Backtracking Algorithmic Complexity Attacks Against a NIDS

Randy Smith Cristian Estan Somesh Jha
Computer Sciences Department
University of Wisconsin-Madison
{smithr,estan,jha}@cs.wisc.edu

Abstract

Network Intrusion Detection Systems (NIDS) have become crucial to securing modern networks. To be effective, a NIDS must be able to counter evasion attempts and operate at or near wire-speed. Failure to do so allows malicious packets to slip through a NIDS undetected. In this paper, we explore NIDS evasion through algorithmic complexity attacks. We present a highly effective attack against the Snort NIDS, and we provide a practical algorithmic solution that successfully thwarts the attack. This attack exploits the behavior of rule matching, yielding inspection times that are up to 1.5 million times slower than that of benign packets. Our analysis shows that this attack is applicable to many rules in Snort's ruleset, rendering vulnerable the thousands of networks protected by it. Our countermeasure confines the inspection time to within one order of magnitude of benign packets. Experimental results using a live system show that an attacker needs only 4.0 kbps of bandwidth to perpetually disable an unmodified NIDS, whereas all intrusions are detected when our countermeasure is used.

1. Introduction

Network Intrusion Detection Systems (NIDS) and Intrusion Prevention Systems (IPS) have become crucial to securing today's networks. Typically, a NIDS residing on the edge of a network performs deep packet inspection on every packet that enters the protected domain. When a packet is matched against a signature, an alert is raised, indicating an attempted intrusion or other misuse.

To be effective in an online environment, packet inspection must be performed at or near wire speed. The consequences of not doing so can be dire: an intrusion *detection* system that fails to perform packet inspection at the required rate will allow packets to enter the network undetected. Worse, an inline intrusion *prevention* system that fails to keep up can cause excessive packet loss.

A NIDS must also guard against evasion attempts which often succeed by exploiting ambiguities in a protocol definition itself. For example, attack mechanisms have relied on ambiguities in TCP to develop evasion techniques using overlapping IP fragments, TTL manipulation, and other transformations [10, 15, 18].

In this paper, we explore NIDS evasion through the use of algorithmic complexity attacks [9]. Given an algorithm whose worst-case performance is significantly worse than its average case performance, an algorithmic complexity attack occurs when an attacker is able to trigger worst-case or near worst-case behavior. To mount evasion attempts in NIDS, two attack vectors are required. The first is the true attack that targets a host inside the network. The second is aimed squarely at the NIDS and serves as a cover by slowing it down so that incoming packets (including the true attack) are able to slip through undetected. Evasion is most successful when the true attack enters the network, and neither it nor the second attack is detected by the NIDS.

We present an algorithmic complexity attack that exploits worst-case signature matching behavior in a NIDS. By carefully constructing packet payloads, our attack forces the signature matcher to repeatedly backtrack during inspection, yielding packet processing rates that are up to 1.5 million times slower than average. We term this type of algorithmic complexity attack a *backtracking attack*. Our experiments show that hundreds of intrusions can successfully enter the network undetected during the course of a backtracking attack against a NIDS. Further, the backtracking attack itself requires very little bandwidth; *i.e.*, a single attack packet sent once every three seconds is enough to perpetually disable a NIDS.

Our countermeasure to the backtracking attack is an algorithmic, semantics-preserving enhancement to signature matching based on the concept of memoization. The core idea is straightforward: whereas the backtracking attack exploits the need of a signature matcher to evaluate signatures at all successful string match offsets, a memoization table can be used to store intermediate state that must otherwise be recomputed. Our defense against the backtracking attack relies on the use of better algorithms that reduce the disparity between worst and average case without changing functionality. Empirical results show that this solution confines the processing times of attack packets to within one order of magnitude of benign packets.

Our result applies directly to Snort [17], a popular open-source package that provides both NIDS and IPS functionality and claims more than 150,000 active users. Snort uses a signature-based architecture in which each signature is composed of a sequence of operations, such as string or regular expression matching, that together identify a distinct misuse. In our experiments, we use Snort over both traces and live traffic. In addition, we provide a practical implementation of the defense by extending Snort's signature matching functionality directly.

In summary, our contributions are two-fold. First, we discuss NIDS evasion through algorithmic complexity attacks. We present a highly effective real attack, the backtracking attack, that yields slowdowns of up to six orders of magnitude and is feasible against the (estimated) tens of thousands of networks monitored by Snort. Second, we present an algorithmic defense, based on the principle of memoization, that confines the slowdown to less than one order of magnitude in general and to less than a factor of two in most cases. We provide a practical implementation of this solution and show its efficacy in a live setup.[1]

We organize the remainder of this paper as follows: Section 2 provides a summary of related work, and Section 3 describes the rule-matching architecture of Snort. In Sections 4 and 5 we present the backtracking attack and the countermeasure, respectively. Section 6 details our experimental results, and Section 7 considers other types of complexity attacks. Section 8 concludes.

2. Related work

To our knowledge, Crosby and Wallach [8, 9] were the first to provide an algorithmic basis for denial of service attacks. They exploit weaknesses in hash function implementations and backtracking behavior in common regular expression libraries to produce worst-case behavior that is significantly more expensive than the average case. The result is denial of service in general and evasion in our context. For their examples, the authors observe that algorithmic attacks against hash tables and regular expressions can be thwarted by better algorithm and data structure selections. Our defense also relies on algorithmic improvements.

The backtracking attack we present falls within the general family of algorithmic attacks, although to the best of our knowledge our method of achieving evasion through backtracking is novel.

[1] We have presented our findings to the Snort developers, who have confirmed the efficacy of the evasion attack and have integrated the solution into their NIDS.

In a systems-oriented approach to addressing resource consumption and other attacks, Lee *et al.* [12] dynamically divide the workload among multiple modules, making adjustments as necessary to maintain performance. Load-shedding is performed as necessary to distribute the load to different modules, or to lower its priority. Alternatively, Kruegel *et al.* [11] have proposed achieving high speed intrusion detection by distributing the load across several sensors, using a scatterer to distribute the load and slicers and reassemblers to provide stateful detection. Still other approaches seek to provide better performance by splitting up (and possibly replicating) a sensor onto multiple cores or processors [6, 26]. These approaches show that allocating more hardware can better protect large networks with large amounts of traffic, but they are not a cost effective way of dealing with algorithmic complexity attacks.

The use of custom hardware has also been proposed for performing high-speed matching [3,5,20,24,25]. The backtracking attack is probably not applicable to these solutions. As our focus is on software-based systems, we do not consider hardware solutions further in this paper.

Both [12] and [13] propose the use of monitors to track the resource usage and performance history of a NIDS. In [12], if a monitor discovers abnormally long processing times, the current operations are aborted and optionally transferred to a lower priority process. For [13], on the other hand, the monitor simply triggers a restart of the NIDS. In the general case, such techniques may provide a useful mechanism for ensuring guaranteed minimum performance rates at the cost of decreased detection accuracy. However, such mechanisms result in periodic lapses in detection capability. Our solution is semantics-preserving, in the sense that it does not sacrifice detection to maintain performance.

Finally, NIDS evasion has been extensively studied in the literature. The earliest work describing evasion was presented by Paxson [13] and Ptacek and Newsham [15]. Handley *et al.* [10] show that normalization combined with stateful analysis to remove protocol ambiguities can foil evasion attempts, although it may affect stream semantics. Shankar and Paxson [19] address semantics by providing an online database of network attributes, such as the hop count from the NIDS to a protected host, that provides the same benefits as normalization without the risk of changing stream semantics. These solutions are orthogonal to the problem discussed in this paper.

3. Rule matching in Snort

Our work is performed in the context of the Snort NIDS. Snort employs a signature-based approach to intrusion detection, defining distinct signatures, or rules, for each misuse to be searched for. Each signature is in turn composed of a sequence of *predicates*, that describe the operations that

Predicate	Description	Type
content:$< str >$	Searches for occurrence of $< str >$ in payload	multiple-match
pcre:$/regex/$	Matches regular expression $/regex/$ against payload	multiple-match
byte_test	Performs bitwise or logical tests on specified payload bytes	single-match
byte_jump	Jumps to an offset specified by given payload bytes	single-match

Table 1. Subset of Snort predicates used for packet inspection. Multiple-match predicates may need to be applied to a packet several times.

```
alert tcp $EXT_NET any -> $HOME_NET 99
 (msg:"AudioPlayer jukebox exploit";
  content:"fmt=";                      //P1
  pcre:"/^(mp3|ogg)/",relative;        //P2
  content:"player=";                   //P3
  pcre:"/.exe|.com/",relative;         //P4
  content:"overflow",relative;         //P5
  sid:5678)
```

Figure 1. Rule with simplified Snort syntax describing a fictional vulnerability.

the signature must perform. Section 3.1 gives an overview of the language used to specify these rules. Section 3.2 describes the algorithm used to match rules against packets.

3.1. Expressing rules in Snort

Snort's rules are composed of a header and a body. The header specifies the ports and IP addresses to which the rule should apply and is used during the classification stage. The body has a sequence of predicates that express conditions that need to succeed for the rule to match. A rule matches a packet only if all predicates evaluated in sequence succeed. Of the predicates that are part of Snort's rule language, we focus on those used to analyze the packet payloads. Table 1 summarizes the relevant rules.

Figure 1 depicts a signature using a simplified version of Snort's rule language. The header of the rule instructs Snort to match this signature against all TCP traffic from external sources to servers in the home network running on port 99. The body of the rule contains three `content` predicates, two `pcre` [14] predicates, and two terms, `msg` and `sid`, used for notification and bookkeeping. The rule matches packets that contain the string `fmt=` followed immediately by `mp3` or `ogg`, and also contain the string `player=`, followed by `.exe` or `.com`, followed by `overflow`.

Predicates have one important side effect: during rule matching a predicate records the position in the payload at which it succeeded. Further, when a predicate contains a `relative` modifier, that predicate inspects the packet beginning at the position at which the previous predicate succeeded, rather than the start of the payload. For example, if predicate `P3` from Figure 1 finds the string `player=` at offset i in the payload, the subsequent `pcre` predicate (`P4`) succeeds only if it matches the packet payload after position i.

3.2. Matching signatures

When matching a rule against a packet, Snort evaluates the predicates in the order they are presented in the rule, and concludes that the packet does not match the rule when it reaches a predicate that fails. To ensure correctness, Snort potentially needs to consider all payload offsets at which `content` or `pcre` predicates can succeed. We term these *multiple-match* predicates. In contrast, predicates `byte_test` and `byte_jump` are *single-match*, meaning that any distinct predicate invocation evaluates the payload once.

In the presence of a multiple-match predicate `P`, Snort must also retry all subsequent predicates that either directly or indirectly depend on the match position of `P`. For example, consider matching the rule in Figure 1 against the payload in Figure 2. The caret (^) in `P2` indicates that `P2` must find a match in the payload immediately after the previous predicate's match position. If Snort considers only `P1`'s first match at offset 4, then `P2` will fail since `P2` is looking for `mp3` or `ogg` but finds `aac` instead. However, if Snort also considers `P1`'s second match at offset 28, `P2` will succeed and further predicates from the rule will be evaluated. Snort explores possible matches by backtracking until either it finds a set of matches for all predicates or it determines that such a set does not exist.

Figure 3 presents a simplified version of the algorithm used by Snort to match rules against packets.[2] All predicates support three operations. When a predicate is evaluated, the algorithm calls `getNewInstance` to do the required initializations. The previous match's offset is passed to this function. The `getNextMatch` function checks whether the predicate can be satisfied, and it sets the offset of the match returned by calls to the `getMatchOffset` predicate. Further invocations of `getNextMatch` return true as long as more matches are found. For each of these matches, all subsequent predicates are re-evaluated, because their outcome can depend on the offset of the match. The rule matching stops when the last predicate succeeds, or when all possible matches of the predicates have been explored. Figure 2 shows the stack at each stage of the algo-

[2]The Snort implementation uses tail calls and loops to link predicate functions together and to perform the functionality described in Figure 3. The algorithm presented here describes the behavior that is distributed throughout these functions.

```
Payload | fmt=aac player=play 000 fmt=mp3 rate=14kbps player=cmd.exe?overflow
Offset  | 01234567890123456789012345678901234567890123456789012345678901234567
        |           1         2         3         4         5         6
```

```
                                                                      (P5,59,67)
                                                          (P4,51,59)  (P4,51,59)
                                              (P3,31,51)  (P3,31,51)  (P3,31,51)
            (P2, 4, f)          (P2,28,31)    (P2,28,31)  (P2,28,31)  (P2,28,31)
(P1, 0, 4)  (P1, 0, 4)  (P1, 0,28)  (P1, 0,28)  (P1, 0,28)  (P1, 0,28)  (P1, 0,28)
```

Figure 2. Packet payload matching the rule in Figure 1 and corresponding stack trace after each call to `getNextMatch` on line 3 of Figure 3.

MatchRule($Preds$):

1 $Stack \leftarrow (Preds[0]$.`getNewInstance`$(0))$;
2 **while** $Stack.size > 0$ **do**
3   **if** $Stack.top$.`getNextMatch`() **then**
4   **if** $Stack.size == Preds.size$ **then return** $True$;
5     $ofst \leftarrow Stack.top$.`getMatchOffset`();
6   `Push`$(Stack, Preds[Stack.size]$.`getNewInstance`$(ofst))$;
7   **else** `Pop`$(Stack)$;
8 **return** $False$;

Figure 3. Rule matching in Snort. The algorithm returns $True$ only if all predicates succeed.

rithm. Each stack record contains three elements: the predicate identifier, the offset passed to `getNewInstance` at record creation, and the offset of the match found by `getNextMatch` (`f` if no match is found). In this example, the algorithm concludes that the rule matches.

4. NIDS evasion via backtracking

The use of backtracking to cover all possible string or regular expression matches exposes a matching algorithm to severe denial of service attacks. By carefully crafting packets sent to a host on a network that the NIDS is monitoring, an attacker can trigger worst-case backtracking behavior that forces a NIDS to spend seconds trying to match the targeted rule against the packet before eventually concluding that the packet does not match. For the rule from Figure 1, `P2` will be evaluated for every occurrence of the string `fmt=` in the packet payload. Furthermore, whenever this string is followed by `mp3`, `P2` will succeed and the matcher will evaluate `P3`, and if `P3` succeeds it will evaluate `P4`. If `fmt=mp3` appears n_1 times, `P3` is evaluated n_1 times. If there are n_2 occurrences of `player=`, `P4` will be evaluated n_2 times for each evaluation of `P3`, which gives us a total of $n_1 \cdot n_2$ evaluations for `P4`. Similarly, if these occurrences are followed by n_3 repetitions of `.exe` or `.com`, `P5` is evaluated $n_1 \cdot n_2 \cdot n_3$ times. Figure 4 shows a packet that has $n_1 = n_2 = n_3 = 3$ repetitions. Figure 5 shows the evaluation tree representing the predicates evaluated by the algorithm as it explores all possible matches when matching Figure 1 against the payloads in Figure 2 and in Figure 4. Our experiments show that with packets constructed in this manner, it is possible to force the algorithm to evaluate some predicates hundreds of millions of times while matching a single rule against a single packet.

The amount of processing a backtracking attack can cause depends strongly on the rule. Let n be the size of a packet in bytes. If the rule has k unconstrained multiple-match predicates that perform $O(n)$ work in the worst case, an attacker can force a rule-matching algorithm to perform $O(n^k)$ work. Thus the following three factors determine the power of a backtracking attack against a rule.

1. *The number of backtracking-causing multiple-match `content` and `pcre` predicates k.* The rule from Figure 1 has $k = 4$ because it has 4 backtracking-causing multiple-match predicates (including `P5` which does not match the attack packet, but still needs to traverse the packet before failing). Note that not all `contents` and `pcres` can be used to trigger excessive backtracking. Often, predicates that have constraints on the positions they match cannot be used by an attacker to cause backtracking. An example of such a predicate is the first `pcre` from Figure 1, predicate `P2`, which has to match immediately after the first `content`.
2. *The size of the attack packets n.* We can use Snort's reassembly module to amplify the effect of backtracking attacks beyond that of a single maximum sized packet. The rule from Figure 1 is open to attacks of complexity $O(n^4)$. When Snort combines two attack packets into a virtual packet and feeds it to the rule-matching engine, n doubles, and the rule-matcher does 16 times more work than for either packet alone.
3. *The total length of the strings needed to match the k predicates.* If these strings are short, the attacker can repeat them many times in a single packet. This influences the constants hidden by the O-notation. Let $s_1,\ldots,s_k$ be the lengths of the strings that can cause matches for the k predicates. If we make their contribution to the processing time explicit we can compute for each string the exact number of repetitions. If we divide the packet into

Payload	`fmt=mp3fmt=mp3fmt=mp3player=player=player=.exe.exe.exe`
Offset	`0123456789012345678901234567890123456789012345678901234`
	`          1         2         3         4         5`

Figure 4. A packet payload that causes rule matching to backtrack excessively.

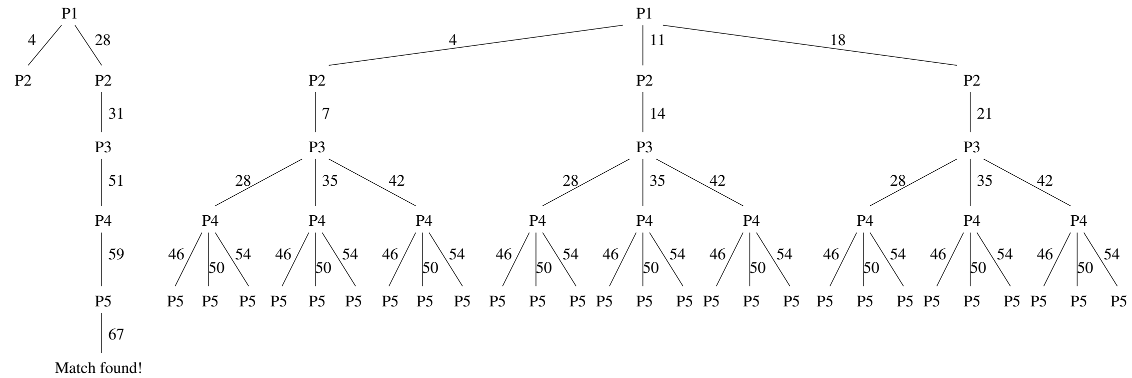

Figure 5. Predicate evaluation trees in Snort. The left tree represents the 6 predicate evaluations performed on the payload in Figure 2, and the right tree shows the 43 evaluations performed for the payload in Figure 4. Numbers on edges indicate payload offsets where a predicate matched.

k equal-sized portions, each filled with repetitions of one of these strings, we obtain $n_i = \lfloor \lfloor n/k \rfloor / s_i \rfloor$. The cost of the attack is $O(\prod_{i=1}^{k} n_i) = O(n^k / (k^k \prod_{i=1}^{k} s_i))$. Other factors such as the amount of overlap between these strings, the length of the strings needed to match predicates that do not cause backtracking, and the details of the processing costs of the predicates also influence the processing cost. These factors remain hidden by the constants inside the O-notation.

Approximately 8% of the 3800+ rules in our ruleset were susceptible to backtracking attacks to some degree. Our focus is on the most egregious attacks, which typically yielded slowdowns ranging from three to five orders of magnitude. We quantify the strength of these attacks experimentally in Section 6.

MemoizedMatchRule($Preds$):

1 $Stack \leftarrow (Preds[0]$.`getNewInstance`$(0))$;
2 $MemoizationTable \leftarrow \emptyset$;
3 **while** $Stack.size > 0$ **do**
4   **if** $Stack.top$.`getNextMatch`() **then**
5   **if** $Stack.size == Preds.size$ **then return** $True$;
6     $ofst \leftarrow Stack.top.getMatchOffset()$;
7   **if** $(Stack.top, ofst) \notin MemoizationTable$ **then**
8       $MemoizationTable \leftarrow MemoizationTable \cup \{(Stack.top, ofst)\}$;
9   `Push`$(Stack, Preds[Stack.size]$.`getNewInstance`$(ofst))$;
10   **else** `Pop`$(Stack)$;
11 **return** $False$;

Figure 6. The memoization-enhanced rule-matching algorithm. Lines 2, 7, and 8 have been added.

5. Memoization, a remedy for backtracking

As illustrated above, rule-matching engines are open to backtracking attacks if they retain no memory of intermediate results, which for Snort are predicate evaluations that have already been determined to fail. Thus, matching engines can be forced to unnecessarily evaluate the same doomed-for-failure predicates over and over again, as Figure 5 indicates.

Figure 6 shows our revised algorithm for rule matching that uses memoization [7, 16]. It is based on the observation that the outcome of evaluating a sequence of predicates depends only on the payload and the offset at which processing starts. The memoization table holds $(predicate, offset)$ pairs indicating for all predicates, except the first, the offsets at which they have been evaluated thus far. Before evaluating a predicate, the algorithm checks whether it has already been evaluated at the given offset (line 7). If the predicate has been evaluated before, it must have ultimately led to failure, so it is not evaluated again unnecessarily. Otherwise, the $(predicate, offset)$ pair is added to the memoization table (line 8) and the predicate is evaluated (line 9). Note that memoization ensures that no predicate is evaluated more than n times. Thus, if a rule has k' predicates performing work at most linear in the packet size n, memoization ensures that the amount of work performed by the rule matching algorithm is at most $O(k' \cdot n \cdot n) = O(k'n^2)$. Figure 7 updates Figure 5 to reflect the effects of memoization. The greyed out nodes in the large tree from Figure 7 correspond to the predicates that would not be re-evaluated when using memoization. For the most damaging backtracking

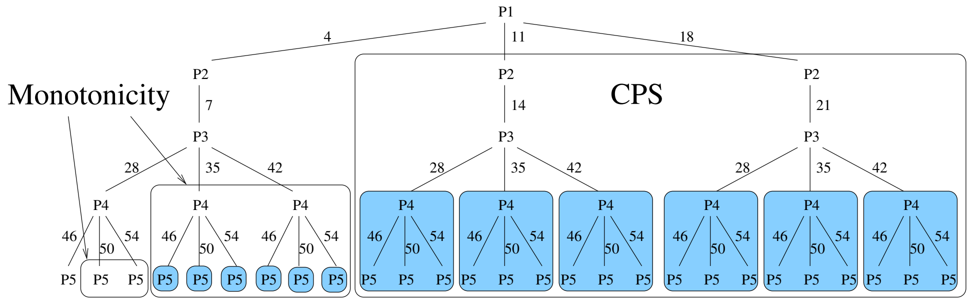

Figure 7. The memoization algorithm performs only 13 predicate evaluations instead of 43 as it avoids the grayed-out nodes. The CPS optimization reduces the number of predicate evaluations to 9, and the monotonicity optimization further reduces the evaluations to 5.

attacks against rules in Snort's default rule set, *memoization can reduce the time spent matching a rule against the packet by more than four orders of magnitude* (with the optimizations from Section 5.1, more than five orders of magnitude).

To implement memoization, we used pre-allocated bitmaps for the memoization table, with a separate bitmap for each predicate except the first. The size of the bitmaps (in bits) is the same as the size v (in bytes) of the largest virtual packet. Thus if the largest number of predicates in a rule is m, the memory cost of memoization is $v(m-1)/8$ bytes. In our experiments, memoization increases the amount of memory used in Snort by less than 0.1%.

A naive implementation of memoization would need to initialize these bitmaps for every rule evaluated. We avoid this cost by creating a small array that holds up to 5 offsets and an index into the array. When a rule is to be evaluated, only the index into the array needs to be initialized to 0. If the number of offsets a predicate is evaluated at exceeds 5, we switch to a bitmap (and pay the cost of initializing it). It is extremely rare that packets not specifically constructed to trigger backtracking incur the cost of initializing the bitmap.

5.1. Further optimizations

We present three optimizations to the basic memoization algorithm: detecting constrained predicate sequences, monotonicity-aware memoization, and avoiding unnecessary memoization after single-match predicates. The first two of these significantly reduce worst case processing time, and all optimizations we use reduce the memory required to perform memoization. Most importantly, all three optimizations are sound when appropriately applied; none of them changes the semantics of rule matching.

Constrained predicate sequences: We use the name *marker* for predicates that ignore the value of the offset parameter. The outcome of a marker and of all predicates subsequent to the marker are independent of where predicates preceding the marker matched. As a result, markers break a rule into sequences of predicates that are independent of each other. We use the name *constrained predicate sequence* (CPS) for a sequence of predicates beginning at one marker and ending just before the next marker. For example, `P3` in Figure 1 looks for the string `player=` in the entire payload, not just after the offset where the previous predicate matches because `P3` does not have the `relative` modifier. Thus the rule can be broken into two CPSes: `P1`-`P2` and `P3`-`P4`-`P5`.

Instead of invoking the rule-matching algorithm on the entire rule, we invoke it separately for individual CPSes and fail whenever we find a CPS that cannot be matched against the packet. The algorithm does not need to backtrack across CPS boundaries. Less backtracking is performed because the first predicate in each CPS is invoked at most once. For the example in Figure 7, detecting CPSes causes the algorithm not to revisit `P1` and `P2` once `P2` has matched, thus reducing the number of predicate invocations from 13 to 9.

Monotone predicates: Some expensive multiple-match predicates used by Snort have the monotonicity property which we define below. For these predicates we use the more aggressive *lowest-offset memoization*. In this optimization, we skip calls to a monotone predicate if it has previously been evaluated at an offset smaller than the offset for the current instance. For example, say we first evaluate a monotone `content` predicate starting at offset 100 that does not lead to a match of the entire rule. Later we evaluate the same predicate starting at offset 200. The second instance is guaranteed to find only matches that have already been explored by the first instance. With basic memoization, after each of these matches of the second instance we check the memoization table and do not evaluate the next predicate because we know it will lead to failure. But, the `content` predicate itself is evaluated unnecessarily. With monotonicity-aware memoization, we do not even evaluate the `content` predicate at offset 200.

The monotonicity property generalizes to some regular expressions too, and it can be defined formally as follows: let S_1 be the set of matches obtained when predicate p is evaluated at offset o_1, and S_2 the matches for starting offset o_2. If for all packets and $\forall o_1 \leq o_2$ we have $S_2 \subset S_1$, then p is monotone. In our example from Figure 1, all `contents` and pcres are monotone with the exception of the first pcre, P2, because it matches at most once *immediately after* the position where the previous predicate matched.

Lowest-offset memoization helps reduce worst case processing because for some predicates the number of worst-case invocations is reduced from $O(n)$ to 1. For the example in Figure 7, this optimization would have eliminated the second and third evaluations for predicates P4, and P5 (and for P3 also if CPSes are not detected). This further reduces the number of predicate instances evaluated from 9 to 5.

Unnecessary memoization: Basic memoization guarantees that no predicate is evaluated more than n times. For some rules with single-match predicates we can provide the same guarantee even if we omit memoizing some predicates. If we employ memoization before evaluating a single-match predicate, but not before evaluating its successor, we can still guarantee that the successor will not be evaluated more than n times (at most once for every evaluation of our single-match predicate). Also, if we have chains of single-match predicates it is enough to memoize only before the first one to ensure that none is evaluated more than n times. Thus, our third optimization is not to perform memoization after single-match predicates, such as byte_test and byte_jump (see Table 1), except when they are followed by a monotone predicate. For our rule set, this optimization reduces by a factor of two the amount of memory used for memoization.

6. Experimental results

We performed empirical evaluations with traces and in a live setting. In Section 6.1, we present measurements comparing backtracking attack packets with traces of typical network traffic. Our results show that three to six orders of magnitude slowdowns achieved with the backtracking attack are reduced to less than one order of magnitude slowdown under memoization. In Section 6.2, we show actual evasion using a non-memoized implementation, and the resulting recovery with the memoized version.

For our experiments we used the Snort NIDS, version 2.4.3, configured to use the Aho-Corasick [2] string matching algorithm. Snort is run on a 2.0 GHz Pentium 4 processor and is loaded with a total of 3812 rules. We instrumented Snort using cycle-accurate Pentium performance counters. When enabled, instrumentation introduced less than 2% overhead to the observed quantities of interest. We found that our measured observations were consistent with the instrumentation results collected in [4].

6.1. Trace-based results

For benign traffic, we obtained two groups of three traces each captured on different days at distinct times. The first

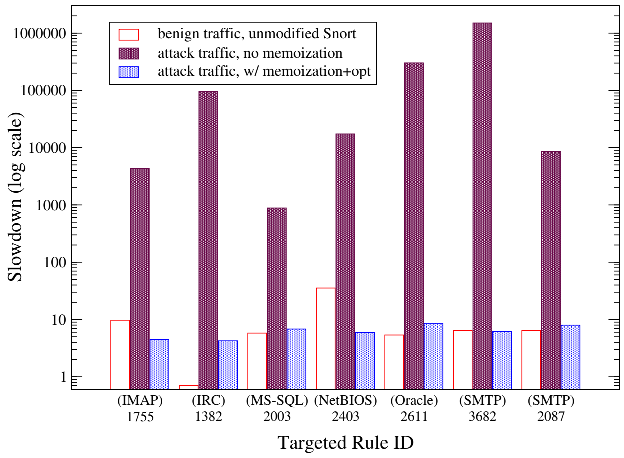

Figure 8. Relative processing times for benign and attack traffic, and attack traffic with memoization. Memoization confines the slowdown to less than one order of magnitude.

group of traces were captured on the link between a university campus and a departmental network with 1,200 desktop and laptop computers, a number of high-traffic servers (web, ftp, ntp), and scientific computing clusters generating high volumes of traffic. These traces are 7 minutes long and range in size from 3.1 GB to just over 8 GB. The second group of traces were captured in front of a few instructional laboratories totaling 150 desktop clients. They are also 7 minutes long and range in size from 816 MB to 2.6 GB.

We created attack traffic by generating flows corresponding to several protocols and supplying payloads that are constructed in a similar manner to the payload construction outlined in Section 4.

In the trace-based experiments, we fed the benign traffic and attack traffic traces into Snort and observed the performance. We performed these experiments with and without memoization enabled. Figure 8 shows the slowdowns experienced due to backtracking attacks targeting several rules and the corresponding defense rates. It summarizes the information in Table 2. In each group, the leftmost bar represents the cost of packet processing for the specified protocol relative to 20.6 s/GB, the combined average packet processing rate in all our traces. For Rule 1382 (IRC), the rate is less than 1, reflecting the fact that the average traffic processing time for IRC traffic is less than the baseline.

The central bar in each group shows the slowdown observed by packets crafted to target the specific rules indicated at the base of each group. The attacks result in processing times that are typically several orders of magnitude slower than the baseline, with the most egregious attack coming in at a factor of 1.5 million times slower. Finally, in the rightmost bar of each group we see the result of each attack repeated with the memoization defense deployed. In

Protocol	Rule ID	Processing time (seconds/gigabyte)				Slowdown w.r.t. avg traffic		Slowdown w.r.t. same protocol	
		Trace traffic	Backtracking attack						
			Original	Basic Memo.	Memo+Opt.	Original	Memo+Opt	Original	Memo+Opt
IMAP	1755	200.6	89,181	1,802	91.9	4,329×	4.46×	444×	0.46×
IRC	1382	14.6	1,956,858	1,170	87.6	94,993×	4.25×	134,031×	6.00×
MS-SQL	2003	119.3	18,206	715	140.4	884×	6.82×	152×	1.17×
NetBIOS	2403	729.7	357,777	57,173	122.0	17,368×	5.92×	490×	0.17×
Oracle	2611	110.5	6,220,768	3,666	174.0	301,979×	8.45×	56,296×	1.57×
SMTP	3682	132.8	30,933,874	2,192	126.4	1,501,644×	6.14×	232,936×	0.95×
SMTP 3682, w/o reassembly			1,986,624	903	103.1	96,438×	5.00×	14,960×	0.78×
SMTP	2087	132.8	175,657	5,123	164.5	8,527×	7.99×	1,323×	1.24×

Table 2. Strength of the backtracking attack and feasibility of the memoization defense. Columns 7-8 shows the overall slowdown under attack when memoization is not and is used. Columns 9-10 shows similar slowdowns with respect to the same protocol.

most cases, Snort performance when under attack is comparable if not better than when not under attack.

Table 2 details the attacks and the defenses quantitatively for several different protocols. For each attack, Columns 1 and 2 give the protocol and the targeted Rule ID to which the attack belongs, respectively. Column 3 shows the average processing time for each protocol. Columns 4 through 6 show the raw processing times for attack packets under an unmodified Snort, Snort with basic memoization, and Snort with fully optimized memoization. Columns 7-8 give overall slowdowns and Columns 9-10 supply the slowdowns on a per-protocol basis. The backtracking attack achieves slowdowns between 3 and 5 orders of magnitude for rules from many protocols. When memoization is employed, the overall slowdown is confined to within one order of magnitude. Per-protocol, memoization confines most attacks to within a factor of two of their normal processing time.

Rows 7 and 8 highlight the impact that reassembly has on the processing time. In this experiment, when reassembly is performed the size of the virtual packet fed to the rule-matching engine is only twice the size of a non-reassembled packet, but the processing time is almost 16× longer.

The effects of the three memoization optimizations can be seen by comparing Columns 5 and 6 in Table 2. The strength of the optimizations varies by protocol, ranging from just under a factor of 10 to just over a factor of 30, excluding the NetBIOS outlier. In the Snort rule set, NetBIOS rules contain many predicates that can be decomposed into constrained predicate sequences. These rules benefit considerably from the optimizations. The accompanying technical report [21] contains the individual contributions of each optimization to the reduction in processing time.

Recall that the attacks applied are all low-bandwidth attacks. Even though the overall slowdown rate using memoization is up to an order of magnitude slower, these rates apply *only* to the attack packets (which are few in number) and not to the overall performance of Snort. Under memoization, processing times for attack packets fall within the normal variation exhibited by benign packets.

In the rightmost column, slowdowns less than 1.0 indicate that with all the optimizations included, Snort was able to process backtracking attack packets more quickly than it could process legitimate traffic. In other words, our optimizations allowed Snort to reject these attack packets more quickly than it otherwise was able since fewer overall predicate evaluations are performed.

6.2. Evading a live Snort

In this section we demonstrate the efficacy of the backtracking attack by applying it to a live Snort installation. We first show successful evasion by applying the attack under a variety of conditions. We then show that with memoization, all the formerly undetected attacks are observed.

Figure 9 shows the topology used for testing evasion for this experiment. To induce denial of service in Snort, we use an SMTP backtracking attack that connects to a Sendmail SMTP server in the protected network. We are using this attack to mask a Nimda [1] exploit normally recognized by Snort. Both the Nimda exploit and its SMTP cover are sent from the same attacking computer. Each Nimda exploit is sent one byte at a time in packets spaced 1 second apart. To simulate real world conditions, we used the Harpoon traffic generator [23] to continuously generate background traffic at 10 Mbps during the experiments.

We measure the effectiveness of the backtracking attack by the number of malicious exploits that can slip by Snort undetected over various time frames. We initiated a new Nimda exploit attempt every second for 5 minutes, yielding 300 overlapping intrusion attempts. Table 3 shows the results. Test 1 is the control: when the backtracking exploit is not performed, Snort recognizes and reports all 300 exploits despite our fragmenting them. In Test 2, we sent two backtracking attack packets every 60 seconds for the duration of the experiment. Snort missed only one-third of the attacks, detecting 222 out of 300 intrusion attempts. In Test 3, we increased the frequency of the backtracking attacks to 2 packets every 15 seconds, dropping the detection

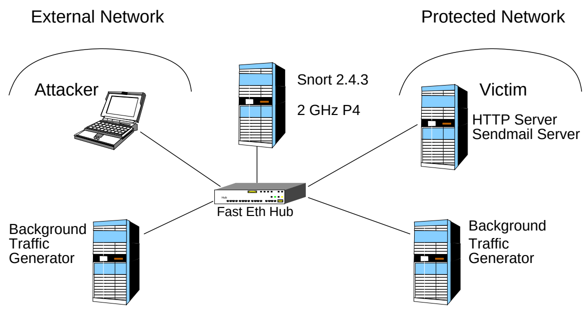

Figure 9. Live Snort evasion environment. Snort monitors a network composed of web and mail servers.

Test	Description of backtrack attack	Exploits detected	Required rate (kbps)
1	Control; no attack	300/300	N/A
2	*two* packets every 60 sec.	220/300	0.4
3	*two* packets every 15 sec.	6/300	1.6
4	*one* packet every 5 sec.	4/300	2.4
5	*one* packet every 3 sec.	0/300	4.0
6	*twenty* packets initially	0/300	0.8
7	*one* packet every 3 sec. (memoization enabled)	300/300	N/A
8	*twenty* packets initially (memoization enabled)	300/300	N/A

Table 3. Summary of live Snort experiments. Without memoization, 300 intrusions pass into the network undetected.

rate to just 2% of the transmitted exploits. Test 4 decreased the detection rate even further, and in Tests 5 and 6 the attacker successfully transmitted all 300 exploits without detection. Aside from high CPU utilization during the attacks and an occasional, sporadic port scan warning directed at the SMTP attack, Snort gave no indication of any abnormal activity or intrusion attempt.

These experiments show that the transmission rate needed to successfully penetrate a network undetected is quite low, with both tests 5 and 6 requiring no more than 4.0 kbps of bandwidth. Test 5, in particular, suggests that perpetual evasion can be achieved through regular, repeated transmissions of backtracking attack packets.

Tests 7 and 8 demonstrate the effectiveness of memoization. These tests repeat Tests 5 and 6 with memoization enabled (including all optimizations). With memoization, Snort successfully detected all intrusions in both tests.

In summary, these experiments validate the results of our trace-based experiments and illustrate the real-world applicability of the backtracking attack. Using carefully crafted and timed packets, we can perpetually disable an IPS without triggering any alarms, using at most 4 kilobits per second of traffic. Correspondingly, the memoization defense can effectively be used to counter such attacks.

7. Discussion

Often, algorithmic complexity attacks and their solutions seem obvious once they have been properly described. Nevertheless, software is still written that is vulnerable to such attacks, which begs the question–how can a NIDS or IPS designer defend against complexity attacks that she has not yet seen? A possible first step is to explicitly consider worst-case performance in critical algorithms and to look at whether it is significantly slower than average case and can be exploited. For example, [9] has shown that in the Bro NIDS, failure to consider worst-case time complexity of hash functions leads to denial of service. With this mindset, we briefly consider mechanisms employed by existing NIDS with an eye towards triggering the worst case.

- Deterministic finite automata (DFA) systems can experience exponential memory requirements when DFA's corresponding to individual rules are combined. In some cases, automata are built incrementally [22] to reduce the footprint of a DFA that cannot otherwise fit in memory. Because each byte of traffic is examined exactly once in a DFA, backtracking does not occur. However, it may be possible for an adversary to construct packets that trigger incremental state creation on each byte of payload, resulting in consistently increased computation costs and potentially leading to memory exhaustion.
- Nondeterministic finite atomata (NFA) systems reduce the memory requirement costs of DFA systems by allowing the matcher to be in multiple states concurrently. In practice, this is achieved either through backtracking or by explicitly maintaining and updating multiple states. In the first case, algorithmic complexity attacks are achieved by triggering excessive backtracking. In the second, the attacker tries to force the NIDS to update several states for each byte processed.
- Predicate-based systems such as Snort can be slowed down if the attacker can cause more predicates to be evaluated than in the average case. We have presented an attack that forces the repeated evaluation of a few predicates many times. In contrast, attacks can be devised that seek to evaluate many predicates a few times. For example, Snort employs a multi-pattern string matcher [2] as a pre-filter to pare down the rules to be matched for each packet. Constructing payloads that trigger large numbers of rules can lead to excessive predicate evaluations.

We have performed preliminary work that combines the second and third observations above to yield packet processing times in Snort that are up to 1000 times slower than average. These results, combined with those of this paper, suggest that left unaddressed, algorithmic complexity attacks can pose significant security risks to NIDS.

8. Conclusions and future work

Algorithmic complexity attacks are effective when they trigger worst-case behavior that far exceeds average-case behavior. We have described a new algorithmic complexity attack, the backtracking attack, that exploits rule matching algorithms of NIDS to achieve slowdowns of up to six orders of magnitude. When faced with these attacks, a real-time NIDS becomes unable to keep up with incoming traffic, and evasion ensues. We tested this attack on a live Snort installation and showed that the protected network is vulnerable under this attack, along with the tens of thousands of other networks protected by Snort.

To counter this attack, we have developed a semantics-preserving defense based on the principle of memoization that brings Snort performance on attack packets to within an order of magnitude of benign packets. Our solution continues the trend of providing algorithmic solutions to algorithmic complexity attacks.

In general, it is not clear how to find and root out all sources of algorithmic complexity attacks. To do so requires knowledge of average- and worst-case processing costs. Without a formal model of computation, such knowledge is difficult to obtain and is often acquired in an ad-hoc manner. Mechanisms for formally characterizing and identifying algorithms and data structures that are subject to complexity attacks can serve as useful analysis tools for developers of critical systems, such as NIDS. We are currently exploring these issues.

References

[1] Cert advisory ca-2001-26 nimda worm, 2001. `http://www.cert.org /advisories /CA-2001-26.html`.

[2] A. V. Aho and M. J. Corasick. Efficient string matching: An aid to bibliographic search. In *Communications of the ACM*, June 1975.

[3] M. Attig and J. W. Lockwood. SIFT: Snort intrusion filter for TCP. In *Hot Interconnects*, Aug. 2005.

[4] J. B. Cabrera, J. Gosar, W. Lee, and R. K. Mehra. On the statistical distribution of processing times in network intrusion detection. In *43rd IEEE Conference on Decision and Control*, Dec. 2004.

[5] C. R. Clark and D. E. Schimmel. Scalable pattern matching for high-speed networks. In *IEEE Symposium on Field-Programmable Custom Computing Machines (FCCM)*, pages 249–257, Napa, California, Apr. 2004.

[6] The Snort network intrusion detection system on the intel ixp2400 network processor. Consystant White Paper, 2003.

[7] T. H. Cormen, C. E. Leiserson, and R. L. Rivest. *Introduction to Algorithms*. MIT Press/McGraw-Hill, 1990.

[8] S. Crosby. Denial of service through regular expressions. In *Usenix Security work in progress report*, Aug. 2003.

[9] S. A. Crosby and D. S. Wallach. Denial of service via algorithmic complexity attacks. In *Usenix Security*, Aug. 2003.

[10] M. Handley, V. Paxson, and C. Kreibich. Network intrusion detection: Evasion, traffic normalization, and end-to-end protocol semantics. In *Usenix Security*, Aug. 2001.

[11] C. Kruegel, F. Valeur, G. Vigna, and R. Kemmerer. Stateful Intrusion Detection for High-Speed Networks. In *Proceedings of the IEEE Symposium on Security and Privacy*, pages 285–293, Oakland, CA, May 2002. IEEE Press.

[12] W. Lee, J. B. D. Cabrera, A. Thomas, N. Balwalli, S. Saluja, and Y. Zhang. Performance adaptation in real-time intrusion detection systems. In *RAID*, Zurich, Switzerland, Oct. 2002.

[13] V. Paxson. Bro: a system for detecting network intruders in real-time. In *Computer Networks (Amsterdam, Netherlands: 1999)*, volume 31, pages 2435–2463, 1999.

[14] PCRE: The perl compatible regular expression library. `http://www.pcre.org`.

[15] T. H. Ptacek and T. N. Newsham. Insertion, evasion and denial of service: Eluding network intrusion detection. In *Secure Networks, Inc.*, Jan. 1998.

[16] T. Reps. “Maximal-munch” tokenization in linear time. *ACM Transactions on Programming Languages and Systems*, 20(2):259–273, 1998.

[17] M. Roesch. Snort - lightweight intrusion detection for networks. In *Proceedings of the 13th Systems Administration Conference*. USENIX, 1999.

[18] S. Rubin, S. Jha, and B. P. Miller. Automatic generation and analysis of NIDS attacks. In *ACSAC ’04*, pages 28–38, Washington, DC, USA, Dec. 2004. IEEE Computer Society.

[19] U. Shankar and V. Paxson. Active mapping: resisting NIDS evasion without altering traffic. In *IEEE Symposium on Security and Privacy*, pages 44–61, May 2003.

[20] R. Sidhu and V. Prasanna. Fast regular expression matching using FPGAs, 2001.

[21] R. Smith, C. Estan, and S. Jha. Algorithmic complexity attacks against Snort. University of Wisconsin Technical Report 1561, Sept. 2006.

[22] R. Sommer and V. Paxson. Enhancing byte-level network intrusion detection signatures with context. In *ACM CCS*, Washington, DC, Oct. 2003.

[23] J. Sommers and P. Barford. Self-configuring network traffic generation. In *Internet Measurement Conference*, pages 68–81, 2004.

[24] I. Sourdis and D. Pnevmatikatos. Fast, large-scale string match for a 10gbps FPGA-based network intrusion detection system. In *International Conference on Field Programmable Logic and Applications*, Sept. 2003.

[25] L. Tan and T. Sherwood. A high throughput string matching architecture for intrusion detection and prevention. In *International Symposium on Computer Architecture ISCA*, June 2005.

[26] T. Vermeiren, E. Borghs, and B. Haagdorens. Evaluation of software techniques for parallel packet processing on multi-core processors. In *IEEE Consumer Communications and Networking Conference*, Jan. 2004.