When are Identification Protocols with Sparse Challenges Safe? The Case of the Coskun and Herley Attack

Hassan Jameel Asghar and Mohamed Ali Kaafar

Data61, CSIRO, Sydney, Australia
{hassan.asghar, dali.kaafar}@data61.csiro.au

December 24, 2015

Abstract

Cryptographic identification protocols enable a prover to prove its identity to a verifier. A subclass of such protocols are shared-secret challenge-response identification protocols in which the prover and the verifier share the same secret and the prover has to respond to a series of challenges from the verifier. When the prover is a human, as opposed to a machine, such protocols are called human identification protocols. To make human identification protocols usable, protocol designers have proposed different techniques in the literature. One such technique is to make the challenges *sparse*, in the sense that only a subset of the shared secret is used to compute the response to each challenge. Coskun and Herley demonstrated a generic attack on shared-secret challenge-response type identification protocols which use sparse challenges. They showed that if the subset of the secret used is too small, an eavesdropper can learn the secret after observing a small number of challenge-response pairs. Unfortunately, from their results, it is not possible to find the safe number of challenge-response pairs a sparse-challenge protocol can be used for, without actually implementing the attack on the protocol and weeding out unsafe parameter sizes. Such a task can be time-consuming and computationally infeasible if the subset of the secret used is not small enough. In this work, we show an analytical estimate of the number of challenge-response pairs required by an eavesdropper to find the secret through the Coskun and Herley attack. Against this number, we also give an analytical estimate of the time complexity of the attack. Our results will help protocol designers to choose safe parameter sizes for identification protocols that employ sparse challenges.

1 Introduction

An identification protocol is a cryptographic protocol through which a prover $\mathcal{P}$ verifies its identity to a verifier $\mathcal{V}$ [1]. A protocol is referred to as a challenge-response type identification protocol when $\mathcal{P}$ and $\mathcal{V}$ share the same secret, and $\mathcal{P}$ responds to a series of challenges from $\mathcal{V}$.[1] Usually, the response is computed from a publicly known function f of the secret s and the challenge c so that the verifier can check if the responses are correct at its end. When the prover $\mathcal{P}$ is a human, the identification protocol is generally known as a human identification protocol. The holy grail of protocol designers in the realm of human identification protocols is to devise a function f that is simple enough

[1] See Section 4 for an example of such protocols.

for a human to *mentally* compute yet requires the eavesdropper to observe a sizeable number of challenge-response pairs to reconstruct the secret.

Another way to make the identification protocol practical for humans is to reduce the size of the challenge c such that the function f is applicable to only a small fraction u of the shared secret s. We refer to such protocols as *sparse-challenge* protocols. Given a generic sparse-challenge protocol, Coskun and Herley [2] showed an attack that exploits the observation that when u is small, *candidates* of the secret that are close to the secret s (in terms of a distance metric) yield similar responses to s when f is applied on them. In a nutshell, their attack samples a large enough subset S' of the set of all possible secrets S such that with high probability there is at least one element (candidate) in S' that is a distance ξ from the target secret s. The attacker can apply the function f on each of the candidates in S' and the observed challenges to weed out those candidates whose responses are further away from the observed responses.[2] For more details, see [2]. We call this attack, the CH attack in short.

The CH attack demonstrates that with small values of u, say ≤ 10, sparse-challenge identification protocols are not secure in the sense that the attack is computationally feasible with a small number of observed challenge-response pairs m. Furthermore, the complexity of the attack can be decreased by observing more challenge-response pairs. However, Coskun and Herley's work leaves open the following question: Suppose a time complexity of 2^λ is considered infeasible for some λ,[3] what value of m is *safe* enough in the sense that if the eavesdropper observes $\leq m$ challenge-response pairs, the CH attack has complexity at least 2^λ? Indeed, this question is important for protocol designers who wish to use higher values of u (perhaps for authentication of pervasive devices if higher values of u are considered infeasible for humans). While it is possible to implement the CH attack to check its feasibility on a given protocol when the sizes of the protocol parameters are small, for larger sizes and large m this may be impractical.

In this paper, we describe an analytical estimate for m that is safe against the CH attack. Against this safe value of m we also describe a *simpler* estimate of the work factor (WF) of the CH attack, compared to the analytically difficult expression obtained by Coskun and Herley in [2], to determine the complexity of the attack. Our results can help protocol designers in setting sizes of the protocol parameter, such that the protocol can safely be used for m rounds (challenge-response pairs) against the CH attack with complexity $\approx 2^\lambda$, where λ can be chosen according to what is considered infeasible.[4]

2 Preliminaries

Let C denote the challenge space, R the response space and S the secret space, all three being finite sets. A member $c \in C$ is called a challenge, $r \in R$ a response and $s \in S$ a secret. We let $\log_2 |S| = \eta$. Then a secret $s \in S$ is represented as a binary string of η bits. A fraction $u < \eta$ are used to compute a function $f : C \times S \to R$.[5] Given two strings x_1 and x_2 of equal length, the Hamming distance, denoted $d(x_1, x_2)$, is the number of positions at which x_1 and x_2 differ. For two secrets $s_1, s_2 \in S$, it follows that $0 \leq d(s_1, s_2) \leq \eta$. If $d(s_1, s_2) = i$ we say that s_1 is a distance-i neighbour of s_2 (and vice versa). We let $s_0 \in S$ denote the *target* secret. Any other element $s \in S - \{s_0\}$ is then called a *candidate* for the secret, or simply a candidate. Given a sequence of challenges $(c_1, \ldots, c_m)$, let $\Gamma(s)$ denote the string of responses $f(c_1, s)||\cdots||f(c_m, s)$, for some $s \in S$. Where there is no ambiguity, we shall denote $\Gamma(s)$ simply by Γ. The

[2]The notion of distance shall be made exact later.

[3]For instance, $\lambda = 80$.

[4]After m rounds, the secret can be renewed.

[5]The verifier in fact samples a random $c \in C$ in each round such that a random u out of η bits of s are in c, which are then used to compute f.

response stream of the target secret s_0 shall be denoted by Γ_0. For integers x and y, we use the convention that the binomial coefficient $\binom{x}{y} = 0$, whenever $x < y$. The following two well-known results will be used later.

Theorem 1 (Central Limit Theorem [3, §8.3, p. 434])**.** *Let $X_1, X_2, \ldots, X_m$ be a sequence of i.i.d. random variables each having mean μ and variance σ^2. Then*

$$\mathbb{P}\left[\frac{\sum_{i=1}^{m} X_i - m\mu}{\sigma\sqrt{m}} \leq a\right] \to \Phi(a) \text{ as } m \to \infty,$$

where $\Phi(\cdot)$ denotes the cumulative distribution function of the standard normal distribution.

Theorem 2 (Hoeffding's Inequality [4, p. 217])**.** *Let $X_1, X_2, \ldots, X_m$ be independent bounded random variables such that X_i falls in the interval $[a_i, b_i]$ with probability one. Then for any $t > 0$,*

$$\mathbb{P}\left[\sum_{i=0}^{m} X_i - \sum_{i=0}^{m} \mathbb{E}[X_i] \geq t\right] \leq \exp\left(-\frac{2t^2}{\sum_{i=0}^{m}(b_i - a_i)^2}\right),$$

and

$$\mathbb{P}\left[\sum_{i=0}^{m} X_i - \sum_{i=0}^{m} \mathbb{E}[X_i] \leq -t\right] \leq \exp\left(-\frac{2t^2}{\sum_{i=0}^{m}(b_i - a_i)^2}\right).$$

The following result will come handy

Proposition 1. *For all $x \in \mathbb{R}, x(1-x) \leq \frac{1}{4}$.*

Let X denote a Bernoulli random variable with distribution $P(x)$, for $x \in X$. We shall denote the probability that $X = 1$ as $\mathbb{P}(X = 1) = P(1) = p$. Let P and Q be two probability distributions, then the Kullback-Leibler divergence of Q from P is

$$D(P \,||\, Q) = \sum_{x \in X} P(x) \log_2 \frac{P(x)}{Q(x)},$$

which for Bernoulli distributions P and Q is,

$$D(P \,||\, Q) = p \log_2 \frac{p}{q} + (1-p) \log_2 \frac{1-p}{1-q},$$

where we use the convention that $\log_2 \frac{0}{a} = 0$ for any $a \geq 0$. The following useful bound bears resemblance to a result from [5, p. 2].

Theorem 3. *Let P and Q be two Bernoulli distributions with $0 < q < p < 1$. Then $D(P \,||\, Q) \leq \frac{1}{\theta}(p-q)^2$, where $\theta = \min_x\{x(1-x)\}$ for all $x \in [q, p]$.*

Proof.

$$\begin{aligned}
D(P \,||\, Q) &= p \log_2 \frac{p}{q} + (1-p) \log_2 \frac{1-p}{1-q} \\
&= p \log_2 \frac{p}{e} \frac{e}{q} + (1-p) \log_2 \frac{1-p}{e} \frac{e}{1-q} \\
&= p \ln \frac{p}{q} + (1-p) \ln \frac{1-p}{1-q} \\
&= p \int_q^p \frac{1}{x} dx - (1-p) \int_q^p \frac{1}{1-x} dx \\
&= \int_q^p \left(\frac{p-x}{x(1-x)}\right) dx \\
&\leq \frac{1}{\theta} \int_q^p (p-x) dx = \frac{1}{\theta}(p-q)^2.
\end{aligned}$$

$\square$

Finally we shall utilize Sanov's theorem.

Theorem 4 (Sanov's Theorem [6, §11.4, p. 362])**.** *Let $X_1, X_2, \dots, X_m$ be i.i.d. with distribution Q and let $Q^m = \prod_i Q(x_i)$. Let E be a set of distributions such that E is the closure of its interior, then*

$$Q^m(E \cap \mathcal{P}_m) = Q^m(E) \to 2^{-mD(P^* || Q)} \text{ as } m \to \infty,$$

where $\mathcal{P}_m$ is the set of all empirical probability distributions with denominator m and P^* is the distribution in E that minimizes $D(P \mid\mid Q)$ for $P \in E$.

Corollary 1. *Let $X_1, X_2, \dots, X_m$ be i.i.d. Bernoulli random variables with distribution Q. Then,*

$$\mathbb{P}\left[\sum_{i=1}^{m} X_i \geq m\alpha\right] = Q^m(E),$$

for $0 < \alpha < 1$, where E is the set of distributions

$$E = \left\{ P \;\middle|\; \sum_{x \in \{0,1\}} P(x) \geq \alpha = P(1) \geq \alpha \right\}.$$

Moreover, the P^* that minimizes $D(P \mid\mid Q)$ is P_α given by $P_\alpha(1) = \alpha$.

Proof. The proof of the first part is in [6, §11.4, p. 361]. For the second part, the distribution P^* that minimizes $D(P \mid\mid Q)$ for $P \in E$ is given by [6, §11.5, p. 364]

$$P^*(x) = \frac{q^x(1-q)^{(1-x)} \exp(\kappa x)}{q \exp(\kappa x) + 1 - q},$$

where κ is chosen such that $P^*(1) = \alpha$. Setting $P^*(1) = \alpha$ in the above, we get

$$\exp(\kappa) = \frac{\alpha}{1-\alpha} \frac{1-q}{q},$$

which after substitution yields $P^*(x) = \alpha^x(1-\alpha)^{(1-x)} = P_\alpha(x)$. $\square$

3 Analysis of the Coskun and Herley Attack

We shall begin this section by giving an overview of the Coskun and Herley (CH) attack. After that we shall derive an estimate of a safe m, i.e., the allowed number of challenge-response pairs. Finally we shall give an estimate of the work factor of the CH attack against this safe m.

3.1 Overview of the CH Attack

We assume an eavesdropper, i.e., a passive adversary, observing challenge-response pairs exchanged between $\mathcal{P}$ and $\mathcal{V}$ who share a target secret $s_0 \in S$. Each challenge-response pair is assumed to come from one *round* of the protocol. Our discussion in this section is from the adversary's viewpoint. Assume we have observed m challenge-response pairs, and we wish to retrieve the target secret s_0. Fix a $\xi \in \{1, \dots, \eta - u\}$. The CH attack is as follows [2, p. 434]:

Attack: `The CH Attack`

1 Sample a random subset S' of S with $2^\eta/\binom{\eta}{\xi}$ candidates.
2 **for** *each candidate* s' *in* S' **do**
3 | Initialize a list with s'.
4 | **for** $i = 0, 1, \ldots, \xi - 1$ **do**
5 | | Compute $d(\Gamma_0, \Gamma)$ of each distance-1 neighbour s of each element in list.
6 | | Retain τ candidates in the list that maximise $d(\Gamma_0, \Gamma)$.
7 | | **if** $d(\Gamma_0, \Gamma) = m$ *for any candidate secret in the list* **then**
8 | | | Output the candidate and halt.

The parameter τ is used to reduce the complexity of the attack. Coskun and Herley use $\tau = 10$ (which we shall also adopt for our simulations). Note that if $\xi = 0$ the complexity of the attack is equivalent to brute-force, and we therefore use $\xi > 0$. The reason for choosing $2^\eta/\binom{\eta}{\xi}$ candidates is to expect at least one distance-ξ neighbour of s_0 to be present. Increasing this to a factor $2^{\eta+q}/\binom{\eta}{\xi}$, where $q \geq 1$ increases the probability of this event [2]. However, $q = 0$ corresponds to least computational overhead, and hence we use this in this paper. A glance at the attack reveals that a large ξ reduces the complexity of the attack (since $2^\eta/\binom{\eta}{\xi}$ decreases as ξ increases), while a small ξ increases the probability of finding the secret (as there are less number of iterations, i.e., step 4, and therefore less chance of error). For large values of ξ, we therefore need a larger value of m to successfully find the secret. Next, we show how to estimate a "safe" m, denoted $\hat{m}$, for a given ξ. By safe we mean that if the adversary observes less than $\hat{m}$ challenge-response pairs, the success probability of the CH attack is low. After estimating an expression for $\hat{m}$, we shall derive a simpler estimate for the work factor (WF) of the CH attack obtained by Coskun and Herley, first for a general m and then in particular for a given value of $\hat{m}$.

3.2 Estimating a Safe m

For reasons of usability, the size of the response space R in a human identification protocol is in general small. Thus, there is a high probability $|R|^{-1}$ that a candidate s has the same response as s_0 on a given challenge. Over m challenges this probability reduces to $|R|^{-m}$. Thus a fraction $\frac{\eta}{|R|^m}$ of the candidates, i.e., elements of S, agree with the response stream Γ_0 of length m. This imposes an *information theoretic* bound on m beyond which we expect only the target secret s_0 to satisfy the target response stream Γ_0. This bound, denoted m_{it}, is given as $\frac{\eta}{\log_2 |R|}$. Thus the attacker has to observe at least m_{it} challenge-response pairs to obtain the unique secret, independent of any attack. Our estimation of a safe m, i.e., $\hat{m}$, for the CH attack shall be meaningful only when it is higher than m_{it}.

Let $s \in S$ be such that $d(s_0, s) = i$. Given a uniformly random $c \in C$, let p_i denote the probability that $f(c, s) = f(c, s_0)$. Then

$$\begin{aligned} p_i &= a_i + (1 - a_i)\frac{1}{|R|} \\ &= a_i\left(1 - \frac{1}{|R|}\right) + \frac{1}{|R|}, \end{aligned} \tag{1}$$

where

$$a_i \stackrel{\text{def}}{=} \frac{\binom{\eta-i}{u}}{\binom{\eta}{u}}. \tag{2}$$

Intuitively, a_i denotes the probability that the u bits of the candidate s chosen to respond to the challenge are the same as those of the secret s_0, which is a distance i from s.

Theorem 5. *Let $i < j < \eta$, then $p_i \geq p_j$, with equality if and only if $\eta - i$ and $\eta - j$ are both less than u.*

Proof. When both $\eta - i$ and $\eta - j$ are less than 0, then by convention $\binom{\eta-i}{u} = \binom{\eta-j}{u} = 0$, and hence $p_i = p_j = \frac{1}{|R|}$. When $\eta - i \geq u$ and $\eta - j < u$, then $a_i \geq 1$, whereas $a_j = 0$, from which it follows that $p_i > p_j$. We are left with the case when $\eta - i > u$ and $\eta - j \geq u$. Since $i < j < \eta$, we have $\eta - i > \eta - j$ which allows us to write

$$\begin{aligned}\binom{\eta-i}{u} &= \frac{(\eta-i)}{(\eta-i-u)} \cdot \frac{(\eta-i-1)}{(\eta-i-u-1)} \cdots \frac{(\eta-j+1)}{(\eta-j+1-u)} \cdot \binom{\eta-j}{u} \\ &> 1 \cdot 1 \cdots 1 \cdot \binom{\eta-j}{u} \\ &= \binom{\eta-j}{u}.\end{aligned}$$

This implies that $a_i > a_j$ and consequently $p_i > p_j$. □

Figure 1 shows the value of p_i as i ranges from 0 to η, for $\eta = 80$, $u = 5$ and $|R| = 4$. Notice how after around $i = \frac{\eta}{2}$ the p_i's are closer to $|R|^{-1} = 0.25$. Now, as before let

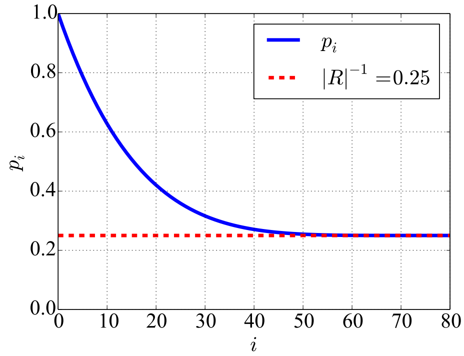


Figure 1: The graph of p_i as i ranges from 0 to η, for $\eta = 80$, $u = 5$ and $|R| = 4$.

$s \in S$ be such that $d(s, s_0) = i$. Then the probability that the response stream of s, i.e., Γ, is a distance γ from Γ_0 is given by [2]

$$\mathbb{P}[d(\Gamma_0, \Gamma) = \gamma \mid d(s_0, s) = i] = b(\gamma, m, p_i), \tag{3}$$

where $b(\gamma, m, p_i)$ is the probability mass function of the binomial distribution given by

$$b(\gamma, m, p_i) = \binom{m}{\gamma} p_i^{\gamma} (1 - p_i)^{m-\gamma}. \tag{4}$$

Let $d(s_0, s) = \xi$, for some $s \in S$. We shall denote such an s by s_ξ. Given s_ξ, let $s_{\xi-1}$ and $s_{\xi+1}$ denote two neighbours of s_ξ such that $d(s_0, s_{\xi-1}) = \xi - 1$ and $d(s_0, s_{\xi+1}) = \xi + 1$. We are first interested in finding an estimate for m, the number of samples (challenge-response pairs) required to distinguish between $\Gamma(s_{\xi-1}) = \Gamma_{\xi-1}$ and $\Gamma(s_{\xi+1}) = \Gamma_{\xi+1}$. Such a value of m ensures that with high probability distance-$(\xi - 1)$ candidates will be retained in the CH attack as opposed to distance-$(\xi + 1)$, which in turn will help the attack to move iteratively closer to the target secret s_0 [2]. If m is such that it is hard to distinguish between the response streams $\Gamma_{\xi-1}$ and $\Gamma_{\xi+1}$ then the CH attack will not converge to the target secret s_0. Therefore this is a necessary condition for the success of the CH attack. We will therefore obtain an expression for a safe m, i.e., $\hat{m}$, through this step.

To do this, first define the Bernoulli random variables

$$X = \begin{cases} 1 & \text{with probability } p_{\xi-1} \\ 0 & \text{otherwise} \end{cases}$$

and,

$$Y = \begin{cases} 1 & \text{with probability } p_{\xi+1} \\ 0 & \text{otherwise} \end{cases}.$$

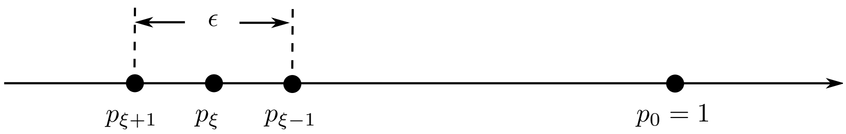
Intuitively, X denotes the indicator random variable which is 1 when $s_{\xi-1}$ has the same response as s_0. Likewise for Y. Further define the random variable $Z = X - Y$. The expected value of Z is given by

$$\mathbb{E}[Z] \stackrel{\text{def}}{=} \mu = \mathbb{E}[X] - \mathbb{E}[Y] = p_{\xi-1} - p_{\xi+1} \stackrel{\text{def}}{=} \epsilon, \tag{5}$$

and the variance (assuming X and Y to be independent) is given by

$$\text{Var}[Z] \stackrel{\text{def}}{=} \sigma^2 = (1)^2\text{Var}[X] + (-1)^2\text{Var}[Y] = p_{\xi-1}(1 - p_{\xi-1}) + p_{\xi+1}(1 - p_{\xi+1}). \tag{6}$$

Note that according to Theorem 5, $\epsilon = p_{\xi-1} - p_{\xi+1} > 0$. This is depicted pictorially below.



Given m i.i.d. random variables Z_i of type Z, we are then interested in the probability

$$\mathbb{P}\left[\sum_{i=0}^{m} Z_i \leq 0\right], \tag{7}$$

which estimates the probability that in m challenges the response stream of $s_{\xi-1}$ is at least as close to s_0 as that of $s_{\xi+1}$. In essence, the probability in Eq. 7 is the *error probability* that given the stream $\Gamma_{\xi+1}$, a *distinguisher* erroneously decides that it belongs to $s_{\xi-1}$. Being asked to give a binary decision, we can assume that the error probability of the distinguisher is less than or equal to $\frac{1}{2}$. We are interested in finding a bound on the number of samples m such that the error probability is close to $\frac{1}{2}$. This implies that the distinguisher will not be able to differentiate between the two response streams $\Gamma_{\xi-1}$ and $\Gamma_{\xi+1}$, and hence with high probability CH attack will fail to output the target secret s_0.

Now, we can write Eq. 7 as

$$\begin{aligned} \mathbb{P}\left[\sum_{i=0}^{m} Z_i \leq 0\right] &= \mathbb{P}\left[\sum_{i=0}^{m} Z_i - m\mu \leq -m\mu\right] \\ &= \mathbb{P}\left[\frac{\sum_{i=0}^{m} Z_i - m\mu}{\sigma\sqrt{m}} \leq -\frac{m\mu}{\sigma\sqrt{m}}\right] \\ &= \mathbb{P}\left[\frac{\sum_{i=0}^{m} Z_i - m\mu}{\sigma\sqrt{m}} \leq -\frac{\sqrt{m}\mu}{\sigma}\right] \\ &\to \Phi\left(-\frac{\sqrt{m}\mu}{\sigma}\right) \quad \text{as } m \to \infty, \end{aligned} \tag{8}$$

where we have applied Theorem 1 in the last step. Fix an error probability δ. Then we want

$$\begin{aligned} \Phi\left(-\frac{\sqrt{m}\mu}{\sigma}\right) &\geq \delta \\ \Rightarrow -\sqrt{m} &\geq \frac{\sigma}{\mu}\Phi^{-1}(\delta) \\ \Rightarrow m &\leq \left(\frac{\sigma}{\mu}\Phi^{-1}(\delta)\right)^2, \end{aligned} \tag{9}$$

where the change in the inequality sign follows since $\Phi^{-1}(\delta) \leq 0$ for $\delta \leq \frac{1}{2}$. Note that we could also use a concentration inequality, such as Hoeffding's inequality (Theorem 2), to obtain

$$\begin{aligned} \mathbb{P}\left[\sum_{i=0}^{m} Z_i \leq 0\right] &= \mathbb{P}\left[\sum_{i=0}^{m} Z_i - \sum_{i=0}^{m} \mathbb{E}[Z_i] \leq -m\epsilon\right] \\ &\leq \exp\left(-\frac{2\epsilon^2}{\sum_{i=0}^{m} 4}\right) \\ &= \exp\left(-\frac{\epsilon^2 m}{2}\right), \end{aligned}$$

where we have used the fact that $a_i = -1$, $b_i = 1$ and $\mathbb{E}[Z_i] = \mu = \epsilon$. But this serves as an upper bound on the tail of the error probability, and will only give us a lower bound

$$m \geq -\frac{2}{\epsilon^2}\ln(\delta), \tag{10}$$

for the minimum number of samples required for a fixed upper bound on the tail of the error probability δ. This is contrary to our purpose, which is to find a "safe" upper bound on m such that the error probability is close to $\frac{1}{2}$.

We define our bound on m from Eq. 9 by fixing $\delta = 0.495$, which gives us

$$\hat{m} \stackrel{\text{def}}{=} \left(\frac{\sigma}{\mu}\Phi^{-1}(0.495)\right)^2 \approx \frac{\sigma^2}{\mu^2}(-0.0125)^2 \approx \frac{\sigma^2}{\mu^2}0.00016 = \frac{0.00016\sigma^2}{\epsilon^2}, \tag{11}$$

where values of ϵ and σ^2 are as given in Eq. 5 and Eq. 6, respectively.

3.2.1 Empirical Evaluation

Since this $\hat{m}$ corresponds to a probability of error close to $\frac{1}{2}$, the CH attack should be unsuccessful with (observed) $m \leq \hat{m}$. Since we rely on some simplifying assumptions to obtain this estimate, such as the independence of X and Y, and normal approximation through the central limit theorem, it is worthwhile to verify this estimate empirically. To do this, we ran the *subroutine* of the CH attack which uses a distance-ξ neighbour of s_0 on a simulated identification protocol (as was done by Coskun and Herley [2]). More specifically, we choose a random $s \in S - \{s_0\}$ such that $d(s, s_0) = \xi$, and run steps 2 to 8 of Attack 1. This relieves us of having to weed through $2^\eta / \binom{\eta}{\xi}$ candidates, which would require considerable time. Thus, a random $s \in S - \{s_0\}$ was chosen each time such that $d(s, s_0) = \xi$. Each new random challenge was simulated by randomly sampling u bits out of η of the (initially randomly chosen) target secret s_0 (thus simulating the fact that a random challenge would contain a random u bits of a secret). The response was generated randomly from the set $\{0, 1, \ldots, |R| - 1\}$. In order to make the responses consistent in case two challenges contain the same u bits of the secret, a hash table was created which had the u-bit string as the key, and the response as the value. For different sets of values of η and u, we ran the CH attack subroutine 25 times once per each value of ξ, starting at 1, until the estimate $\hat{m}$ was greater than $10,000$ challenge-response pairs. The results are shown in Figure 2. The information theoretic lower bound $m_{\text{it}} = \frac{\eta}{\log_2 |R|}$ is also indicated in the plots. We can see from the plots that with higher values of η and u, i.e., Figures 2e, 2f, 2h, 2i, 2k and 2l, the success probability is 0 against $\hat{m}$ obtained through Eq. 11. For Figure 2l, we extend the *cut-off* value of $\hat{m}$ to $60,000$ challenge-response pairs. As is shown, the success rate of the CH attacks is still 0. We remark that while we have chosen $\delta = 0.495$, any value of δ close to 0.5 should suffice. For instance, our simulations also found $\delta = 0.490$ to be a safe choice. Of course, $\delta = 0.490$ gives a higher value of $\hat{m}$ versus $\delta = 0.495$ for a fixed value of ξ. Lowering δ further, say to 0.400, is not recommended as a safe choice. Figure 3 shows why.

3.3 Estimating the Work Factor

In this section we obtain a simple analytical estimate for the work factor (WF) of the CH attack. Note that we are interested in finding a value of λ such that $2^\lambda \approx \text{WF}$. Therefore, an estimate that is off by a couple of powers of 2 is sufficient for our purposes. Whenever we shall plot WF, it will be in $\log_2$-scale. The work factor of the CH attack is given by [2, p. 434][6]

$$\text{WF} = \frac{1}{\binom{\eta}{\xi}} \left(2^\eta + \tau\eta\xi \sum_{i=0}^{\eta} \binom{\eta}{i} \sum_{j=mp_\xi}^{m} \binom{m}{j} p_i^j (1-p_i)^{m-j} \right), \tag{12}$$

where again τ is a threshold set at 10 by Coskun and Herley, and mp_ξ is the boundary chosen such that candidates with response streams that are a distance less than mp_ξ from the target secret's response stream are discarded [2]. Define

$$\text{WF}_1 = \frac{2^\eta}{\binom{\eta}{\xi}}, \tag{13}$$

which we call the brute-force term as in [2]. Let

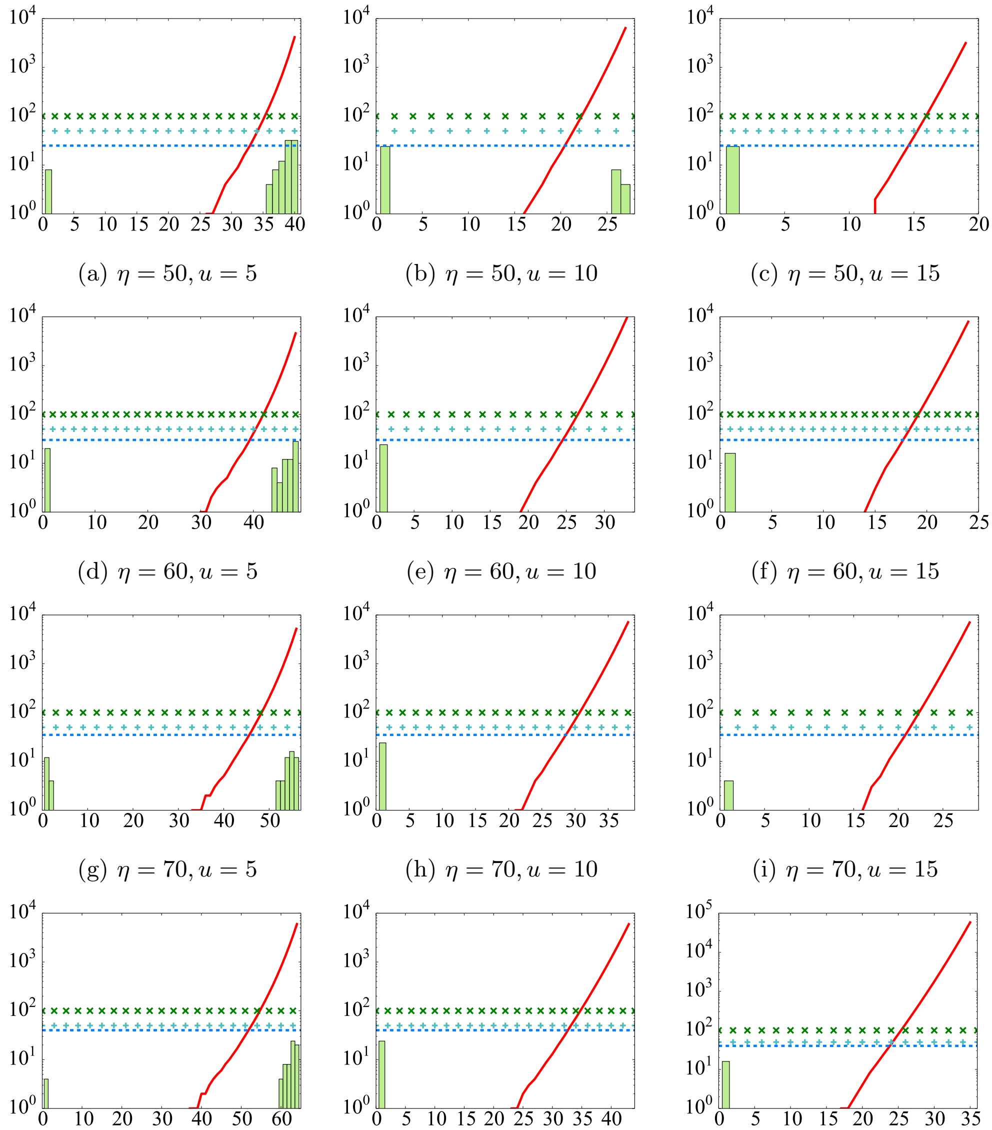
$$\alpha_i \overset{\text{def}}{=} \sum_{j=mp_\xi}^{m} \binom{m}{j} p_i^j (1-p_i)^{m-j}, \tag{14}$$

and define the right hand term

$$\text{WF}_2 = \frac{\tau\eta\xi}{\binom{\eta}{\xi}} \sum_{i=0}^{\eta} \binom{\eta}{i} \alpha_i. \tag{15}$$

Given a value of $\hat{m}$ for a fixed value of ξ calculated through Eq. 11, one can directly measure WF through Eq. 12. However, there are two problems with this approach. First as $\hat{m}$ grows larger, the α_i's are computationally expensive to compute. Indeed, Coskun and Herley only showed an estimate of the work factor for $u = 5$ [2, p. 437]. Secondly, not much insight is possible from this rather crude expression of WF. We therefore explore this expression further.

[6] To be precise, the work factor should be multiplied by m (the number of observed challenge-response pairs). However, as in [2] we ignore this and assume that our fundamental unit of complexity is the evaluation of the response stream Γ of a given candidate s for the secret, where the size of Γ, i.e., m, could be a variable.



(a) $\eta = 50, u = 5$ (b) $\eta = 50, u = 10$ (c) $\eta = 50, u = 15$

(d) $\eta = 60, u = 5$ (e) $\eta = 60, u = 10$ (f) $\eta = 60, u = 15$

(g) $\eta = 70, u = 5$ (h) $\eta = 70, u = 10$ (i) $\eta = 70, u = 15$

(j) $\eta = 80, u = 5$ (k) $\eta = 80, u = 10$ (l) $\eta = 80, u = 15$

Figure 2: The values of $\hat{m}$ according to Eq. 11 and the success percentage of the CH attack against increasing values of ξ (x-axis) and . $|R|$ is fixed at 4. The y-axis is log-scaled. Legend: — $\hat{m}$; - - $m_{\mathtt{it}}$; ▭ success percentage; $\times\times$ 100% success boundary; $++$ 50% success boundary.

First, we find an estimate for the α_i's. Note that in essence, α_i is the proportion of the binomial $\binom{\eta}{i}$ retained. Define Bernoulli random variables

$$X_i = \begin{cases} 1 & \text{with probability } p_i \\ 0 & \text{otherwise} \end{cases} . \tag{16}$$

Let $X_{i,1}, X_{i,2}, \ldots, X_{i,m}$ denote i.i.d Bernoulli random variables where each $X_{i,j}$ is of type X_i above. First *fix* an $i > \xi$, where $\xi > 0$. Then, from Theorem 4 and Corollary 1

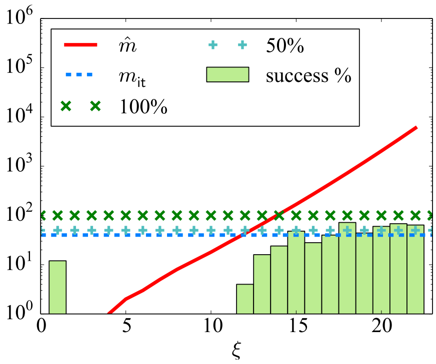


Figure 3: The values of $\hat{m}$ according to Eq. 11 and the success percentage of the CH attack against increasing values of ξ when $\eta = 80$, $u = 15$ and $\delta = 0.400$ in Eq. 11. Note high success rate of the CH attack.

we have

$$\mathbb{P}\left[\sum_{j=1}^{m} X_{i,j} \geq mp_\xi\right] \to 2^{-mD(P_\xi||P_i)} \text{ as } m \to \infty$$

From this we may estimate the $\alpha_i \approx 2^{-mD(P_\xi||P_i)}$ when $i \geq \xi$ and $\alpha_i \approx 1 - 2^{-mD(P_\xi||P_i)}$ when $i < \xi$ (see for instance Figure 4). However, as indicated by Figure 1, for larger

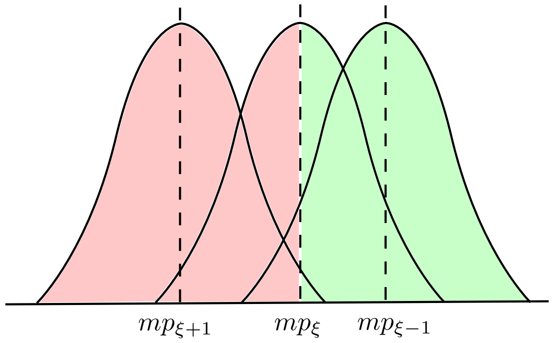
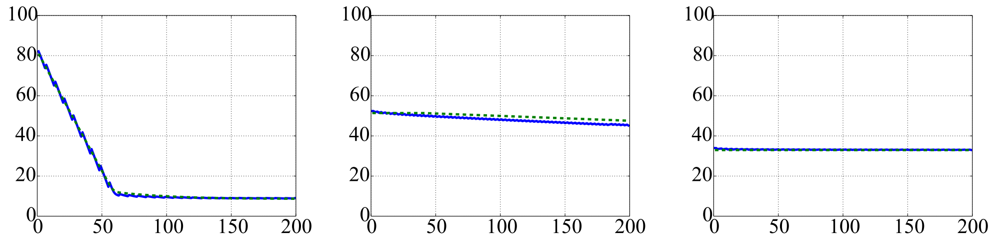


Figure 4: The portions of the binomials retained in the CH attack where: portion retained; portion discarded. Note that $\alpha_\xi = 0.5$, $\alpha_i > 0.5$ for $i < \xi$ and $\alpha_i < 0.5$ for $i > \xi$.

values of ξ, say $\xi > \frac{\eta}{2}$, the difference in probabilities is very small for the neighbours of s_ξ, and therefore our bound of α_i, which is based on large deviations assumption, will err. One way to compensate for this is to instead use the normal approximation of the α_i's as $\Phi(-\frac{\sqrt{m}\mu_{X_i}}{\sigma_{X_i}})$ for the nearby neighbours of s_ξ when $\xi > \frac{\eta}{2}$ in a manner similar to Section 3.2, where $\mu_{X_i} = p_i$ and $\sigma^2_{X_i} = p_i(1-p_i)$. A less messier way is to upper bound the α_i's by 0.5 for $i > \xi$ and lower bound by 0.5 for $i < \xi$, indicating that these sums are not expected to exceed these limits (see Figure 4). This gives us the following estimate

$$\alpha_i \approx \begin{cases} \max\{1 - 2^{-mD(P_\xi||P_i)}, 0.5\} & \text{if } i < \xi \\ 0.5 & \text{if } i = \xi \\ \min\{0.5, 2^{-mD(P_\xi||P_i)}\} & \text{if } i > \xi \end{cases}. \tag{17}$$

With this estimate of the α_i's we denote the corresponding estimate of WF_2 by $\hat{\mathrm{WF}}_2$. Figure 5 shows the actual work factor WF_2 against our estimate $\hat{\mathrm{WF}}_2$. Note that $\hat{\mathrm{WF}}_2$ is expected to be a better estimate than simply approximating all the α_i's by the standard normal estimate (not just the neighbours of α_ξ), since the standard normal estimate is poor for larger deviations. To show this, we illustrate WF_2 against $\hat{\mathrm{WF}}_2$ and the standard normal estimate, denoted $\tilde{\mathrm{WF}}_2$, in Figure 6 for $\eta = 800$. Notice how $\tilde{\mathrm{WF}}_2$ is many orders of magnitude off. To get high precision values, we implemented $\hat{\mathrm{WF}}_2$ using the `mpmath` Python library [7] with a precision of 200 decimal places.



(a) $\eta = 80, u = 15, \xi = 1$ (b) $\eta = 80, u = 15, \xi = 10$ (c) $\eta = 80, u = 15, \xi = 60$

Figure 5: Actual work factor WF_2 versus the estimated $\hat{\mathrm{WF}}_2$ when $\eta = 80$, $u = 15$ and $\xi \in \{1, 10, 60\}$ against m in the range $[1, 200]$. Legend: — $\log_2 \mathrm{WF}_2$; -- $\log_2 \hat{\mathrm{WF}}_2$.

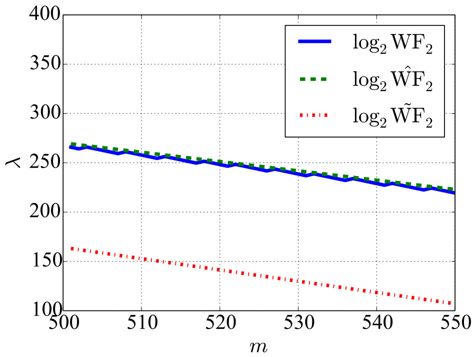


Figure 6: Work factor WF_2 and its estimates for $\eta = 800$, $u = 25$ and $\xi = 10$. The λ in the y-axis indicates power of 2. Notice how $\tilde{\mathrm{WF}}_2$ is off the mark.

We can now see the evolution of the estimate of the work factor $\hat{\mathrm{WF}}$ against $\hat{m}$ (using $\delta = 0.495$ in Eq. 11) as shown in Figure 7, which is obtained by replacing WF_2 by $\hat{\mathrm{WF}}_2$ in Eq. 12. The figure indicates that the computational complexity of the CH attack is minimized at $\xi \approx \frac{\eta}{2}$. We now show that this is true in general for suitably large values of u. In the process, we also obtain a simplified expression of $\hat{\mathrm{WF}}$. Fix a $0 < \beta < 1$ such that $u = \beta n$. We want to show that for a suitably large β (say $\frac{1}{10}$), the α_i's are at least 0.5 for all $i \in \{1, \ldots, \xi, \ldots, \eta - u\}$. From the definition of the α_i's in Eq. 17 it is obvious that for $i \leq \xi$, $\alpha_i \geq 0.5$. For $i > \xi$, showing that $\alpha_i \geq 0.5$ is the same as showing that $2^{-\hat{m}D(P_\xi || P_i)} \geq 2^{-1} \Rightarrow -\hat{m}D(P_\xi \mid\mid P_i) \geq -1 \Rightarrow \hat{m}D(P_\xi \mid\mid P_i) \leq 1$. First observe that since $p_i \geq |R|^{-1} > 0$ for all $i \in \{0, 1, \ldots, \eta\}$, and $p_\xi < 1$ for all $\xi > 1$, we have according to Theorem 5, $0 < p_i < p_\xi < 1$. Then from Theorem 3 and by setting $z = \Phi^{-1}(\delta)$ in

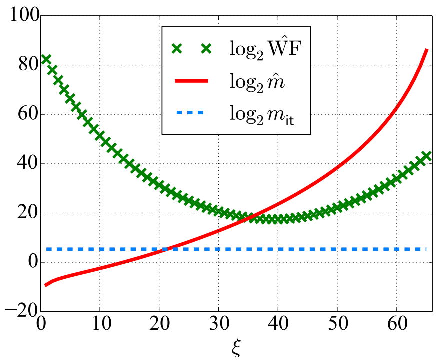


Figure 7: The total work factor $\hat{\mathrm{WF}}$ for $\eta = 80$ and $u = 15$ against $\hat{m}$ for $\delta = 0.495$.

Eq. 11, we get

$$\begin{aligned}
\hat{m}D(P_\xi \mid\mid P_i) &= \frac{\sigma^2 z^2}{\epsilon^2} D(P_\xi \mid\mid P_i) \\
&\le \frac{\sigma^2 z^2}{(p_{\xi-1} - p_{\xi+1})^2} \frac{1}{\theta} (p_\xi - p_i)^2 \\
&= \frac{\sigma^2 z^2}{\theta} \frac{(p_\xi - p_i)^2}{(p_{\xi-1} - p_{\xi+1})^2} \\
&\le \frac{z^2}{2\theta} \frac{(p_\xi - p_i)^2}{(p_{\xi-1} - p_\xi)^2},
\end{aligned}$$

where we have used the fact that $\sigma^2 \le \frac{1}{2}$ (by applying Proposition 1 on Eq. 6). Now for $i > \epsilon$

$$(p_\xi - p_i)^2 \le \left(p_\xi - \frac{1}{|R|} \right)^2 = \left(\frac{\binom{\eta-\xi}{u}}{\binom{\eta}{u}} \left(1 - \frac{1}{|R|} \right) \right)^2,$$

and

$$\begin{aligned}
(p_{\xi-1} - p_\xi)^2 &= \left(\frac{\binom{\eta-\xi+1}{u} - \binom{\eta-\xi}{u}}{\binom{\eta}{u}} \left(1 - \frac{1}{|R|} \right) \right)^2 \\
&= \left(\left(\frac{\eta - \xi + 1}{\eta - \xi + 1 - u} - 1 \right) \frac{\binom{\eta-\xi}{u}}{\binom{\eta}{u}} \left(1 - \frac{1}{|R|} \right) \right)^2 \\
&= \left(\frac{u}{\eta - \xi + 1 - u} \right)^2 \left(\frac{\binom{\eta-\xi}{u}}{\binom{\eta}{u}} \left(1 - \frac{1}{|R|} \right) \right)^2.
\end{aligned}$$

Substituting these values and simplifying we obtain

$$\hat{m}D(P_\xi \mid\mid P_i) \le \frac{z^2}{2\theta} \left(\frac{\eta - \xi + 1 - u}{u} \right)^2.$$

Now, substituting $u = \beta n$, we get

$$\begin{aligned}
\left(\frac{\eta - \xi + 1 - u}{u} \right)^2 &= \frac{1}{\beta^2} \left(1 - \frac{\xi}{\eta} + \frac{1}{\eta} - \beta \right)^2 \\
&\le \frac{1}{\beta^2} (1 - \beta)^2 \quad (\text{since } \xi \ge 1).
\end{aligned}$$

Also, for all $i \in \{1, \dots, \eta - u\}$, the p_i that minimizes $p_i(1-p_i)$ corresponds to $i = 1$. Now,

$$\begin{aligned}
p_1 &= \frac{\binom{\eta-1}{u}}{\binom{\eta}{u}} \left(1 - \frac{1}{|R|}\right) + \frac{1}{|R|} \\
&= \frac{\eta - u}{\eta} \left(1 - \frac{1}{|R|}\right) + \frac{1}{|R|} \\
&= (1-\beta) \left(1 - \frac{1}{|R|}\right) + \frac{1}{|R|} \quad \text{(substituting } u = \beta\eta) \\
&= 1 - \beta \left(1 - \frac{1}{|R|}\right).
\end{aligned}$$

Let θ_i correspond to the θ in Theorem 3 with the associated interval $[p_\xi, p_i]$. If we let $\theta = \min_i\{\theta_i\}$, then

$$\begin{aligned}
\theta &= p_1(1-p_1) \\
&= \left(1 - \beta\left(1 - \frac{1}{|R|}\right)\right) \beta \left(1 - \frac{1}{|R|}\right) \\
&= \beta^2 \left(\frac{1}{\beta} - \left(1 - \frac{1}{|R|}\right)\right) \left(1 - \frac{1}{|R|}\right) \\
&> \beta^2 \left(\frac{1}{\beta} - 1\right) \frac{|R|-1}{|R|} \quad \text{(since } |R|^{-1} > 0) \\
\Rightarrow \frac{1}{\theta} &< \frac{1}{\beta^2} \frac{\beta}{1-\beta} \frac{|R|}{|R|-1} \\
&\leq \frac{2}{\beta^2} \frac{\beta}{1-\beta} \quad \text{(since } |R| \geq 2).
\end{aligned}$$

Substituting these results into the expression for $\hat{m}D(P_\xi \mid\mid P_i)$ we finally obtain

$$\begin{aligned}
\hat{m}D(P_\xi \mid\mid P_i) &\leq \frac{z^2}{2} \frac{2}{\beta^2} \frac{\beta}{1-\beta} \frac{1}{\beta^2} (1-\beta)^2 \\
&= \frac{z^2}{\beta^4} \beta(1-\beta) \\
&\leq \frac{1}{4} \frac{z^2}{\beta^4},
\end{aligned}$$

where the last inequality follows from Proposition 1. The above is less than or equal to 1 if $\beta \geq \sqrt{\frac{|z|}{2}}$. Since according to Eq. 11, we chose $|z| \approx 0.0125$, it follows that $\beta \geq 0.08$ suffices as a choice.[7] From this it implies that $2^{-\hat{m}D(P_\xi || P_i)} \geq 0.5$. Therefore $\hat{\mathrm{WF}}_2$ for $\hat{m}$ is at least

$$\hat{\mathrm{WF}}_2 = \frac{\tau\eta\xi}{\binom{\eta}{\xi}} \sum_{i=0}^{\eta} \binom{\eta}{i} \alpha_i \geq \frac{\tau\eta\xi}{\binom{\eta}{\xi}} \frac{1}{2} 2^\eta = \frac{\tau\eta\xi}{\binom{\eta}{\xi}} 2^{\eta-1}. \tag{18}$$

Note that as the binomial sum has a maximum value of 2^η, the above work factor is close to the maximum, since for $\hat{m}$

$$\hat{\mathrm{WF}}_2 \leq \frac{\tau\eta\xi}{\binom{\eta}{\xi}} 2^\eta.$$

[7] For $\delta = 0.490$, $\beta = 0.12$ does the job.

The above expression has a minimum at around $\xi = \frac{\eta}{2}$.[8] This is true since $\binom{\eta}{\xi} = \mathcal{O}(\eta^{\xi})$ for $\xi \leq \frac{\eta}{2}$ and $\binom{\eta}{\xi} = \mathcal{O}(\eta^{\eta-\xi})$ for $\xi > \frac{\eta}{2}$, and the terms $\frac{\xi}{\eta^{\xi}}$ and $\frac{\xi}{\eta^{\eta-\xi}}$ have a minimum when $\xi = \frac{\eta}{2}$. Substituting WF_2 with the value of $\hat{\mathrm{WF}}_2$ obtained through Eq. 18 in Eq. 12, we see that the total work factor WF is dominated by WF_2 when $m = \hat{m}$ (and $u \geq \beta\eta$), since

$$\mathrm{WF} \approx \frac{(2 + \tau\eta\xi)}{\binom{\eta}{\xi}} 2^{\eta-1} \approx \frac{\tau\eta\xi}{\binom{\eta}{\xi}} 2^{\eta-1}.$$

We summarise our findings in the following heuristic theorem.

Theorem 6. *Let p_i be as defined by Eq. 1 for $i \in \{0, 1, \ldots, \eta\}$. Further, let $\hat{m} = \left(\frac{\sigma z}{\epsilon}\right)^2$ where*

$$\epsilon^2 = (p_{\xi-1} - p_{\xi+1})^2 \text{ and } \sigma^2 = p_{\xi-1}(1 - p_{\xi-1}) + p_{\xi+1}(1 - p_{\xi+1}),$$

and $z = \Phi^{-1}(\delta)$ for some $\delta \in (0, \frac{1}{2})$. Then if $u \geq \beta\eta$, where $\beta = \sqrt{\frac{|z|}{2}}$, the work factor of the Coskun and Herley attack is

$$WF \approx \frac{\tau\eta\xi}{\binom{\eta}{\xi}} 2^{\eta-1},$$

which is minimum when $\xi = \frac{\eta}{2}$, for $1 \geq \xi \geq \eta - u$ and $u \leq \frac{\eta}{2}$. If $u > \frac{\eta}{2}$ the minimum is achieved at $\xi = \eta - u$.

We reiterate that we are interested in finding λ such that $2^{\lambda} \approx \mathrm{WF}$, and hence an estimate that is far from the true value by a few of powers of 2 is sufficient.

4 Case Study: Setting Parameter Sizes for Identification Protocols

Suppose we a prover $\mathcal{P}$ and a verifier $\mathcal{V}$ share a secret s which is a set of k indexes out of n. All indexes are from the set $\{1, \ldots, n\}$, n being a positive integer. We denote this set of indexes by $[n]$. Consider the following identification protocol between $\mathcal{P}$ and $\mathcal{V}$:

Protocol: `An Identification Protocol`

$\mathcal{V}$ sends a challenge c to $\mathcal{P}$ which is a random subset of indexes from $[n]$ of cardinality l, $l < n$, such that each element is associated with a random *weight* from $\mathbb{Z}_4$.
$\mathcal{P}$ initializes $r_1 \leftarrow 0$.
for *each index i in c* **do**
    if $i \in s$ **then**
        $\mathcal{P}$ updates $r_1 \leftarrow r_1 + w$, where w is the weight associated with i.
$\mathcal{P}$ computes $r_2 \leftarrow r_1 \bmod 4$.
if $r_2 \in \{0, 1\}$ **then**
    $\mathcal{P}$ returns $r \leftarrow 0$ as its response.
else if $r_2 \in \{2, 3\}$ **then**
    $\mathcal{P}$ returns $r \leftarrow 1$ as its response.
$\mathcal{V}$ accepts $\mathcal{P}$ if the response r is correct.

The protocol above is known as the Foxtail protocol with window [8]. In an actual protocol, the above process is repeated a number of times per authentication session such that the probability of randomly guessing the response (without knowing anything

[8] We ignore the detail of η being odd or even.

about the secret!) is low. But for our purposes we ignore this detail, and assume that there is one round per authentication session. The term *window* alludes to the l-element challenge presented to the prover. Since a fraction of the secret s is expected to be present in each challenge, CH attack can be applied to the protocol to find the secret. The question arises, with a given set of values of protocol parameters (l, k, n), how many rounds can the above protocol be used for, i.e., the value of m, so that the CH attack has complexity of about 2^λ for a fixed λ? We first estimate u as follows

$$u = \frac{\mathbb{E}[|s \cap c|]}{k}\eta = \frac{\mathbb{E}[|s \cap c|]}{k}\log_2\binom{n}{k} = \frac{lk}{n}\frac{1}{k}\log_2\binom{n}{k} = \frac{l}{n}\log_2\binom{n}{k}, \tag{19}$$

where $\frac{lk}{n}$ is the expected value of the hypergeometric distribution. If we choose k and n to be such that $\eta = \log_2\binom{n}{k} \approx 80$ (e.g., $n = 180$ and $k = 18$), and choose $l = 40$, then we get $u \approx 0.22 \times 80 \approx 18$ bits. We can now choose a δ to obtain a value of $\hat{m}$ using Theorem 6 which in turn gives us a value for λ, i.e., $\log_2 \mathrm{WF}$. If a work factor of 2^{25} is considered infeasible, i.e., $\lambda = 25$, then with $\delta = 0.495$, we can use the protocol for $m \leq 993 \approx 1{,}000$ (which corresponds to $\xi = 25$). Choosing $\delta = 0.490$ allows us to use the protocol for $m \leq 3{,}972 \approx 4{,}000$ rounds (again corresponding to $\xi = 25$).

5 The CH Attack is not always Optimal

Consider the variant of the above protocol in which the weights are from $\mathbb{Z}_2$ and the response is simply $r \leftarrow r_1 \bmod 2$. Then we can use Gaussian elimination to find the secret s after n observations, by constructing the $n \times n$ binary matrix H whose rows are constructed from $m = n$ observed challenges and using the n-element response vector $\mathbf{r}$ obtained from the responses. On the other hand, the CH attack requires a much larger number of observations m to be feasible, provided l and n are not too small. For instance, if $(l, k, n) = (40, 18, 180)$, as above, with $m \approx 262 > n$, the work factor of the CH attack is $\approx 2^{58}$. Gaussian elimination on the other hand is a polynomial time algorithm that would yield the secret after $m \approx n = 180$ observations.

Of course, for Gaussian elimination to yield a unique solution we require the n rows (or equivalently, columns) of H to be linearly independent. We show here that this is likely to happen with high probability. Note that the total number of possible challenges in this protocol are

$$|C| = \sum_{i=0}^{l}\binom{n}{i},$$

which can be arrived at by observing that there are $\binom{n}{i}$ possible binary vectors with Hamming weight i. From this we could use a counting argument to get the number of possible combinations of n linearly independent vectors, by iteratively discarding any linearly dependent choices for the vector i given $i-1$ linearly independent vectors. However, this is not straightforward, as the linear combination of any two vectors in C might not be a vector in C (e.g., two vectors with Hamming weight l which differ in at least one position). Asymptotic results, however, suggest that if l is large enough we are likely to construct a full rank $n \times n$ matrix H after observing $m > n$ challenges, where the difference $m - n$ is bounded from below by a constant. To be precise, we reword the corollary from [9] as a theorem using our notation.

Theorem 7. *Let $m > n$. If $l > \ln n + \omega(1)$ and $m - n \geq \omega(1)$, then almost every set of m uniformly random vectors from C have n linearly independent vectors.*[9]

[9] Recall that $f = \omega(g)$ means that $\frac{f(n)}{g(n)}$ tends to infinity as n grows to infinity.

As an example, for the case under consideration, i.e., $n = 180$ and $l = 40$, a simple Python script using the `Sage` library[10] returned a full rank $n \times n$ matrix H, each row of which was randomly sampled from C, at a success rate of 0.2969 (over 10,000 repetitions). In contrast, $l = 2$ returned 0 such incidences. The fact that the success rate is high for reasonably large values of l is not surprising. For instance, a necessary condition for the matrix H to be linearly independent is to have no zero columns. Observe that the probability of an element from a vector $c \in C$ being zero is given by

$$\frac{1}{2} \cdot \frac{l}{n} + 1 \cdot \left(1 - \frac{l}{n}\right) = 1 - \frac{l}{2n},$$

which follows since either the element could be absent from the l chosen elements, or it could be present but with weight 0. Let B_i be the probability that the ith column is *not* a zero column vector, then

$$\mathbb{P}(B_i) = 1 - \left(1 - \frac{l}{2n}\right)^n.$$

Let A be the event that the matrix H has no zero column vectors, then

$$\begin{aligned}
\mathbb{P}(A) &= \bigcap_{i=1}^{n} \mathbb{P}(B_i) \\
&\geq \left(\sum_{i=1}^{n} \mathbb{P}(B_i)\right) - n + 1 \\
&= 1 - n\left(1 - \frac{l}{2n}\right)^n
\end{aligned}$$

which is close to 1 for sufficiently large l and n (the inequality above is the application of Benferroni's inequality [3, §7, p. 426]). For instance, when $n = 180$, $l = 15$ yields $\mathbb{P}(A) > 0.91$.

6 Related Work

Baignères, Junod and Vaudenay [10] show a more general result which estimates the number of samples required by an optimal distinguish between two probability distributions (not necessarily Bernoulli) that are close to each other. Barring constant factors, the estimate from them and the two estimates given in Eq. 9 and Eq. 10 all yield $m \sim \frac{1}{\epsilon^2}$, which can be used as a *rough* guide for the number of samples required in our case.

We note that somewhat similar to our estimation of m, the work in [8, §5] attempts to bound the safe number of rounds against counting based statistical attacks on human identification protocols introduced by Yan et al. [11]. However, the resultant figures for the number of safe rounds against these counting based attacks are erroneous, since they do not treat the associated probability as an error probability, and wrongly calculate the required samples by fixing the one-sided cumulative distribution function of the standard normal distribution at 0.6.

Research on human identification protocols dates back to the work of Matsumoto and Imai in [12]. Juels and Weis [13] noted parallels between humans and resource-constrained devices, that both suffer from low computational and memory capabilities, and proposed a variant of the Hopper and Blum (HB) human identification protocol [14] to be used for identification of such devices. To the best of our knowledge, to date, this

[10] http://www.sagemath.org/

is the only human identification protocol whose application as an identification protocol for resource-constrained devices has been extensively studied. It is an interesting area of research to analyze other human identification protocols, such as the sum of k mins protocol [14], for their suitability as identification protocols for resource constrained devices. While some of the human identification protocols in literature are based on ad hoc design [12, 15, 16], there are others whose security is based on the hardness of interesting mathematical problems [17, 18, 19, 20, 21]. There are also some theoretical advances in generic attacks on human identification protocols [2, 11, 8, 22]. These results may result in other human identification protocols being proposed for identification of resource-constrained devices.

7 Conclusion

We have shown estimates for the number of sessions that should be allowed before secret renewal in challenge-response type identification protocols, which use a fraction of the secret to respond to a challenge, to be safe against the Coskun and Herley attack. We have also shown how we can estimate the work factor of this attack against the number of allowable sessions in a computationally efficient and protocol independent manner. These estimates are empirically straightforward to obtain without the need to implement the Coskun and Herley attack on a given protocol to check its complexity against different sets of values of protocol parameters. This work can benefit protocol designers to set parameter sizes both in the field of human identification protocols or identification protocol for resource constrained devices in which the use of "sparse" challenges is desirable.

References

[1] Berry Schoenmakers. Lecture Notes Cryptographic Protocols. Version 1.1, http://www.win.tue.nl/~berry/2WC13/LectureNotes.pdf, 2015.

[2] Baris Coskun and Cormac Herley. Can "Something You Know" Be Saved?. In Tzong-Chen Wu, Chin-Laung Lei, Vincent Rijmen, and Der-Tsai Lee, editors, *ISC '08*, pages 421–440. Springer, 2008.

[3] Sheldon M. Ross. *A First Course in Probability*. Prentice Hall, 4 edition, 2002.

[4] Stéphane Boucheron, Gábor Lugosi, and Olivier Bousquet. Concentration Inequalities. In *Advanced Lectures in Machine Learning*, pages 208–240. Springer, 2004.

[5] Robert L. Wolpert. Markov, Chebychev and Hoeffding Inequalities. Lecture notes, https://stat.duke.edu/courses/Spring09/sta205/lec/hoef.pdf, 2009.

[6] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. John WIley and Sons, Hoboken, New Jersey, USA, 2nd edition, 2006.

[7] Fredrik Johansson et al. *mpmath: a Python library for arbitrary-precision floating-point arithmetic (version 0.18)*, December 2013. http://mpmath.org/.

[8] Hassan Jameel Asghar, Shujun Li, Ron Steinfeld, and Josef Pieprzyk. Does Counting Still Count? Revisiting the Security of Counting based User Authentication Protocols against Statistical Attacks. In *20th Annual Network and Distributed System Security Symposium*, NDSS '13. The Internet Society, 2013.

[9] Nathan Linial and Dror Weitz. Random vectors of bounded weight and their linear dependencies. http://dimacs.rutgers.edu/~dror/pubs/rand_mat.pdf, 2000.

[10] T. Baignères, P. Junod, and S. Vaudenay. How far can we go beyond linear cryptanalysis? In P.J. Lee, editor, *Advances in Cryptology* - ASIACRYPT *'04*, volume 3329 of *Lecture Notes in Computer Science*, pages 432–450. Springer-Verlag, 2004.

[11] Qiang Yan, Jin Han, Yingjiu Li, and Robert H. Deng. On Limitations of Designing Leakage-Resilient Password Systems: Attacks, Principals and Usability. In *19th Annual Network and Distributed System Security Symposium*, NDSS '12. The Internet Society, 2012.

[12] T. Matsumoto and H. Imai. Human identification through insecure channel. In *EUROCRYPT '91*, pages 409–421. Springer-Verlag, 1991.

[13] Ari Juels and Stephen A. Weis. Authenticating Pervasive Devices with Human Protocols. In *Proceedings of the 25th Annual International Conference on Advances in Cryptology*, CRYPTO '05, pages 293–308, Berlin, Heidelberg, 2005. Springer-Verlag.

[14] N. J. Hopper and M. Blum. Secure Human Identification Protocols. In *ASIACRYPT '01*, pages 52–66. Springer-Verlag, 2001.

[15] Daphna Weinshall. Cognitive Authentication Schemes Safe Against Spyware (Short Paper). In *IEEE Symposium on Security and Privacy*, SP '06, pages 295–300. IEEE Computer Society, 2006.

[16] M. Lei, Y. Xiao, S. V. Vrbsky, and C.-C. Li. Virtual Password using Random Linear Functions for On-line Services, ATM Machines, and Pervasive Computing. *Computer Communications*, 31(18):4367–4375, 2008.

[17] Hassan Jameel Asghar, Josef Pieprzyk, and Huaxiong Wang. A New Human Identification Protocol and Coppersmith's Baby-step Giant-step Algorithm. In *Proceedings of the 8th International Conference on Applied Cryptography and Network Security*, ACNS '10, pages 349–366, Berlin, Heidelberg, 2010. Springer-Verlag.

[18] L. Sobrado and J.-C. Birget. Graphical Passwords. *The Rutgers Scholar*, 4, 2002.

[19] S. Wiedenbeck, J. Waters, L. Sobrado, and J.-C. Birget. Design and Evaluation of a Shoulder-Surfing Resistant Graphical Password Scheme. In *AVI '06*, pages 177–184. ACM, 2006.

[20] S. Li and H.-Y. Shum. Secure Human-Computer Identification (Interface) Systems against Peeping Attacks: SecHCI. IACR's Cryptology ePrint Archive: Report 2005/268, http://eprint.iacr.org/2005/268, 2005.

[21] Jeremiah Blocki, Manuel Blum, Anupam Datta, and Santosh Vempala. Human Computable Passwords. arXiv preprint arXiv:1404.0024 http://arxiv.org/pdf/1404.0024.pdf, 2014.

[22] Hassan Jameel Asghar, Ron Steinfeld, Shujun Li, Mohamed Ali Kâafar, and Josef Pieprzyk. On the linearization of human identification protocols: Attacks based on linear algebra, coding theory, and lattices. *IEEE Transactions on Information Forensics and Security*, 10(8):1643–1655, 2015.