A summary of recent and old results on the security of the Diffie-Hellman key exchange protocol in finite groups

Ionuț Florescu

Abstract

Regarding fundamental protocols in cryptography, the Diffie-Hellman (Diffie and Hellman, 1976) public key exchange protocol is one of the oldest and most widely used in today's applications. Consequently, many specific cryptographic implementations depend on its security. Typically, an underlying (finite dimensional) group is selected to provide candidates for the key. The study of the security of the exchange as depending on the structure of the underlying group is even today poorly understood, with the most common approaches relying on the security of the Discrete Logarithm problem or on the size of the group. Recent developments bring to attention that the relationship is not necessarily valid and that more research is needed that will relate the underlying structure of the group and the security of the Diffie-Hellman exchange. In this chapter we describe the problem in detail, we present the relationship with the previously studied Discrete Logarithm and Computational Diffie-Hellman problems, we expose the various concepts of security, and we introduce a new statistical concept specifically designed to serve the assessment of the security of the exchange.

Keywords and phrases: public key cryptography, statistical indistinguishability, group theory, prime subgroups, group structure.

Introduction

A key exchange protocol, is any algorithm through which two parties *A* and *B* agree on a common key K_{AB} .Once the key is established, any further information shared between the parties is encoded, transmitted and decoded using the key K_{AB} . The protocol is secure if any third party *C* finds it extremely hard (almost impossible) to identify the key.

In a public key exchange protocol the two parties agree on a common key pooled from a set *S* while communicating over an insecure channel. The difference is that all the information exchanged over the insecure channel as well as the set of possible keys *S* is known by the perpetrator *C*. If *C* cannot tell apart K_{AB} from any other value in the set *S* guarantees that it is computationally unfeasible to gain "any" partial information on the key.

The Diffie-Hellman key exchange protocol Diffie and Hellman (1976) is a primary example of a public key exchange protocol. In its most basic form, the protocol chooses a finite cyclic group $(G,\cdot)$ of order *N*, with generator *g*, where $\cdot$ denotes the group operation. In what follows we chose the multiplicative operation to denote the operation in the group, and thus the group *G* is generated by the powers of *g* (i.e., $G = \{g, g^2, \ldots, g^N\}$), symbolically *G*=<*g*>. Note that *G*, *g* and *N* are public information.

The participants in the information transfer *A* and *B* each randomly chose an integer $a \in \{1,2,\ldots,N\}$ and $b \in \{1,2,\ldots,N\}$ independently. Then *A* computes g^a, *B* computes g^b and exchange these elements of *G* over the insecure channel. Since each of *A* and *B* knows their respective values chosen (*a* and *b)* they can both compute g^{ab}, which or a publicly known derivation K_{AB} of that becomes the public key.

Any method of converting g^{ab} to K_{AB} is publicly known, and the security of the key K_{AB} is directly dependent on the security of g^{ab} thus, most articles consider g^{ab} as the established key of the exchange.

In the cryptology literature there are two concepts of security – the core security and the concept of semantic security which leads to various security models. The semantic security and the related concepts come under the name of "provable security" Koblitz and Menezes (2004). The core security of the Diffie-Hellman key exchange protocol depends on the discrete logarithm problem, the computational Diffie-Hellman problem and the decision Diffie-Hellman problem. In this article we are concerned with the core security of the exchange. We give a brief introduction to the discrete logarithm problem and the computational Diffie-Hellman problem, for more on these a reader can look at Koblitz and Menezes (2004) or Stinson (2005).

In the present work we will be concerned with the practical security of this protocol. We will investigate the various concepts of security and the known relationships between them. We interpret security in a probabilistic manner and devise a statistical test that will "assess" the security of the exchange in a given group. Our main objective is to find a test that would determine given two cyclic groups G_1 and G_2 with similar orders but perhaps different structures whether or not the security of the key exchange is the same using either group.

Background

Traditionally the study of the security of the exchange was restricted to the verification of the following assumptions:

The Discrete Logarithm Assumption (DL): *For a cyclic group G, generated by g, we are given g and* g^n, $n \in N$, *the challenge is to compute n.*

The Computational Diffie-Hellman Assumption (CDH): *Given* g, g^a, g^b *it is hard to compute* g^{ab}.

Whether or not these assumptions are true in a given group are called the respective problems. For example we say that the Discrete Logarithm problem is hard in a given group if the DL assumption is satisfied in that group.

Clearly, if these assumptions are not satisfied then *C*, an adversary[1], can gain access to the key g^{ab}. The relationship between these two assumptions has been extensively studied. It is clear that the CDH assumption will not be satisfied in a group where finding the solution to the discrete logarithm problem is easy. In Maurer and Wolf (1999), Boneh and Lipton (1996), the authors show that in several settings the validity of the CDH assumption and the hardness of the Discrete Logarithm problem are in fact equivalent.

Unfortunately, the DL and the CDH assumptions are not enough to ensure security of the Diffie-Hellman key exchange protocol. Even if these assumptions are true, the eavesdropper *C* may still be able to gain useful information about g^{ab}. For example, if *C* can predict 90% of the bits in g^{ab} with high probability then for all intents and purposes the key exchange protocol is broken. Moreover, there exist protocols where the knowledge of even one bit will break its security (Casino electronic games). With the current state of knowledge we cannot be confident that assuming only CDH, a scenario like the one described above does not exist (Boneh (1998)).

[1] There are various concepts of adversary in cryptographic literature, the power and authority they have. In this article we assume that our adversary is a passive eavesdropper.

It is clear that both assumptions are necessary for the security of the exchange but are they sufficient? It was evident that a new assumption needed to be formulated.

The Decision Diffie-Hellman Assumption (DDH) *Given* g, g^a, g^b *and an element* $z \in G$ *it is hard to decide whether or not the key* g^{ab} *is more likely to be equal to z or to another random element of* G .

In this form the DDH assumption constitutes a sufficient condition for the security of the Diffie-Hellman key exchange protocol since it directly assesses the established key.

Furthermore, Joux and Nguyen (2003) construct groups based on elliptic curves where the DDH assumption is not satisfied while the CDH and the DL problems are proven to be equivalent and hard. This fact shows that the notions are not equivalent and prompts the necessity to directly check the validity of the DDH assumption for a given group.

The DDH assumption is assumed, either implicitly or explicitly in many cryptographic systems and protocols. Applications include: the many implementations of the DH key exchange itself (e.g., Diffie et al. (1992)), the El-Gamal encryption scheme El-Gamal (1984), the undeniable signatures algorithm Chaum and van Antwerpen (1989), Feldsman's verifiable secret sharing protocol Feldman (1987), Pedersen (1991), and most recently an implementation to the SSH file transfer protocol (Friedl et al., 2006). For a much more detailed list we point to Naor and Reingold (1997).

The DDH assumption in the form presented above is a little vague because of the use of the predicate, "hard to decide". Surprisingly, attempts to make the DDH assumption explicit were not made until late after its formulation in Diffie and Hellman (1976). The first ventures Boneh and Lipton (1996) use standard cryptographic machinery (Yao (1982); Goldwasser and Micali (1984)), to express the assumption in terms of *computational indistinguishability*. Put in this traditional cryptographic form it was discovered quickly by Stadler (1996) and independently Naor and Reingold (1997) that if one assumes the existence of a polynomial time probabilistic algorithm which distinguishes the real key g^{ab} from the other possible values even with a very small probability[2] (for all the possible inputs), then another polynomial time algorithm can be constructed from the first which will output g^{ab} with a very large (almost one) probability. The only requirement is that the size of the group is known, requirement lessened by Boneh (1998) which only requires finiteness of the group.

All this work points toward a more specific definition based entirely on the notion of *statistical significance.* Indeed, this fact materialized in a series of papers Canetti et al. (1999); Canetti et al. (2000); Friedlander and Shparlinski (2001); Vasco et al. (2004), which call this new form of the assumption the Diffie Hellman Indistinguishability assumption (DHI). We note that Gennaro et al. (2004); Joux and Nguyen (2003) use the same form except it continues to call it DDH. We point the reader to Håstad et al. (1999) for a detailed discussion on the concept of statistical significance versus computational significance; in the context of pseudo-random number generation.

In order to introduce this assumption we give the definition of a discrete uniformly distributed random variable.

---

[2] but not negligible. For the sake of completeness we give here the whole definition. It is presented in the footnote since it is not relevant to our approach at all. Suppose that the group G where the exchange takes place has order N and $n = \log_2 N$. It is said that a probabilistic algorithm A decides on the right key with small (non-negligible) probability if there exist a polynomial expression $p(\cdot)$ such that for any $r \in G$:

$$\left|\Pr ob(A \quad outputs \quad g^{ab}) - \Pr ob(A \quad outputs \quad r)\right| > \frac{1}{p(n)}$$

Definition 1: *We say that a variable X has a discrete uniform distribution on the elements of a set* $S=\{a_1,a_2,\ldots,a_n\}$ *if it can take any value in the set S equally likely, i.e.,* $\Pr ob(X=a_i)=\frac{1}{N}$ *for any* $i\in\{1,2,\ldots,N\}$. *We will use the notation DU(S) to denote this distribution.*

For the purpose of studying the security of the Diffie Hellman exchange we use:

The Diffie-Hellman Indistinguishability Assumption (DHI) *Given* g, g^a, g^b *the distribution of* g^{ab} *is indistinguishable from the Discrete Uniform distribution on the elements of G (DU(G)).*

The notion of indistinguishability initially used was the traditional computational one. However, herein we use the usual statistical notion: two variables are indistinguishable if they have essentially the same distribution, or formally put, X_1 and X_2 are indistinguishable if their distribution functions $F_i(x)=\Pr ob(X_i\le x)$ with *i=1,2* have the property: $F_1(x)=F_2(x),\quad \forall x\in R\setminus(A_1\cup A_2)$, where A_1, A_2 are the sets which contain the discontinuity points of F_1, respectively F_2.

In our specific case the state space is finite, therefore the distribution functions F_1 and F_2 are just step functions with jumps in a compact set included in the real axis *R*, thus using the right continuity of the distribution functions, the usual definition translates here in equality everywhere. We conclude that in our context, statistical indistinguishability means that the variables have the same distribution.

This formulation is perfectly natural for a statistician who tries to express the DDH formulation presented above. We note that our version of the DHI assumption requires that the conditional distribution $g^{ab}\,|\,g^a, g^b, g$ be uniform while the previous articles Canetti et al. (1999); Canetti et al. (2000); Friedlander and Shparlinski (2001); Vasco et al. (2004); Gennaro et al. (2004); Joux and Nguyen (2003) require that the distribution of the entire triple $\left(g^{ab}, g^a, g^b\,|\,g\right)$ be Discrete Uniform on the elements of $G^3=G\times G\times G$ ($DU\left(G^3\right)$). Given an outcome (*x*,*y*,*z*) we may write using the simple multiplicative rule:

$$\Pr ob(g^{ab}=z, g^a=x, g^b=y\,|\,g)=\Pr ob(g^{ab}=z\,|\,g^a=x, g^b=y, g)\Pr ob(g^a=x, g^b=y\,|\,g)\quad (1)$$

Under the original condition that *a* and *b* are $DU(\{1,\ldots,N\})$ and using the fact that *g* is a generator for *G*, the distribution of $g^a, g^b\,|\,g$ is $DU\left(G^2\right)$, thus the two formulations are perfectly equivalent.

In general it is known that *statistical indistinguishability* implies *computational indistinguishability*, but the reverse is not in general true, Goldreich (2001). The following lemma states the same result in our specific case using the assumptions presented in this section: DHI and DDH.

Lemma 1 *In a group G of order N, if the DHI assumption is true then the DDH assumption is true as well.*

Proof. Assume that DHI is true in *G*. Then for given g^a, g^b, the probability $\Pr ob(g^{ab}=z\,|\,g^a, g^b)=\frac{1}{N}$ for any $z\in G$. This is the hardest possible scenario in the DDH assumption and hence the DDH assumption is satisfied.

This lemma says that in any group *G*, DHI is a stronger[3] condition than that of the DDH assumption. If we look at the statements in the two assumptions we find that DHI provides a measure of hardness over the DDH assumption via the uniform distribution.

[3] or at least as strong

__Testing the Diffie-Hellmann Assumption__

For every pair *(a,b)* there exist unique values $\left(g^a, g^b\right)$ and correspondingly a unique key g^{ab} the question is: where does the probability enters into the picture?
Note that in this paper we are not considering the distribution of $(g^{ab} \big| g^a = x, g^b = y)$. That would be irrelevant. We are looking at the distribution $(g^{ab} \big| g^a, g^b)$ and consider $\left(g^a, g^b\right)$ a pair of random variables.

Look at this issue from the following perspective. Since *a* and *b* span the whole $G^2 = G \times G$ range we can see the key assigning process as a mapping from G^2 to G. Accordingly, studying the distribution of $(g^{ab} \big| g^a, g^b)$ and comparing it with the uniform distribution on the elements of G amounts to checking whether or not some subsets of values in G are more likely to be the key than others as the pair *(a,b)* span the group. In other words if there exists a subset of G^2 such that the resulting key from that subset puts higher probabilities on certain values then the conditional distribution $(g^{ab} \big| g^a, g^b)$ will not be uniform.

This approach is identical with the approach of Canetti et al (1999). Canetti et al (2000), Friedlander and Shparlinski (2001), Gennaro et al. (2004) and many other papers. The approach in these papers is to consider the triple (g^a, g^b, g^{ab}) and its distribution in G^3. Because the key g^{ab} uniquely corresponds to $\left(g^a, g^b\right)$ the points (g^a, g^b, g^{ab}) determine a two-dimensional surface included in a space of three dimensions. However, if this surface is randomly scattered in the cube G^3 then we can say that the key is secure. In other words no region in G^3 contains more points than any other similarly sized region in G^3. Owing to the relation (1) the two approaches are equivalent.

Why did we choose our approach and not the more traditional (and established) trivariate (joint) distribution approach? The reason is that our approach is more convenient for checking. It essentially amounts to checking whether a one dimensional distribution is close to the *DU(G)* while the more traditional approach would require verifying three-dimensional distributions.

For the one-dimensional distribution we have a well defined way to establish measure of information present in data: the entropy measure. We note that the same measure exists for 3-dim distributions but it is more cumbersome to use in practical problems. We shall now introduce some notations and definitions.

Let *X*, *Y*, and *Z* be three discrete random variables taking values in the sets $\{x_1, x_2, \ldots, x_n\}$, $\{y_1, y_2, \ldots, y_m\}$, $\{z_1, z_2, \ldots, z_l\}$ respectively.

Denote $p(x_i, y_j, z_k) = \Pr ob\left(X = x_i, Y = y_j, Z = z_k\right)$, the joint probability function corresponding to (*X*,*Y*,*Z*). We continue by using notations $p(x_i \mid y_j)$, $p(x_i, y_j \mid z_k)$, etc. for the conditional probability functions of *X*|*Y*, (*X*,*Y*)|*Z*, etc. Furthermore, assume that for all $k \in \{1,2,\ldots,l\}$ the marginal distribution $p(z_k) = \Pr ob(Z = z_k) \neq 0$ to avoid complications conditioning on a set of measure zero.

Definition 2 (Entropy): *We define the joint and conditional* **measures of uncertainty**.

$$H(X,Y) = -\sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{k=1}^{l} p(x_i, y_j, z_k) \log p(x_i, y_j) \qquad (2)$$

$$H(X,Y \mid Z) = -\sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{k=1}^{l} p(x_i, y_j, z_k) \log p(x_i, y_j \mid z_k) \qquad (3),$$

with the convention 0(−∞)=0.

In the above definition we choose to work with the natural logarithm, however any other basis will be equivalent for our purpose due to the constant in the usual definition of the entropy function (see Shannon (1948)).

Remark 3 *In the definition of the entropy functions (2) and (3) we did not use the structure of the group G in any way, only the relative frequency of the elements in the group. This fact makes the methods based on the entropy function well suited for comparison between diverse groups. We will take advantage of this feature later in this work.*

The idea is to use the entropy function (3) in the sense of Kullback-Leibler divergence (Kullback and Leibler, 1951) as a measure of departure from the entropy calculated under the hypothesis of Uniform distribution. Specifically, using earlier notation, we wish to construct a statistical test that will check the validity of the following hypotheses:

$$H_0: \text{The distribution of } (g^{ab} | g^a, g^b) \text{ is DU(G)}$$
$$H_a: \text{The distribution of } (g^{ab} | g^a, g^b) \text{ is NOT DU(G)} \qquad (4)$$

Let us denote the elements of G as $\{g_1, g_2, ..., g_N\}$. Suppose we can look at all the possible triples (g^a, g^b, g^{ab}) when $a,b \in \{1,2,\dots,N\}$ take all the possible values. Clearly, there are N^2 such possible triples and assuming that a and b are chosen at random, each such triple will have probability $1/N^2$. The last element in the triple g^{ab} will get mapped into N possible values (the elements of G). Thus, some values in G will be repeated. For an element $g_k \in G$ denote m_k the number of times g_k appears for g^{ab} among all the N^2 triples. We have then $\sum_k m_k = N^2$. For any pair (g^a, g^b) that corresponds to $g^{ab} = g_k$ we can then calculate the following conditional probability as:

$$\Pr ob\left(g^a = g_i, g^b = g_j \middle| g^{ab} = g_k\right) = \frac{1}{m_k} 1_A(g_i, g_j, g_k),$$

where A is the set of all possible N^2 triples (g^a, g^b, g^{ab}), and we have used the notation $1_A(x)$ to denote the indicator function of the set $A \subset \Omega$, i.e., $1_A : \Omega \to \{0,1\}$ is given by:

$$1_A(x) = \begin{cases} 1 & if \quad x \in A \\ 0 & if \quad x \notin A \end{cases}$$

We can continue:

$$H\left(g^a, g^b \middle| g^{ab}\right) = -\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{k=1}^{N} p(g_i, g_j, g_k) \log p(g_i, g_j | g_k)$$

Error! Bookmark not defined. $$= -\sum_{k=1}^{N}\sum_{i,j=1}^{N} \frac{1}{N^2} \log \frac{1}{m_k} 1_A(g_i, g_j, g_k)$$

Error! Bookmark not defined. $$= -\sum_{k=1}^{N} \frac{m_k}{N^2} \log \frac{1}{m_k} \qquad (5)$$

Under the null hypothesis H_0, the distribution of $(g^{ab}|g^a, g^b)$ is uniform, therefore we should have all the m_k multiplicities equal. This automatically implies that m_k=N for all k's and then the entropy function in (5) is:

$$H\left(g^a, g^b \middle| g^{ab}\right) = -\sum_{k=1}^{N} \frac{1}{N} \log \frac{1}{N} = \log N$$

The testing statistics is defined as:

$$T_N = H\left(g^a, g^b \middle| g^{ab}\right) - \log N = \sum_{k=1}^{N} \frac{m_k}{N^2} \log m_k - \log N \tag{6}$$

This test is based on the whole set of values in G^2. Accordingly, if the value of the test equals zero then the null hypothesis H_0 is true, any other value of the test will support the alternative hypothesis. We summarize this result in the following:

Lemma 4 (Testing Procedure) *Using the previous notations if* $T_N = 0$ *then the DHI assumption is satisfied in a given group G.*

Remark 3 certainly applies for this testing procedure.

Remark 4 *In practice if we wish to calculate* T_N *we have to calculate all the possible values for* (g^a, g^b) *and this will take longer than an exhaustive search. Thus calculating* T_N *is not practical, instead we would have to estimate it. However, estimating statistics means that we have to calculate the distribution associated with the test statistic. We detail the estimation in the following.*

Assume that we can obtain a sample of n pairs $\{(a_i, b_i)\}_{i\in\{1,\ldots,n\}}$ from $\{1,2,\ldots,N\}\times\{1,2,\ldots,N\}$. For each pair in the sample we calculate the triple $(g^{a_i}, g^{b_i}, g^{a_i b_i})$. Let A_n be the set of all the triplets in the sample.

Using (5) we calculate an estimate of $H\left(g^a, g^b \middle| g^{ab}\right)$ using:

$$\hat{p}_n(g_i, g_j, g_k) = \frac{k_{ijk}}{n} 1_{A_n}(g_i, g_j, g_k) \tag{7}$$

$$\hat{p}_n(g_i, g_j \mid g_k) = \frac{k_{ijk}}{n} 1_{A_n}(g_i, g_j, g_k)$$

where once again m_k denotes the multiplicity of g_k, but in the given sample of n observations. We took into account the possibility of obtaining repeated observations in the sample by multiplying with the factor k_{ijk}; which represents the number of times we see the same observation (g_i, g_j, g_k) in our sample.

The test statistic is:

$$T_n = -\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} \hat{p}(g_i, g_j, g_k) \log \hat{p}(g_i, g_j | g_k) - \log n \tag{8}$$

Now we need to investigate the distribution of T_n under the null hypothesis H_0. Under the null m_k's are the multiplicities of g_k's in a sample of size n drawn from the set $\{g_1, \ldots g_1, g_2, \ldots, g_2, \ldots, g_N, \ldots, g_N\}$ where each element in the group G is repeated N times.

Let us denote by $M_1, M_2, \ldots, M_N$ the multiplicities of the elements $\{g_1, ..., g_N\}$ in a sample of size *n*. It is not hard to show that the joint probability distribution of $(M_1, ..., M_N)$ is the so called *multivariate hypergeometric distribution*:

$$\Pr ob(M_1 = m_1, ..., M_N = m_N) = \frac{\binom{N}{m_1}\binom{N}{m_2}\cdots\binom{N}{m_N}}{\binom{N^2}{n}}$$

The test statistic under H_0 is:

$$T_N = \sum_{k=1}^{N} \frac{M_k}{n} \log M_k - \log n \tag{9}$$

If we would be able to calculate the distribution of T_n knowing that $(M_1, ..., M_N)$ is distributed as a multivariate hypergeometric random vector then we would be in position to reach the conclusion of the test of uniformity (4) by calculating the p-value of the test statistic (8) using this distribution.

Finding the distribution of the test statistic under H_0 in (9) is however not an easy task. This is the reason we propose the use of permutation testing for which knowledge of this distribution is not necessary.

The permutation testing procedure generates samples $(M_1, ..., M_N)$ from the Multivariate hypergeometric distribution. For each sample, it calculates the corresponding value of the test statistic under the null hypothesis as in (9). These values are obtained from the assumption that H_0 is true; this allow us to calculate the empirical distribution of our sample statistic T_n under the null hypothesis. The *p*-value of our test is given by the proportion of values as extreme or more than the one calculated in (8) using the group *G*.

A small *p*-value is evidence against the null hypothesis in (4) that the sample comes from a uniform distribution. We summarize the procedure bellow:

Testing procedure to determine validity of DHI for a group G

i. We take a sample of size *n* and we calculate the test statistic as in (8).

ii. We generate many test statistic values under the hypothesis H_0 using (9), and then we construct their empirical distribution.

iii. We calculate the *p*-value of the test as the proportion of values in the empirical distribution found in (ii) lower than the test value found using *G* in (i).

iv. If the *p*-value is small we reject the DHI assumption. If the *p*-value is big we did not find evidence that the DHI is not satisfied in the given group *G*.

Extension to two or more groups

We note that the absolute value of the test $|T_N|$ and its estimate $|T_n|$ represent a measure of departure from the Discrete Uniform distribution. The bigger the estimate the further is the distance from the uniform distribution and the weaker is the validity of the DHI assumption. Remark 3 also tells us that the nature of the group operation is irrelevant for the testing procedure. Therefore, we can use the test as a tool to compare the strength of the Diffie-Hellman key exchange protocol in two or more groups. In order to do so the order of the groups in the comparison needs to be similar and, more importantly, the sample size on the basis of which we calculate the permutation test needs to be the same. We take advantage of the ability to compare different groups in the next section.

Testing the Diffie-Hellmann indistinguishability assumption in the multiplicative group Z_p^*

We note that the simplest groups $(Z_p,+)$ cannot be tested with the current procedure since the discrete logarithm problem is trivial in these groups. We are going to look at the efficiency of the testing procedure for the finite groups included in Z_p^* with the multiplicative operation. We present the following examples as a way for checking the validity of the testing procedure.

Example 1 (A group where the DDH assumption does nothold.) *Consider $G = Z_p^*$ with p prime. It is known that computing Legendre symbol in this group gives a distinguisher against DDH (Gennaro et al. (2004)).*

Example 2 (A group where the DDH assumption is conjectured tohold) *We currently do not know any DDH distinguisher for a prime order subgroup of Z_p^*. Therefore, given p and q prime with q divisor of p–1 it is conjectured that in a subgroup of order q of Z_p^* the DDH assumption holds.*

We start with a given group *G* and using the test presented in the previous section we will test for the validity of the DHI assumption in that group *G*. This should provide a strong indication towards the security of the Diffie-Hellman key exchange protocol in that group.

The rate of convergence of the testing procedure

Firstly, we investigate is the rate of convergence for our test. To do this we need to calculate the true value of T_N and thus we have to look at small groups.

We plot in Figure 1 the evolution of the test values with the size of the sample. This figure suggest that to get a good estimate for T_N the sample size will depend on the size of the group, for example we need a larger sample size for Z_{11903}^* than we need for Z_{1193}^*.

The figure points out another interesting fact. Following example 5.1 we know that Z_p^* is not secure. It is also conjectured that some groups are more secure than others. Looking at the problem from that perspective, for which groups are more easily broken using the Legendre symbol, it is also assumed that by increasing the size of the group one can make the group more secure.

We can see from the image presented that the second assertion is not true. Just increasing the size of the group does not make it more secure. Remembering that a smaller relative distance corresponds to closeness to the Discrete uniform distribution on the elements of *G*, we see from the Figure 1 that while Z_{11903}^*, the largest group, is the most secure of the three, the situation between the other two groups is not what we would have expected looking at the size of the group alone. Even though Z_{2131}^* is the larger group (almost twice the size), it is also less secure from the DHI assumption perspective than Z_{1193}^*. This indicates that the choice of the group *G* rather than its size is essential for the security of the Diffie-Hellman key exchange protocol.

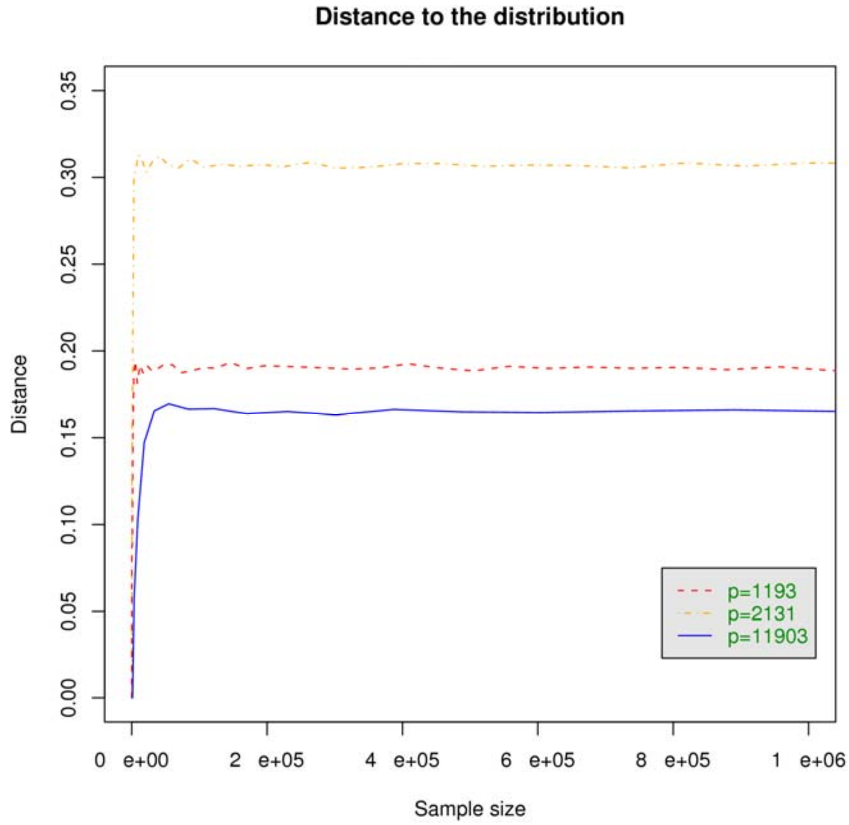


Figure 1: Comparison of the test values for different sample sizes and Z_p^*'s

Comparison of the DHI assumption across groups

Next we wanted to give an indication of groups that are more secure than others. It is known that considering only the Legendre symbol criterion the safest groups among Z_p^* are the ones obtained when p is a safe prime i.e., of the form p=2q+1 where q is another prime Menezes et al. (1996). We shall call any such group a *safe group*.

We tested this theory for a large set of Z_p^* groups with varying p's. We looked at all primes between 2000 and 4000, and again for primes between 9000 and 11000. The reason for the two separate segments of primes is that we expect some sort of consistency between them. We show the distribution of the test values for these groups separated into safe and not safe primes in Figures 2 and 3.

First, we notice that the behavior of primes in the range 2000 to 4000 is very similar with the primes for the higher range 9000 to 11000. Second, in both ranges we see the same conclusion applies, the safe prime groups are more secure than any other groups. However, the test estimate obtained for each of the safe prime groups is significantly different from zero therefore there is no safe group in the ranges given for which the DHI assumption is verified. This seems to confirm the assertion in the Example 5.1.

Next, we look to Example 5.2. We will use our test for the prime subgroups of each of the safe primes in the range 9000 to 11000. More specifically, we look at each Z_p^* with p a safe prime, and we construct the prime subgroup of order q in each such group. Then we test the DHI assumption in each subgroup thus constructed. The values obtained for the distances are plotted in the upper histogram of Figure 4. *We mention that the behavior of the test values for primes between 2000 and 4000 was very similar, and for space consideration we omit the corresponding plot.* All the values are obtained using the same sample size $n = 8 \times 10^6$. The reason for this particular value is that while the groups themselves are in the range 9000 to 11000, the subgroups are of order 4500 to 5500.

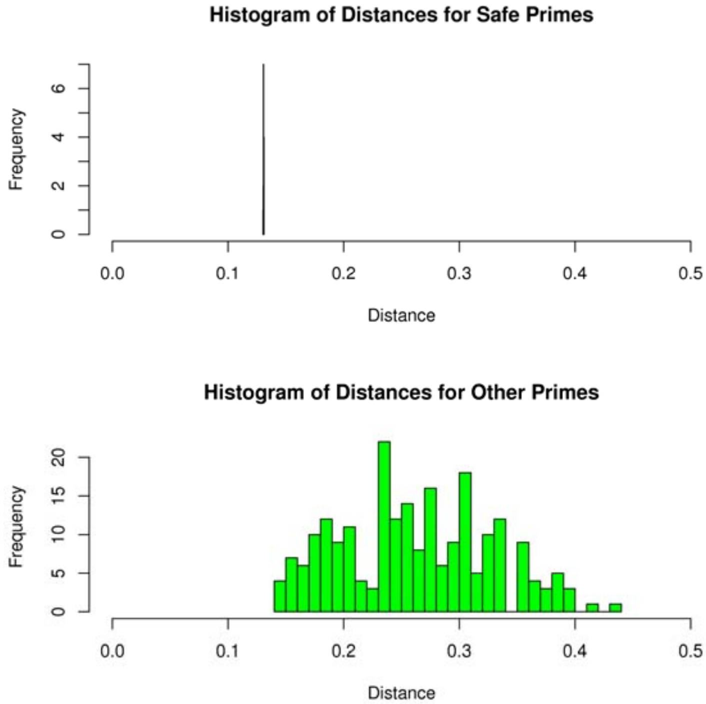


Figure 2: Histogram of all the test values for Z_p^* with 2000<p<4000. Values closer to zero represent safer groups for DH exchange.

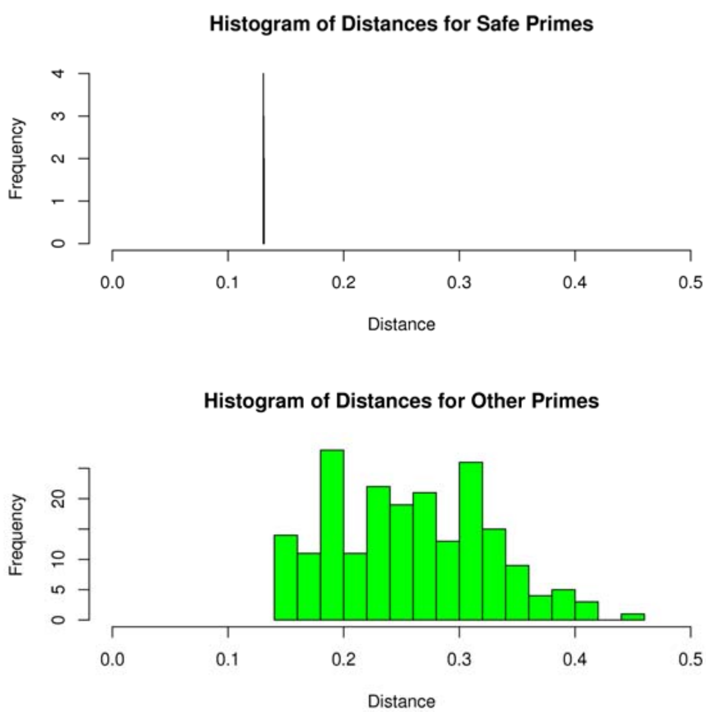


Figure 3: Histograms of test values obtained for Z_p^* with 9000<p<11000. Values closer to zero represent safer groups for DH exchange.

It is remarkable to see that these subgroups are clearly safer for the DH exchange than any other groups plotted in the picture. The results seem to confirm the conjecture in the Example 5.2. *The actual test of uniformity was rejected, but we needed a very large sample size almost equal to the maximum value* N^2.

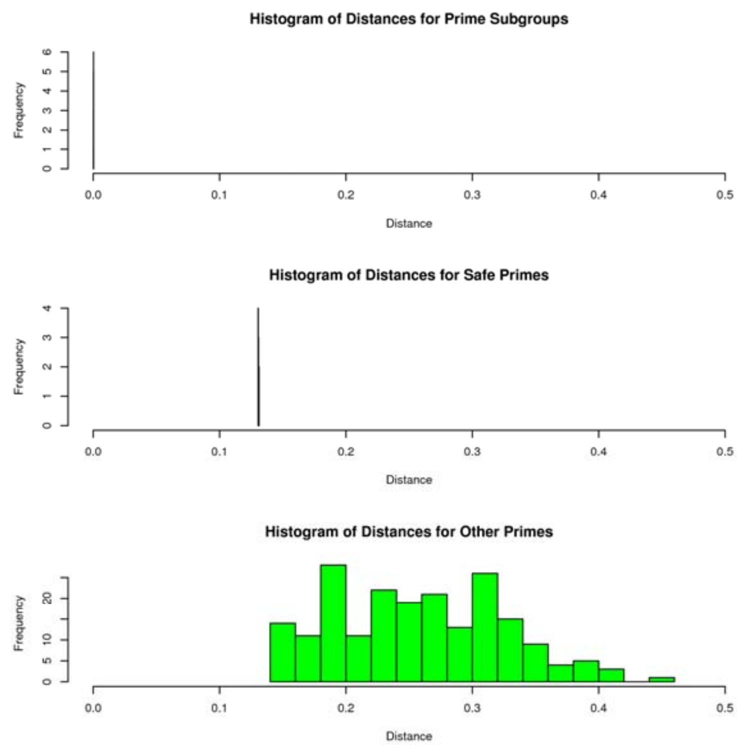


Figure 4: Comparing values of the test for different type of groups when 9000<p<11000. On top, we plot values for prime subgroups of Z_p^* when p is a safe prime. Middle, we plot values for Z_p^* when p is a safe prime. On bottom, we plot values for all the other groups Z_p^* in the range given. Values closer to zero represent better groups for DH exchange.

For a better comparison we plotted in Figure 5 only the values obtained for the prime subgroups of the Z_p^* with p a safe prime (top) and the histogram of the values obtained for the Z_p^* groups, p a safe prime between 9000 and 11000 (bottom).

It is remarkable the closeness of these values to each other considering that the order of the group varies between 9000 and 11000 a 20% variation in size. This is an encouraging fact, which suggests that for even larger p's we will see the same sort of consistency in the values. This will imply that groups with the same operational structure will have similar behavior from the point of view of the Diffie-Hellman security. However, there is a variation in the values as illustrated in the Figure 6 plotting the histogram of the values obtained for the prime subgroup of Z_p^* groups, with p a safe prime varying between 9000 and 11000.

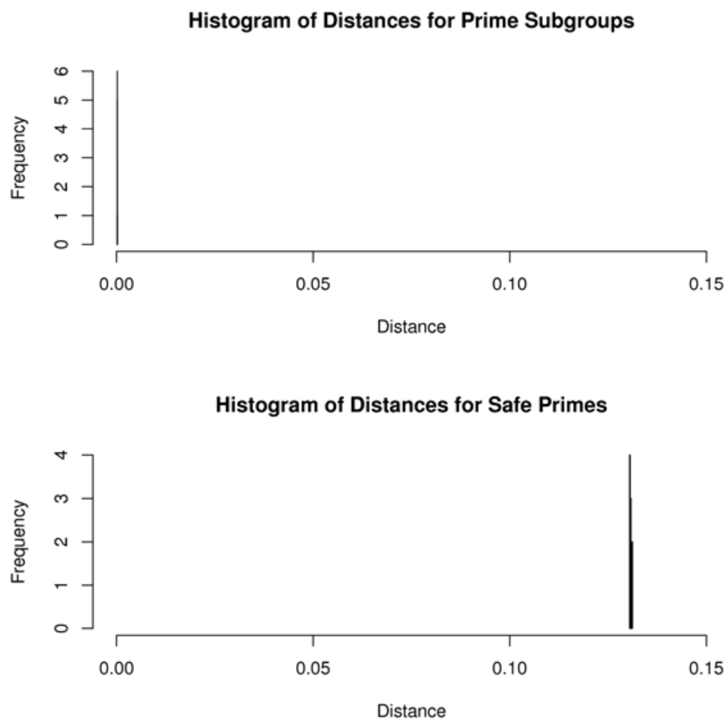


Figure 5: A more detailed comparison of the previous image (Fig. 4). We compare the prime subgroups with the corresponding safe groups. Values closer to zero represent safer groups for DH exchange.

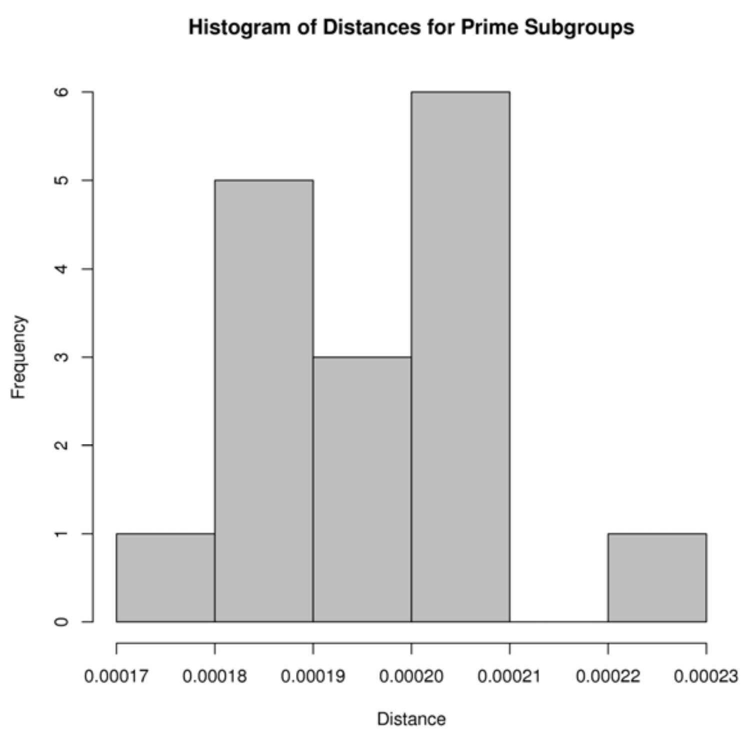


Figure 6: A blowup of the histogram of the values for the prime subgroups in the safe primes. Note the values are close to zero but not equal to zero.

Furthermore, in Figure 7 we plot the test values versus the size of the group from which the prime subgroup originated. We can immediately see that as the size of the group increases the exchange tends to become more secure as measured by our test. This increases our belief in the test results and brings evidence that as the size of the group increases **and the structure of the underlying group remains the same** the security of the exchange increases as well.

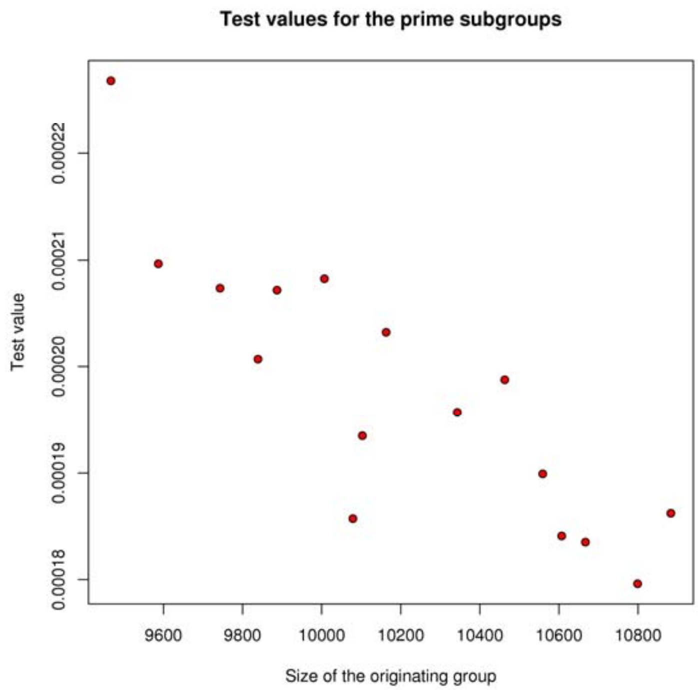


Figure 7: Here we are showing that as the size of the group increases the values of the test do in fact decrease. Here we plot the corresponding values for the histogram in Figure 6. We note that the test values for this part were estimated using the same sample size (8 million) to insure that the test values are comparable and that variability of the test values is the same regardless of the size of the group.

A look at the relationship with the Discrete Logarithm problem

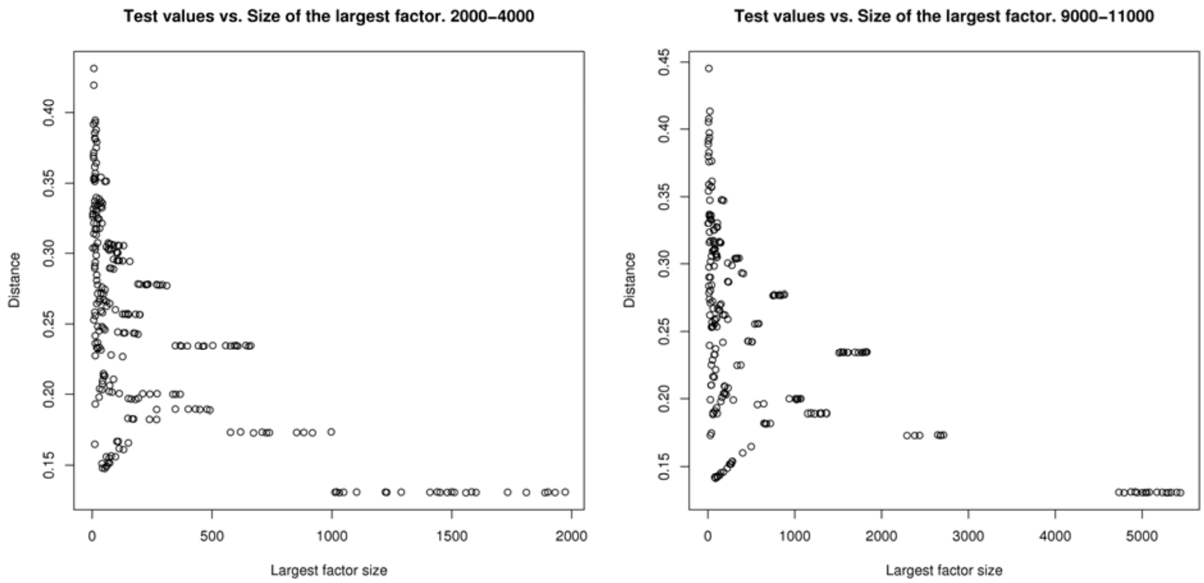
In this section we study in more depth the distribution of the test values as *p* varies in the ranges considered. We calculate the test values for each such group (for such small groups we do not need to estimate or construct statistical distributions for the test values), with the idea to compare the groups themselves from the perspective of the test and identify (if possible) patterns. To our knowledge this is the first approach of this kind.

We first look to $(Z_p^*,\cdot)$, for *p* primes in the two ranges $p\in(2000,4000)$ and $p\in(9000,11000)$. We plot the test values in the Figure 8.

It is known – due to the existence of the Pohlig-Hellman algorithm[4] that in all of these groups the Discrete Logarithm problem is easy and therefore the Diffie Hellman exchange should be breakable. It is also *conjectured* that the actual security depends on the size of the largest factor in the decomposition of $p-1$[5]. For this reason it is believed that the "most secure" groups among $(Z_p^*,\cdot)$ are the ones generated by the safe primes.

[4]which computes the Legendre symbol in these groups and therefore gives a distinguisher against DDH (see Genaro et al. (2004))

[5]This is due to the nature of the algorithm

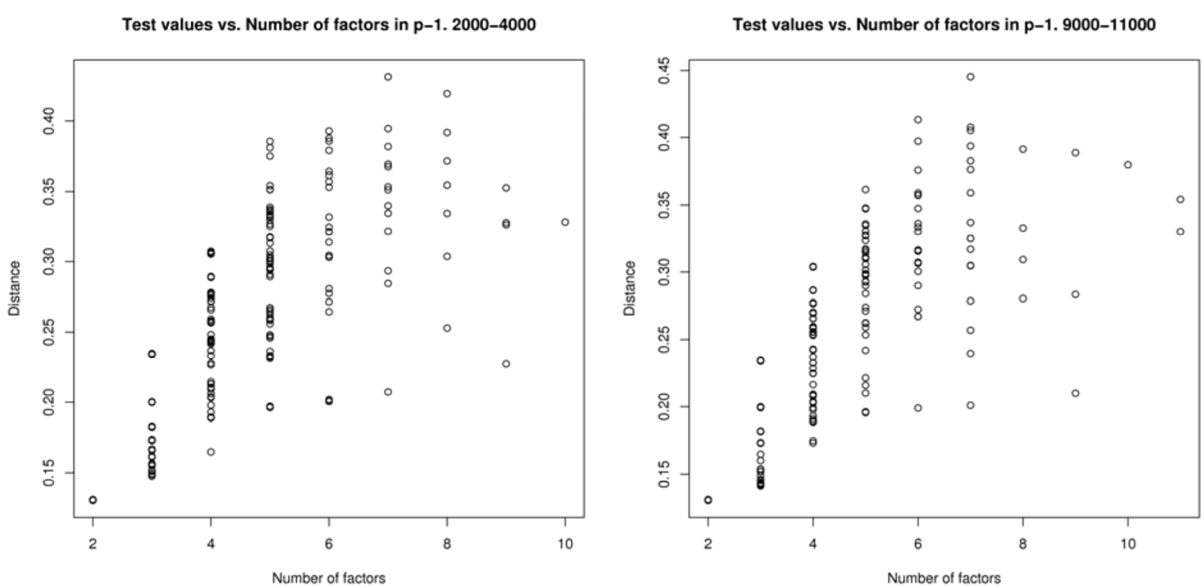


(a) Test values for $p \in (2000,4000)$ (b) Test values for $p \in (9000,11000)$

Figure 8: On the *x* axis we plot the biggest factor in the decomposition of p–1. Values closer to zero on the *y*-axis represent safer groups for DH exchange.

In our case we consider the security of the Diffie Hellman exchange. Since the Discrete Logarithm assumption is a necessary condition for this security we expect that our test will identify these assertions and by using our test we will be able to answer questions related to the DL problem.

Figure 8 presents the test values vs. the size of the largest prime factor in the decomposition of $p-1 = q_1 q_2 ... q_k$. We can immediately see that the structure of the test values for the two ranges is very similar. In both of these images, points corresponding to values closer to 0 on the *y* axis represent groups that are more secure for the DH exchange.



(a) Test values for $p \in (2000,4000)$ (b) Test values for $p \in (9000,11000)$

Figure 9: On the *x* axis we plot the number of factors in the decomposition of p–1. Values closer to zero represent safer groups for DH exchange.

Note that while the points in the lower right corner of the image correspond indeed to the safe primes and they are clearly more secure than the other groups as the popular belief would tell us, we can also see that there exist certain groups which have a small factor (lower left corner) and yet they are comparably secure.

This is investigated further in the Figure 9 where we plot the test values obtained for each group versus the number of factors in the decomposition of $p-1$. While we can see more clearly now that the groups corresponding to the safe primes (x=2 factors in the plot) are indeed more secure than all the other groups, we also find that generally as the number of factors in the decomposition increases the security decreases.

Once again we remark the closeness of the two plots in Figure 9. Based on the two pair of plots it would seem that both the number of factors and the size of the largest factor are important elements when considering the security of the exchange.

But now we are dealing with a statistical problem: trying to relate two determining factors to the variable that quantifies the security of the exchange. There probably exist other factors that are important but let us concentrate on these two for the current work. We know from the statistical theory that if there would be no interaction between the number of factors in the decomposition and the size of the largest factor then we should see points inside each category close to parallel lines. For exemplification we plotted in Figure 10 the same image as in Figure 9(b), but with the points separated by the number of factors in each group. We eliminated the safe groups from the comparison and we only made the picture for $p\in(9000,11000)$ since for the other range the image looks very similar.

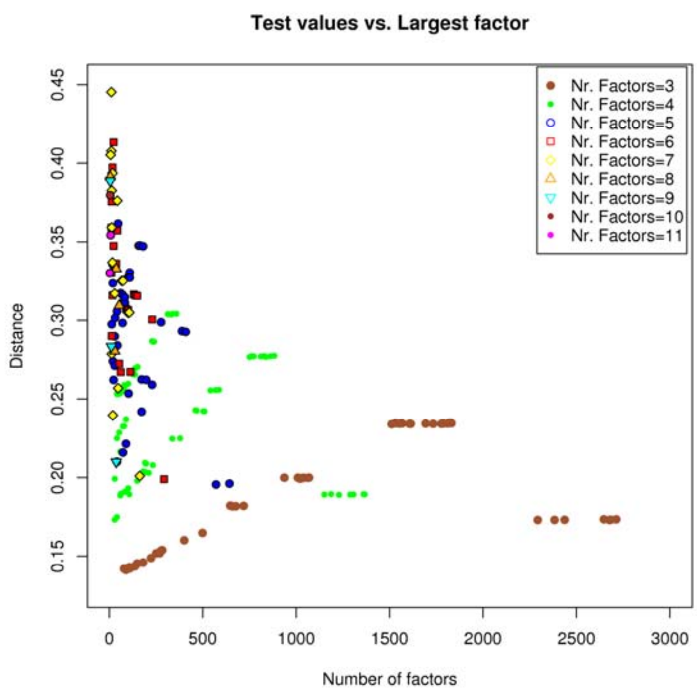
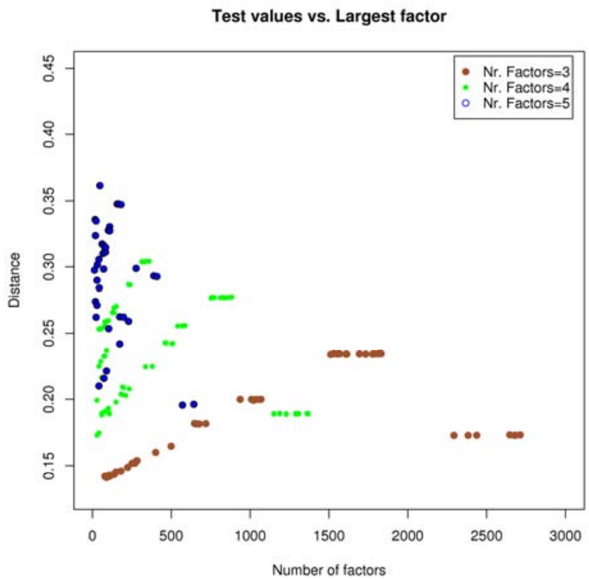
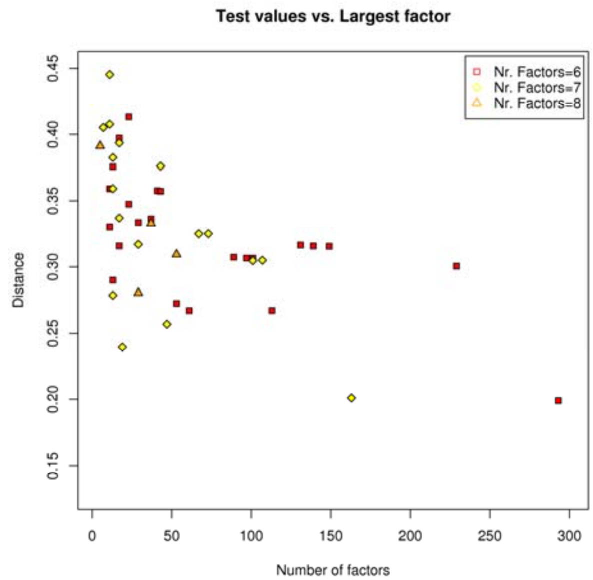


Figure 10: This is the same image as Figure 9(b) but with points corresponding to number of factors in the decomposition of $p-1$ identified.

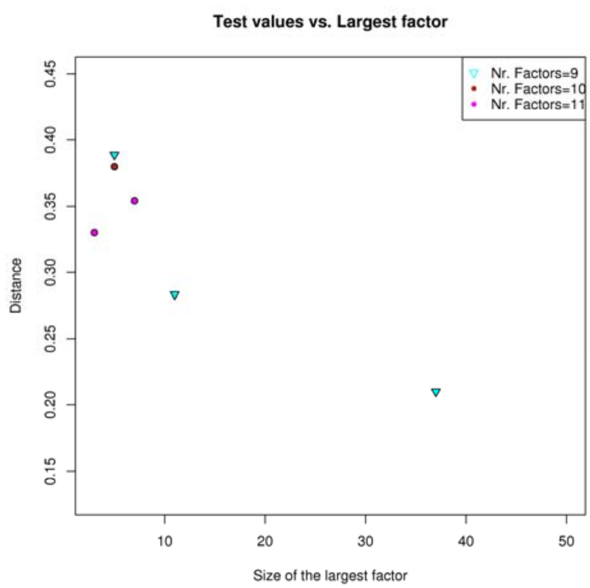
We can start to see that there must be interaction between the two factors. To exemplify better we separated the points depending on the number of factors and we plotted them in Figure 11. We see better that the determining elements for the security of the DH exchange seem to be correlated (they are interacting).



(a) Test values for 3, 4, and 5 factors in the decomposition of $p-1$



(b) Test values for 6, 7, and 8 factors in the decomposition of $p-1$



(c) Test values for 9, 10, and 11 factors in the decomposition of $p-1$

Figure 4: If the two determining elements (number of factors and the size of the largest factor in the decomposition of $p-1$ are independent we should see points of the same color close to parallel lines.

As an example of such discrepancy the group Z_{9473} which is of the order $9473-1=2\times2\times2\times2\times2\times2\times2\times2\times37$, and whose biggest factor in the decomposition is 37 is more secure (test value 0.2) than both groups: Z_{9421} and Z_{9781}, whose decompositions of $p-1$ are $9421-1=2\times2\times3\times5\times157$ and $9781-1=2\times2\times3\times5\times163$ respectively[6].

[6]the test values obtained for these two later groups are very close to each other 0.3475897 and 0.3476914.

Next we analyze statistically the relationship between the test values that quantify the strength of the relationship and the size of the largest factor in the decomposition of $p-1$ (treated as a quantitative variable) and the number of factors in the same decomposition (treated as a categorical variable). We included interaction terms in the model and we present the ANOVA table in Table 1.

Table 2 presents the estimated coefficients of the regression lines for each level, and for each, a test of whether the mean is actually zero. There are 218 primes between 9,000 and 11,000. We note that there was only one prime within the range whose $p-1$ decomposition had 10 factors thus the interaction for that level could not be estimated.

Table 1: ANOVA table for the relationship between the size of the largest factor, the number of factors, and the test values

Factors	**df**	**Deviance**	**df**	**Residual Deviance**
Largest factor size	1	0.37997	216	0.77629
Number of factors	9	0.41977	207	0.35652
Interaction term	8	0.08881	199	0.26772
Error			217	1.15626

Table 2: Effects for each level of factor. The semicolon denotes the levels of the interaction term. The individual factors have to be included even though they appear not significant since the interaction is.

Factor levels	Estimate	Std. Error	t-values	p-values
(Intercept)	1.322e-01	2.221e-01	0.595	0.552573
Largest factor	-2.718e-07	4.353e-05	-0.006	0.995025
Nr factors=3	2.794e-02	2.223e-01	0.126	0.900116
Nr factors=4	1.098e-01	2.222e-01	0.494	0.621938
Nr factors=5	1.766e-01	2.223e-01	0.795	0.427804
Nr factors=6	2.202e-01	2.224e-01	0.990	0.323292
Nr factors=7	2.406e-01	2.225e-01	1.082	0.280778
Nr factors=8	2.462e-01	2.253e-01	1.093	0.275831
Nr factors=9	2.459e-01	2.248e-01	1.094	0.275293
Nr factors=10	2.475e-01	2.249e-01	1.100	0.272574
Nr factors=11	1.798e-01	2.329e-01	0.772	0.440964
maxFact:N.fact=3	2.375e-05	4.399e-05	0.540	0.589972
maxFact:N.fact=4	-7.348e-06	4.505e-05	-0.163	0.870586
maxFact:N.fact=5	-1.260e-04	5.954e-05	-2.115	0.035637 *
maxFact:N.fact=6	-4.137e-04	1.140e-04	-3.629	0.000362 ***
maxFact:N.fact=7	-9.027e-04	2.112e-04	-4.274	2.98e-05***
maxFact:N.fact=8	-1.609e-03	1.060e-03	-1.518	0.130544
maxFact:N.fact=9	-4.757e-03	1.525e-03	-3.119	0.002087 **
maxFact:N.fact=10	NA	NA	NA	NA
maxFact:N.fact=11	6.041e-03	1.297e-02	0.466	0.641823
Signif. codes: ’***’ = 0.001; ’**’= 0.01; ’*’ = 0.05; ’.’ = 0.1				

We can see very clearly from the table that the interaction between the two factors analyzed is significant. We do not present the results for the other range of primes studied 2000–4000 since they are entirely similar.

So what is the conclusion to be drawn from these numbers?

These numbers show that the interaction between the number of factors in the decomposition of p−1 and the size of the largest factor in the decomposition is statistically significant for the security of the Diffie-Hellman security as quantified by our test.

In plain terms, it would seem natural that as the size of the largest factor in the decomposition increases the group becomes more complex and therefore it is more secure. Likewise, as the number of factors in the decomposition increases, there are more equations to solve modulo each factor therefore having a larger number intuitively would also increase the security.

However, as the results in the table show that is not necessarily so, and since the interaction between the two is significant the combination of the two factors is important and the seemingly logical statements presented are not necessarily true.

Future Trends

The papers studying statistical aspects of the distribution of the key of the Diffie-Hellmann exchange are generally concerned with the limiting distribution as the size of the underlying group converges to infinity. However, in practice we do not work with infinite groups and the question of how fast the key distribution converges to infinity is valid and of significant interest. We hope with convinced the reader that the rate of convergence is not uniform across types of the groups and that some group structures lead to a much faster convergence than others.

However it would be much more interesting if we could follow the analysis and observe similar conclusions for very large primes, typically used in cryptography (of the order comparable with 2^{1024}). The use of our testing procedure, ad-literam as in the current work prevents us from analyzing such large groups directly.

In the future we plan to investigate directions of circumventing the permutation testing approach, thus eliminating the need for the sample generation process and transforming the methodology into a practical procedure applicable to big size groups. For this purpose several directions are possible. One direction is to approximate the distribution of the test in (9) with a multinomial distribution, then use a multivariate normal distribution as a second approximation. This would give us an approximate distribution of the test statistic under the null hypothesis, which should allow us to calculate the *p*-value of the test directly without the need of the permutation testing.

Another direction is to put together outcomes into coarser groups and look at the distribution of these groups of outcomes. This idea is similar with the approach of Canetti et al. (1999) and Banks et al. (2006), and will allow us to speed up the procedure in order to apply it to much larger groups.

A third direction is to look at the distribution of the binary representation of prime subgroups of a large group and compare the new resulting groups.

If the computing power suffices or if any of these directions would prove valid the resulting test procedure will allow a comparison between the prime subgroup of a large $\left(Z_p^*,\cdot\right)$ which we asserted to be secure and a similarly sized finite group defined using elliptical curves. This would answer a question of undeniable importance: are the groups constructed using elliptical curves potentially more secure than simpler structure groups?

Conclusion

This paper **does not** break or gives an algorithm to break the Diffie-Hellman exchange. What we do is analyze empirically how hard would it be to break the exchange, *on average, on any random inputs drawn from the underlying group*. The groups under study were small in order (very far from the typical cryptographic groups used in practice), but we give compelling evidence that the security of the exchange tends to be dependent on the structure of the underlying groups. That structure can be recovered and rediscovered over and over as the group size increases.

We have studied the relationship between the security of the Diffie Hellman public key exchange protocol and the structure of the underlying group. We looked at groups were the protocol is provable not secure (of the type $\left(Z_p^*,\cdot\right)$). We have found compelling evidence that breaking it (in the sense of actually finding the key) is dependent not only on the size of the largest factor in the decomposition of *p-1* but also on the number of terms in the decomposition. Furthermore, the relationship is not straightforward (as either one increases the security increases) since the interaction between these two determining factors is statistically significant. This means that it is entirely possible to have a group with large prime factor in the decomposition and a large number of terms in the decomposition of *p-1* and yet to be easier to break (on average for random inputs) than another groups where both these factors are smaller but they interact in a different way.

We show using statistical arguments that the prime subgroups of the groups of type $\left(Z_p^*,\cdot\right)$ are the most secure groups we have studied. Furthermore, if one assumes that the structure of the group from which the subgroups are drawn remains the same, increasing the group's size indeed translates into increasing the security of the Diffie-Hellman exchange as well.

References

Banks, W., J. Friedlander, S. Konyagin, and I. Shparlinski (2006). Incomplete exponential sums and Diffie-Hellman triples. *Math. Proc. Cambridge Philos. Soc. 140*, 193–206.

Boneh, D. (1998). The Decision Diffie-Hellman problem. *Lecture Notes in Computer Science 1423*, 48–63.

Boneh, D. and R. J. Lipton (1996). Algorithms for black-box fields and their application to cryptography (extended abstract). In *CRYPTO '96: Proceedings of the 16th Annual International Cryptology Conference on Advances in Cryptology*, London, UK, pp. 283–297. Springer-Verlag.

Canetti, R., J. Friedlander, S. Konyagin, M. Larsen, D. Lieman, and I. Shparlinski (2000). On the statistical properties of Diffie-Hellman distributions. *Israel Journal of Mathematics 120*(part A), 23–46.

Canetti, R., J. Friedlander, and I. Shparlinski (1999). On certain exponential sums and the distribution of Diffie-Hellman triples. *J. London Math. Soc. 59*, 799–812.

Chaum, D. and H. van Antwerpen (1989). Undeniable signatures. In *CRYPTO '89: Proceedings on Advances in cryptology*, New York, NY, USA, pp. 212–216. Springer-Verlag New York, Inc.

Diffie, W. and M. Hellman (1976). New directions in cryptography. *IEEE Transactions on Information Theory 22*(6), 644–654.

Diffie, W., P. C. V. Oorschot, and M. J. Wiener (1992). Authentication and authenticated key exchanges. *Des. Codes Cryptography 2*(2), 107–125.

El-Gamal, T. (1984). *Cryptography and logarithms over finite fields*. Ph. D. thesis, Elec. Eng. Dept., Stanford Univ., Stanford, CA.

Feldman, P. (1987). A practical scheme for non- interactive verifiable secret sharing. In *Proc. of the 28th FOCS*, pp. 427–437. IEEE.

Friedl, M., N. Provos, and W. Simpson (2006). Diffie-Hellman group exchange for the secure shell (SSH) transport layer protocol. Internet proposed standard RFC 4419.

Friedlander, J. and I. Shparlinski (2001). On the distribution of Diffie-Hellman triples with sparse exponents. *SIAM Journal on Discrete Mathematics 14*, 162–169.

Gennaro, R., H. Krawczyk, and T. Rabin (2004). Secure hashed diffie-hellman over non-ddh groups. In *Advances in Cryptology - EUROCRYPT 2004*, Lecture Notes in Computer Science, pp. 361–381. Springer Berlin / Heidelberg.

Goldreich, O. (2001). *Foundations of Cryptography: Basic Techniques*, Volume 1. Cambridge University Press.

Goldwasser, S. and S. Micali (1984). Probabilistic encryption. *Journal of Computer and System Sciences 28*, 270–299.

Håstad, J., R. Impagliazzo, L. A. Levin, and M. Luby (1999). A pseudorandom generator from any one-way function. *SIAM J. Comput. 28*(4), 1364–1396.

Joux, A. and K. Nguyen (2003). Separating Decision Diffie-Hellman from Computational Diffie-Hellman in cryptographic groups. *Journal of Cryptology 16*, 239–247.

Koblitz, N. and A. J. Menezes (2004). Another look at "Provable Security". Technical report, http://eprint.iacr.org/2004/152.

Kullback, S. and R. A. Leibler (1951). On information and sufficiency. *Annals of Mathematical Statistics* (22), 79–86.

Maurer, U. M. and S. Wolf (1999). The relationship between breaking the diffie–hellman protocol and computing discrete logarithms. *SIAM J. Comput. 28*(5), 1689–1721.

Menezes, A. J., S. A. Vanstone, and P. C. V. Oorschot (1996). *Handbook of Applied Cryptography*. CRC Pr Llc.

Naor, M. and O. Reingold (1997). Number-theoretic constructions of efficient pseudo-random functions. In *FOCS '97: Proceedings of the 38th Annual Symposium on Foundations of Computer Science (FOCS '97)*, Washington, DC, USA, pp. 458. IEEE Computer Society.

Pedersen, T. P. (1991). Distributed provers with applications to undeniable signatures. In *Advances in Cryptology - EUROCRYPT '91: Workshop on the Theory and Application of Cryptographic Techniques*, Lecture Notes in Computer Science, Brighton, UK, pp. 221–242.

Rokhlin, V. A. (1967). Lectures on the entropy theory of measure-preserving transformations. *Russian Mathematical Survey 22*(5), 1–52.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal 27*, 379–423, 623–656.

Stadler, M. (1996). Publicly verifiable secret sharing. In *Advances in Cryptology - EUROCRYPT '96*, Volume 1070 of *Lecture Notes in Computer Science*, pp. 190–199.

Stinson, D. R. (2005). *Cryptography: Theory and Practice* (3 ed.), Volume 36 of *Discrete Mathematics and Its Applications*. University of Waterloo, Ontario, Canada: CRC. Press Online.

Vasco, M. I. G., M. Näslund, and I. Shparlinski (2004). New results on the hardness of Diffie-Hellman bits. In *Proc. Intern. Workshop on Public Key Cryptography*, Volume 2947 of *Lect. Notes in Comp. Sci.*, Singapore, pp. 159–172. Springer-Verlag.

Yao, A. C. (1982). Theory and application of trapdoor functions. In *Proceedings of the 23rd IEEE Symposium on Foundations of Computer Science*, pp. 80–91.

Key Terms and Their Definitions

Generator of a cyclic group: is an element *g* such that all the elements of the group are generated by successive applications of the group operation to *g* itself. Not all the elements in a group are generators.

Subgroup of a group: a set of elements from the initial group which together form a smaller goup structure included in the original group (i.e, the operation stays in the subgroup, the identity and the inverse elements are in the subgroup) . An example is the trivial subgroup $\{\hat{1}\}$.

Prime group: a group that contains no subgroups except for the trivial subgroup. A *prime subgroup* is a subgroup of a group that contains no further subgroups except for the trivial subgroup. An example is $\{\hat{1},\hat{4}\}$ included in $\left(Z_5^*,\cdot\right)$

p-value of a test: the probability of obtaining as extreme or more extreme values as the result of the experiment assuming that the null hypothesis is true. Numbers close to *0* are evidence against the null hypothesis (it is unlikely to see such numbers if the null hypothesis would be true).

Statistically indistinguishable random variables: are two or more random variable whose distribution is identical almost everywhere (with the possible exception of a set of probability measure zero).

Cryptographic key: a piece of information that controls the operation of a cryptographic algorithm.

Encryption key: a piece of information used to specify the particular transformation of plaintext into ciphertext, or vice versa during the encryption/decryption process.