Optimality Results on the Security of Lookup-Based Protocols

Sjouke Mauw[1,2], Jorge Toro-Pozo[1], and Rolando Trujillo-Rasua[2]

[1] CSC, University of Luxembourg
[2] SnT, University of Luxembourg
6, rue Richard Coudenhove-Kalergi, L-1359 Luxembourg
{sjouke.mauw, jorge.toro, rolando.trujillo}@uni.lu

Abstract. Distance-bounding protocols use the round-trip time of a challenge-response cycle to provide an upper-bound on the distance between prover and verifier. In order to obtain an accurate upper-bound, the computation time at the prover's side should be as short as possible, which can be achieved by precomputing the responses and storing them in a lookup table. However, such lookup-based distance bounding protocols suffer from a trade-off between the achieved security level and the size of the lookup table. In this paper, we study this security-memory trade-off problem for a large class of lookup-based distance bounding protocols; called *layered protocols*. Relying on an automata-based security model, we provide mathematical definitions for different design decisions used in previous lookup-based protocols, and perform general security analyses for each of them. We also formalize an interpretation of *optimal trade-off* and find a non-trivial protocol transformation approach towards optimality. That is to say, our transformation applied to any layered protocol results in either an improved or an equal protocol with respect to the optimality criterion. This transformation allows us to provide a subclass of lookup-based protocol that cannot be improved further, which means that it contains an optimal layered protocol.

Keywords: Distance bounding · RFID · Security · Mafia-fraud · Relay attack

1 Introduction

Secure physical proximity checking in wireless technologies is a well-established field within computer security. As the speed of light represents a hard limit on the speed of radio waves, it has been used to accurately compute an upper bound on the distance between a transmitter (e.g., a satellite or an RFID reader) and a receiver (e.g., a GPS receiver or an RFID tag). The equation is simple: the distance to the receiver is half the round-trip time (RTT) multiplied by the speed of light. Nonetheless, this computation ought to be embedded into a cryptographic protocol in order to ensure the authenticity of the receiver and the integrity of the distance calculation. Such a cryptographic protocol is called a *distance bounding protocol* [4], that is, an authentication protocol that also determines an upper bound on the distance between the protocol's participants. In this setting, the transmitter is called *verifier* and the receiver is called *prover*.

Like most cryptographic protocols, a distance bounding protocol consists of a series of challenge-response rounds, with the peculiarity that some of these rounds are used to compute round-trip times. Replying to such a round-trip challenge should be a computationally inexpensive task, because otherwise the round-trip time will become tainted by the prover's computation time as illustrated in Figure 1.

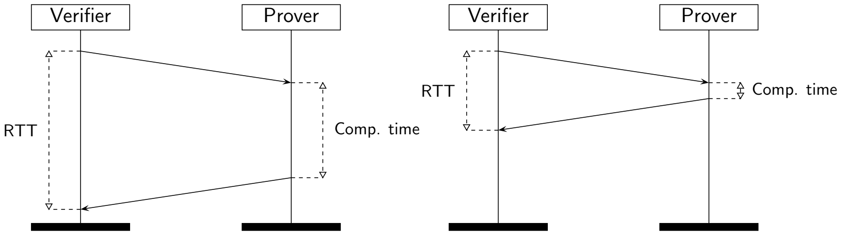


Fig. 1: The impact of the prover's computation time on the round-trip-time (RTT) measurement.

A fundamental approach to computational efficiency is precomputation. In distance bounding, this technique was first proposed by Hancke and Kuhn [8] in 2005, and later improved in [2, 7, 9, 10, 12–14]. These protocols, named lookup-based protocols in [12], contain an initial precomputation phase in which the verifier and the prover secretly agree on a lookup table. This phase is followed by n round-trip-time measurements realized by an n-fold challenge-response process. For each challenge, the prover determines his response by looking it up in the table. The protocol finishes successfully if all responses are correct with respect to the precomputed lookup table and all round-trip-times are below a given time threshold. Simplicity and efficiency make this type of protocols appealing for battery-less or constrained devices, such as RFID tags. Moreover, their security can be arbitrarily improved by increasing n, with little or no computational overhead [8].

Despite the simplicity of lookup-based protocols, none of them have been proven optimal in terms of security, given a practical threshold on the size of the lookup table. This is a security-memory trade-off problem which resembles the classical time-memory trade-off problem occurring in other application domains, such as cryptanalysis, rainbow tables, and dynamic programming. A recent work [12] shows that if an exponential number of values (in terms of n) can be precomputed, then the tree-based protocol proposed in [3] is optimal. However, an exponentially large lookup table is unattainable in practice.

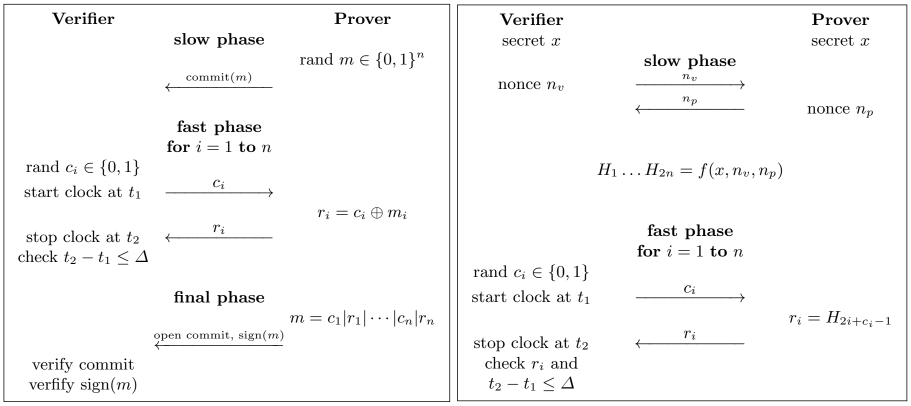
Contributions. In this article, we study the security-memory trade-off problem for a large class of lookup-based protocols; called *layered protocols*. The class of layered protocols is not trivial, as it contains all lookup-based distance bounding protocols proposed to date, except for the Poulidor protocol [14]. Our findings can be listed as follows.

- First, we rely on the automata-based model proposed in [12] (Section 3) to provide mathematical definitions of common design decisions used in lookup-based protocols (Section 4). For example, many lookup-based protocols use a fixed and known data access structure for the lookup table. Others, such as the Poulidor protocol [14], use a randomized indexing technique. We formalize both techniques by exploiting equivalence relations between automata, and perform general security analyses for those layered protocols using either of the two techniques (Section 4).
- Second, we formalize an interpretation of the security-memory trade-off problem for layered protocols, and provide a non-trivial protocol transformation approach towards optimality (Section 5). In more detail, for every protocol P with size s (a measure of memory), we show how to obtain another protocol P' of the same size and with equal or higher resistance to pre-ask attacks. We also prove that the proposed transformation has the full class of layered protocols as domain. We can thus conclude that an optimal layered protocol is within the image set of our transformation.

2 Related Work

The earliest distance bounding protocol based on RTT measurements was proposed by Brands and Chaum [4] in 1993 (see Figure 2a). Their protocol (BC) assumes a verifier and a prover, each armed with a public/private key pair. By considering n to be the number of RTT measurements, the prover commits to a random sequence of bits $m_1 \cdots m_n$, after which the fast phase or RTT measurement phase starts. The fast phase consists of n rounds of a single bit-exchange. At the ith round, the verifier sends a random bit-challenge c_i and the prover instantly replies with $c_i \oplus m_i$. Proximity checking fails if the computed RTT at any round is above a given threshold. Authentication, on the other hand, is verified during a final phase where the prover opens the commit and signs all exchanges during the fast phase.

In 2005, Hancke and Kuhn (HK) proposed a different design for distance bounding protocols [8], where proximity checking and authentication are performed together during the fast phase. Because each round during the fast phase is quick and inexpensive, this makes it possible to improve the security with very-low computational overhead by simply increasing n. Hancke and Kuhn's protocol, depicted in Figure 2b, works as follows. First, the parties exchange one nonce each (n_v and n_p). The two nonces and a shared secret key x are used by both parties as input to a pseudo-random function f, whose output is a sequence of $2n$ bits $H_1 \cdots H_{2n}$. As usual, the fast phase consists of n rounds. At the i-th round, the verifier generates a random bit-challenge c_i. If the bit-challenge is 0, the prover replies H_{2i-1}, otherwise H_{2i}. The protocol succeeds if all responses are correct and all round-trip times are below a predefined threshold.



(a) Brands and Chaum's protocol (b) Hancke and Kuhn's protocol

Fig. 2: Two distance bounding protocols.

The difference between HK and BC approaches becomes apparent if we consider the role of the sequences $m = m_1 \ldots m_n$ and $H = H_1 \ldots H_{2n}$ in their corresponding protocols. Revealing m allows an adversary to successfully pass the proximity checking phase in Brands and Chaum's protocol, yet the adversary cannot impersonate the prover as the prover's signature is required after the fast phase.

Hancke and Kuhn's protocol, instead, fails to provide proximity checking and authentication if H is revealed, as the two properties rely on the secrecy of H. Both approaches have their merits and shortcomings, resulting in the publication of two different types of distance bounding protocols: those based on Brands and Chaum's approach (e.g., [3–5, 11]) and those based on Hancke and Kuhn's approach (e.g., [2, 7, 9, 10, 14]).

A common feature of HK-based protocols is that, during the fast phase, the prover uses a simple lookup operation to compute the correct reply to the verifier's challenge; hence their name *lookup-based protocols*. The drawback of lookup-based protocols, as shown in [1, 12], is its low resistance to *pre-ask attacks*. A pre-ask attack is a sophisticated version of the so-called *mafia-fraud attack*, introduced by Desmedt, Goutier, and Bengio in 1987 [6].

For example, a pre-ask attack against Hancke and Kuhn's proposal might be the following. The adversary queries the legitimate prover with all-zero challenges to learn half of the sequence H. If the adversary then moves sufficiently close to the verifier, he can use this knowledge to impersonate the prover. By using his knowledge of half of H, the adversary will provide the correct response for the verifier's challenges that are equal to 0. For the challenges equal to 1, the adversary simply replies with random bits. With this strategy, the adversary will pass Hancke and Kuhn's protocol with probability $(3/4)^n$. In contrast, a pre-ask attack against Brands and Chaum's protocol succeeds with probability $(1/2)^n$.

Resisting pre-ask attacks and reducing the size of the lookup table has been the aim of many lookup-based protocols proposed in the last ten years [2, 7, 9–14]. However, these two goals, namely improving security and reducing size, seem to be in conflict in lookup-based protocols. This is a security-memory trade-off problem which has not been formally addressed yet. In the remainder of this article we study this problem for a subclass of lookup-based protocols.

3 Preliminaries

In this section we present the MTT security model introduced in [12] to study lookup-based distance bounding protocols.

3.1 The MTT Model

The security model is based on state-labeled Deterministic Finite Automata (DFAs) of the form $(\Sigma, \Gamma, Q, q_0, \delta, \ell)$, where Σ is a set of input symbols, Γ is a set of output symbols, Q is a set of states, $q_0 \in Q$ is the initial state, $\delta \colon Q \times \Sigma \to Q$ is a transition function, and $\ell \colon Q \to \Gamma$ is a labeling function. Given input and output symbol sets Σ and Γ, respectively, we use $\mathbf{U}_{\Sigma,\Gamma}$ to denote the universe of all DFAs over Σ and Γ.

Definition 1 (Lookup-based protocol). *A* lookup-based distance bounding protocol*, lookup-based protocol for short, with input set Σ and output set Γ is a finite non-empty subset of $\mathbf{U}_{\Sigma,\Gamma}$.*

A restriction imposed by Definition 1 on a lookup-based protocol is that it must be formed by automata with the same input and output sets. The reason is that Σ and Γ define the alphabets of the verifier's challenges and prover's responses, respectively.

Given an automaton $A = (\Sigma, \Gamma, Q, q_0, \delta, \ell)$ and a current state $q \in Q$, a lookup operation is regarded as a transition to a new state $q' = \delta(q, c)$ where $c \in \Sigma$ is a verifier's challenge. The corresponding response for such challenge is the output symbol attached to the new state q', i.e., $\ell(q')$. Successive lookup operations form a path in the automaton. We thus introduce additional notation for a sequence of input symbols $c = c_1 \cdots c_i \in \Sigma^i$:

- $\hat{\delta}(c) = q_0$ if $i = 0$, i.e., c is an empty string, otherwise $\hat{\delta}(c) = \delta(\hat{\delta}(c_1 \ldots c_{i-1}), c_i)$. In a nutshell, $\hat{\delta}(c)$ returns the state reached by the input sequence c.
- $\hat{\ell}(c) = \ell(\hat{\delta}(c))$, which is the output symbol attached to the state reached by the sequence c.
- $\Omega_A(c)$ is used to represent the sequence of labels attached to the states $\hat{\delta}(c_1), \hat{\delta}(c_1 c_2), \ldots, \hat{\delta}(c_1 c_2 \cdots c_i)$ in that order, i.e., $\Omega_A(c) = r_1 \cdots r_i \in \Gamma^i$, where $r_j = \hat{\ell}(c_1 \cdots c_j),\ \forall j \in \{1, \ldots, i\}$.

The model in [12] abstracts away from the precomputation phase in lookup-based protocols by considering the following execution model.

Definition 2 (Execution model). *A* correct execution *of a lookup-based protocol P with $n > 0$ rounds is a triple (A, C, R), where A is an automaton randomly selected from P, C is an input sequence randomly selected from $\varSigma^n$ and R is an output sequence from $\varGamma^n$ such that $R = \varOmega_A(C)$.*

The outcome of the precomputation phase is considered to be a random automaton $A \in_R P$ within the set of automata defining the lookup-based protocol P. The input sequence $C = c_1 \cdots c_n$ corresponds to the verifier's challenges, and the correct replies $R = r_1 \cdots r_n$ must satisfy that $\hat{\ell}(c_1) = r_1, \hat{\ell}(c_1c_2) = r_2, \cdots, \hat{\ell}(c_1 \cdots c_n) = r_n$.

3.2 Layered Protocols

The set-of-automata representation of many existing lookup-based protocols, e.g., [2, 8, 9, 11, 12], satisfies that, given an automaton $A = (\varSigma, \varGamma, Q, q_0, \delta, \ell)$ and two input sequences x and y, $\hat{\delta}(x) = \hat{\delta}(y)$ implies that the sizes of x and y are equal. Formulated differently, given any of the automata defining a protocol, two sequences of different size do not end in the same state. Protocols satisfying this property are called *layered protocols* [12].

Definition 3 (Layered protocol). *A lookup-based protocol P is* layered *if, for every $(\varSigma, \varGamma, Q, q_0, \delta, \ell) \in P$ and every $x, y \in \varSigma^*$, $\hat{\delta}(x) = \hat{\delta}(y) \implies |x| = |y|$.*

Layered protocols guarantee that the prover's answers at different rounds rely on different states regardless of the verifier's challenges. Given an automaton $(\varSigma, \varGamma, Q, q_0, \delta, \ell)$ in a layered protocol, we thus consider Q to be partitioned in $Q_0 \cup Q_1 \cup Q_2 \cdots$ where $Q_i = \{\hat{\delta}(x) \mid x \in \varSigma^i\}$.

For the sake of simplicity, and because most distance bounding protocols consider bit exchanges during the fast phase, we assume that the input and output symbol sets are binary, i.e., $\varSigma = \varGamma = \{0, 1\}$. Consequently, unless otherwise specified, all DFAs considered in the remainder of this article belong to the universe $\mathbf{U}_{\varSigma,\varGamma}$ with $\varSigma = \varGamma = \{0, 1\}$.

4 Security Analysis through Equivalence Relations

While modeling lookup-based protocols by a set of automata, we have found that most lookup-based protocols defined in literature share some structural properties. For instance, the automata for the HK protocol all have a similar layered structure with two nodes in each layer [12]. We will exploit such similarities when reasoning about protocols. In this section, we will introduce two equivalence relations on automata that express relevant similarities and we will define closure and consistency with respect to these equivalences. Based on these relations, we provide a general security analysis of layered protocols.

4.1 Equivalence Relations between Automata

A common feature of many lookup-based protocols is that all their automata have the same *shape* and differ only in the symbols attached to the states. This property, which we name *state-label-insensitive*, is satisfied by HK [8], KA [10], Tree-based [2], PUF [9], and Uniform [12], amongst others. The design decision represented by the state-label-insensitive property makes it easy for participants in a protocol to agree on the shape or structure of the lookup table, while only requiring randomness on the precomputed values.

Definition 4 (State-label-insensitive). *The* state-label-insensitive *relation $\sim_S$ is a symmetric binary relation on $\mathbf{U}_{\Sigma,\Gamma}$, defined by $(\Sigma, \Gamma, Q, q_0, \delta, \ell) \sim_S (\Sigma, \Gamma, Q', q_0', \delta', \ell')$ if and only if $Q = Q'$, $q_0 = q_0'$ and $\delta = \delta'$.*

A few lookup-based protocols contain automata that are not related according to $\sim_S$, e.g., the Poulidor protocol [14]. In this protocol, the authors designed a mechanism to prevent (to some extent) an adversary from knowing which state is being used by the prover at any round of the fast phase. The idea is simple, the probability of two automata sharing the same transition function must be negligible. Such a mechanism seems to improve the resistance to pre-ask attacks as the adversary cannot easily use knowledge acquired in previous rounds to succeed in the current round of the fast phase.

Even though two automata in the Poulidor protocol can have different transition functions, they still preserve a slightly weaker structural property than the above mentioned state-label-insensitive property. If we ignore the edge labels of the automata, i.e., if we only look at the structure of the underlying graph of the automata, the transition functions of the automata in the Poulidor protocol are identical. We provide a formal definition for this structural relation next.

Definition 5 (Label-insensitive). *The* label-insensitive *relation $\sim_L$ is a symmetric binary relation on $\mathbf{U}_{\Sigma,\Gamma}$, defined by $(\Sigma, \Gamma, Q, q_0, \delta, \ell) \sim_L (\Sigma, \Gamma, Q', q_0', \delta', \ell')$ if and only if $Q = Q'$, $q_0 = q_0'$ and $\{\delta(q, c) \mid c \in \Sigma\} = \{\delta'(q, c) \mid c \in \Sigma\}$ for every $q \in Q$.*

A fundamental operator in binary relations is the *closure* with respect to a given property.

Definition 6 (Closure). *Let P be a lookup-based protocol and $\sim_R \subseteq \mathbf{U}_{\Sigma,\Gamma} \times \mathbf{U}_{\Sigma,\Gamma}$ be a binary relation. The* closure *of P with respect to $\sim_R$, denoted by $\overline{P}^R$, is the minimal superset of P such that $\forall (A, A') \in \sim_R \colon A \in \overline{P}^R \implies A' \in \overline{P}^R$.*

We observe that most existing lookup-based protocols can be easily defined by using the closure operator. As an example, we show next a representation of the Poulidor protocol by using the closure w.r.t. the label-insensitive relation.

Example 1 (Poulidor protocol). Consider the automaton $A = (\Sigma, \Gamma, Q, q_0, \delta, \ell)$ with $\Sigma = \Gamma = \{0, 1\}$, where $Q = \{0, 1, \ldots, 2n-1\}$, $q_0 = 0$, $\delta(q, c) = (q + c + 1) \pmod{2n}$ and $\ell(q) = 0$, $\forall q \in Q$. The Poulidor protocol, up to n rounds, can be defined as a set of automata that is the closure of $\{A\}$ with respect to $\sim_L$, i.e., $\overline{\{A\}}^L$.

We say that a protocol P is *closed under* $\sim_R$ if $P = \overline{P}^R$. Moreover, if $\forall A, A' \in P\colon A \sim_R A'$, then P is said to be *consistent with respect to* $\sim_R$. The Poulidor protocol, for example, is closed with respect to $\sim_S$ and $\sim_L$, and it is consistent with respect to $\sim_L$. However, it is not consistent with respect to $\sim_S$.

4.2 Resistance to Pre-ask Attacks of Layered Protocols

Avoine et al. concluded in [1] that in the context of DB protocols without a final authentication phase, the most efficient adversarial strategy when conducting a mafia-fraud attack is the pre-ask strategy.

In a recent work [12], it was proven that, if a protocol is layered and closed under $\sim_S$ (called random-labeled in [12]), then there exists a deterministic optimal strategy to execute a pre-ask attack against the protocol. This strategy consists in replying to the verifier's challenges with exactly the same sequence of responses obtained from the prover in the pre-ask session. The next proposition is based on this result.

Proposition 1 (Mafia success probability). *Let P be a layered protocol with $n > 0$ rounds. For every $x \in \{0,1\}^n$, let E^x be the event that $\Omega_A(x) = \Omega_A(c)$ for a random automaton $A \in P$ and a random input sequence $c \in \{0,1\}^n$. If P is closed under $\sim_S$, then the probability of success of an optimal pre-ask attack against P, denoted by $\mathcal{M}(P)$, can be computed as follows:*

$$\mathcal{M}(P) = \max_{x \in \{0,1\}^n} \{\Pr(E^x)\}.$$

Proof. The proposition easily follows from the definition of optimal strategy, and the underlying assumption that the adversary can find out $x \in \{0,1\}^n$ such that $\Pr(E^x) \geq \Pr(E^y)$ for every $y \in \{0,1\}^n$. □

In Lemmas 1 and 2 below, we make Proposition 1 more precise by providing formulas to compute $\Pr(E^x)$ in protocols that are consistent and closed with respect to $\sim_S$ and $\sim_L$, respectively. We do so by considering, for a given automaton, the meeting points between the input sequence used by the adversary during the pre-ask session and the challenges sent by the verifier.

Definition 7 (Meeting points set). *Given an automaton $A = (\Sigma, \Gamma, Q, q_0, \delta, \ell)$ and two input sequences $x, y \in \Sigma^n$, the* meeting points set $\mathcal{I}(A, x, y)$ *is the set* $\left\{i \in \{1, \ldots, n\} \mid \hat{\delta}(x_1 \cdots x_i) = \hat{\delta}(y_1 \cdots y_i)\right\}$.

Lemma 1. *Let P be a layered protocol with $n > 0$ rounds and A any automaton in P. If P is consistent and closed with respect to $\sim_S$, then:*

$$\mathcal{M}(P) = \max_{x \in \{0,1\}^n} \left\{ \frac{1}{4^n} \sum_{y \in \{0,1\}^n} 2^{|\mathcal{I}(A,x,y)|} \right\}.$$

Proof. Let $x, y \in \{0,1\}^n$ be two random input sequences. As P is consistent with respect to $\sim_S$, we have that $\mathcal{I}(A,x,y) = \mathcal{I}(A',x,y)$ for every $A' \in P$. In addition, given that P is closed under $\sim_S$, there are $2^{|Q|-(n-|\mathcal{I}(A,x,y)|)}$ automata A' in P such that $\Omega_{A'}(x) = \Omega_{A'}(y)$. Therefore, for a random automaton $A' \in P$, we have:

$$\Pr\left(\Omega_{A'}(x) = \Omega_{A'}(y)\right) = \frac{2^{|Q|-(n-|\mathcal{I}(A,x,y)|)}}{2^{|Q|}} = \frac{2^{|\mathcal{I}(A,x,y)|}}{2^n}. \qquad (1)$$

Now, define the event E^x that $\Omega_{A'}(x) = \Omega_{A'}(c)$ for a random input sequence $c \in \{0,1\}^n$ and a random automaton $A' \in P$ (as in Proposition 1). Hence, by the law of total probability we have:

$$\Pr(E^x) = \sum_{y \in \{0,1\}^n} \Pr(E^x \mid c = y) \Pr(c = y). \qquad (2)$$

But, $\Pr(c = y) = \frac{1}{2^n}$ and $\Pr(E^x \mid c = y) = 2^{|\mathcal{I}(A,x,y)|}/2^n$ (see Equation 1). Therefore, by applying these results in Equation 2 we obtain:

$$\Pr(E^x) = \frac{1}{4^n} \sum_{y \in \{0,1\}^n} 2^{|\mathcal{I}(A,x,y)|}. \qquad (3)$$

Finally, by using Equation 3 in Proposition 1 we obtain the expected result. □

Lemma 2. *Let P be a layered protocol with $n > 0$ rounds and A any automaton in P. If P is consistent and closed with respect to $\sim_L$, then:*

$$\mathcal{M}(P) = \frac{1}{8^n} \sum_{x,y \in \{0,1\}^n} 2^{|\mathcal{I}(A,x,y)|}.$$

Proof. For every $x \in \{0,1\}^n$, we define the event E^x that $\Omega_{A'}(x) = \Omega_{A'}(c)$ for a random automaton $A' \in P$ and a random input sequence $c \in \{0,1\}^n$. Let $\{P_1, P_2, \ldots, P_k\}$ be the equivalence classes of P with respect to $\sim_S$, i.e., $\forall B, B' \in P\colon B \sim_S B' \implies \exists i \in \{1, \ldots, k\}\colon B \in P_i \wedge B' \in P_i$. Since P is consistent and closed with respect to $\sim_L$ we derive that $\forall i, j \in \{1, \ldots, k\}\colon |P_i| = |P_j|$. Moreover, for every $i \in \{1, \ldots, k\}$ the protocol defined by the set of automata P_i is consistent and closed with respect to $\sim_S$.

Let x be a sequence in $\{0,1\}^n$ and $r \in \{1, \ldots, k\}$ such that $A \in P_r$ and let A' be a random automaton in P. According to the law of total probability we have:

$$\Pr(E^x) = \sum_{i=1}^{k} \Pr(E^x | A' \in P_i) \Pr(A' \in P_i) = \frac{1}{k} \sum_{i=1}^{k} \Pr(E^x | A' \in P_i). \qquad (4)$$

Consider the sets $R_i = \left\{ y \in \{0,1\}^n \mid \forall j \leq n\colon \hat{\delta}_r(x_1 \cdots x_j) = \hat{\delta}_i(y_1 \cdots y_j) \right\}$, for every $i \in \{1, \ldots, k\}$, where δ_i is the common transition function of the automata in P_i; remark that P_i is consistent and closed with respect to $\sim_S$. Thus, for every $i \in \{1, \ldots, k\}$ and every $y \in R_i$, we have:

$$\Pr(E^x | A' \in P_i) = \Pr(E^y | A' \in P_r). \tag{5}$$

Therefore, by iterating i over $\{1, \ldots, k\}$ and y on the corresponding R_i, we get:

$$\sum_{i=1}^{k} \sum_{y \in R_i} \Pr(E^x | A' \in P_i) = \sum_{i=1}^{k} \sum_{y \in R_i} \Pr(E^y | A' \in P_r). \tag{6}$$

Let $C_y = \{i \in \{1, \ldots, k\} \mid y \in R_i\}$ for every $y \in \{0,1\}^n$. By symmetry, $|R_i| = |R_j|$, $\forall i, j \in \{1, \ldots, k\}$ and $|C_y| = |C_z|, \forall y, z \in \{0,1\}^n$. Let $t = |R_i|$ and $c = |C_y|$. The left-hand and right-hand sides of Equation 6 are respectively equivalent to:

$$\sum_{i=1}^{k} \sum_{y \in R_i} \Pr(E^x | A' \in P_i) = t \sum_{i=1}^{k} \Pr(E^x | A' \in P_i), \text{ and} \tag{7}$$

$$\sum_{i=1}^{k} \sum_{y \in R_i} \Pr(E^y | A' \in P_r) = c \sum_{y \in \{0,1\}^n} \Pr(E^y | A' \in P_r). \tag{8}$$

Thus, from Equations 4, 6, 7 and 8 we obtain $\Pr(E^x) = \frac{c}{tk} \sum_{y \in \{0,1\}^n} \Pr(E^y | A' \in P_r)$. But, given that P_r is consistent and closed with respect to $\sim_S$ we derive that $\Pr(E^y | A' \in P_r) = \frac{1}{4^n} \sum_{z \in \{0,1\}^n} 2^{|\mathcal{I}(A,y,z)|}$ (see Equation 3). Therefore:

$$\Pr(E^x) = \frac{c}{4^n tk} \sum_{y,z \in \{0,1\}^n} 2^{|\mathcal{I}(A,y,z)|}. \tag{9}$$

Observe that the right-hand side of Equation 9 does not depend on x, therefore $\mathcal{M}(P) = \Pr(E^x)$, $\forall x \in \{0,1\}^n$. On the other hand, $\sum_{i=1}^{k} |R_i| = \sum_{y \in \{0,1\}^n} |C_y|$ which gives $tk = 2^n c$. By using this in Equation 9 we complete the proof. □

The closed formulas from the two lemmas above will be used in the next section to prove that, for every layered protocol P, there exists a protocol $P' \subseteq P$ (i.e. a protocol formed by a subset of the automata in P) such that its closure under $\sim_L$ results in a protocol with equal or better resistance to pre-ask attacks than the original protocol P.

5 A Protocol Transformation towards Optimality

As pointed out in the previous section, all lookup-based protocols found in the literature are closed and consistent with respect to either $\sim_S$ or $\sim_L$. Both design principles have been shown effective by comparing different protocol designs, but it is not clear whether they must be applied in general. In this section, we give, to the best of our knowledge, the first formal proof that those design principles are indeed well-founded. We do so by providing a protocol transformation that uses the closure with respect to both $\sim_S$ and $\sim_L$, and results in a better or equal protocol with respect to the resistance to pre-ask attacks.

Theorem 1. *Let P be a layered protocol with $n > 0$ rounds, then:*

$$\mathcal{M}(P) \geq \mathcal{M}\left(\overline{P}^{S}\right).$$

Proof. Let $x \in \{0,1\}^n$ be the input sequence selected by the adversary to query the prover in the pre-ask session. Consider the following pre-ask strategy, for a given $z \in \{0,1\}^n$ and the responses from the prover $y \in \{0,1\}^n$: at the i-th round, the adversary will reply to the verifier's challenges c with the sequence $y \oplus \neg z$. In other words, the adversary will reply with $\Omega^z_A(c) = \Omega_A(x) \oplus \neg z$, where A is the selected automaton for the execution. Let $\mathcal{M}^z(P)$ be the probability that the adversary succeeds with z in P, i.e., $\mathcal{M}^z(P) = \Pr(\Omega_A(c) = \Omega^z_A(c))$, for a random $A \in P$ and a random $c \in \{0,1\}^n$. Therefore,

$$\mathcal{M}^z(P) = \frac{1}{2^n|P|} \sum_{A \in P} \sum_{c \in \{0,1\}^n} D\left(\Omega_A(c), \Omega^z_A(c)\right), \tag{10}$$

where $D(u,v)$ is 1, if $u = v$ or 0, otherwise. Let us assume that $\overline{P}^{S} \neq P$, otherwise the theorem holds straightforwardly. Hence, $\mathcal{M}^z\left(\overline{P}^{S}\right) = a \cdot \mathcal{M}^z(P) + b \cdot \mathcal{M}^z\left(\overline{P}^{S} - P\right)$, where $a = \frac{|P|}{\left|\overline{P}^{S}\right|}$ and $b = \frac{\left|\overline{P}^{S} - P\right|}{\left|\overline{P}^{S}\right|}$. Now, assume that P is more resistant than $\overline{P}^{S}$ to pre-ask attacks, i.e. $\forall z \in \{0,1\}^n \colon \mathcal{M}^z\left(\overline{P}^{S}\right) \geq \mathcal{M}^z\left(P\right)$ and there exists at least one value z such that the inequality is strict. Therefore, $\forall z \in \{0,1\}^n \colon \mathcal{M}^z\left(\overline{P}^{S} - P\right) \geq \mathcal{M}^z(P)$ (and at least for one z the inequality is strict). This gives us:

$$\sum_{z \in \{0,1\}^n} \mathcal{M}^z(\overline{P}^{S} - P) > \sum_{z \in \{0,1\}^n} \mathcal{M}^z(P). \tag{11}$$

Our goal is to reach a contradiction. To do so, consider the set $B^{c,z}$ of automata A in P such that $\Omega_A(c) = \Omega^z_A(c)$. Hence, from Equation 10 we have:

$$\mathcal{M}^z(P) = \frac{1}{2^n|P|} \sum_{c \in \{0,1\}^n} |B^{c,z}|. \tag{12}$$

Now, observe that $\forall z, z' \in \{0,1\}^n \colon z \neq z' \implies B^{c,z} \bigcap B^{c,z'} = \emptyset$. Besides, $\forall (A,c) \in P \times \{0,1\}^n \colon \Omega_A(c) = \Omega^z_A(c)$ where $z = \Omega_A(x) \oplus \Omega_A(c)$. This gives that for every $c \in \{0,1\}^n$ it holds that $\{B^{c,z} \mid z \in \{0,1\}^n\}$ is a partition of P which gives $\forall c \in \{0,1\}^n \colon \sum_{z \in \{0,1\}^n} |B^{c,z}| = |P|$. Therefore, by applying this in Equation 12 we derive $\sum_{z \in \{0,1\}^n} \mathcal{M}^z(P) = 1$. Analogously, we obtain that $\sum_{z \in \{0,1\}^n} \mathcal{M}^z(\overline{P}^{S} - P) = 1$, which is in contradiction with the inequality in Equation 11. □

A consequence of Theorem 1 is the following. If P is optimal in terms of resistance to pre-ask attacks, then $\overline{P}^{S}$ is optimal as well. This result is relevant, as the closure $\overline{P}^{S}$ only differs from P on the distribution of the precomputed values, while it keeps the same structure of the lookup table.

The second design principle is defined as the property of the protocol to be consistent and closed with respect to $\sim_L$. In the case that P is already closed under $\sim_S$ (i.e., it satisfies the first principle), we prove, in Theorem 2 below, that there exists a subset of P whose closure w.r.t. $\sim_L$ is at least as resistant to pre-ask attacks as the former.

Theorem 2. *Let P be a layered protocol with $n > 0$ rounds and $\{P_1, \ldots, P_k\}$ be the equivalence classes of P with respect to $\sim_S$. Let $j \in \{1, \ldots, k\}$ such that $\forall i \in \{1, \ldots, k\} \colon \mathcal{M}\left(\overline{P_i}^{L}\right) \geq \mathcal{M}\left(\overline{P_j}^{L}\right)$. If P is closed under $\sim_S$, then:*

$$\mathcal{M}(P) \geq \mathcal{M}\left(\overline{P_j}^{L}\right).$$

Proof. Given that P is closed under $\sim_S$, P_i is consistent and closed with respect to $\sim_S$, for every $i \in \{1, \ldots, k\}$. Now, let A be a random automaton in P. As in Proposition 1, for every $x \in \{0,1\}^n$, we define the event E^x that $\Omega_A(x) = \Omega_A(c)$ for a random input sequence $c \in \{0,1\}^n$. By the law of total probability we have that for every $x \in \{0,1\}^n$:

$$\Pr(E^x) = \sum_{i=1}^{k} \Pr(E^x \mid A \in P_i) \Pr(A \in P_i). \tag{13}$$

From $\mathcal{M}(P) = \max_{x \in \{0,1\}^n} \Pr(E^x)$ we deduce that $\mathcal{M}(P) \geq \frac{1}{2^n} \sum_{x \in \{0,1\}^n} \Pr(E^x)$. By substituting Equation 13 in such inequality, and inverting the order of the sums we obtain:

$$\mathcal{M}(P) \geq \sum_{i=1}^{k} \left(\frac{1}{2^n} \sum_{x \in \{0,1\}^n} \Pr\left(E^x \mid A \in P_i\right) \right) \Pr(A \in P_i). \tag{14}$$

On the other hand, we derive from Lemmas 1 and 2 that, for every $i \in \{1, \ldots, k\}$, $\frac{1}{2^n} \sum_{x \in \{0,1\}^n} \Pr\left(E^x \mid A \in P_i\right) = \mathcal{M}\left(\overline{P_i}^{L}\right) \geq \mathcal{M}\left(\overline{P_j}^{L}\right)$ as P_i is consistent and closed with respect to $\sim_S$. By applying this result in Equation 14 and given that $\sum_{i=1}^{k} \Pr(A \in P_i) = 1$, we complete the proof. □

Our protocol transformation towards optimality consists of successive applications of Theorems 1 and 2. We observe that, by only considering the structure of the underlying graph of the automata, both transformations either preserve or simplify a protocol. We capture this notion of structural complexity of a protocol by the following notion of *size*.

Definition 8 (Size). *The* size *of a lookup-based protocol P, denoted by $\mathcal{S}(P)$, is the number of states of the largest automaton in P.*

In a nutshell, the size of a protocol is determined by the largest automaton that can be used during a fast phase. Remark that the number of states is a standard measure of size in automata theory.

Corollary 1 (Optimal Trade-off). *Given a positive integer number s, consider the set S of layered lookup-based protocols with size less or equal than s, i.e., $S = \{P \subseteq \mathbf{U}_{\Sigma,\Gamma} \mid \mathcal{S}(P) \leq s\}$. Let $O \subseteq S$ be the set of protocols that are* optimal *in terms of resistance to pre-ask attacks, within the set S, i.e., $O = \{P \in S \mid \nexists P' \in S\colon \mathcal{M}(P') < \mathcal{M}(P)\}$. If S is not empty, then there exists a protocol P in O that is consistent and closed with respect to $\sim_L$.*

Proof. The proof comes straightforwardly from Theorems 1 and 2. Let $P \in O$, from Theorem 1 we have $\mathcal{M}(P) \geq \mathcal{M}(\overline{P}^S)$. Because $\mathcal{S}(P) = \mathcal{S}\left(\overline{P}^S\right)$, then $\overline{P}^S \in O$. On the other hand, because of Theorem 2, there exists P_1 consistent with respect to $\sim_S$ such that $P_1 \subseteq \overline{P}^S$ and $\mathcal{M}(\overline{P}^S) \geq \mathcal{M}(\overline{P_1}^L)$. Besides, $\mathcal{S}\left(\overline{P_1}^L\right) = \mathcal{S}(P_1) \leq \mathcal{S}\left(\overline{P}^S\right) = \mathcal{S}(P) \leq s$. Therefore $\overline{P_1}^L \in S$ and $\overline{P_1}^L \in O$. □

Corollary 1 is a useful result towards finding an optimal layered protocol, as it reduces the search space to the subclass of protocols that are consistent and closed with respect to $\sim_L$. It is worth noticing that a consistent protocol with respect to $\sim_L$ imposes a rather strong structural property on a protocol, that is, all the automata in the protocol are equal if we ignore the edge and state labels. This rules out, for example, protocol composition as a technique to obtain an optimal protocol, where a protocol composition is simply the union of the sets of automata defining the two protocols, e.g. the union of Hancke and Kuhn's protocol [8] and the Uniform protocol [12].

We conclude by stressing that we have focused on non-trivial transformations in lookup-based protocols, while there exist others that can be easily included in our analyses. For example, the security-memory trade-off of every lookup-based protocol can be improved by simply removing all unreachable states in its automata representation. This consideration corresponds to a rather trivial design principle whereby states that are not reachable from the initial state can be removed, as they are not used in the protocol execution. Similarly, it can be easily proven that two equivalent protocols up to isomorphism are equal in size and resistance to pre-ask attacks.

6 Conclusions

We have studied layered protocols, a subclass of lookup-up based distance bounding protocols that contains most lookup-based protocols proposed to date. Relevant structural properties of this type of protocols that have been used in previous work in a rather intuitive way, have been formalized in this article. As a result, we developed a general security analysis that applies to all layered protocols. We have also addressed the security-memory trade-off problem in lookup-based protocols. Our results indicate that there exists an optimal layered protocol that is consistent and closed with respect to $\sim_L$, if an optimal protocol exists at all. Our future work will be oriented towards finding sufficient conditions for a layered protocol to be optimal. We plan to also extend this study to those lookup-based protocols that are not layered.

References

1. Avoine, G., Bingöl, M.A., Kardas, S., Lauradoux, C., Martin, B.: A framework for analyzing RFID distance bounding protocols. Journal of Computer Security 19(2), 289–317 (2011)
2. Avoine, G., Tchamkerten, A.: An efficient distance bounding RFID authentication protocol: Balancing false-acceptance rate and memory requirement. In: Proc. 12th International Conference on Information Security (ISC'09). LNCS, vol. 5735, pp. 250–261. Springer (2009)
3. Boureanu, I., Mitrokotsa, A., Vaudenay, S.: Secure and lightweight distance-bounding. In: Proc. 2nd International Workshop on Lightweight Cryptography for Security and Privacy (LightSec'13). LNCS, vol. 8162, pp. 97–113. Springer-Verlag, Gebze, Turkey (May 2013)
4. Brands, S., Chaum, D.: Distance-bounding protocols. In: Proc. Advances in Cryptology (EUROCRYPT'93). LNCS, vol. 765, pp. 344–359. Springer (1994)
5. Bussard, L., Bagga, W.: Distance-bounding proof of knowledge to avoid real-time attacks. In: Security and Privacy in the Age of Ubiquitous Computing, IFIP TC11 20th International Conference on Information Security (SEC2005). IFIP International Federation for Information Processing, vol. 181, pp. 223–238. Springer-Verlag, Chiba, Japan (June 2005)
6. Desmedt, Y., Goutier, C., Bengio, S.: Special uses and abuses of the Fiat-Shamir passport protocol. In: Proc. Advances in Cryptology (CRYPTO'87). LNCS, vol. 293, pp. 21–39. Springer (1988)
7. Gürel, A.O., Arslan, A., Akgün, M.: Non-uniform stepping approach to RFID distance bounding problem. In: Proc. 5th International Workshop on Data Privacy Management (DPM'10), and 3rd International Conference on Autonomous Spontaneous Security (SETOP'10). LNCS, vol. 6514, pp. 64–78. Springer (2011)
8. Hancke, G.P., Kuhn, M.G.: An RFID distance bounding protocol. In: Proc. First International Conference on Security and Privacy for Emerging Areas in Communications Networks (SecureComm-2005) , Athens, Greece, 5-9 September, 2005. pp. 67–73. IEEE Computer Society, Washington, DC, USA (2005)
9. Kardas, S., Kiraz, M.S., Bingöl, M.A., Demirci, H.: A novel RFID distance bounding protocol based on physically unclonable functions. In: Proc. 7th International Conference on Radio Frequency Identification: Security and Privacy Issues (RFIDSec'11). LNCS, vol. 7055, pp. 78–93. Springer (2012)
10. Kim, C.H., Avoine, G.: RFID distance bounding protocols with mixed challenges. IEEE Transactions on Wireless Communications 10(5), 1618–1626 (2011)
11. Kim, C.H., Avoine, G., Koeune, F., Standaert, F., Pereira, O.: The Swiss-Knife RFID distance bounding protocol. In: Proc. 11th International Conference on Information Security and Cryptology (ICISC'08), LNCS, vol. 5461, pp. 98–115. Springer (2008)
12. Mauw, S., Toro-Pozo, J., Trujillo-Rasua, R.: A Class of Precomputation-based Distance-bounding Protocols. In: Proc. 1st IEEE European Symposium on Security and Privacy – EuroS&P'16. Saarbrücken, Germany (2016)
13. Munilla, J., Peinado, A.: Distance bounding protocols for RFID enhanced by using void-challenges and analysis in noisy channels. Wireless Communications and Mobile Computing 8(9), 1227–1232 (2008)
14. Trujillo-Rasua, R., Martin, B., Avoine, G.: The Poulidor distance-bounding protocol. In: Proc. 6th International Conference on Radio Frequency Identification: Security and Privacy Issues (RFIDSec'10). LNCS, vol. 6370, pp. 239–257. Springer (2010)