



**HAL**  
open science

# Evaluating the Security of a DNS Query Obfuscation Scheme for Private Web Surfing

Dominik Herrmann, Max Maass, Hannes Federrath

► **To cite this version:**

Dominik Herrmann, Max Maass, Hannes Federrath. Evaluating the Security of a DNS Query Obfuscation Scheme for Private Web Surfing. 29th IFIP International Information Security Conference (SEC), Jun 2014, Marrakech, Morocco. pp.205-219, 10.1007/978-3-642-55415-5\_17. hal-01370367

**HAL Id: hal-01370367**

**<https://inria.hal.science/hal-01370367v1>**

Submitted on 22 Sep 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# Evaluating the Security of a DNS Query Obfuscation Scheme for Private Web Surfing<sup>\*</sup>

Dominik Herrmann, Max Maaß, and Hannes Federrath

University of Hamburg, Computer Science Department, Germany

**Abstract.** The Domain Name System (DNS) does not provide query privacy. Query obfuscation schemes have been proposed to overcome this limitation, but, so far, they have not been evaluated in a realistic setting. In this paper we evaluate the security of a random set range query scheme in a real-world web surfing scenario. We demonstrate that the scheme does not sufficiently obfuscate characteristic query patterns, which can be used by an adversary to determine the visited websites. We also illustrate how to thwart the attack and discuss practical challenges. Our results suggest that previously published evaluations of range queries may give a false sense of the attainable security, because they do not account for any interdependencies between queries.

## 1 Introduction

DNS is an integral part of the Internet infrastructure. Unfortunately, it does not offer privacy, i. e., the so-called resolvers (recursive nameservers) can see all queries sent to them in the clear. Resolvers can learn about users' habits and interests, which may infringe their privacy if the resolver is not run by a trusted party, but by a third party such as Google, whose resolver 8.8.8.8 serves more than 130 billion queries per day on average [11]. The discussions about limiting tracking via cookies spurred by the "Do not track" initiative may result in DNS queries becoming the next target for tracking and profiling purposes [7]. According to [12] behavior-based tracking based on DNS queries may be feasible.

Integrating mechanisms for confidentiality into DNS is difficult because of the need for compatibility with existing infrastructure. Fundamental changes to the protocol are implemented very slowly, as previous attempts have shown: Although the initial DNSSEC security extensions have been proposed in 1999 [9], the majority of users still can not profit from their benefits today. Unfortunately, DNSSEC does not address privacy issues due to an explicit design decision [1].

Currently, there is no indication that facilities for privacy-preserving resolution will be integrated into the DNS architecture in the short term. Previous research efforts have focused on interim solutions, i. e., add-ons and tools that enable users who care for privacy to protect themselves against profiling and tracking efforts. The objective consists in designing and evaluating suitable privacy enhancing techniques in such a way that users do not have to rely on or

---

<sup>\*</sup> This paper is based on the BSc thesis [14] of the second author.

trust the existing DNS infrastructure. The “range query” scheme by Zhao et al. [17] is one of those efforts. The basic idea consists in *query obfuscation*, i. e., sending a set of dummy queries (hence the term “range”) with random hostnames along with the actual DNS query to the resolver.

So far the security of range query schemes has only been analyzed within a simplistic theoretical model that considers the obtainable security for *singular queries*. In this paper we study the security offered by range queries for a more complex real-world application, namely *web surfing*, which is one of the use cases Zhao et al. envision in [17]. In contrast to singular queries, downloading websites typically entails a number of inter-related DNS queries. Our results indicate that the range query scheme offers less protection than expected in this scenario, because dependencies between consecutive queries are neglected.

The main contribution of this paper is to **demonstrate that random set range queries offer considerably less protection than expected in the web surfing use case**. We demonstrate that a curious resolver (the adversary) can launch a semantic intersection attack to disclose the actually retrieved website with high probability. We also show how the effectiveness of the attack can be reduced, and we identify a number of challenges that have to be addressed before range query schemes are suitable for practice.

The paper is structured as follows. In Sects. 2 and 3 we review existing work and fundamentals. Having described our dataset in Sect. 4, we continue with theoretical and empirical analyses in Sects. 5 and 6. We study countermeasures in Sect. 7 and discuss our results in Sect. 8. We conclude in Sect. 9.

## 2 Related Work

The basic DNS range query scheme was introduced by Zhao et al. in [17]; there is also an improved version [18] inspired by private information retrieval [6]. Although the authors suggest their schemes especially for web surfing applications, they fail to demonstrate their practicability using empirical results.

Castillo-Perez and Garcia-Alfaro propose a variation of the original range query scheme [17] using multiple DNS resolvers in parallel [3, 4]. They evaluate its performance for ENUM and ONS, two protocols that store data within the DNS infrastructure. Finally, Lu and Tsudik propose PPDNS [13], a privacy-preserving resolution service that relies on CoDoNs [15], a next-generation DNS system based on distributed hashtables and a peer-to-peer infrastructure, which has not been widely adopted so far.

The aforementioned publications study the security of range queries for singular queries issued independently from each other. In contrast, [10] observes that consecutively issued queries that are dependent on each other have implications for security. They describe a timing attack that allows an adversary to determine the actually desired website and show that consecutive queries have to be serialized in order to prevent the attack.

## 3 Fundamentals

### 3.1 Random Set DNS Range Query Scheme

In this paper we focus on the basic “random set” DNS range query scheme as introduced in [17]. Zhao et al. stipulate that each client is equipped with a large database of valid domain names (**dummy database**). Each time the client wants to issue a DNS query to a resolver, it randomly draws (without replacement)  $N - 1$  *dummy names* from the database, and sends  $N$  queries to the resolver in total. When all replies have been received from the resolver, the replies for the dummy queries are discarded and the desired reply is presented to the application that issued the query.

Zhao et al. claim that this strategy leaves the adversary with a chance of  $\frac{1}{N}$  to guess the desired domain name. The value of  $N$  is a security parameter, which is supposed to be chosen according to the user’s privacy expectations and performance needs.

### 3.2 Query Patterns

The semantic intersection attack exploits the fact that typical websites embed content from multiple servers, causing clients to issue a burst of queries for various domain names in a deterministic fashion, whenever they visit the site. For example, visiting *google.com* will also trigger a DNS request for *ssl.gstatic.com*, as the site includes some resources from that domain. We call the set of domain names that can be observed upon visiting a site its *query pattern*  $p$ , i. e.,  $p(\text{google.com}) = \{\text{google.com}, \text{ssl.gstatic.com}\}$ . In Sect. 4, we will show that many popular websites do have query patterns that can be used for this attack.

Using range queries, each individual query from a pattern  $p$  is hidden in a set of  $N - 1$  randomly chosen queries, leading to  $|p|$  sets, each containing  $N$  queries, being sent to the resolver in order to retrieve all the domain names required to visit the corresponding website. We refer to  $N$  as the *block size* of the range query scheme and to each individual range query as a *block*.

Note that the client uses standard DNS queries to deliver the range query, because it uses a conventional DNS resolver, i. e., a single range query with a block size of  $N$  causes  $N$  individual DNS queries.

### 3.3 The Semantic Intersection Attack

An adversary, who is in possession of a database that contains the query patterns for a set of websites he is interested in (**pattern database**), can check whether one of these patterns can be matched to consecutive query blocks received by the client. As all the dummy names are drawn independently from each other from the dummy database, it is quite unlikely that the client will draw the pattern of a different website by chance. Therefore, the adversary can be optimistic that he will only find a single pattern in the set of consecutive range queries he receives from the client, i. e., the pattern of the actually desired website.

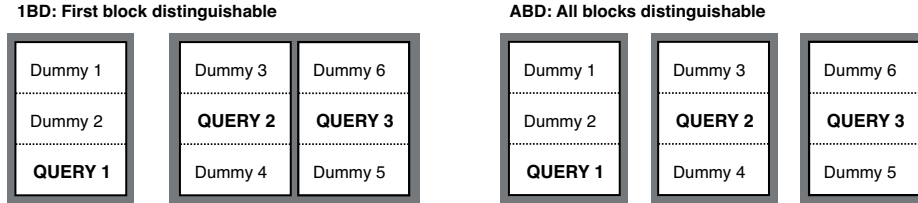


Fig. 1. Distinguishability of blocks for the resolver

From the viewpoint of the adversary there are two different scenarios, depending on how well the adversary can distinguish consecutive blocks (cf. Fig. 1). The adversary may either be able to identify all the queries that belong to the first block, but be unable to determine which of the remaining queries belongs to which of the remaining blocks (**1BD**, 1st block distinguishable), or be able to distinguish all individual blocks, i. e., be able to determine for all queries to which block they belong (**ABD**, all blocks distinguishable). The difference between the 1BD and the ABD scenario becomes evident by considering the following example. When a user visits the site <http://www.rapecrisis.org.uk>, her browser will issue a query for [www.rapecrisis.org.uk](http://www.rapecrisis.org.uk). Moreover, it will issue two additional queries, for [twitter.com](http://twitter.com) and [www.rapecrisislondon.org](http://www.rapecrisislondon.org), once the HTML page has been parsed. For illustrative purposes we assume that range queries with  $N = 3$  are used. In the **ABD scenario** the adversary might, for instance, observe a first block of queries for ([cnn.com](http://cnn.com), [www.rapecrisis.org.uk](http://www.rapecrisis.org.uk), [img.feedpress.it](http://img.feedpress.it)), then a second block for ([github.com](http://github.com), [twitter.com](http://twitter.com), [s.ebay.de](http://s.ebay.de)), and finally a third block for ([www.rapecrisislondon.org](http://www.rapecrisislondon.org), [yimg.com](http://yimg.com), [conn.skype.com](http://conn.skype.com)). In contrast, in the **1BD scenario** the adversary might observe a first block with ([cnn.com](http://cnn.com), [www.rapecrisis.org.uk](http://www.rapecrisis.org.uk), [img.feedpress.it](http://img.feedpress.it)) and a second block with ([github.com](http://github.com), [twitter.com](http://twitter.com), [www.rapecrisislondon.org](http://www.rapecrisislondon.org), [s.ebay.de](http://s.ebay.de), [yimg.com](http://yimg.com), [conn.skype.com](http://conn.skype.com)).

The first block is distinguishable in both scenarios, because the web browser has to resolve the *primary domain name* in order to learn the IP address of the main web server. This IP address is received within the replies that belong to the first block of queries. After the browser has downloaded the HTML file from the main web server, it will issue queries for the *secondary domain names* in order to retrieve all embedded content hosted on other web servers.

Given a *pattern database*  $DB$  that contains primary and secondary domain names of websites, the adversary proceeds as follows in order to carry out the intersection attack in the **ABD scenario**:

1. From  $DB$  the adversary selects all patterns, whose primary domain name is contained in the first block, obtaining the set of candidates  $C$ .
2. The adversary selects all patterns with length  $|p|$ , which is the number of observed blocks, from  $C$  to obtain  $C_{|p|}$ .
3. For each pattern  $q$  in  $C_{|p|}$  the adversary performs a *block-wise set intersection*:  $q$  is a *matching pattern*, if all of its domain names are dispersed among the blocks in a plausible fashion, i. e., iff

- (a) each block contains at least 1 element from  $q$ , and
- (b) each element of  $q$  is contained in at least 1 block, and
- (c)  $q$  can be completely assembled by drawing one element from each block.

In the **1BD scenario** the adversary has to use a different approach, because there are only two blocks observable:

1. From the pattern database the adversary selects all patterns, whose primary domain name is contained in the first block, thus obtaining the set of candidate patterns  $C$ .
2. For each pattern  $q$  in  $C$  the adversary performs a *block-wise set intersection*:  $q$  is a *matching pattern*, if all of its secondary domain names are contained within the second block.

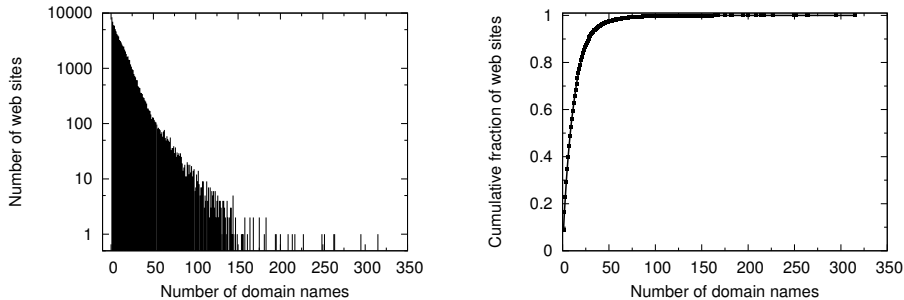
Note that due to caching, the adversary cannot reliably determine  $|p|$  in the 1BD scenario. Due to variations in the lookup time of different domain names, the stub resolver on the client may already receive replies (and cache the results) for some domain names before all range queries have been submitted to the resolver. However, if the range query client happens to draw one of the cached domain names as a dummy, the stub resolver will not send another query, but answer it immediately from its cache. As a result, some queries will not reach the adversary and the effective size of consecutive blocks will vary. Therefore, the adversary cannot easily determine  $|p|$  in the 1BD scenario in order to filter the set  $C$ . For now, we neglect the fact that caching may also affect the desired queries (cf. Sect. 8 for a discussion of this issue).

In the remainder of the paper we focus on the **1BD scenario**, which we deem to be more realistic than the ABD scenario. Contemporary web browsers issue the queries for the secondary queries in parallel. Thus, when the range query client constructs range queries for each of the desired domain names, the individual queries of all the blocks will be interleaved, causing uncertainty about the composition of the individual blocks. On the other hand, the ABD scenario is relevant for range query schemes that submit all queries contained in a block in a single message. We will consider the effect of this approach in Sect. 6.4.

## 4 Dataset

In order to evaluate the feasibility of the semantic intersection attack, we performed probabilistic analyses and implemented a simulator that applies the attack to the patterns of actual websites. For this purpose we obtained the query patterns of the top 100,000 websites of the “Alexa Toplist” (<http://www.alexa.com>) with the headless Webkit-based browser PhantomJS (<http://phantomjs.org>).<sup>1</sup> As PhantomJS was not able to reach and retrieve all of the websites contained in the Toplist at the time of the data collection (May 2013)

<sup>1</sup> The source code of our crawler and simulator as well as all experimental data is available at <https://github.com/Semantic-IA>



**Fig. 2.** Histogram and cumulative distribution of pattern lengths

the cleaned dataset contains  $|P| = 92,880$  patterns and  $|Q| = 216,925$  unique queries. The average pattern length (*mean value*) is 13.02 with a standard deviation of 14.28. The distribution of pattern lengths as displayed in Fig. 2 shows that, while patterns of the length 1 are frequent, patterns of higher lengths make up the majority of the dataset. The longest pattern consists of 315 queries.

## 5 Probabilistic Analysis

Before we carry out any practical evaluation using our simulator, we want to get an expectation of the likelihood of **ambiguous results**, which occur if the client happens to draw all the domain names of another website from the dummy database while the range queries needed for the desired website are assembled. If the client draws all domain names of a different pattern by chance and distributes the individual names among the blocks in a plausible fashion, the adversary will observe two patterns: the pattern of the actually desired website as well as the **random pattern**.

### 5.1 Modeling the Probability of Ambiguous Results

In the 1BD scenario an ambiguous result occurs if the primary domain name of a random pattern (the domain name of the corresponding website) is selected as a dummy in the first block, and all remaining elements of the pattern are contained in the union of the remaining blocks.<sup>2</sup> The probability for an ambiguous result can be modeled as a series of hypergeometric distributions. A hypergeometric distribution  $h(k|N; M; n)$  describes the probability of drawing  $k$  elements with a specific property when drawing  $n$  elements out of a group of  $N$  elements, of which  $M$  have the desired property:

$$h(k|N; M; n) := \frac{\binom{M}{k} \binom{N-M}{n-k}}{\binom{N}{n}} \quad (1)$$

<sup>2</sup> In the 1BD scenario the query distribution between the remaining blocks is irrelevant, as long as all needed queries occur at least once in the union of the blocks.

First, we need to obtain the probability to draw the first element of a pattern of the correct length  $n$  into the first block of queries. As the variables of the hypergeometric distribution overlap with those we use to describe the properties of a range query, we substitute them for their equivalents in our range query notation.  $N$  is equal to  $|Q|$ , the number of names in the dummy database.  $M$  equals to the number of patterns of the correct length, which we will write as  $|P_n|$ . In our case, the parameter  $n$  of the hypergeometric distribution corresponds to  $N - 1$ , as we will draw  $N - 1$  dummy names into the first block. By substituting these values into Eq. 1, we obtain the probability  $p(n, k)$  of drawing exactly  $k$  beginnings of patterns of the length  $n$ :

$$p(n, k) := \frac{\binom{|P_n|}{k} \binom{|Q| - |P_n|}{(N-1) - k}}{\binom{|Q|}{N-1}} \quad (2)$$

In addition to that, we need to determine the probability of drawing the remaining  $k * (n - 1)$  queries into the second block, which contains the remaining  $(n - 1) * (N - 1)$  randomly drawn dummy names in the 1BD scenario. To complete our  $k$  patterns, we need to draw  $k * (n - 1)$  specific dummy names. The probability of success is described by the function  $q(n, k)$ , which is given in Eq. 3.

$$q(n, k) := \frac{\binom{n-1}{n-1}^k \binom{|Q| - (n-1) * k}{(n-1) * (N-1) - (n-1)}}{\binom{|Q|}{(n-1) * (N-1)}} = \frac{\binom{|Q| - (n-1) * k}{(n-1) * (N-1) - (n-1) * k}}{\binom{|Q|}{(n-1) * (N-1)}} \quad (3)$$

The two probabilities  $p(n, k)$  and  $q(n, k)$  can now be combined to receive the probability of drawing  $k$  complete patterns of the correct length  $n$ :

$$P(n, k) := p(n, k) * q(n, k) \quad (4)$$

In this context, the expected value of  $P(n, k)$  for different values of  $n$  is of interest, as it describes the average number of patterns we expect to see. The expected value, in general, is defined as:

$$E(X) := \sum_{i \in I} (x_i p_i) \quad (5)$$

In our case,  $x_1$  is  $k$ , as it describes the number of patterns, and  $p_i$  equals  $P(n, k)$  as the probability of drawing  $k$  patterns, i. e., the expected value is

$$E(n) := 1 + \sum_{k=1}^{N-1} (P(n, k) * k) \quad (6)$$

We are adding 1 to the result, as the original pattern will always be present. Equation 6 will only calculate the expected value for patterns of a specific length. However, as the adversary does not know the length of the pattern with certainty in the 1BD scenario, we have to consider patterns of any length. For that, we have to use a modified variant of Eq. 3:



**Table 1.** Expected avg. number of detected patterns  $F(N)$  for varying block sizes  $N$

$N$	10	50	100
$F(N)$	1.35	2.93	4.83

$$q(n, k, M) := \frac{\binom{|Q| - (n-1)*k}{(M-1)*(N-1) - (n-1)*k}}{\binom{|Q|}{(M-1)*(N-1)}} \quad (7)$$

In Eq. 7,  $n$  is the length of the random pattern, while  $M$  is the length of the original pattern. Accordingly, we modify Eq. 4 and Eq. 6:

$$P(n, k, M) := p(n, k) * q(n, k, M) \quad (8)$$

$$E(M) := 1 + \sum_{n=1}^M \sum_{k=1}^{N-1} (P(n, k, M) * k) \quad (9)$$

Finally, to determine the expected mean value of the number of detected patterns given a specific block size  $N$ , we calculate

$$F(N) = \frac{1}{|P|} * \sum_{M=1}^L (E(M) * |P_M|) \quad (10)$$

where  $L$  is the length of the longest pattern, and  $|P_M|$  the number of patterns having length  $M$ .

## 5.2 Analytical Result

The results (cf. Table 1) indicate that an adversary will, on average, detect only very few random patterns. As expected, the privacy expectation for singular queries ( $\frac{1}{N}$ ) does not apply to the web surfing scenario.

Note that for reasons of conciseness we have provided a slightly simplified model, which disregards overlaps between patterns. Actually, the adversary must expect to find a slightly *higher* number of patterns, because a domain name that is contained within multiple patterns has to be drawn only once to be detected as part of all patterns. Nevertheless, the analysis is instructive and provides us with a baseline for the empirical evaluations that we will describe in the following.

## 6 Evaluation

In order to evaluate the effectiveness of the semantic intersection attack in a realistic scenario, we developed a simulator that enables us to efficiently test different attack strategies and various assumptions about the knowledge of the adversary. In the following we present results for the 1BD scenario.

*Methodology* Given a dataset the simulator will generate range queries for all the patterns from the dataset and perform the semantic intersection attack. We are interested in the influence of two factors on the effectiveness of the attack, namely the *block size*  $N$ , and the *size of the dummy database*  $|Q|$  that contains the dummy names. If the range query scheme was to be used in practice, these two factors could be easily influenced by the user. Thus, it is worthwhile to analyze their effect on the attainable privacy.

In the following, we will use the metric of  **$k$ -identifiability**, which is derived from the well-known metric  $k$ -anonymity [16]: A set of consecutively observed range queries is said to be  $k$ -identifiable, if the adversary finds *exactly*  $k$  matching patterns of websites in his pattern database. For conciseness we will show the cumulative distribution of the fraction of  $k$ -identifiable patterns, i. e., the fraction of patterns that are  $k$ -identifiable or less than  $k$ -identifiable.

### 6.1 Results of Experiment 1: Variation of Block Size

For the purpose of this analysis, we consider three different block sizes:  $N = 10$ ,  $N = 50$ , and  $N = 100$ . [10] has shown that the median latency exceeds 1200 ms for a block size of  $N = 100$ , rendering higher values impractical for practical use.

Based on the result of Sect. 5, we expect to receive some, but not many ambiguous results, i. e., instances where the whole pattern of a different website appears in a set of consecutively observed range queries by chance. Intuitively, the larger the block size, the more random patterns will occur. Accordingly, we expect the effectiveness of the attack to degrade with increasing block sizes.

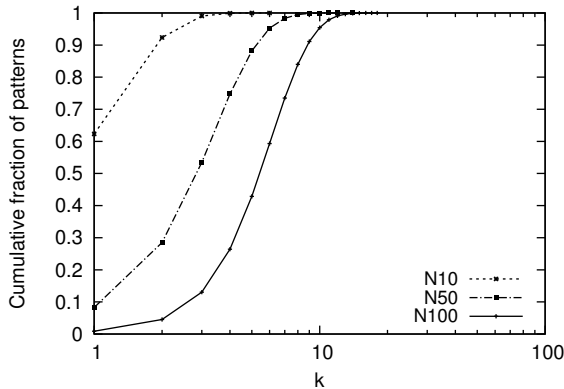
As can be seen in Table 2 and Fig. 3, the smallest block size provides little privacy, with 62 % of patterns being 1-identifiable. Consequently, the median of the observed  $k$ -identifiability values is 1. 99 % of patterns are 5-identifiable or better. No pattern is more than 6-identifiable. For a larger block size of  $N = 50$ , only 8 % of patterns are 1-identifiable, but the cumulative distribution quickly approaches 100 %. All patterns are 14-identifiable or less, and the median of all observed  $k$ -identifiability values is 3, i. e., for 50 % of the websites the adversary can narrow down the actually desired site to a set of 3 or less sites, which is far smaller than the baseline probability of  $\frac{1}{50}$  for finding the desired domain name in the first block. As expected,  $N = 100$  is most effective: 0.8 % of patterns are 1-identifiable, but still 43 % of patterns are at most 5-identifiable.

Generally, we can observe diminishing returns when the block size is increased. While the increase from  $N = 10$  to 50 leads to 54 % less 1-identifiable patterns, adding another 50 queries per block only decreases the fraction by 7.2 percentage points. The same is true for the maximum  $k$ -identifiability, which increases by eight and four, respectively. On overall, the results indicate that range queries provide far less privacy than suggested by Zhao et al. in the web surfing scenario.

*1BD-improved* We also considered an improved attack algorithm that guesses the length of the desired patterns based on the total number of observed queries in the second block, resulting in a range of possible pattern lengths. This allows the adversary to reject all patterns that do not fall into this range. As a result

**Table 2.** Results for varying block sizes  $N$  given the whole dummy database

$N$	$S$	1-identifiable	$\leq 5$ -identifiable	median( $k$ )	max( $k$ )
10	216,925	62 %	99 %	1	6
50	216,925	8 %	88 %	3	14
100	216,925	1 %	43 %	6	18



**Fig. 3.** Distribution of  $k$ -identifiability for varying block sizes  $N$  (whole database)

80 % ( $N = 100$ ) and 94 % ( $N = 10$ ) of all patterns are 1-identifiable. Due to space constraints, we are unable to adequately cover the calculations to estimate the length in this paper, but we have released an implementation including the relevant documentation in the source code repository (see Footnote 1).

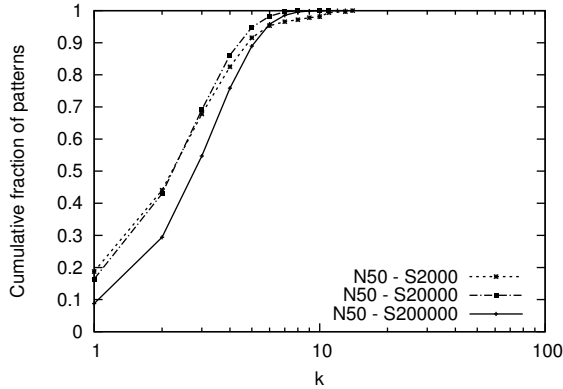
## 6.2 Results of Experiment 2: Variation of Dummy Database

Generating and maintaining a dummy database is a non-trivial task for the client, which gets harder the larger the database is supposed to be. Accordingly, the importance of the size of the dummy database is of interest. We assume that the client’s dummy database is always a subset of the pattern database of the adversary, because, in general, the adversary will have access to more resources than the client, and collecting patterns scales very well.

We compare the effectiveness of three different database sizes ( $S = 2000$ , 20,000 and 200,000). The domain names are chosen by drawing patterns from the full pattern database (without replacement) and adding all domain names of each pattern to the dummy database. This process continues until exactly  $S$  unique domain names have been found. We select full patterns to increase the chance that the client randomly chooses a full pattern when drawing dummies. We used a fixed block size of  $N = 50$  for this experiment.

**Table 3.** Results for varying dummy database sizes  $S$  given the block size  $N = 50$

$N$	$S$	1-identifiable	$\leq 5$ -identifiable	median( $k$ )	max( $k$ )
50	2,000	19%	92%	3	14
50	20,000	16%	95%	3	11
50	200,000	9%	88%	3	13



**Fig. 4.** Distribution of  $k$ -identifiability for varying dummy database sizes  $S$ ;  $N = 50$

Fig. 4 shows that the differences are quite small on overall. Thus, the biggest effect of varying the database is the change in the percentage of 1-identifiable patterns: The percentage of 1-identifiable patterns drops by three percentage points when the dummy database size is increased from  $S = 2,000$  to  $S = 20,000$ , and by another 7 points on the second increase to  $S = 200,000$ . The observed changes have a much smaller effect than the variation of the block size; however, regardless of these results, a larger database is always desirable to prevent other attacks, such as the enumeration of the client’s database.

### 6.3 Effect of Pattern Length on Site Identifiability

Now that we know the effect of varying the block size, the composition of the different  $k$ -identifiabilities is of interest. With this information, we can determine **whether websites with longer or shorter patterns are more at risk** to be identified. Intuitively, shorter patterns should generally have lower  $k$ -identifiabilities, as comparatively few dummies are drawn to obfuscate them, decreasing the chance of drawing a whole pattern. Conversely, longer patterns should generally achieve higher  $k$ -identifiabilities, as they use a higher number of dummy domain names. We will now test this hypothesis by analyzing the composition of the different  $k$ -identifiabilities, using the results of our simulation with a block size of  $N = 50$  and the full dummy database ( $S = 216,925$ ).

**Table 4.** Number of patterns  $n_k$ , mean length  $\overline{|p|}$  and standard deviation SD aggregated by resulting  $k$ -identifiability ( $N = 50$ ,  $S = 216,925$ )

$k$	1	2	3	4	5	6	7	8	9	$\geq 10$
$n_k$	7,693	18,790	23,184	19,784	12,497	6,532	2,875	1,077	336	121
$\overline{ p }$	10.59	11.43	12.52	13.54	14.43	15.45	16.22	17.65	17.09	19.47
SD	12.16	13.24	13.65	14.55	15.02	16.14	16.65	17.71	15.35	19.68

As can be seen in Table 4, the mean pattern length rises almost linearly with increasing  $k$ -identifiability, which supports our hypothesis. The standard deviation exhibits a similar behavior, albeit with a slightly lower and less uniform growth rate. We could reproduce this result for other block and database sizes. The correlation is more distinct for larger block sizes. Smaller block sizes do not show this behavior as clearly, as the range of  $k$ -identifiabilities is too small to show any distinct trend.

#### 6.4 Results of Experiment 3: ABD Scenario

So far, we concentrated on the 1BD scenario (cf. Sect. 3.3). We will now consider the ABD scenario by repeating the experiment from Sect. 6.1, simulating an adversary that can distinguish individual blocks: In the ABD scenario the adversary is able to 1-identify between 87% ( $N = 100$ ) and 97% ( $N = 10$ ) of all domain names, vastly improving on the results of 1BD (1% and 62%, respectively).

The increased accuracy is due to two effects: Firstly, in the ABD scenario the adversary can derive  $|p|$ , the length of the obfuscated pattern, and filter the set of candidate patterns accordingly (cf. Sect. 3.3). Secondly, the probability that another matching pattern is drawn from the dummy database by chance is much smaller when it has to meet the three ABD conditions.

The contribution of these two effects to the overall effectiveness obtained for ABD can be analyzed by reviewing the results obtained for the baseline (1BD) in comparison to 1BD-improved (cf. Sect.6.1) and ABD: The results for 1BD-improved, which filters candidate patterns using a vague estimation of  $|p|$ , already show a significant increase: For  $N = 50$  the fraction of 1-identifiable sites is 83% for 1BD-improved, while it is only 8% for 1BD. On the other hand, the fraction of 1-identifiable websites obtained for ABD, where matching patterns have to meet the additional conditions and the exact value of  $|p|$  is known, rises only by another 6 percentage points (reaching 89%) compared to 1BD-improved.

While this sort of analysis can not conclusively prove that the effect of filtering by length is larger than the effect of filtering via the ABD conditions, we note that the additional benefit of these conditions is comparatively small when the adversary can estimate the length of the obfuscated pattern.

This result indicates that range query schemes that are supposed to provide privacy in a web surfing scenario have to be devised and implemented in a way that the adversary cannot infer the length of the obfuscated query pattern.

**Table 5.** Statistics for varying block sizes  $N$  using the pattern-based dummy construction strategy

$N$	$S$	1-identifiable	$\leq$ 5-identifiable	median(k)	max(k)
10	216,925	0 %	0 %	10	10
50	216,925	0 %	0 %	50	50
100	216,925	0 %	0 %	100	100

## 7 Countermeasures

Having shown the weaknesses of the range query scheme against a pattern-based attack strategy, we will now discuss possible countermeasures. First, we will discuss and evaluate a pattern-based dummy selection strategy. Afterwards, we will consider other strategies that could be used to hinder the adversary.

### 7.1 Pattern-Based Dummy Selection Strategy

In the original dummy selection strategy, the client sampled the dummies independently and randomly from his dummy database. In contrast, the client will now draw *whole patterns* from his database. When querying the resolver for a desired pattern, the client will draw  $N - 1$  random patterns of the same length and use them as dummies. If not enough patterns of the correct length are available, the client will combine two shorter patterns to obtain a concatenated pattern with the correct length. Intuitively, this approach ensures that the adversary will always detect  $N$  patterns. The results of our evaluation, shown in Table 5, confirm this conjecture. All patterns are exactly  $N$ -identifiable.

However, in real-world usage scenarios, the length of the pattern the client is about to query cannot be known in advance. As the dummies for the first element of the pattern have to be chosen before the query can be sent, the client has no way to be sure of the pattern length of the desired website, as these values may change over time when a website changes. This leads to uncertainty about the correct length of the dummy patterns. A wrong choice of pattern length may be used by the adversary to identify the original pattern. Future research could study more sophisticated dummy selection strategies, drawing from experience gained in the field of obfuscated web search [2].

### 7.2 Other Countermeasures

As described in the previous section the pattern-based dummy selection strategy is subject to practical limitations. We will briefly cover other countermeasures that may be used to improve the privacy of clients. This list is not exhaustive.

The first option is to use a variable value for  $N$  that changes on each block. This will raise the difficulty of determining the length of the original pattern, as long as the adversary cannot distinguish individual blocks. This change would render 1BD-improved useless, as it depends on a fixed number of chosen dummies

per block (although similar optimizations could be found that would still improve on the performance of the trivial algorithm). However, this would not impact the performance of the ABD algorithm, as it does not rely on uniform block sizes.

Another improvement that may make the pattern-based strategy more feasible would be to round up the length of the target pattern to the next multiple of a number  $x > 1$ . The additional queries (“padding”) could be chosen randomly, or by choosing patterns of the correct length.

Finally, other privacy-enhancing techniques, such as mixes and onion routing [5, 8], can be employed to counter monitoring and tracking efforts. However, these general-purpose solutions are not specifically designed for privacy-preserving DNS resolution and may introduce significant delays into the resolution process.

## 8 Discussion

We designed our experimental setup to stick as closely as possible to reality. However, for reasons of conciseness and clarity we have neglected some effects. In the following we will discuss whether they affect the validity of our conclusions.

Firstly, the results are implicitly biased due to a closed-world assumption, i. e., our results have been obtained on a dataset of limited size. However, as the Toplist of Alexa contains a large variety of websites we are confident that the results are valid for a large fraction of sites in general. Moreover, we have only evaluated the effectiveness of the attack for the *home pages*; the evaluation of the attack on individual sub-pages is left for future work.

Secondly, while we considered the effects of caching of dummy queries in the 1BD scenario, we disregarded caching of the desired queries: The client may still have (parts of) a pattern in his local cache, resulting in incomplete patterns being sent to the resolver. However, the adversary may adapt to caching by remembering the TTL of all responses he sent to a client and matching the patterns against the union of the received domain names and the cached entries.

Moreover, an adversary who wants to determine all websites a user visits needs the patterns of all websites on the Internet. Such a database would be non-trivial to generate and maintain. However, a *reactive* adversary may visit any domain name he receives a query for and store the pattern for that domain name in its pattern database, making a slightly delayed identification possible.

Finally, we disregarded changing patterns as well as DNS prefetching techniques, which cause longer and more volatile patterns. However, a determined adversary will have no problems in addressing these issues.

## 9 Conclusion

We demonstrated that random set range queries offer considerably less protection than expected in the use case of web surfing. Our attack exploits characteristic query patterns, which lead to greatly reduced query privacy compared to the estimations made by Zhao et al. in their original work. Moreover, we proposed

and evaluated an improved range query scheme using query patterns to disguise the original pattern. We encourage researchers to consider the effects of semantic interdependencies between queries when designing new schemes for query privacy, as the rising pervasiveness of social networking buttons, advertising and analytics makes singular queries less and less common for web surfing.

## References

1. Arends, R., Austein, R., Larson, M., Massey, D., Rose, S.: DNS Security Introduction and Requirements. RFC 4033 (Mar 2005)
2. Balsa, E., Troncoso, C., Díaz, C.: OB-PWS: Obfuscation-Based Private Web Search. In: Symposium on Security and Privacy. pp. 491–505. IEEE (2012)
3. Castillo-Perez, S., García-Alfaro, J.: Anonymous Resolution of DNS Queries. In: On the Move to Meaningful Internet Systems. pp. 987–1000. Springer, LNCS 5332 (2008)
4. Castillo-Perez, S., García-Alfaro, J.: Evaluation of Two Privacy-Preserving Protocols for the DNS. In: International Conference on Information Technology: New Generations (ITNG 2009). pp. 411–416. IEEE (2009)
5. Chaum, D.: Untraceable electronic mail, return addresses, and digital pseudonyms. *Communications of the ACM* 24(2) (1981)
6. Chor, B., Goldreich, O., Kushilevitz, E., Sudan, M.: Private Information Retrieval. In: Symposium on Foundations of Computer Science. pp. 41–50. IEEE (1995)
7. Conrad, D.: Towards Improving DNS Security, Stability, and Resiliency (2012), [http://internetsociety.org/sites/default/files/bp-dnsresiliency-201201-en\\_0.pdf](http://internetsociety.org/sites/default/files/bp-dnsresiliency-201201-en_0.pdf)
8. Dingedine, R., Mathewson, N., Syverson, P.: Tor: The Second-Generation Onion Router. In: USENIX Security Symposium. pp. 303–320 (2004)
9. Eastlake, D.: Domain Name System Security Extensions. RFC 2535 (Mar 1999)
10. Federrath, H., Fuchs, K.P., Herrmann, D., Piosecny, C.: Privacy-Preserving DNS: Analysis of Broadcast, Range Queries and Mix-Based Protection Methods. In: European Symposium on Research in Computer Security (ESORICS 2011). pp. 665–683. Springer, LNCS 6879 (2011)
11. Google Online Security Blog: <http://googleonlinesecurity.blogspot.com/2013/03/google-public-dns-now-supports-dnssec.html> (accessed 2014/03/18) (2013)
12. Herrmann, D., Banse, C., Federrath, H.: Behavior-based Tracking: Exploiting Characteristic Patterns in DNS Traffic. *Computers & Security* 39A, 17–33 (Nov 2013)
13. Lu, Y., Tsudik, G.: Towards Plugging Privacy Leaks in the Domain Name System. In: International Conference on Peer-to-Peer Computing. pp. 1–10. IEEE (2010)
14. Maaß, M.: Schnittmengenangriffe auf DNS Range Queries. Bachelor Thesis, University of Hamburg, <http://nbn-resolving.de/urn:nbn:de:gbv:18-228-7-1989> (2013)
15. Ramasubramanian, V., Sire, E.: The Design and Implementation of a Next Generation Name Service for the Internet. In: SIGCOMM. pp. 331–342. ACM (2004)
16. Sweeney, L.: k-Anonymity: A Model for Protecting Privacy. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* 10(5), 557–570 (2002)
17. Zhao, F., Hori, Y., Sakurai, K.: Analysis of Privacy Disclosure in DNS Query. In: International Conference on Multimedia and Ubiquitous Engineering. pp. 952–957. IEEE (2007)
18. Zhao, F., Hori, Y., Sakurai, K.: Two-Servers PIR Based DNS Query Scheme with Privacy-Preserving. In: International Conference on Intelligent Pervasive Computing. pp. 299–302. IEEE (2007)