus on specific events of interest and obtain various

levels of information about these events. At the same time, it should expose minimal details about the methodology to an average user. Finally it should generate meaningful and easy-to-interpret event reports, instead of streams of statistics.

These design considerations form a guideline of our system design and drive each stage of our system implementation. In the rest of the section, we will discuss the overall architecture of the real-time profiling system, its functional modules and key implementation details that achieve the design goals.

Figure 3 The architecture of real-time traffic profiling system (see online version for colours)



3.2 System architecture

Figure 3 depicts the architecture of the profiling system that is integrated with an 'always-on' monitoring system, an event analysis engine and a feedback channel of generating ACL entries of routers for blocking unwanted traffic. The flow-level information used by the profiling system is generated from continuous packet or flow monitoring systems that capture packet headers on a high-speed internet link via an optical splitter and a packet capturing device, i.e., DAG card. The monitoring system aggregates packets into 5-tuple flows and exports the flow records for a given time interval into disk files. In general, the profiling system could obtain flow records through two ways: shared disk access or file transfer over socket, and the option depends on the locations of the profiling and monitoring systems. The first way works when both systems run on the same machine. If they are located in different machines, the monitoring system transfers the files of flow records in the last time interval to our traffic profiling system. In order to improve the efficiency of the profiling system, we use distinct process threads to carry out multiple tasks in parallel. For example, one thread continuously reads flow records in the current time interval T_i from the monitoring systems, while another thread profiles flow records that are complete for the previous time interval T_{i-1}.

The real-time traffic profiling system consists of five functional modules (shadowed boxes), namely, 'cluster construction', 'adaptive thresholding', 'behaviour profiling' and 'profile tracking'. Their functions are briefly summarised below:

- *Cluster construction:* This module has two initialisation tasks. First it starts to load a flow table (`FTable`) in a time interval T into memory from disk files once the profiling system receives a signal indicating `FTable` is ready. The second task is to group flows in `FTable` associated with the same feature values (i.e., cluster keys) into *clusters*.
- *Adaptive thresholding:* This module analyses the distribution of *flow counts* in the four feature dimensions, and computes a threshold for extracting significant clusters along each dimension.
- *Behaviour profiling:* This module implements a combination of behaviour classification and structural modelling that builds behaviour profiles in terms of communication patterns of significant end hosts and applications.
- *Profile tracking:* This module examines all behaviour profiles built from the profiling system from various aspects to find interesting and suspicious network events.
- *Event analysis engine:* This module analyses a *behaviour profile database*, which includes current and historical behaviour profiles of end hosts and network applications reported by the *behaviour profiling* and *profile tracking* modules in the profiling system.

3.3 Key implementation details

In this section we focus on some key aspects of our implementation for achieving aforementioned design goals.

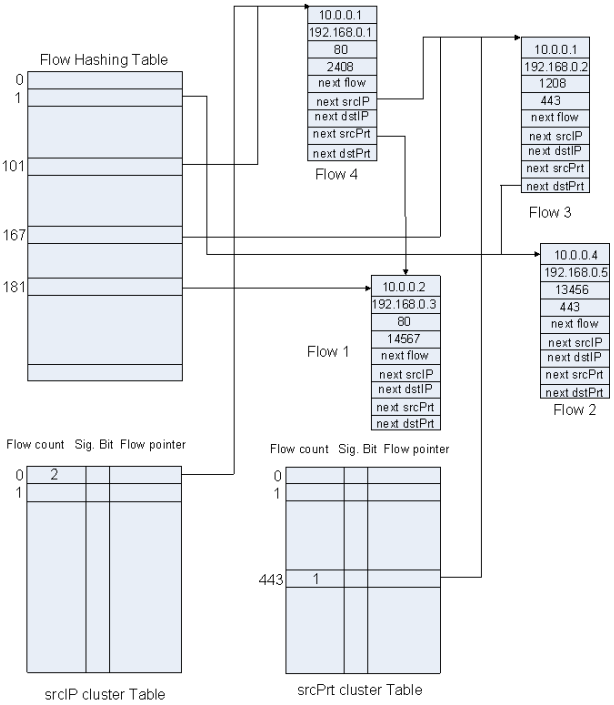
3.3.1 Data structures

High speed backbone links typically carry a large amount of traffic flows. Efficiently storing and searching these flows is critical for the *scalability* of our real-time profiling system. We design two efficient data structures, namely `FTable` and `CTable` for efficient storage and fast lookups during cluster extraction and behaviour modelling.

Figure 4 illustrates the data structure of `FTable` and `CTable` with an example. `FTable`, an array data structure, provides an index of 5-tuple flows through a widely-used hashing function, $FH = srcip \hat{} dstip \hat{} srcport \hat{} dstport \hat{} proto\%$ $(FTableEntries - 1)$, where $FTableEntries$ denotes the maximum entries of `FTable`. For example, in Figure 4, flow 1 is mapped to the entry 181 in `FTable`, while flow 2 is mapped to the entry 1. In case of hashing collision, i.e., two or more flows mapping to the same table entry, we use a linked list to manage them. In our experiments, the (average) collision rate of this flow hashing function is below 5% with $FTableEntries = 2^{20}$. While constructing clusters, the naive approach would be to make four copies of 5-tuple flows, and then group each flow into four clusters

along each dimension. However, this method dramatically increases the memory cost of the system since the flow table typically has hundreds or millions of flows in each time interval. Instead of duplicating flows, which is expensive, we add four flow pointers (i.e., next `srcIP`, next `dstIP`, next `srcPrt` and next `dstPrt`) in each flow. Each flow pointer will link the flows sharing the same feature value in the given dimension. For example, the `next srcIP` pointer of *flow 4* links to `flow 3` since they share the same `srcIP` *10.0.0.1*. Similarly, the `next srcPrt` pointer of *flow 4* links to `flow 1` since they share the same `srcPrt` *80*. However, the question is how to quickly find the 'old' flows of the same clusters when adding a new flow in the flow table.

Figure 4 Data structure of flow table and cluster table (see online version for colours)



To address this problem, we create another data structure, `CTable`, which links the first flow of each cluster in `FTable`. Since there are four types of clusters, we create four instances of `CTable` for managing clusters along four dimensions. Considering `srcPrt` and `dstPrt` dimensions with 65536 possible clusters (ports), we use an array with a size of 65536 to manage the clusters for each of these two dimensions. The index of the array for each port is the same as the port number. For `srcIP` and `dstIP` dimensions, we use a simple hashing function that performs a bitwise exclusive OR (XOR) operation on the first 16 bits and the last 16 bits of `IP` address to map each `srcIP` or `dstIP` into its `CTable` entry. When adding a new flow, e.g., `flow 3` in Figure 4, in the given `dstPrt`, we first locate the first flow (`flow 2`) of the cluster `dstPrt` 443, and make the next `dstPrt` pointer of `flow 3` to `flow 2`. Finally the first flow of the cluster `dstPrt` 443 is updated to `flow 3`. This process is similar for the cluster srcPrt 1208, as well as the clusters `srcIP` *10.0.0.1* and `dstIP` *192.168.0.2*.

In addition to pointing to the first flow in each cluster, each `CTable` entry also includes flow count for the cluster and significant bit for marking significant clusters. The former maintains flow counts for cluster keys. As discussed in Section 2, the flow count distribution will determine the adaptive threshold for extracting significant clusters.

3.3.2 Space and time complexity of modules

The space and time complexity of modules essentially determines the CPU and memory cost of the profiling system. Thus, we quantify the complexity of each module in our profiling system. For convenience, Table 1 shows the definitions of the notations that will be used in the complexity analysis.

Table 1 Notations used in the paper

Notation	*Definition*
F	Set of 5-tuple flows in a time interval
i	Dimension id (0/1/2/3 = `srcIP`/`dstIP`/`srcPort`/`dstPort`
C_i	Set of clusters in dimension i
S_i	Set of significant clusters in dimension i
c_i	A cluster in dimension i
s_i	A significant cluster in dimension i
r_f	Size of a flow record
r_v	Size of the volume information of a cluster
r_b	Size of behaviour information of a significant cluster
r_s	Size of dominant states of a significant cluster

The time complexity of cluster construction is $O\left(|F|+\sum_{i=0}^{3}|C_i|\right)$ for `FTable` and `CTable` constructions. Similarly, the space complexity is $O\left(|F|*s_{fr}+\sum_{i=0}^{3}\left(|C_i|*r_v\right)\right)$.

The time complexity of adaptive thresholding is $\sum_{i=0}^{3}\left(|C_i|*e_i\right)$. This module does not allocate additional memory, since its operations are mainly on the existing `CTable`. Thus, the space complexity is zero.

The time complexity of behaviour profiling is $O\left(\sum_{i=0}^{3}\sum_{j=0}^{|S_i|}|S_j|\right)$, while the space complexity is $O\left(\sum_{i=0}^{3}\left[|S_i|*\left(r_b+r_s\right)\right]\right)$. The output of this step is the behaviour profiles of significant clusters, which are recorded into a database along with the timestamp for further analysis.

Due to a small number of significant clusters extracted, the computation complexity of profile tracking is often less than the others in two or three orders of magnitude, so we will not consider its time and space requirement for simplicity.

3.3.3 Parallelisation of input and profiling

In order to improve the efficiency of the profiling system, we use mechanisms for paralleling tasks in multiple modules, such as continuously importing flow records in the current time interval T_i, and profiling flow records that are complete for the previous time interval T_{i-1}. Clearly, the parallelisation could reduce the time cost of the profiling system. The disadvantage of doing so is that we have to maintain two set of `FTable` and `CTable` for two time intervals.

3.3.4 Event analysis engine

To discover interesting or anomalous network events, we build an event analysis engine with three aspects:

1 temporal behaviour analysis

2 feature dimension correlation

3 event configurations.

The objective of temporal behaviour analysis is to characterise temporal properties of behaviour classes as well as individual clusters from the behaviour profile database that records behaviour profiles built from the profiling system. Prior work in Lakhina et al. (2005) and Xu et al. (2005a) have demonstrated that temporal properties could help distinguish and classify behaviour classes. Feature dimension correlation attempts to find the correlation between clusters from various dimensions to detect emerging exploit and worm activities (Kim and Karp, 2004; Singh et al., 2004) that often trigger new clusters from `srcIP`, `dstIP` and `dstPrt` dimensions.

We develop a simple *event configuration* language that enables network operators or security analysts to extract information on events of interest from behaviour profiles for network management or troubleshooting. To express the policy, we use four distinct fields: dimension, event type, filter and description. The options of these fields include:

- Dimension $\in \{srcIP, dstIP, srcPrt, dstPrt, all\}$
- Event type $\in \{rare, deviant, exploit, unusual\ service, ports, all\}$
- Filter $\in \{high, frequency, high\ intensity, matching\ selected\ ports, ports, others\}$
- Discription $\in \{full, summary\}$

For example, if a network operator wants to monitor *rare* behaviour of `srcIP` end hosts, she could use the rule `srcIP` (dimension) – rare (event type) – all (filter) – full (description), which expresses the policy of reporting *full* profiles of *all* `srcIP` clusters with *rare* behaviour. Similarly, we could construct other filter rules using the combinations of all available options.

In the next section, we will demonstrate the operational feasibility of this system by performing extensive benchmarking of CPU and memory costs using packet-level traces from OC-48 backbone links. To evaluate the robustness of the system, we also test the system against anomalous traffic patterns under DoS attacks or worm outbreaks.

4 Performance evaluation

In this section, we first conduct performance benchmarking of CPU and memory cost of the profiling system using a variety of packet traces from OC-48 backbone links. Subsequently, we evaluate the performance bottlenecks of the system under anomalous traffic patterns such as those caused by DoS attacks and worm outbreaks.

4.1 Benchmarking

We measure CPU usage of the profiling process by using a system call, namely, *getrusage*(), which queries actual system and user CPU time of the process. The system call returns with the resource utilisation including *ru_utime* and *ru_stime*, which represent the user and system time used by the process, respectively. The sum of these two times indicates the total CPU time that the profiling process uses. Let T denote the total CPU time, and T_l, T_a and T_p denote the CPU usage for the modules of cluster construction, adaptive thresholding and behaviour profiling, respectively. Then we have

$$T = T_l + T_a + T_p \tag{3}$$

Similarly, we collect memory usage with another system call, *mallinfo*(), which collects information of the dynamic memory allocation. Let M denote the total memory usage, and Ml, Ma, and Mp denote the memory usage in three key modules. Then we have

$$M = M_l + M_a + M_b \tag{4}$$

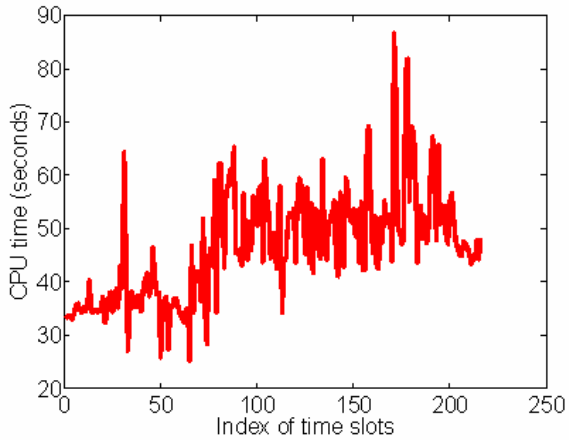
In order to track the CPU and memory usages of each module, we use these two system calls before and after the module. The difference of the output becomes the actual CPU and memory consumption of each module. Next, we show the CPU time and memory cost of profiling system on three OC-48 links during a continuous 18 hours with average link utilisation of 209 Mbps, 86 Mbps, and 78 Mbps. For convenience, let L_1, L_2 and L_3 denote three links, respectively.

Table 2 Total CPU and memory cost of the real-time profiling system on 5-min flow traces

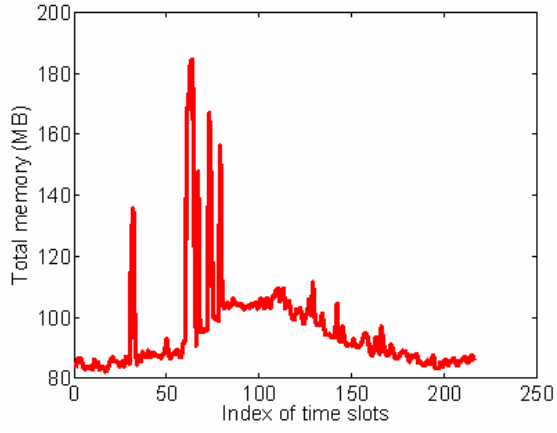
Link	*Util.*	*CPU time (sec)*			*Memory (MB)*		
		min	*avg*	*max*	*min*	*avg*	*max*
L_1	207 Mbps	25	46	65	82	96	183
L_2	86 Mbps	7	11	16	46	56	71
L_3	78 Mbps	7	12	82	45	68	842

Table 2 shows a summary of CPU time and memory cost of the profiling system on L_1 to L_3 for 18 consecutive hours. It is not surprising to see that the average CPU and memory costs for L_1 are larger than the other two links due to a higher link utilisation. Figure 5 shows the CPU and memory cost of the profiling system on all 5-min intervals for L_1 (the link with the highest utilisation). For the majority of time intervals, the profiling system requires less than 60 seconds (1 minute) of CPU time and 150 MB of memory using the flow records in 5-min time intervals for L_1.

Figure 5 CPU and memory cost of the real-time profiling system on flow records in 5-min time interval collected in L_1 for 18 consecutive hours (a) CPU time (b) memory cost (see online version for colours)



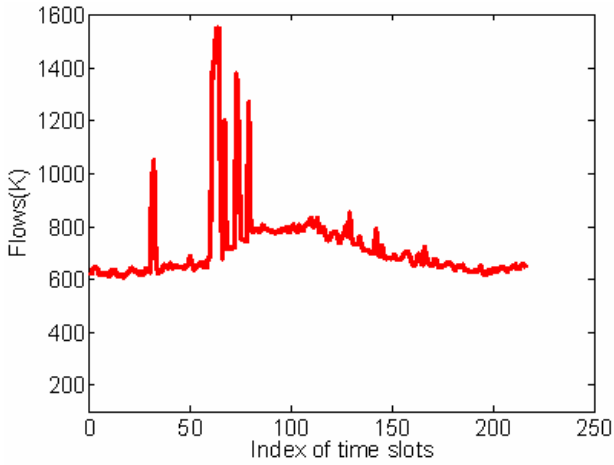
(a)



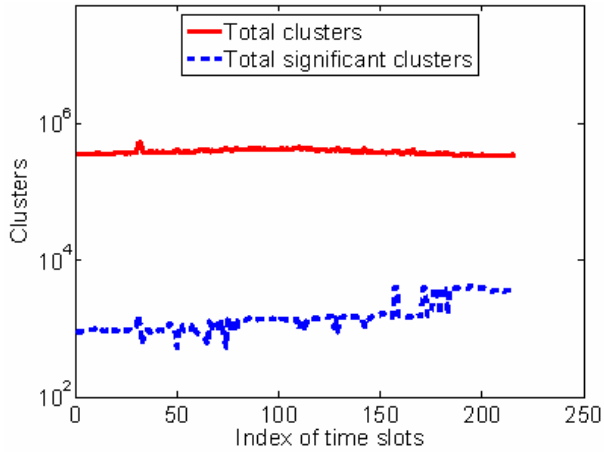
(b)

Figure 6(a) further illustrates the number of flow records over time that ranges from 600 K to 1.6 M, while Figure 6(b) shows the number of all clusters as well as the extracted significant clusters. It is very interesting to observe the similar patterns in the plot of memory cost [Figure 5(b)] and that of the flow count over time [Figure 6(a)]. This observation leads us to analyse the correlation between these two measurements. By examining the breakdown of the memory cost, we find that M_l in the cluster construction module accounts for over 98% of the total memory consumptions. Recall that the space complexity of this module is larger than the others by two or three orders of magnitude, and dominated by the factor in terms of size of flow table $|F|$. The scatter plot of $|F|$ vs. M_l in Figure 7 reflects the linear relationship between them. Therefore, this strong correlation suggests that the memory cost of the profiling system is mainly determined by the number of flow records collected by the monitoring system in the given time interval.

Figure 6 Input of flow traces in 5-min time interval collected in L_1 for 18 consecutive hours (a) size of `FTable` (b) number of clusters (see online version for colours)



(a)



(b)

Figure 7 Correlation between memory cost in cluster constructions and size of `FTable` (see online version for colours)

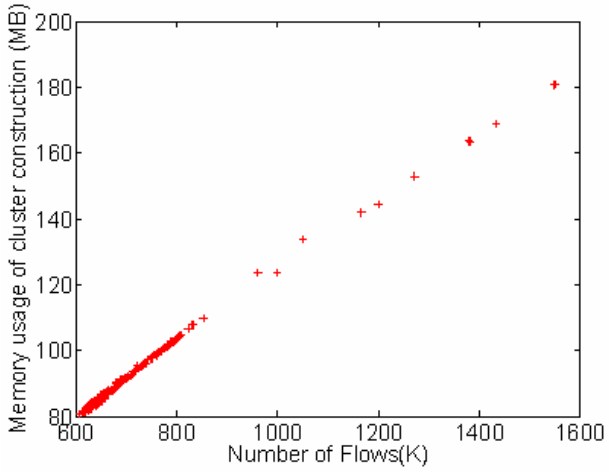
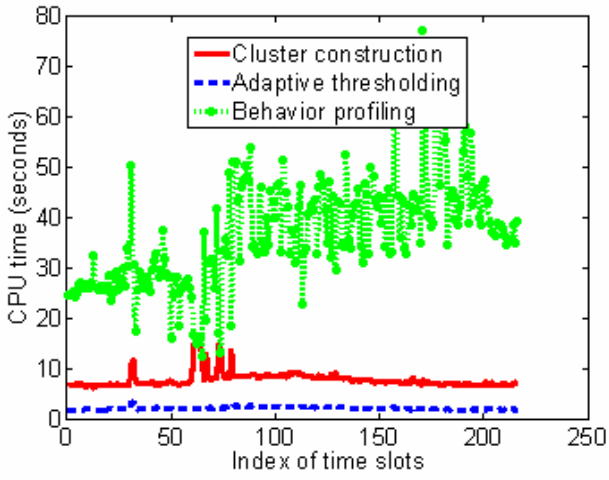


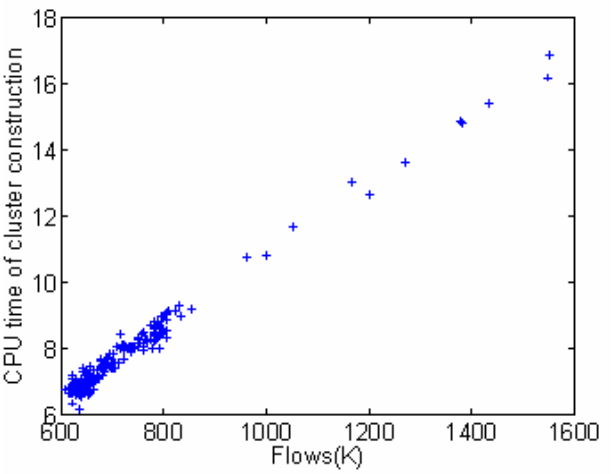
Figure 8(a) shows a breakdown in CPU usage of the various modules in the profiling system, and suggests that cluster construction and behaviour profiling account for a large fraction of CPU time. Similar to the space complexity, the time complexity in cluster construction is also determined by $|F|$. The linear relationship demonstrated by the scatter plot of $|F|$ vs. T_l in Figure 8[b] confirms this complexity analysis. Figure 8(c) shows the scatter plot of the number of significant clusters vs. CPU time in behaviour profiling. Overall, we observe an approximately linear relationship between them. This suggests that the CPU cost in behaviour profiling is largely determined by the number of significant clusters whose behaviour patterns are being analysed.

Figure 8 Breakdown of CPU time and correlations (a) breakdown of CPU time (b) correlation between size of `FTable` and CPU time in cluster construction (c) correlation between size of `FTable` and CPU time in behaviour profiling (see online version for colours)

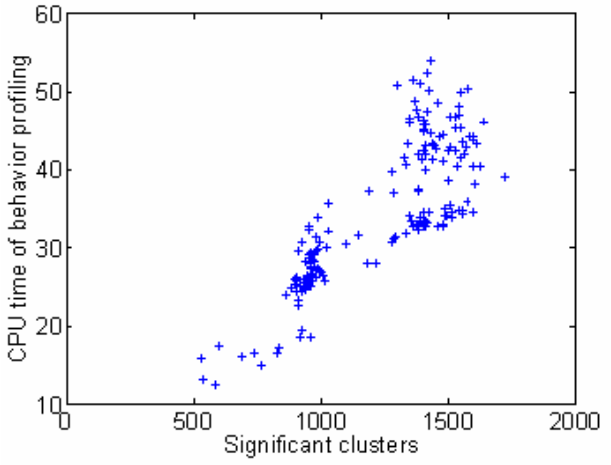


(a)

Figure 8 Breakdown of CPU time and correlations (a) breakdown of CPU time (b) correlation between size of `FTable` and CPU time in cluster construction (c) correlation between size of `FTable` and CPU time in behaviour profiling (continued) (see online version for colours)



(b)



(c)

Table 3 Total CPU and memory cost with various time granularities

Link	*Util.*	*CPU time (sec)*			*Memory (MB)*		
		min	*avg*	*max*	*min*	*avg*	*max*
L_1	1 min	5	12	31	26	36	76
L_2	2 min	10	22	42	31	46	89
L_3	5 min	25	46	65	82	96	183
L_1	10 min	35	82	129	91	151	208
L_1	15 min	45	88	152	145	218	267

To understand how performance is affected by time granularity, we also evaluate the system on L_1 using 1 min, 2 min, 10 min and 15 min flow traces. The results are shown in Table 3. In general, the CPU time and memory cost increase as the length of the time interval. On the other hand, the CPU time of the profiling system is always less than the time interval T. In addition, the average memory cost for 5 min, 10 min and 15 min are 96 MB, 151 MB, and 218 MB, respectively which are within the affordable range on commodity PCs. These results clearly suggest that our real-time profiling system satisfies the *scalability* requirement raised in the previous section.

In summary, the average CPU and memory costs of the real-time profiling system on 5 min flow records collected from an OC-48 link with a 10% link utilisation are 60 seconds and 100 MB, respectively. Moreover, the CPU time is largely determined by the number of flow records as well as that of significant clusters, and the memory cost is determined by the number of flow records. During these monitoring periods, these links are not fully utilised, so we can not extensively measure the performance of the real-time profiling system for a highly loaded link. Next, we will test the profiling system during sudden traffic surges such as those caused by DoS attacks, flash crowds, and worm outbreaks that increase the link utilisation as well as the number of flow records.

4.2 Stress test

The performance benchmarking of CPU and memory costs demonstrates the operational feasibility of our traffic profiling system during normal traffic patterns. However, the profiling system should be robust during atypical traffic patterns, such as DoS attacks, flash crowds and worm outbreaks (Jung et al., 2002; Kandula et al., 2005; Moore et al., 2003; Zou et al., 2003). In order to understand the system performance during these incidents, we inject packet traces of three known DoS attacks and simulated worm outbreaks by superposing them with backbone traffic.

We use the packet traces of three DoS attacks with varying intensity and behaviour studied in Hussain et al. (2003). All of these attacks are targeted on a single destination IP address. The first case is a multiple-source DoS attack, in which hundreds of source IP addresses send 4200 ICMP echo request packets per second for about 5 minutes. The second case is a TCP SYN attack lasting 12 minutes from random IP addresses that send 1100 TCP SYN packets per second. In the last attack, a single source sends over 24 K *ip-proto* 255 packets per second for 15 minutes. In addition to DoS attacks, we simulate the SQL slammer worm on 25th January 2003 (Moore et al., 2003) with an internet worm propagation simulator used in Zou et al. (2003). In the simulation experiments, we adopt the same set of parameters in Zou et al. (2003) to obtain similar worm simulation results, and collect worm traffic monitored in a 2^{20} IP space.

For each of these four anomalous traffic patterns, we replay packet traces along with backbone traffic, and aggregate synthetic packets traces into 5-tuple flows. For simplicity, we still use 5 minutes as the size of the time interval, and run the profiling system against the flow records collected in an interval. Table 4 shows a summary on flow traces of the first 5-minute interval for these four cases. The flow, packet and byte counts reflect the intensity of attacks or worm propagation, while the link utilisation indicates the impact of such anomaly behaviours on internet links. For all of these cases, the profiling system is able to successfully generate event reports in less than 5 minutes.

Table 4 Synthetic packet traces with known DoS attacks and worm simulations

Anomaly	*Flows*	*Packets*	*Bytes*
DoS-1	2.08 M	18.9 M	11.8 G
DoS-2	1.80 M	20.7 M	12.5 G
DoS-3	16.5 M	39.8 M	16.1 G
Worm	18.9 M	43.0 M	23.6 G
Link utilisation	*CPU time*	*Memory*	*Details*
314.5 Mbps	45 seconds	245.5 MB	Distributed DoS attacks from multiple sources
333.4 Mbps	59 seconds	26631 MB	Distributed DoS attacks from random sources
430.1 Mbps	210 seconds	1.75 GB	DoS attacks from single source
629.2 Mbps	231 seconds	2.01 GB	Slammer worm simulations

For the synthetic traffic, the link utilisation ranged from 314.5 Mbps to 629.2 Mbps. We run the profiling system on flow traces after replaying synthetic packets and collect CPU and memory cost of each time interval, which is also shown in Table 4. The system works well for low intense DoS attacks in the first two cases. However, due to intense attacks in the last DoS case (DoS-3) and worm propagations, the CPU time of the system increases to 210 and 231 seconds, but still under the 5 minute interval. However, the memory cost jumps to 1.75 GB and 2.01 GB indicating a performance bottleneck. This clearly suggests that we need to provide practical solutions to improve the robustness of the system under stress. In the next section, we will discuss various approaches, including traditional sampling techniques and new profiling-aware filtering techniques towards this problem, and evaluate the tradeoff between performance benefits and profiling accuracy.

5 Sampling and filtering

In this section, we first adopt traditional sampling techniques to address performance bottleneck during sudden traffic surges as caused by severe DoS attacks or worm outbreaks. After evaluating its strength and limitation, we

propose a simple yet effective profiling-aware filtering algorithm that not only reduces memory cost, but also retains profiling accuracy.

5.1 Random sampling

Random sampling is a widely-used simple sampling technique in which each object, flow in our case, is randomly chosen based on the same probability (also known as sampling ratio μ). Clearly, the number of selected flows is entirely decided by the sampling ratio μ. During the stress test in the last section, the profiling system requires about 2 GB memory when the number of flow records reach 16.5 M and 18.9 M during DoS attacks and worm outbreaks. Such high memory requirement is not affordable in real-time since the machine installed with the profiling system could have other tasks as well, e.g., packet and flow monitoring. As a result, we attempt to set 1 GB as the upper bound of the memory cost. Recall that in the performance benchmarking, we find that memory cost is determined by the number of flow records. Based on their linear relationship shown in Figure 7 we estimate that flow records with a size of 10 M will require approximately 1 GB memory. Thus, 10 M is the desirable limit for the size of the flow records.

Using the limit of flow records, l, we could configure the sampling ratio during sudden traffic increase as $\mu = \frac{l}{|F|}$. As a result, we set the sampling ratios in the last DoS attacks and worm outbreaks as 60% and 55%, respectively, and randomly choose flows in loading flow tables in the *cluster construction* module. Table 5 shows the reduction of CPU time and memory consumptions with the sampled flow tables for both cases.

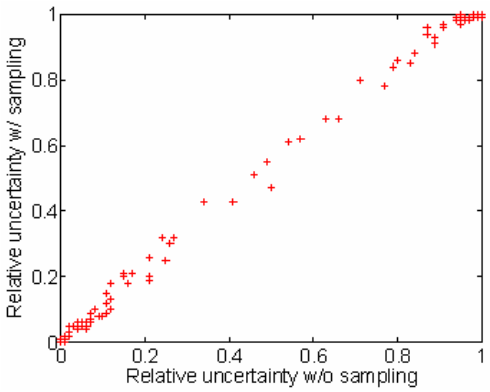
Table 5 Reduction of CPU time and memory cost using the random sampling technique

Case	μ	*Size of FTable*	*CPU time*	*Memory*
DoS	66%	10 M	89 seconds	867 MB
Worm	55%	10 M	97 seconds	912 MB

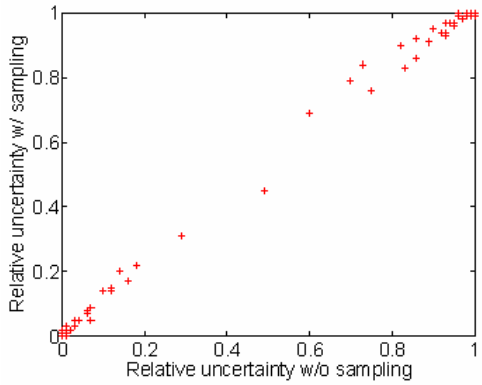
Table 6 Reduction of significant clusters and behaviour accuracy

Dim.	*Significant clusters without sampling*	*Significant clusters with sampling*	*Clusters with same behavior classes*
srcPrt	23	4	3
dstPrt	6	5	4
srcIP	47	31	29
dstIP	233	140	125
Total	309	180	161

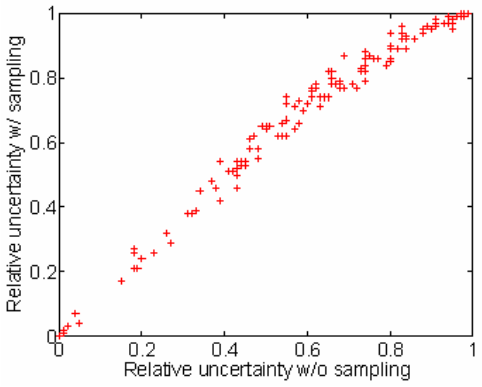
Figure 9 Feature distribution of free dimensions for 140 `dstIP` clusters with and without random sampling (a) `srcPort` (b) `dstPort` (c) `srcIP` (see online version for colours)



(a)



(b)



(c)

On the other hand, random sampling has substantial impact on behaviour accuracy. First, the set of significant clusters from four feature dimensions are smaller than that without sampling, which is caused by the changes of the underlying cluster size distribution after flow sampling. Table 6 shows the number of significant clusters extracted along each dimension without and with sampling for the DoS case. In total, among 309 significant clusters without sampling, 180 (58%) of the *most significant* clusters are extracted with random sampling. Secondly, the behaviour of a number of extracted clusters are altered, since flow sampling changes the feature distribution of free dimensions as well as the

behaviour classes for these clusters. As shown in the last column of Table 6, 161 out of 180 significant clusters with random sampling are classified with the same behaviour as those without sampling. In other words, the behaviour of 19 (10.5%) extracted significant clusters are changed as a result of random sampling. Figure 9 shows the feature distributions of free dimensions for 140 `dstIP` clusters with and without random sampling. The deviations from the diagonal line indicate the changes of feature distribution and the behaviour due to flow sampling. We also perform random sampling on the synthetic flow traces in the case of worm outbreak, and the results of sampling impact on cluster extractions and behaviour accuracy are very similar.

In summary, random sampling could reduce the CPU time and memory cost during sudden traffic surges caused by DoS attacks or worm outbreaks. However, random sampling reduces the number of interesting events, and changes behaviour classes of a number of significant clusters. Such impact could have become worse if finer sampling granularities are selected. Thus, it becomes very necessary to develop a profiling-aware algorithm that not only reduces the size of flow tables, but also retains the (approximately) same set significant clusters and their behaviour.

5.2 Profiling-aware filtering

A key lesson from random sampling is that the clusters associated with DoS attacks are usually very large in flow count, and hence consume a large amount of memory and CPU time (Argyraki and Cheriton, 2005). In addition, profiling such behaviour does not require a large number of flows, since the feature distributions very likely remain the same even with a small percentage of traffic flows. Based on this insight, we develop a profiling-aware filtering solution that limits the size of very large clusters, and adaptively samples on the rest of clusters when the system is faced with sudden explosive growth in the number of flows.

Algorithm 1 describes the details of the profiling-aware sampling algorithm. First, we choose two *watermarks* (L and H) for the profiling system. L represents the moving average of flow tables over time, and H represents the maximum size of flow tables that system will accept. In our experiments, we set $H = 10M$, which is estimated to require 1 GB memory cost. In addition, we set the maximum and minimum sampling ratios, i.e., $\mu_{\max}$ and $\mu_{\min}$. The actual sampling μ will be adaptively decided based on the status of flow table size. Specifically, the sampling ratio becomes thinner as the size of flow table increases. For simplicity, let `ftable` denote the size of flow table. If `ftable` is below L, the profiling system accepts every flow. In contrary, if `ftable` is equal to H, the system will stop reading flows and exit with a warning signal.

If `ftable` is equal to L or certain intermediate marks, i.e., $L + i * D$, where D is the incremental factor and $i = 1, 2, \ldots, (H - L)/D - 1$, the system computes the RU of each dimension and evaluates whether there is one or a few dominant feature values along each dimension. In our experiments, we set $D = 1M$ as the incremental factor. The existence of such values suggests that certain types of flows dominate current flow tables, and indicates anomalous traffic patterns. Thus, the system searches these values and marks them as significant clusters for flow filtering. Subsequently, any flow, which contains a feature value marked with *significant*, will be filtered, since such flow will not likely contribute much for the behaviour of the associated clusters. Instead, we should give preference to flows that belong to other small clusters. On the other hand, the system could not accept all of these flows with preference after $ftable$ exceeds L watermark. As a result, each of these flows is added with the adaptive sampling ratio

$$\mu = \mu_{\max} - i * \frac{\mu_{\max} - \mu_{\min}}{(H - L)/D - 1}. \tag{5}$$

5.2.1 Evaluations

We run the profiling system on the flow tables in the cases of DoS attack and worm outbreaks (cf. Table 6) with the profile-aware filtering algorithm. Similar to random sampling, this sampling solution could also reduce CPU time and memory cost due to a small size of flow table. However, the profiling-aware sampling has two advantages compared with random sampling. Firstly, the set of clusters extracted using this algorithm is very close to the set without sampling. For example, in the case of DoS attack, the system obtains 41 `srcIP` clusters, 210 `dstIP` clusters, 21 `srcPrt` clusters and 6 `dstPrt` cluster, respectively. Compared with 58% of significant clusters extracted in random sampling, our profiling-aware algorithm could extract over 90% of 309 original clusters that are selected without any sampling.

Algorithm 1 A Profiling-aware filtering algorithm

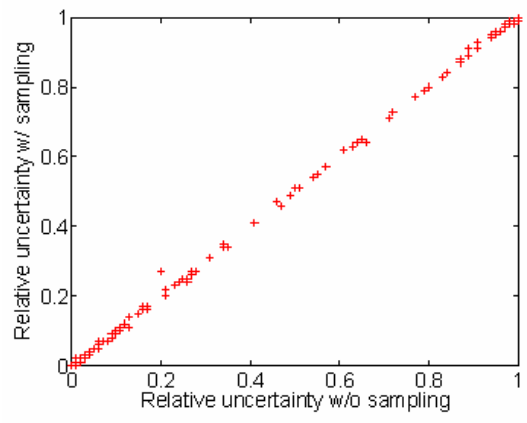
```
1   Parameters: L, H, D, μ_max, μ_min, β*
2   Initilisation: I = (H − L)/D
    δ_μ = (μ_max − μ_min)/I
    μ = μ_max
    i = 0
    ftable = 0
3     srcPrt = 0; dstPrt = 1
4     srcIP = 2; dstIP = 3
5     while next flow do
6     if (ftable < L) then
7         Insert flow into FTable (Flow Table)
8         ftable++
9         continue;
10    else
11        if (ftable > H) then
```

```
12          Stop reading flows
13          exit
14        end if
15      end if
16      if (ftable == L + i * D) then
17        ru_0 = Relative_Uncertainty(FTable, srcPrt)
18        ru_1 = Relative_Uncertainty(FTable, dstPrt)
19        ru_2 = Relative_Uncertainty(FTable, srcIP)
20        ru_3 = Relative_Uncertainty(FTable, dstIP)
21        for (dim = 0; dim ≤ 3; dim++) do
22          ru = ru_dim
23          while (ru ≤ β*) do
24            remove feature value with highest
              probability
25            mark feature value as significant
26            ru = Relative_Uncertainty(FTable, dim)
27          end while
28        end for
29        μ = μ_max − i * δ_μ
30        i++
31      end if
32      if ((ftable ≥ L) && (ftable ≤ H)) then
33        if (any feature value in flow is marked
          significant) then
34          Filter flow flow
35        else
36          Insert flow into FTable with probability μ
37          if (flow is selected then
38            ftable++
39          end if
40        end if
41      end if
42    end while
```

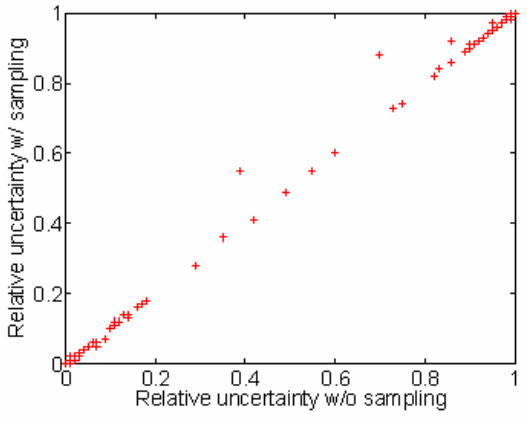
Secondly, the behaviour accuracy of significant clusters are also improved. Specifically, among 41 `srcIP`'s, 210 `dstIP`'s, 21 `srcPrt`'s, and 6 `dstPrt`'s significant clusters, only 3 `dstIP`'s and 1 `srcPrt` exhibit changes in behaviour classes. This finding highly suggests that the profiling-aware profiling algorithm successfully retains the feature distributions of those clusters and behaviours.

Figure 10 shows the feature distribution of free dimensions of 210 `dstIP` clusters, extracted both without sampling and with profiling-aware filtering algorithm. In general, the feature distributions of all free dimensions for almost all clusters after filtering are approximately the same as those without sampling. The outliers deviant from the diagonal lines correspond to feature distributions of three clusters whose behaviour has changed. Upon close examinations, we find that flows in these clusters contain a mixture of web and ICMP traffic. The latter are the dominant flows in DoS attacks, so they are filtered after the size of flow table reaches L in the profiling-aware filtering algorithm. The filtered ICMP flows in these clusters explain the changes of the feature distributions as well as the behaviour classes.

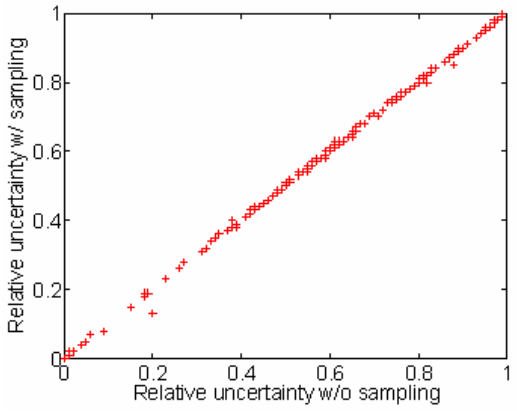
Figure 10 Feature distribution of free dimensions for 210 `dstIP` clusters with and without profiling-aware sampling (a) `srcPort` (b) `dstPort` (c) `srcIP` (see online version for colours)



(a)



(b)



(c)

In the *worm* case, the profiling-aware filtering algorithm also successfully reduces CPU and memory cost of the profiling system, while maintaining high profiling accuracy in terms of the number of extracted significant clusters and the feature distributions of these clusters. Thus, the profiling-aware filtering algorithm can achieve a significant reduction of CPU time and memory cost during anomalous traffic patterns while obtaining accurate behaviour profiles of end hosts and network applications.

6 Reducing unwanted exploit traffic

The prevalent exploit behaviour indicates the severity of worm or exploits spread and the magnitude of infected hosts (cf. Yegneswaran et al., 2003; Pang et al., 2004; SNORT, 2009; Paxson, 1999; Staniford et al., 2002)). Given the exploit traffic identified through our real-time profiling system, we devise several heuristic rules of blocking strategies in the profiling system, and then evaluate their efficacy in reducing unwanted traffic (Xu et al., 2005b). In order to determine which sources to block traffic from, we first identify all sources that exhibit exploit behaviour. We then devise simple rules to select some or all of these sources as candidates for blocking. Instead of blocking all traffic from the selected sources, we consider blocking traffic on only the ports that source seek to exploit. This is because exploit hosts may indeed be sending a mixture of legitimate and exploits traffic. For example, if an infected host behind a NAT box is sending exploit traffic, then we may observe a mixture of legitimate and exploit traffic coming from the single IP address corresponding to the NAT box.

For our evaluation, we start with the following benchmark rule. If a source is profiled as an exploit source during any five minute interval, then all traffic from this source on vulnerable ports is blocked from then on. In general, the benchmark rule could block about 80% of unwanted traffic from these exploit sources. In other words, this rule may still not block all traffic from the source due to two reasons. First, the source might already have been sending traffic, perhaps legitimate, prior to the time-slot in which it exhibited the exploit profile. Second, as explained above, only ports on which we see exploit traffic are considered to be blocked.

While this benchmark rule is very aggressive in selecting sources for blocking, the candidate set of source/port pairs to be added to ACL list of routers may grow to be very large across all links in a network. ACL is a scarce resource of network routers, thus we devise and evaluated other blocking rules that embody more restrictive criteria that an exploit source must satisfy in order to be selected for blocking. We introduce three metrics, *cost*, *effectiveness* and *wastage* to evaluate the efficacy of these rules. The cost refers to the overhead incurred in a router to store and lookup the ACLs of blocked sources/ports. For simplicity, we use the total number of sources/ports as an index of the overhead for a blocking rule. The effectiveness measures the reduction of unwanted traffic in terms of flow, packet and byte counts compared with the benchmark rule. The resource wastage refers to the number of entries in ACLs that are never used after creations. Based on the evaluations, we find that the rule of blocking aggressive sources (An ACL entry is created if and only if the source has an average intensity of at least 300 flows per minute.) to be very effective.

We apply this blocking strategy and the benchmark rule in the real-time profiling system and replay a 24 hour packet trace collected from an OC-48 internet backbone link. The benchmark rule achieves the optimal performance, but has the largest cost, i.e., 3756 blocking entries, 34.8% of which are never used after creation. The selected blocking strategy reduces 84.3% unwanted flows, 80.4% unwanted packets, and 72.7% unwanted bytes, respectively. The strategy has a cost of 1789 ACL entries and over 83% of entries are applied at least once to block similar exploit traffic from the same attacker. We believe that that this simple rule is cost-effective when used to block the aggressive or frequent sources that send a majority of self-propagating exploit traffic, in particular, in the early stage of a malware outbreak, to hinder their spread. The similar observations also hold for other internet backbone links.

In summary, our experiment results demonstrate that blocking the most offending sources in the real-time profiling system is reasonably cost-effective and can potentially stop self-propagating malware in their early stage of outburst.

7 Conclusions

This paper explores the feasibility of designing, implementing and utilising a real-time behaviour profiling system for high-speed internet links. We first discuss the design requirements and challenges of such a system and present an overall architecture that integrates the profiling system with always on monitoring systems and an event analysis engine. Subsequently, we demonstrate the operational feasibility of building this system through extensive performance benchmarking of CPU and memory costs using a variety of packet traces collected from OC-48 backbone links. To improve the robustness of this system during anomalous traffic patterns such as DoS attacks or worm outbreaks, we propose a simple yet effective filtering algorithm to reduce resource consumptions while retaining high profiling accuracy. Through simple yet effective ACL rules, we demonstrate the application of the real-time profiling system to reduce a significant amount of unwanted exploit traffic.

Acknowledgements

The work is supported in part by the National Science Foundation grants CNS-0626808 and CNS-0626812, a Sprint ATL gift grant, a University of Minnesota DTC DTI grant, and an Arizona State University New College SRCA grant.

References

Argyraki, K. and Cheriton, D. (2005) 'Active internet traffic filtering: real-time response to denial-of-service attacks', in *Proceedings of USENIX Annual Technical Conference*, April.

Cavallo, R. and Klir, G. (1979) 'Reconstructability analysis of multi-dimensional relations: a theoretical basis for computer-aided determination of acceptable systems models', *International Journal of General Systems*, Vol. 5, pp.143–171.

Hussain, A., Heidemann, J. and Papadopoulos, C. (2003) 'A framework for classifying denial of service attacks', in *Proceedings of ACM SIGCOMM*, August.

Iannaccone, G. (2005) 'CoMo: an open infrastructure for network monitoring – research agenda', Intel Research Technical Report, February.

Iannaccone, G., Diot, C., Graham, I. and McKeown, N. (2001) 'Monitoring very high speed links', in *Proceedings of ACM SIGCOMM Internet Measurement Workshop*, November.

Jordan, M. (2004) 'Graphical models', *Statistical Science, Special Issue on Bayesian Statistics*, Vol. 19, pp.140–155.

Jung, J., Krishnamurthy, B. and Rabinovich, M. (2002) 'Flash crowds and denial of service attacks: characterization and implications for CDNs and web sites', in *Proceedings of International World Wide Web Conference*.

Kandula, S., Katabi, D., Jacob, M. and Berger, A. (2005) 'Botz-4-sale: surviving organized DDoS attacks that mimic flash crowds', in *Proceedings of Symposium on NSDI*, May.

Karagiannis, T., Papagiannaki, K. and Faloutsos, M. (2005) 'BLINC: multilevel traffic classification in the dark', in *Proceedings of ACM SIGCOMM*, August.

Keys, K., Moore, D. and Estan, C. (2005) 'A robust system for accurate real-time summaries of internet traffic', in *Proceedings of ACM SIGMETRICS*, June.

Kim, H. and Karp, B. (2004) 'Autograph: toward automated, distributed worm signature detection', in *Proceedings of USENIX Security Symposium*, August.

Krippendorff, K. (1986) *Information Theory: Structural Models for Qualitative Data*, Sage Publications.

Lakhina, A., Crovella, M. and Diot, C. (2005) 'Mining anomalies using traffic feature distributions', in *Proceedings of ACM SIGCOMM*, August.

Moore, D., Paxson, V., Savage, S., Shannon, C., Staniford, S. and Weaver, N. (2003) 'Inside the slammer worm', *IEEE Security and Privacy*, July.

Pang, R., Yegneswaran, V., Barford, P., Paxson, V. and Peterson, L. (2004) 'Characteristics of internet background radiation', in *Proceedings of ACM SIGCOMM Internet Measurement Conference*, October.

Paxson, V. (1999) 'Bro: a system for detecting network intruders in real-time', *Computer Networks*, Vol. 31, pp.2435–2463, December.

Singh, S., Estan, C., Varghese, G. and Savage, S. (2004) 'Automated worm fingerprinting', in *Proceedings of Symposium on Operating Systems Design and Implementation*, December.

SNORT (2009) available at http://www.snort.org/.

Staniford, S., Hoagland, J. and McAlerney, J. (2002) 'Practical automated detection of stealthy portscans', *Journal of Computer Security*, Vol. 10, pp.105–136.

Xu, K., Wang, F., Bhattacharyya, S. and Zhang, Z-L. (2007) 'A real-time network traffic profiling system', in *Proceedings of International Conference on Dependable Systems and Networks*, June.

Xu, K., Zhang, Z-L. and Bhattacharyya, S. (2005a) 'Profiling internet backbone traffic: behavior models and applications', in *Proceedings of ACM SIGCOMM*, August.

Xu, K., Zhang, Z-L. and Bhattacharyya, S. (2005b) 'Reducing unwanted traffic in a backbone network', in *Proceedings of Steps to Reducing Unwanted Traffic on the Internet Workshop (SRUTI)*, July.

Yegneswaran, V., Barford, P. and Ullrich, J. (2003) 'Internet intrusions: global characteristics and prevalence', in *Proceedings of ACM SIGMETRICS*, June.

Zou, C., Gao, L., Gong, W. and Towsley, D. (2003) 'Monitoring and early warning for internet worms', in *Proceedings of ACM CCS*, October.

Zwick, M. (2004) 'An overview of reconstructability analysis', *International Journal of Systems & Cybernetics*.

Notes

1 A preliminary version of this paper appeared in the *Proceedings of IEEE/IFIP International Conference on Dependable Systems and Networks 2007*.