Leveraging Traffic Repetitions for High-Speed Deep Packet Inspection

INFOCOM 2015 Paper #54

Abstract— **Deep Packet Inspection (DPI) plays a major role in contemporary networks, and specifically, in datacenters of content providers, scanned data may be highly repetitive. Most DPI engines are based on identifying signatures in packet payload. This pattern matching process is expensive both in memory and CPU resources, and therefore, often becomes the bottleneck of the entire application.**

This paper shows how DPI can be accelerated by leveraging repetitions in the inspected traffic. We first show that such repetitions exist in many traffic types and present a mechanism that allows skipping repeated data instead of scanning it again. In its slow path, frequently repeated strings are identified and stored in a dictionary along with some succinct information for accelerating the DPI process. In the mechanism's data path, each time the scanning algorithm encounters a string from the dictionary, it skips it and recovers to the correct state had this word been scanned byte by byte.

Our solution achieves significant performance boost, especially when data is of the same content source (e.g. same website). Our experiments show that for such cases, our solution achieves throughput gain of $1.25-2.5$ times the original throughput, when implemented in software.

I. Introduction

Content providers, such as Internet Service Providers (ISPs), Google, and Netflix maintain datacenters to host their content, or their customers' content. Usually, such providers also maintain monitoring appliances such as network intrusion detection systems (NIDS), content filtering (such as parental control services), spam filtering, and more. All these appliances scan the payload of packets in a process known as Deep Packet Inspection (DPI). In addition, providers sometimes use *Layer 7 routing*, which relies as well on scanning the application layer header, and is performed using similar techniques.

Perhaps the most significant technique used in today's DPI engines is *signature matching*, in which the payload of the packet is compared against a predetermined set of *patterns* (with exact strings or regular expressions), which should alert on protocol non-compliance, viruses, spam, intrusions, and so on. Signature matching is a well-established subject in Computer Science since the seventies, and usually involves a *memoryless* scanning of the packets. For example, the widely-used Aho-Corasick algorithm builds a *Deterministic Finite Automaton* (DFA) to represent the set of patterns; each byte of the packet causes a transition in that DFA, and a pattern is found if the DFA transits to an accepting state in the automaton. Evidently, when scanning a byte using the Aho-Corasick algorithm, only the current state of the automaton is used. Informally speaking, this implies that no information of other packets, or different fragments of the same packet, is used to enhance the scanning process. Specifically, even if *the same packet arrives at the DPI engine many times*, the engine will *always scan it from scratch*.

On the other hand, a closer look at Internet traffic, and specifically HTTP traffic, clearly indicates many repetitions. Such repetitions can be classified either as *full repetitions*, in which the entire object (e.g., image, stylesheet, javascript) appears several times, or *partial repetitions*, in which only shorter fragments (e.g., shared HTML code) appear in many packets or sessions.

In content providers' networks, most of the data is highly similar and many times it is simply the same files, or files with minimal modifications, that are being sent over the network. Moreover, recent trends in content providers' networks include *Software Defined Networking* (SDN), where routing is based on multiple, arbitrary header fields. Several suggestions to make SDNs aware of application layer information has been proposed [1], and thus we envision that DPI will get higher attention as a new bottleneck for such networks. Another interesting direction of content providers' networks is *Network Function Virtualization* (NFV), where network functions such as monitoring appliances are virtualized for higher flexibility and scalability. In some cases, these virtual appliances scan traffic from a closed set of servers or even a single server that serves several virtual machines. Thus, the similarity between pieces of data to be scanned is relatively very high. Moreover, using SDN one can make traffic flow so that similar traffic (from similar sources) flow to the same monitoring appliances.

Our paper presents a mechanism that uses such repetitions efficiently in order *to accelerate the signature matching component of the DPI engine*. Our mechanism is based solely on modifications to the signature matching algorithm, and thus does not involve any change to the inspected traffic and does not require any cooperation from any other component in the network. Conceptually, it is divided to two parts: a *slow path* that samples the traffic and creates a *dictionary* with the fixed-length popular strings (which we call *grams*), and a *data path* that scans the traffic byte by byte and checks the dictionary for matches; if a gram is found in the dictionary, the data path *skips* the gram and adjusts its state according to an information saved along this gram.

Specifically, our solution is based on the DFA-based Aho-Corasick algorithm. In the slow path, we save the state of the automaton after scanning the saved gram from the initial automaton's state. In the data path, we show that after skipping

a gram, one should continue scanning from that saved state.[1] To accelerate the data path operations, we use a *bloom filter* that represents the set of grams in the dictionary. Since bloom filters are compact data structures, they reside in fast memory and therefore, reduce the overhead presented by our mechanism in case there is no match in the dictionary. We further note that our mechanism is generic and can be implemented either in software or in hardware. In software implementation, the data path is implemented as a thread, while the slow path is implemented as another thread, possibly with lower priority. In a typical multi-core, multi-threaded environment, our solution uses a single slow-path thread that gets packet samples and calculates dictionaries, and many data-path threads (possibly on many cores), each inspecting different packets (or different connections). On a hardware implementation, on the other hand, we can parallel the operation in finer granularity (for example, checking the bloom filter in parallel with scanning a byte), which can lead to a significant performance boost. Section IV presents a model for the performance gain by our mechanism. As we shall show, this gain depends on various system parameters (such as memory access time), traffic parameters (such as the amount of repetitions), and mechanism parameters (such as the length of grams). We then measure these parameters in our software implementation (Section VI) and show its performance boost. Finally, by those measurements we deduce what was the performance gain had our mechanism implemented in hardware.

One of the significant challenges in implementing our mechanism is deciding which grams should be saved in the dictionary at a given time. We chose to implement a variation of the algorithm suggested in [2] which is able to efficiently find the most popular strings of variable length. We then chop the strings to fixed-length grams and those in the dictionary, as fixed-length grams are easier to handle in the data path. Naturally, the performance boost gained by our mechanism depends on the inspected traffic. We provide analysis and experimental results for several use-cases that describe real-life situations in which DPI is used, and discuss the potential speedup of our mechanism when scanning such traffic.

II. Related Work

Deep packet inspection (DPI) relies on a string matching algorithm, which is an essential building block for numerous other applications as well. Therefore, it has been extensively studied [3]. Some of the fundamental algorithms are Boyer-Moore [4], Aho-Corasick [5] and Wu-Manber [6]. The seminal algorithm of Aho-Corasick (AC) is the de-facto standard for pattern matching in bump-in-the-wire. The AC algorithm constructs a Deterministic Finite Automaton (DFA) for detecting all occurrences of a given set of patterns by processing the input in a single pass. The input is inspected byte by byte. We describe the algorithm in details in section III-A. The string matching algorithm is often a bottle-neck of the system.

[1]Small modifications, which are explained in Section III-B, are required to avoid missing patterns in these skips.

There is an extensive research on accelerating the DPI process, both using hardware and software implementations. The hardware implementations [7]–[11] usually use some special-purpose hardware such as FPGA or a CAM/TCAM. These solutions are usually hard to reprogram, and it is usually complex to update their signatures set. They also tie the engine to a specific type of hardware, which might harden embedding of these solutions. On the other hand, software implementations [12]–[18] are easy to apply, to reprogram and to update, yet have obvious performance disadvantage being implemented on a general purpose system. All of these works *are orthogonal to our work* in a sense that all of them can be applied on top of our engine to further accelerate the DPI process.

Web traffic has many repetitions as we detail below. In this paper we leverage these repetitions to accelerate the DPI process. Another approach that also leverages traffic repetitions is de-duplication. Network data de-duplication is used to reduce the number of bytes that must be transferred between endpoints, resulting in reducing the required bandwidth [19]–[28]. In these works, authors find a redundancy of 35%-45% in general traffic and up to 90% redundancy in web traffic, depends on the type of the traffic.

This paper presents an algorithm that accelerates the DPI process by leveraging the repetitions in plain, non de-duplicated traffic. Leveraging repetitions in DPI engines is entirely different than de-duplication. De-duplication requires extensions and modifications in both server and client sides, while a DPI engine scans traffic on the route between them and cannot force de-duplication or assume it is used. Furthermore leveraging repetition in DPI requires finding the repetitions on the fly, and repetitions can be short. Note that these requirements do not exist for de-duplication solutions.

The work presented at [29] provides a limited solution to accelerate the DPI process using Aho-Corasick algorithm. In this work, a repetition is defined as a repeated string that also starts at the same state in the DFA. Thus, this approach only works when scanning several copies of the exact same string, or same strings are stored over and over along with different starting states. However, this approach can miss a repeated string. Furthermore, this approach only checks sequential strings of fixed length, thus the solution is limited and can only take advantage of repetition of big chunks such as 256-1280 bytes.

III. Enhanced Aho-Corasick Algorithm

At the heart of our solution is a modification to the widely-deployed Aho-Corasick signature matching algorithm. Our modification enhances the algorithm so that it will be able to skip previously-scanned bytes, which are saved in a dictionary along with some auxiliary information. In this section we first briefly describe the Aho-Corasick algorithm and its properties, and then describe the required modifications to the algorithm. We prove that although the modified algorithm skips bytes, it detects the same patterns as the original algorithm.

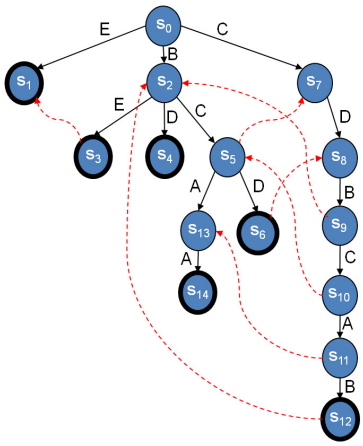


Fig. 1. The Aho-Corasick trie corresponding to the signature set *E,BE,BD,BCD,BCAA,CDBCAB*}. Solid black edges correspond to *forward* transitions, while dashed red edges correspond to *failure* transitions.

A. Background: The Aho-Corasick Algorithm

The Aho-Corasick (AC) algorithm [5] matches multiple signatures simultaneously, by constructing a trie that represents the signatures set. Usually, this trie is then converted into a *deterministic finite automaton* (DFA) for better performance and then, with this DFA on its disposal, the text is processed in a single pass.

Specifically, the trie construction is done in two phases. First, all the signatures are added from the root as chains, where each state corresponds to one byte. When signatures share a common prefix, they also share the corresponding set of states in the trie. The edges of the first phase are called *forward transitions*. In the second phase, *failure transitions* are added to the trie. These edges solve situations where, given an input byte b and a state s, there is no forward transition from s using b. In such a case, the trie should follow the failure transition to some state s' and take a forward transition from there. This process is repeated until a forward transition is found or until the root is reached, leading to possible *failure paths*.

Figure 1 shows an example for an AC trie. Let the *label* of a state s, denoted by $L(s)$, be the concatenation of bytes along the path from the root to s. Furthermore, let the depth of a state s be the length of the label $L(s)$. The failure transition from s, $f(s)$, is always to a state s', whose label $L(s')$ is the longest suffix of $L(s)$ among all other trie states.

The trie is traversed starting from root. When the traversal goes through an *accepting state*, it indicates that some signatures are a suffix of the input; one of these signatures always corresponds to the label of the accepting state. Note that the unique structure of the trie promises that the converted DFA has exactly the same number of states, but much more transitions, to take care of all possible inputs without failure transitions.

The correctness of the AC algorithm essentially stems from the following simple property (see, e.g., [18, Property 2]):

TABLE I

SAMPLE DICTIONARY: EACH STRING IS ASSOCIATED WITH THE DFA STATE REACHED BY SCANNING IT FROM ROOT.

	string	*saved state*
	BYTAFGBC	s_5
	CABXTHGH	s_0

Property 1 *Let $b_1,\ldots b_n$ be the input, and let $s_0,\ldots,s_n$ be the sequence of states the AC algorithm goes through, after scanning the bytes one by one (s_0 is the root of the DFA). For any $i \in \{1,\ldots,n\}$, $L(s_i)$ is a suffix of $b_1,\ldots,b_i$; furthermore, it is the longest such suffix among all other states of the DFA.*

B. Enabling Skips within the Execution of the Aho-Corasick Algorithm

To enable skipping repeating data we add to the Aho-Corasick algorithm an auxiliary *dictionary* that contains (popular) strings. We explain in Section IV how dictionaries are created, and how they are accessed from the data-path. In this section, we will show how our modified algorithm uses the dictionaries in order to skip bytes during execution without missing signatures.

1) Scanning the dictionary: We assume that the dictionary is a set of strings. For each string, separately, we initiate a DFA scan from the initial state s_0. If a match is found by the end of the string, we delete the string from the dictionary[2]. Otherwise, we save the state reached by the DFA after scanning this string along with the string itself.

2) Scanning the data: During DFA traversal, for each input byte, the algorithm checks whether it can skip subsequent bytes using one of the strings in the dictionary. More formally, let $(b_1,\ldots,b_n)$ denote the data; when scanning byte b_i, the algorithm looks for the gram $gram_k(b_i) = (b_i, b_{i+1},\ldots,b_{i+k-1})$. If x is found, the algorithm proceeds in two steps.

First, it performs a *left-margin resolution*, in which we start scanning the bytes $(b_i, b_{i+1}, b_{i+2},\ldots,b_{i+k-1})$ one by one, until when scanning a bytes b_{i+j} we reach a state in the automaton whose depth is less than or equal to j.

Then, if b_{i+k-1} was not reached in the left-margin resolution step, the algorithm transits to the state which was saved along with $gram_k(b_i)$ and continues scanning from byte b_{i+k}.

3) Correctness proof: The correctness of our algorithm stems from the fact that after skipping $gram_k(b_i)$, the algorithm transits to the same state as if $gram_k(b_i)$ was scanned byte by byte. In addition, we need to ensure that if some signature is detected when $gram_k(b_i)$ was scanned byte by byte, it will also be detected in our algorithm.

Theorem 1 *Let the traffic be $(b_1,\ldots,b_n)$ and let $(s_0,\ldots,s_n)$ be the sequence of states the traditional Aho-Corasick algorithm goes through, after scanning the bytes one by one (starting from the root of the DFA). Assume that our modified algorithm scans the traffic up to byte b_i, it is in state s_i*

[2]In practice, this rarely happens and does not have any effect on the overall system performance

TABLE II

EXAMPLE OF SCANNING PROCESS FOR INPUT STRING *CDBCABYTAFGBCD*.

b_i	C	D	B	C	A	B	Y	**T**	**A**	**F**	**G**	**B**	**C**	D
dictionary hit/miss	miss	miss	miss	miss	miss	hit	-	-	-	-	-	-	-	miss
s after scanning b_i	s_7	s_8	s_9	s_{10}	s_{11}	s_{12}	s_0	-	-	-	-	-	s_5	s_6
depth	1	2	3	4	5	6	0	-	-	-	-	-	2	3
j (left-margin res.)	-	-	-	-	-	1	2	-	-	-	-	-	-	-

and it found the string $gram_k(b_i) = (b_i, b_{i+1}, \ldots, b_{i+k-1})$ *in the dictionary. Let* z_{i+k} *be the state of our algorithm after scanning byte* b_{i+k}*. Then, (i)* $s_{i+k} = z_{i+k}$*, (ii) if there are one or more accepting states in states* $(s_i, \ldots, s_{i+k})$*, the left margin resolution does not end before scanning byte* $b_{i+j'}$*, for which* $s_{i+j'}$ *is the last accepting state in* $(s_i, \ldots, s_{i+k})$*.*

Proof: We distinguish between two cases: If the left-margin resolution does not end before reaching byte b_{i+k} then our modified algorithm operates exactly the same as the traditional algorithm and therefore reaches the same state.

Otherwise, let j be the index in which the left-margin resolution ends. By construction this implies that the depth of s_{i+j} is at most j, which implies that the depth of s_{i+k} is at most k (each transition in the automaton increases the depth by at most 1).

By Property 1, $L(s_{i+k})$ is the longest suffix of $(b_1, \ldots, b_{i+k})$ among all states. Since its depth is at most k, it implies that $L(s_{i+k})$ is in fact the longest suffix of $gram_k(b_i) = (b_i, \ldots, b_{i+k})$. On the other hand, by applying Property 1 on the Aho-Corasick scan of $gram_k(b_i)$ (which was performed while scanning the dictionary), we get that $L(z_{i+k})$ is also the longest suffix of $gram_k(b_i)$, which implies that $L(s_{i+k}) = L(z_{i+k})$, and therefore, $s_{i+k} = z_{i+k}$.

Similarly, assume that $j < j'$, then the depth of $s_{i+j'}$ is smaller than j', which implies that the length of the signature corresponding to $s_{i+j'}$ is smaller than j'. By Property 1, this signature is a suffix of $(b_i, \ldots, b_{i+j'})$, and therefore it is fully-contained in $gram_k(b_i)$. This contradicts the construction of the dictionary, in which strings that contain signatures are deleted. ■

C. Motivating Example

We demonstrate the insights behind our algorithm using the following motivating example. Assume that the patterns set is {*E,BE,BD,BCD,BCAA,CDBCAB*}, whose corresponding Aho-Corasick automaton is depicted in Fig. 1. In addition, we assume that the dictionary contains the strings depicted in Table I. For each such string, the resulting state in the independent Aho-Corasick scan is also saved.

There are two kinds of matches that involves strings from the dictionary:

1) Signatures whose prefix is a suffix of a string in the dictionary, For example, the prefix *BC* of the signatures *BCD* and *BCAA* is a suffix of the first string.
2) Signatures whose suffix is a prefix of a string in the dictionary, For example, the suffix *CAB* of the signature *CDBCAB* is a prefix of the second string.

Assume that the input traffic is *CDBCABYTAFGBCD*. The first five bytes did not yield any dictionary match and the Aho-Corasick is in state s_{11}. Next, string *BYTAFGBC* is in dictionary. Since the depth of s_{11} is 5, which is greater than 0, we continue the scan with *B*. The new current state is s_{12}, whose depth is $6 > 1$ and therefore we continue to the next character, *Y*. After that, the current state is s_0, whose depth is less than 2. Thus, the left-margin resolution is completed, and we can skip to the saved state s_5. The algorithm skips the rest of the strings' bytes (in this case k-2 bytes) and continue the scan with the byte *D*. Then, the algorithm reaches the accepting state s_6 and finds the signature *BCD*. The flow of the example is presented in Table II and the skipped characters are marked in bold.

IV. SYSTEM DESIGN

Our system is divided into two components: the *slow path* and the *data path*.

A. The Slow Path

The slow path is responsible of creating a dictionary of repeated fixed-length strings (namely, k-grams, where k is the length of the strings). As explained in Section III, for each stored k-gram, we initiate an Aho-Corasick scan from the initial state s_0 and save the DFA state in the end of this scan. This information is sufficient for the data path to adjust its state after skipping that gram.

We note that while our dictionaries aim to store the most popular k-grams, they suffer from inherent inaccuracies, which sometime reduce the performance gains by our mechanisms; our experiments show, however, that these inaccuracies are not significant. Naturally, the most important reason for such inaccuracies is that the dictionary is built on *offline slightly outdated data*. In addition, in a typical multi-core environment, the slow path runs on a single core and gets only *samples* of the packets. Finally, we use off-the-shelf approximate heavy-hitters algorithm [2], which finds popular k-grams, but sometimes not the optimal dictionary.

Many heavy-hitter algorithms use a sliding window and store all popular k-grams. However, this results in a *dictionary pollution*, in which $m-k+1$ substrings of length k of a very popular string of length m are stored in the dictionary, while our mechanism never access all but m/k of them.[3] The algorithm presented in [2] solves this problem by trying to

[3]For example, assume the string `abcdefgh` is very popular in the traffic, and $k = 4$. The dictionary holds the following 4-grams: `abcd`, `bcde`, `cdef`, `defg`, and `efgh`. Most of the time, the data path uses the 4-grams `abcd` and then `efgh` in order to skip over the long popular string.

concatenate k-grams to longer strings, resulting in heavy hitters of variable length (that is, the parameter k is then the minimal length of the heavy hitters). Since our data path works on fixed-size grams better, we split each heavy hitter string of length m to $\lfloor m/k \rfloor$ consequential k-grams.

The resulting dictionary is stored as an open hash table, where colliding keys are chained. Keys are added in the order of popularity, such that the most popular key is first in the chain, to improve lookup time on average.

B. The Data Path

The data path uses a sliding window of length k to extracts k-grams from the data. For each k-gram, the algorithm searches the dictionary and retrieves the corresponded entry, in case a match is found. If there is no match, one byte is scanned with the Aho-Corasick algorithm, the window slides one byte and the process repeats itself with the next k bytes of the data. If there is a match, *left margin resolution* is performed (see Section III-B): The matched k-gram is scanned byte by byte until reaching a state whose depth is smaller or equal to the position of the last-scanned byte in the gram. Then, the data path adjusts its state to the stored state in the corresponding dictionary entry and advances to the end of the k-gram. Namely, if the k-gram has started in the i-th byte of the traffic, the next byte to be scanned will be the $(i+k)$-th one.

Since the dictionaries might not reside in fast memory or cache, and therefore, may require slower access operations, we first query a *Bloom filter* [30] to ensure that the gram is in the dictionary. A Bloom filter is a compact set representation (in our case, the set is all the grams in the dictionary) that enables efficient approximated set membership queries[4]; thus, in case the gram is not in the dictionary, the overhead of our mechanism is reduced by one order of magnitude (see Table III for exact numbers). We note that Bloom filters sometimes generate *false positives*, which in our case imply redundant accesses to the dictionary. However, this only results in a performance penalty as the dictionary-miss is detected immediately afterwards. Since the false positive rate is very small, this performance penalty is usually insignificant. Algorithm 2 describes the data path.

a) Hardware Implementation Analysis: The data path can be implemented in hardware to utilize parallelism: In such implementation, a dictionary lookup can be done in parallel to Aho-Corasick scan, and once a k-gram is found in the dictionary, a skip can be made.

In addition to parallelism, another benefit of a hardware implementation is that the Bloom filter and dictionary data structures can be put into faster memories. Current SRAM chips, which are limited to at most few megabytes, operate with access latency of about 1-10 nanoseconds, compared to DRAM chips that provide access latency of more than 60 nanoseconds [32]. The data structures for Bloom filter and dictionary hash table are very small relatively to the AC DFA (while AC DFA can get to tens or hundreds of megabytes, Bloom filter and dictionary hash table can take up to a few megabytes in the worst-case scenario). Thus, they can be located on an SRAM chip, while most of the AC DFA must reside in DRAM.

[4]We have implemented the Bloom filter as a bit array. We used the Intel on-chip instruction `crc32q` as the hash function. If more than one hash function is used, we have used different random seeds for CRC to achieve independent hash functions as described in [31]. Upon Bloom filter hit, the same hash computation is later used to access the dictionary hash-table.

```
function ScanGram(B = (b_0, b_2, ..., b_{n-1}), n)
    cur_s ← s_0
    i ← 0
    while i < n do
        gram ← (b_i, b_{i+1}, ..., b_{k-1})
        h ← Hash(gram)
        j ← 0
        if h ∈ BloomFilter then
            entry ← Dictionary[h]
            if gram = entry.gram then
                while cur_s.depth > j do
                    cur_s ← AcScanByte(cur_s, gram[j])
                    i ← i + 1
                    j ← j + 1
                i ← i + (k − j)
                cur_s ← entry.state
            else
                cur_s ← AcScanByte(cur_s, gram[j])
                i ← i + 1
        else
            cur_s ← AcScanByte(cur_s, gram[j])
            i ← i + 1
```

Fig. 2. The data path algorithm

V. Analysis

In this section we analyze the various parameters that influence the performance of our system. Given traffic of length n bytes, let k be the length of grams. Let b_i be the i-th byte of the traffic, and let $gram_k(b_i) = (b_i, \ldots, b_{i+k-1})$. We denote the dictionary as the set D and the Bloom filter that represents it as $BF(D)$. With slight abuse of notations, $x \in BF(D)$ if the Bloom filter indicates that the gram x is in D.

We first classify the bytes of the traffic according to the way our system scans the bytes. Specifically, byte b_i is a *miss byte* if the algorithm queries the dictionary and find out that $gram_k(b_i) \notin D$. A byte is a *left-margin byte* if the algorithm scans this byte as part of a left-margin resolution of a matched gram. Finally, a byte is a *skipped byte* if it is neither a miss byte nor a left-margin byte. For ease of presentation, we refer to left-margin and skipped bytes collectively as *in-gram bytes*. Finally, we call byte b_i a *hit byte* if the algorithm queries the dictionary and found out that $gram_k(b_i) \in D$; notice that a hit byte can be either a skipped byte or a left-margin byte.

We define p as the probability that a byte is an in-gram byte. This immediately implies that the number of miss bytes is $n(1-p)$ and that the number of hit bytes is $np\frac{1}{k}$, since only

the first byte of each matched gram is a hit byte. We further note that the number of Bloom filter queries when processing the traffic is exactly $n\left(1-p+\frac{p}{k}\right)$. We then define c as the average number of Aho-Corasick scan iterations until the left margin resolution is completed, thus the number of left-margin bytes is given by $n \cdot c \cdot \frac{p}{k}$.

The *false positive rate* of the Bloom filter is defined as follows:

$$\text{FPR}_{BF(D)} = \Pr[x \in BF(D) \text{ and } x \notin D].$$

We note that unlike previous parameters, $\text{FPR}_{BF(D)}$ does not depend on the inspected traffic and is solely a parameter describing the accuracy of the Bloom filter. Since the Bloom filter is called $n\left(1-p+\frac{p}{k}\right)$ times, the number of times it will result in a false positive is $\text{FPR}_{BF(D)} \cdot n\left(1-p+\frac{p}{k}\right)$. When it is clear from the context, we will refer to $\text{FPR}_{BF(D)}$ as FPR.

We next quantify the processing times of each operation:

- AC - The average processing time of a single byte scan using the Aho-Corasick algorithm.
- DICT - The average processing time of accessing the dictionary and retrieving the entry of a specific k-gram.
- BF - The average query time of a Bloom filter query.

We now attach the different processing time for each type of a byte: A miss byte requires $\text{AC}+\text{BF}$ time, a hit byte requires $\text{BF}+\text{DICT}$ time, a left-margin byte requires AC time, and skipped bytes do not require any processing time. In addition, each Bloom filter query that results in a false positive imposes a penalty of DICT. Note that some of the bytes belong to two categories and their processing time is the sum of both terms.

Putting all terms together implies that the average per-byte processing time is:

$$\begin{aligned}&\tfrac{1}{n}\left(n(1-p)(\text{AC}+\text{BF})+n\tfrac{p}{k}(\text{BF}+\text{DICT})+\right.\\&\left.+n\cdot c\cdot\tfrac{p}{k}\cdot\text{AC}+\text{FPR}\cdot n\left(1-p+\tfrac{p}{k}\right)\text{DICT}\right)=\\&(1-p)(\text{AC}+\text{BF})+\tfrac{p(c\cdot\text{AC}+\text{BF}+\text{DICT})}{k}+\text{FPR}(1-p+\tfrac{p}{k})\text{DICT}\end{aligned}$$

This processing time immediately yields that in order to accelerate the regular signature matching process (whose average per-byte processing time is AC), the ratio of in-gram bytes, denoted $p_{\min}$, should be at least:

$$p_{\min} > \frac{k(\text{BF}+\text{FPR}\cdot\text{DICT})}{(k-1)(\text{BF}+\text{FPR}\cdot\text{DICT})+(k-c)\text{AC}-\text{DICT}}$$

A. Hardware Implementation Analysis

When the data path is implemented in hardware, Bloom filter and dictionary lookups can be done in parallel with Aho-Corasick scans. This implies a significant reduction in scanning of miss bytes: $\max\{\text{AC},\text{BF}\}$ instead of $\text{AC}+\text{BF}$ in a software implementation. In addition, left-margin resolution of a gram can be done in parallel with the corresponding hit byte processing. When no bytes are scanned in the left boundary of a gram, still one Aho-Corasick scan is performed (since the BF query result is not known). This effectively results in slightly higher value of average number of in-gram bytes scan, denoted by c'; note that $c' \leq c+1$.

$$\begin{aligned}&(1-p)(\max\{\text{AC},\text{BF}+\text{FPR}(1+\tfrac{p}{k(1-p)}\text{DICT}\})+\\&\tfrac{p}{k}\max\{c'\cdot\text{AC},\text{BF}+\text{DICT}\})\end{aligned}$$

For all reasonable parameter values, the hardware implementation outperforms the naïve Aho-Corasick implementation that does not leverage traffic repetitions.

VI. Experimental Results

Our experimental code for the data path is based on the multi-pattern matching code of Snort [33] intrusion detection and prevention system, that implements the Aho-Corasick DFA. We added to the basic DFA code the ability to receive a dictionary, to build a Bloom filter and a dictionary hash table, and to perform skips according to our mechanism. In our experiments we limit the number of k-grams in the dictionary to about $45K$. We found that in most cases, this is enough to achieve high skip ratio while keeping the dictionary lookup process relatively fast. For this number of elements, we use a Bloom filter with one hash function and 0.5M-1M bits. For HTTP traffic, we use Snort's pattern-set ($\sim 4K$ patterns).

For performance evaluation we use a system with Intel Sandybridge Core i7 2600 CPU with 32 KB L1 data cache (per core), 256 KB L2 cache (per core), and 8 MB L3 cache (shared among cores). The system runs Linux Ubuntu 11.10.

A. Traffic Sources

We use the following traffic traces for the experiment:

- **Popular Websites:** we crawled several worldwide and local popular websites and downloaded pages up to depth 2. We repeated this process every 1.5 hours to track changes in HTTP responses. For our experiments we only considered HTML content.
- **General HTML Traffic:** HTML responses from a set of HTTP traffic traces, as described in Section VI-B.
- **Cache-Miss Attack Traffic:** In a cache-miss attack, attacker sends a large amount of similar patterns, multiple times, in order to make the AC DFA get out of its locality area in the cache, and thus have much more cache misses [34]. These traces mix general HTTP traffic with cache-miss attack packets, in increasing attack intensity (bandwidth), as described in [35].

B. HTTP Content Characteristics

In order to examine the potential gain from our mechanism on general HTTP and HTML traffic we analyze the characteristics of HTTP content. We use a 9 GB trace collected from a campus wireless network. The trace contains 348,094 HTTP flows, where an HTTP flow is defined by a request and its corresponding response.

We noticed that when we analyze the entire HTTP traffic as a whole, there are relatively not so many repetitions. However, different content-types behave very differently. For example, partial repetitions are very popular in *text/html* or *text/plain*

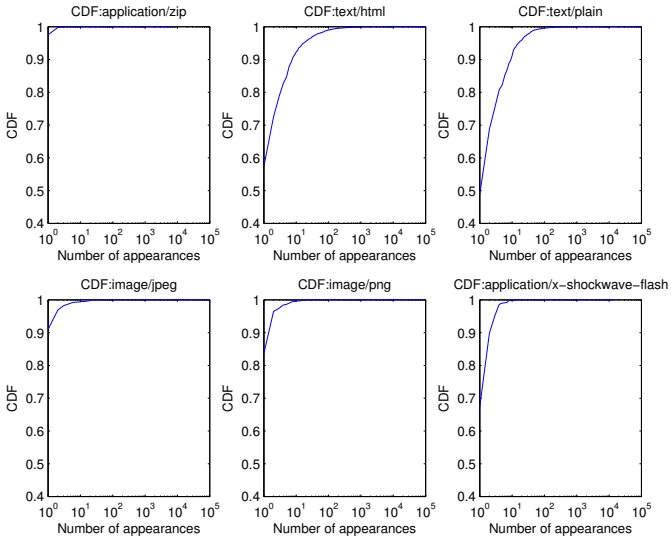


Fig. 3. CDF of Number of Appearances of 16 Bytes Sequences by Content-Type

types, and entire files repetitions are mainly found in images. Some of the types do not contain repetitions at all (except some random strings), e.g. *application/zip*. It is clear that we have to treat each type differently.

For each content type, we count the repetitions of each k-gram in the data. Figure 3 presents the cumulative distribution functions (CDFs) of the number of repetitions per type for $k = 16$. As presented in the figure, for type *application/zip*, almost 100% of the strings appear only once and a few strings appear up to 10 times. However for the type *text/html* 40% of the strings appear more than once, 10% appear more than 10 times and some of them appear more than $10,000$ times. Similar numbers can be found for *text/plain*. Figure 4 present the total repetitions of the different types. The types in the figure are sorted by their popularity in our traces (also presented by the gray line). The potential skip ratio is calculated as the number of bytes in all grams that appear more than once, divided by the total number of bytes of the specific type. This is presented in the figure by the black right bar at each pair. The left white bar presents the potential skip ratio out of the whole data, i.e. the black bar times the gray line at each point. As we can see from the figure, 90% of the HTML data and 85% of the plain data can be potentially skipped, which means that our mechanism has high potential for performance improvement, either in software or in hardware.

There are more types in which high percentage of the traffic can be skipped, e.g. `xml`, `asm` and `c`, but their popularity in our traces is low, and therefore their impact on the global skip ratio is almost negligible. By skipping only HTML and plain data, we achieve more than 35% skip ratio out of all data.

C. Potential Performance Analysis

To asses the potential gain of our mechanism we first isolate each component of the model described in Section V.

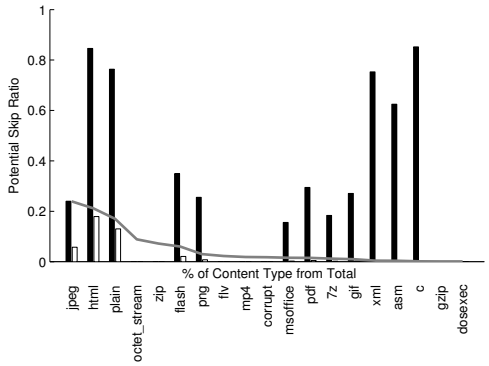


Fig. 4. Skip ratio per content type when using grams of width 32 bytes.

TABLE III
SAMPLE MEASUREMENTS FOR MODEL COMPONENTS.

Traffic	Component	Rate
Popular Website (youtube.com)	BF	2.8 ns/byte
	AC	$4.2-4.3$ ns/byte
	DICT	$27-28$ ns/byte
	p (p_{min})	79% (53%)
	FPR	2.46%
	c	1.87 bytes
Attack Traffic (100% intensity)	BF	$2.7-2.8$ ns/byte
	AC	$5.6-5.8$ ns/byte
	DICT	$28-40$ ns/byte
	p (p_{min})	84% (13%)
	FPR	$3.2-3.7$%
	c	10.1 bytes
General HTML Traffic	BF	$2.02-2.3$ ns/byte
	AC	$4.35-4.5$ ns/byte
	DICT	$25-28$ ns/byte
	p (p_{min})	47% (63%)
	FPR	3.39%
	c	1.8 bytes

We measure times for each operation separately (e.g. Bloom filter lookup, Aho-Corasick DFA lookup, dictionary hash table search) in units of nanoseconds per input byte. Each operation is isolated and timed using a different timer, in separate runs. Note that the different components have different values with each traffic source as traffic induces different AC behavior and different dictionaries. That is, different traffic traces also affect the size of dictionary and Bloom filter differently.

Table III shows sample measurements for different traffic sources we used. Also, for each traffic source, we show the value of p_{min} required for performance gain in the software implementation. By plugging these numbers into the corresponding equations in Section V (for software or hardware implementations), we can retrieve the model prediction of speedup for each traffic source, for both software and hardware implementation, as shown in Table IV. Note that due to compiler and CPU optimizations, the values in Table III are only rough estimations. This explains the differences we got between actual results and model's predictions. Also, note that when implementing the solution in hardware, Bloom filter and dictionary data structures may be put into a faster memory, not being subject to cache replacement, and thus provide better

TABLE IV

MODEL PREDICTED SPEEDUPS FOR SOFTWARE AND HARDWARE IMPLEMENTATIONS, AND ACTUAL SPEEDUP ACHIEVED BY OUR SOFTWARE IMPLEMENTATION, FOR VARIOUS TRAFFIC SOURCES.

Traffic	Potential Software Speedup	Actual Software Speedup	Potential Hardware Speedup
Popular Website (youtube.com)	62%	53%	256%
Attack Traffic (100%)	112%	117%	235%
General HTML Traffic	−16%	−15%	145%

rates. However, for the comparison we assume that rates are equal, and if faster memory is used then the potential speedups on hardware are even faster than those displayed here.

D. Speedup with Software Implementation

Figure 5 shows the actual speedup that our software implementation achieved on traffic from three worldwide and local[5] popular websites. All websites we tested gained a positive speedup in all the experiments we performed. Achieved speedup was very close to the model prediction.

Our mechanism also improves DPI performance when the system is under a cache-miss attack [34], [35]. Such attacks can drop the throughput of the AC DFA by a factor of 7 [34]. As displayed in Table III, the potential number of bytes to skip (p) is very high in such cases. During an attack, the depth in the AC DFA is deeper, as a result of the attack itself. Thus, in most cases, left boundary scan (c) will be longer. However, as can be deduced from the analytical model, there is still a very high potential for throughput gain using our technique. Figure 6 shows the actual speedup achieved using our software implementation on traffic with various attack intensities.

E. Determining the Dictionary Width

The width of grams in the dictionary, denoted k, is an important parameter of our technique. On one hand, when using fixed width dictionary, the larger k is, the longer are the skips we can do when a gram is in the dictionary. On the other hand, if k is too large, the amount of grams that can be put into the dictionary is reduced. Also, our experiments show that variable width dictionary does not always perform better, due to the longer dictionary lookup process.

Figure 7 shows how throughput of the software implementation changes with k on a fixed width dictionary ($k = 0$ means no dictionary is used). In this example, the traffic is of a cache-miss attack of 33% intensity, and $k = 32$ gives the highest speedup, of 33%.

F. Dictionary Creation and Update

Dictionary is first computed in the slow path when a first chunk of data is available, as described in Section IV-A. Dictionary can be computed again at each predefined interval, on the new incoming data. In our experiments, we computed a new dictionary every 10MB-20MB.

When polling the same site repeatedly every predefined interval, we create the dictionary based on several samples together (in our experiments, we polled websites every 90 minutes and created a new dictionary every six hours, on four different samples).

In most cases, a dictionary that was computed once provides a steady speedup for long time, and thus it is not necessary to compute a new dictionary frequently. For example, in the popular websites we studies, we found that even when not updating the dictionary for *days*, potential skip and actual speedup remain almost as they were if we computed a new dictionary over and over again. Figure 8 shows the speedup of the software implementation when scanning youtube.com traffic, similarly to Figure 5(b), but this time, dictionary is only being updated every 6 hours (after computing the first dictionary) and 72 hours (after computing another dictionary, three days later).

[5]The details of the news website were omitted to ensure the anonymity required in the double-blind process.

VII. CONCLUSIONS

In this work we show how repetitions in network traffic can be used to enhance DPI performance. We analyze the potential improvement using a simple, yet accurate, model, and demonstrate the effectiveness of our mechanism in a set of experiments.

Our mechanism changes the legacy Aho-Corasick algorithm, adding a dictionary of repeating data. The slow path of our mechanism uses an off-the-shelf algorithm to recognize repeating strings and create dictionaries from them. Then, in the data path, instead of simply traverse the Aho-Corasick DFA, at each step, we first try to find whether a skip can be done, and if so, avoid scanning the string again.

We show that on certain common traffic types, for various use-cases, our mechanism achieves very high performance gain, when implemented in software or in hardware. We believe that our approach can improve the throughput of DPI in network middleboxes, cloud services and SDN.

REFERENCES

[1] Z. A. Qazi, J. Lee, T. Jin, G. Bellala, M. Arndt, and G. Noubir, "Application-awareness in SDN," in *SIGCOMM*, 2013, pp. 487–488.

[2] Y. Afek, A. Bremler-Barr, and S. L. Feibish, "Automated signature extraction for high volume attacks," in *ANCS*, 2013, pp. 147–156.

[3] B. W. Watson and G. Zwaan, "A taxonomy of keyword pattern matching algorithms," Eindhoven University of Technology, Tech. Rep. 27, 1992.

[4] R. Boyer and J. Moore, "A fast string searching algorithm," *Commun. of the ACM*, pp. 762 – 772, Oct 1977.

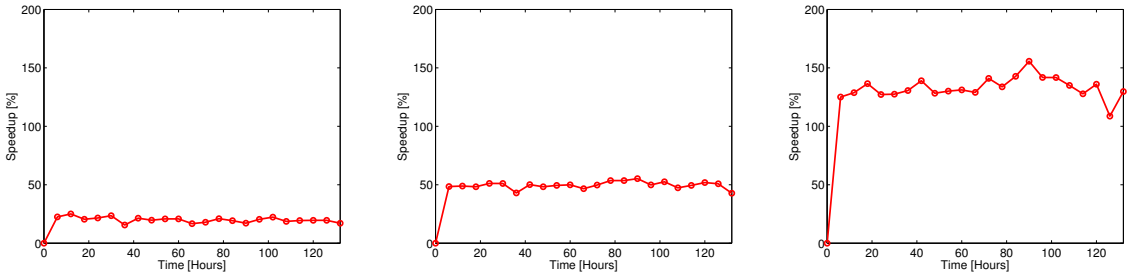
[5] A. Aho and M. Corasick, "Efficient string matching: an aid to bibliographic search," *Commun. of the ACM*, pp. 333–340, 1975.

[6] S. Wu and U. Manber, "A fast algorithm for multi-pattern searching," University of Arizona, Tech. Rep. TR-94-17, May 1993.

[7] Z. K. Baker and V. K. Prasanna, "Time and area efficient pattern matching on fpgas," in *FPGA*, 2004, pp. 223–232.

[8] C. Clark, W. Lee, D. Schimmel, D. Contis, M. Kon, and A. Thomas, "A hardware platform for network intrusion detection and prevention," in *NP*, 2004.

[9] J. Lee, S. H. Hwang, N. Park, S.-W. Lee, S. Jun, and Y. S. Kim, "A high performance NIDS using FPGA-based regular expression matching," in *SAC*, 2007, pp. 1187–1191.



(a) amazon.com (b) youtube.com (c) Local news website

Fig. 5. Actual speedup achieved by our software implementation on different popular websites traffic.

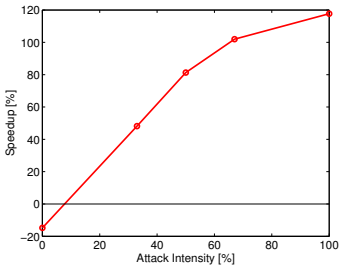


Fig. 6. Speedup achieved by the software implementation on cache-miss attack traffic with different attack intensities.

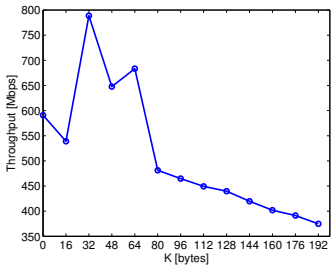


Fig. 7. Throughput change under a cache-miss attack of 33% intensity, with different widths of grams in the dictionary, k.

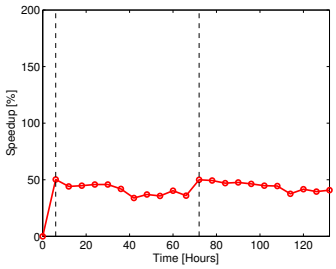


Fig. 8. Speedup of software implementation when scanning traffic from youtube.com, with dictionary update at the beginning and then only after 72 hours. Dashed lines indicate dictionary updates.

[10] C. R. Meiners, J. Patel, E. Norige, E. Torng, and A. X. Liu, "Fast regular expression matching using small tcams for network intrusion detection and prevention systems," in *USENIX Security*, 2010, pp. 8–8.

[11] S. Dharmapurikar and J. Lockwood, "Fast and scalable pattern matching for network intrusion detection systems," *Selected Areas in Communications, IEEE Journal on*, vol. 24, no. 10, pp. 1781–1792, Oct 2006.

[12] D. P. Scarpazza, O. Villa, and F. Petrini, "Exact multi-pattern string matching on the cell/b.e. processor," in *CF*, 2008, pp. 33–42.

[13] D. L. Schuff, Y. R. Choe, and V. S. Pai, "Conservative vs. optimistic parallelization of stateful network intrusion detection," in *PPoPP*, 2007, pp. 138–139.

[14] Y. Afek, A. Bremler-Barr, Y. Harchol, D. Hay, and Y. Koral, "MCA^2: multi-core architecture for mitigating complexity attacks," in *ANCS*, 2012, pp. 235–246.

[15] S. Kumar, S. Dharmapurikar, F. Yu, P. Crowley, and J. Turner, "Algorithms to accelerate multiple regular expressions matching for deep packet inspection," in *SIGCOMM*, 2006, pp. 339–350.

[16] D. Ficara, S. Giordano, G. Procissi, F. Vitucci, G. Antichi, and A. Di Pietro, "An improved dfa for fast regular expression matching," *SIGCOMM Comput. Commun. Rev.*, vol. 38, no. 5, pp. 29–40, Oct 2008.

[17] A. Bremler-Barr and Y. Koral, "Accelerating multi-patterns matching on compressed http traffic," in *INFOCOM*, 2009, pp. 397–405.

[18] A. Bremler-Barr, S. T. David, D. Hay, and Y. Koral, "Decompression-free inspection: DPI for shared dictionary compression over HTTP." in *INFOCOM*, 2012, pp. 1987–1995.

[19] N. T. Spring and D. Wetherall, "A protocol-independent technique for eliminating redundant network traffic," in *SIGCOMM*, 2000, pp. 87–95.

[20] A. Anand, C. Muthukrishnan, A. Akella, and R. Ramjee, "Redundancy in network traffic: findings and implications," in *SIGMETRICS*, 2009, pp. 37–48.

[21] B. Agarwal, A. Akella, A. Anand, A. Balachandran, P. Chitnis, C. Muthukrishnan, R. Ramjee, and G. Varghese, "Endre: An end-system redundancy elimination service for enterprises." in *NSDI*, 2010, pp. 419–432.

[22] E. Zohar, I. Cidon, and O. O. Mokryn, "The power of prediction: cloud bandwidth and cost reduction," *SIGCOMM Comput. Commun. Rev.*, vol. 41, no. 4, pp. 86–97, Aug 2011.

[23] M. Burrows, D. J. Wheeler, M. Burrows, and D. J. Wheeler, "A block-sorting lossless data compression algorithm," Tech. Rep., 1994.

[24] L. Fan, P. Cao, J. Almeida, and A. Z. Broder, "Summary cache: a scalable wide-area web cache sharing protocol," *IEEE/ACM Trans. Netw.*, vol. 8, no. 3, pp. 281–293, Jun 2000.

[25] A. Wolman, G. M. Voelker, N. Sharma, N. Cardwell, A. Karlin, and H. M. Levy, "On the scale and performance of cooperative web proxy caching," in *SOSP*, 1999, pp. 16–31.

[26] A. Wolman, G. Voelker, N. Sharma, N. Cardwell, M. Brown, T. Landray, D. Pinnel, A. Karlin, and H. Levy, "Organization-based analysis of web-object sharing and caching," in *USITS*, 1999, pp. 3–3.

[27] J. Ziv and A. Lempel, "A universal algorithm for sequential data compression," *IEEE Trans. Information Theory*, vol. 23, no. 3, pp. 337–343, May 1977.

[28] J. C. Mogul, F. Douglis, A. Feldmann, and B. Krishnamurthy, "Potential benefits of delta encoding and data compression for http," *SIGCOMM Comput. Commun. Rev.*, vol. 27, no. 4, pp. 181–194, Oct 1997.

[29] G. S. Shenoy, J. Tubella, and A. González, "Improving the performance efficiency of an ids by exploiting temporal locality in network traffic," in *MASCOTS*, 2012, pp. 439–448.

[30] B. H. Bloom, "Space/time trade-offs in hash coding with allowable errors," *Commun. ACM*, vol. 13, no. 7, pp. 422–426, Jul 1970.

[31] F. Hao, M. Kodialam, and T. V. Lakshman, "Building high accuracy bloom filters using partitioned hashing," in *SIGMETRICS*, 2007, pp. 277–288.

[32] D. Levinthal, "Performance analysis guide for intel core i7 processor and Intel Xeon 5500 processors," http://software.intel.com/sites/products/collateral/hpc/vtune/performance_analysis_guide.pdf.

[33] "Snort," http://www.snort.org.

[34] A. Bremler-Barr, Y. Harchol, and D. Hay, "Space-time tradeoffs in software-based deep packet inspection," in *HPSR*, 2011, pp. 1–8.

[35] Y. Afek, A. Bremler-Barr, Y. Harchol, D. Hay, and Y. Koral, "MCA^2: multi-core architecture for mitigating complexity attacks," in *ANCS*, 2012, pp. 235–246.