Computationally Analyzing the ISO 9798–2.4 Authentication Protocol

Britta Hale and Colin Boyd

Norwegian University of Science and Technology (NTNU), Trondheim, Norway
{britta.hale,colin.boyd}@item.ntnu.no

Abstract. We provide a computational analysis of the ISO 9798–2.4 mutual authentication standard protocol in the model of Bellare and Rogaway. In contrast to typical analyses of standardized protocols, we include the optional data fields specified in the standard by applying the framework of Rogaway and Stegers. To our knowledge this is the first application of the Rogaway–Stegers technique in a standardized protocol. As well as a precise definition of the computational security properties achieved by the protocol, our analysis supplies concrete security requirements for the cryptographic primitive applied, which are absent from the protocol standard. We show that a message authentication code can be used to replace the encryption primitive if desired and that if authenticated encryption is applied it must be strongly unforgeable.

Keywords: ISO 9798, Bellare–Rogaway model, real-world protocol analysis

1 Introduction

Because it is widely agreed that authentication protocols are difficult to design correctly, standardized authentication protocols are very useful for practitioners. Today, there are many such protocols available from a variety of different standards bodies; some of these, such as the well known TLS and SSH protocols, are widely deployed. Among its 9798 series of standards, the ISO have standardized a suite of authentication protocols. Like most standardized authentication protocols, the 9798 protocols are not defined in a fully formal way. Effectively, this can lead to a number of undesirable consequences, such as difficulty in establishing exactly what properties the protocols aim to achieve, doubts regarding whether the achieved aims are actually achieved, and uncertainty about how to correctly implement the protocols securely.

Recognizing the value of a formal analysis, Basin *et al.* [2] analyzed protocols in the ISO 9798–2 standard and found a number of potential weaknesses leading to a revision of the standard. However, such a symbolic analysis omits consideration of the specific cryptographic primitives used, instead assuming an idealized encryption function. Accordingly, implementors cannot be sure whether any chosen cryptographic primitive will satisfy the requirements for security. One of the motivations for our work is to provide computational proofs for one of the 9798–2 protocols which have so far been lacking.

In this paper we focus on the ISO 9798–2.4 protocol (9798–2, section 6.2.2 Mechanism 4 of the standard). This protocol is particularly interesting because it aims at an advanced level of authentication while at the same time has the potential to fit in the Bellare–Rogaway '93 model. First, it is a mutual authentication protocol, as opposed to a unilateral one. Second, unlike some other ISO 9798 protocols, it does not rely upon a time-stamps for freshness – instead using nonces. Finally, it does not require that confidence be placed in a third party (TTP), whereas several of the other ISO 9798 protocols do. Hence, the ISO 9798–2.4 standard is more suited than other 9798 protocols to modeling in this manner. ISO 9798–2.4 is presented in §2.

Choice of cryptographic primitives. ISO 9798–2.4 protocol makes use of an encipherment algorithm with a shared symmetric encipherment key. Per the standard, the encipherment algorithm used is required to be able to detect "forged or manipulated data" and authenticated encryption is recommended for its implementation. However, any formal definition or technical description of such properties is missing from the standard. We observe that in order to achieve entity authentication there is no requirement to use encryption at all. Therefore, so as to obtain security under maximal efficiency, we will show in our analysis that a message authentication code (MAC) algorithm can be safely implemented in place of the protocol's encipherment function. Using a MAC is arguably an improvement on the standardized protocol recommendation since it will generally result in a more efficient protocol than when applying authenticated encryption. Simultaneously, we recognize that, strictly speaking, such an improved protocol no longer conforms to the standard definition. Therefore we later show that authenticated encryption, in a formally defined sense, can also provide the required properties by a simple reduction in which we only use the authentication properties of the authenticated encryption algorithm.

Optional text fields. Like most of the protocols in the ISO 9798–2 standard, the ISO 9798–2.4 protocol includes optional text fields which can be chosen in any way desired by the protocol implementor. Potentially, this flexibility is a very useful feature since it allows users to include data which is authenticated by the protocol as an additional service to obtaining entity authentication. However, computational models for protocol analysis do not usually allow such flexibility in the protocols which they analyze. In fact, any change to the analyzed protocol can potentially introduce weaknesses and any security proof will become invalid. In 2009, Rogaway and Stegers [13] introduced the notion of a *partially specified protocol* in order to deal with exactly this problem. Concisely, their model allows the adversary to actively choose the extra data, but the adversary only wins if it changes the data while the parties still accept at the end of the protocol. We apply this technique to the ISO 9798–2.4 protocol to obtain a computational proof of security no matter how the free text is chosen. Rogaway and Stegers illustrated their technique with an academic protocol – as far as we are aware ours is the first example of application of the technique in a standardized protocol.

Contributions. We regard the following as the main contributions of this paper:

– a computational proof for the ISO 9798–2.4 protocol;

– the first application of the Rogaway–Stegers framework to a standardized protocol;
– concrete advice on appropriate primitives to ensure that the ISO 9798–2.4 protocol is provably secure.

Outline. The rest of this paper is structured as follows. In the next section we explain informally the ISO 9798–2.4 protocol, using the language employed in the standard. Section 3 describes the formal Bellare–Rogaway model used in this paper. Section 4 presents the analysis of ISO 9798–4.2 in the BR model on the assumption that the primitive used is a MAC rather than an authenticated encryption algorithm. Extending this analysis, Section 5 includes consideration of associated data into the security assessment by applying the framework of Rogaway and Stegers. In Section 6 we show that our security results will hold when using authenticated encryption, as informally stated in the standard, instead of a MAC.

2 ISO 9798–2.4

Notationally, let $Text_i$ be an optional text field, $\mathcal{E}_K$ an "encipherment function" between A and B [6, p. 4], d_K the corresponding decipherment function, I_B a unique identifier of the initiating party, and R_i a random nonce. In implementation situations where a reflection attack on the protocol is impossible, the distinguisher I_B is optional [6, p. 7]. Moreover, the symbol $||$ is employed to denote the concatenation of strings when order is specified (see [7] for further details on implementation). As presented in the standard, Figure 1 shows ISO 9798–2.4 with two-party three-pass authentication.

$$B \to A : R_B || Text_1$$

$$A \to B : Text_3 || \mathcal{E}_K(R_A || R_B || I_B || Text_2)$$

$$B \to A : Text_5 || \mathcal{E}_K(R_B || R_A || Text_4)$$

Fig. 1. ISO 9798–2 Protocol Mechanism 4 Three Pass Authentication

Per the ISO 9798 standard, $\mathcal{E}_K$ must have the property that "enables the recipient ... to detect forged or manipulated data" [6, p. 4]. Furthermore, it is recommended that authenticated encryption is used [6, p. 4].

Trade-offs between security and efficiency demand heavy consideration and it is desirable to find the least computationally costly implementation of $\mathcal{E}_K$ for which the protocol is secure. The chosen encipherment function will be a

critical factor in the security proof presented in §4. While it is recommended that authenticated encryption (AE) be used for $\mathcal{E}_K$, this may in fact not provide optimal efficiency.

Predominantly, many popular implementations of authenticated encryption use a composition of a symmetric encryption scheme and a message authentication code (MAC) [3, p. 3]. Schemes which apply this composition method have, until recently, precluded the authentication of associated data, such as that which appears in the fields $Text_2$, $Text_4$ above, and even these non-composition AE schemes contain some MAC function in computation [11,8,12]. Accordingly, any AE scheme used on a message m will be no more efficient than a MAC on the same message. Consequently, we consider $\mathcal{E}_K$ concretized as a MAC function. Although this will not preserve confidentiality, just integrity, the scheme is designed for authentication only and not key exchange. Hence it turns out that a MAC is sufficient for security.

Of special note for consideration is the unique identifier I_B, which is addressed in the ISO standard by the following remark:

> When present, distinguishing identifier I_B is included... to prevent a so-called reflection attack. Such an attack is characterized by the fact that an intruder 'reflects' the challenge R_B to B pretending to be A. The inclusion of the distinguishing identifier I_B is made optional so that, in environments where such attacks cannot occur, it may be omitted. The distinguishing identifier I_B may also be omitted if unidirectional keys ... are used.

Analysis of the protocol in this paper will consider the protocol version which includes the unique identifier I_B. The alternative protocol with I_B omitted can still be proven secure in both the core and Rogaway–Stegers frameworks. Details of the required adjustments to the proofs can be found in Appendix B.

3 BR Model

In the seminal model introduced by Bellare and Rogaway [4] (henceforth referred to as the BR model), the security of a mutual authentication scheme is established on the session individuality of matching conversations. Basically, oracles for principals A and B should acquire matching conversations if and only if they both accept.

3.1 Adversary

In BR model, immense power is allowed to the adversary [4]. He is allowed to read, modify, replay, and delete messages – he is also allowed to provide his own messages to corresponding parties. Principals may engage in multiple sessions at once and the adversary may start up new sessions at his choosing.

Oracle calls allowed to the adversary are as follows:

– `Send`. Adversarial ability to request any instance $\Pi^s_{A,B}$ of a principal A to send a message to an instance of another principal B. In addition to learning the outgoing message, the adversary also learns whether or not it was accepted [4, p. 9].
– `Corrupt`. Adversarial ability to take over any principal A, obtain all of its private keys, and compute $\mathcal{E}_K$ under any symmetric key K that belongs to it.

While the `Send` query is used in the BR model, it should be noted that the query `Corrupt` is not. However, this research will employ a `Corrupt` query because it is reasonably within the realm of a real adversary's capabilities. Actually, it has become a generally accepted practice to allow this query since BR was published. Any instance will be considered *fresh* if neither its nor its partner's principal, if the partner exists, have been the subject of a `Corrupt` query by an adversary, and the instance has accepted.

Since ISO 9798–2.4 is designed for mutual authentication in a symmetric setting, there is no need for a `(Session) Key Reveal` query – the `Corrupt` query allows the adversary to access the symmetric key when such a query is desired.

3.2 Matching Conversations

Below is the definition of matching conversations, per the BR model. In short, matching conversations will be the requirement for the definition of a secure mutual authentication scheme in the model being considered for ISO 9798–2.4. Alternative definitions for determining uniqueness of a session have been applied in other research since the BR model was introduced, including using unspecified session identifiers [10]. Still later, more fully specified session identifiers that capture similar information to that of BR (e.g. [9]) have been utilized. Due to the simplicity of matching conversations and the natural session-identifier format that they epitomize, they will be employed in this work to capture partnering between sessions.

Definition 1. Matching Conversations [4]. *Fix a number of moves $R = 2\rho - 1$ and an R-move protocol Π. Run Π in the presence of an adversary E and consider two oracles, $\Pi^s_{i,j}$ and $\Pi^t_{j,i}$ that engage in conversations K and K', respectively. Let τ_l be time increments, α_l be messages sent by $\Pi^s_{i,j}$, and β_l be messages sent by $\Pi^t_{j,i}$.*

1. *Responder oracle has a conversation matching the conversation of an initiator oracle.*
   K' is in a matching conversation with K if there exists $\tau_0 < \tau_1 < \ldots < \tau_R$ and $\alpha_1, \beta_1, \ldots, \alpha_\rho, \beta_\rho$ such that K is prefixed by

$$< \tau_0, \lambda, \alpha_1 >, < \tau_2, \beta_1, \alpha_2 >, < \tau_4, \beta_2, \alpha_3 >, \ldots, < \tau_{2\rho-4}, \beta_{\rho-2}, \alpha_{\rho-1} >, \\ < \tau_{2\rho-2}, \beta_{\rho-1}, \alpha_\rho >$$

   and K' is prefixed by

$$< \tau_1, \alpha_1, \beta_1 >, < \tau_3, \alpha_2, \beta_2 >, < \tau_5, \alpha_3, \beta_3 >, \ldots, < \tau_{2\rho-3}, \alpha_{\rho-1}, \beta_{\rho-1} >.$$

2. *Initiator oracle has a conversation matching the conversation of a responder oracle.*
   K is in a matching conversation with K' if there exists $\tau_0 < \tau_1 < \ldots < \tau_R$ and $\alpha_1, \beta_1, \ldots, \alpha_\rho, \beta_\rho$ such that K' is prefixed by

$$< \tau_1, \alpha_1, \beta_1 >, < \tau_3, \alpha_2, \beta_2 >, < \tau_5, \alpha_3, \beta_3 >, \ldots, < \tau_{2\rho-3}, \alpha_{\rho-1}, \beta_{\rho-1} >, \\ < \tau_{2\rho-1}, \alpha_\rho, * >,$$

   and K is prefixed by

$$< \tau_0, \lambda, \alpha_1 >, < \tau_2, \beta_1, \alpha_2) >, < \tau_4, \beta_2, \alpha_3 >, \ldots, < \tau_{2\rho-4}, \beta_{\rho-2}, \alpha_{\rho-1} >, \\ < \tau_{2\rho-2}, \beta_{\rho-1}, \alpha_\rho >.$$

3.3 Secure Mutual Authentication

Concisely, the BR model prescribes that entities accept if and only if the session transcripts (conversations) match. This is presented below.

Definition 2. *Secure Mutual Authentication [4]. Let* $\mathtt{No-Matching}^E(k)$ *be the event that there exists an uncorrupted oracle $\Pi^s_{i,j}$ which accepted and there is no uncorrupted oracle $\Pi^t_{j,i}$ which engaged in matching conversation with $\Pi^s_{i,j}$. The protocol Π is a* secure mutual authentication *protocol if for any polynomial time adversary E,*

1. *Matching conversations $\Rightarrow$ acceptance.*
   If oracles $\Pi^s_{i,j}$ and $\Pi^t_{j,i}$ have matching conversations, then both oracles accept.
2. *Acceptance $\Rightarrow$ matching conversations.*
   The probability of $\mathtt{No-Matching}^E(k)$ *is negligible.*

In the BR model, the network is viewed as a 'benign adversary' whose actions are restricted to choosing an initiator oracle $\Pi^s_{i,j}$ and responder oracle $\Pi^t_{j,i}$, "faithfully conveying each flow from one oracle to the other" [4, p. 10], starting with the initiator. Such an adversary is deterministic. Thus the adversary has the power to determine i, j, s, and t, but must use these in any protocol execution with parameter k. While this is mostly of interest in a key-exchange setting, it is noted here to highlight strength of the model; i.e. if adversary behaves according to the protocol, with eavesdropping, it gains no additional advantage.

Pursuant to the definition of matching conversations is the uniqueness of matching partners. Bellare and Rogaway [4, p. 13] show that the probability of having multiple matching partners is negligible in this model .

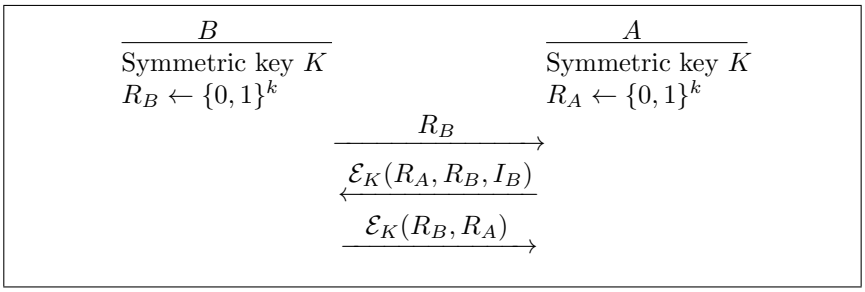

Fig. 2. ISO 9798–2.4 Protocol Core

4 Security of ISO 9798–2.4

In the following section, the security of Π will be considered in the case where $\mathcal{E}_K$ is implemented as a MAC function, with $\mathcal{E}_K(m) = (m, \text{MAC}(m))$, for a message m. Assessment will be performed in the BR model. Additionally, this proof focuses on the protocol core – the associated data in the ISO 9798–2.4 protocol, $Text_i$, will not be considered until §5. Figure 2 summarizes the core.

Theorem 41 *Let Π be the core of the ISO 9798–2.4 protocol implemented with a strongly unforgeable MAC algorithm*[1] $\mathcal{E}_K(M) = \text{MAC}_K(M) = (M, T)$*, as in Definition 4. Let E be a polynomial-time adversary against the mutual authentication scheme, running in time t and asking q queries. Then the advantage of E can be reduced to the advantage of an adversary against the MAC, running in time $t_F \approx t$ and asking $q_F = q$ queries:*

$$\mathbf{Adv}_{\Pi}^{MA}(E) \leq 2p^2S \cdot \mathbf{Adv}_{\Pi}^{\text{MAC}}(F) + \tfrac{q^2}{2^{k+1}}.$$

where S is the number of sessions and p is the number of principals.

Proof (Proof with $\mathcal{E}_K$ implemented as a MAC).

Ideas from this proof follow from other proofs for entity authentication [4,5]. When Definition 2 is satisfied, the proof will be complete.

If the principals possess matching conversations, then they will both accept, by the protocol definition – hence satisfying the first condition of the definition of a secure mutual authentication protocol is trivial. Correspondingly, the remainder of this proof will target the second case; that acceptance implies matching conversations.

Adversarial advantage, $\mathbf{Adv}_{\Pi}^{\text{MA}}(E)$, will be defined as the probability that it can succeed in persuading an oracle to accept without a matching conversation.

[1] Strong-unforgeability is required since an adversary's ability to produce a different, yet valid, MAC on a protocol message flow would trivially result in principals accepting without matching conversations. Essentially, the tag would still verify correctly even though each instance held a different conversation transcript.

Let NC represent the event that two different instances accept with the same nonce pair. Then we can derive the following:

$$\begin{aligned}\mathbf{Adv}^{\mathrm{MA}}(E) &\leq \Pr[\neg \text{ Match.Conv }] \\ &\leq \Pr[\mathrm{NC}] + \Pr[\neg \text{ Match.Conv } | \neg \mathrm{NC}]\end{aligned}$$

Let E be an adversary that attempts to inveigle acceptance from an oracle without that oracle being in matching conversation with a partner oracle. Let q be a polynomial bound on the number of oracle calls allowed to E.

Nonce Collision

As q calls are allowed to E and nonces are selected independently, the birthday bound yields

$$\Pr[\mathrm{NC}] \leq q^2/2^{k+1}. \tag{1}$$

Matching Conversations without Nonce Collision

E will succeed with probability equal to the sum of the probability of success against at least one initiator oracle (i.e. gets an initiator oracle to accept without any other oracle in matching conversation with it) and the probability of success against at least one responder oracle (but no initiator oracle). Thus the proof follows two cases.

Description of F: In the protocol game, E will get an oracle $\Pi^s_{i,j}$ for a principal i to accept with non-negligible probability, without the existence of an oracle $\Pi^t_{j,i}$ in matching conversation with $\Pi^s_{i,j}$. Using this fact, F's goal is to compute a valid MAC for a message m where m has not been queried from the MAC oracle.

F starts the game and initiates E on input 1^k. F selects a pair i, j at random from the set of all principals $\{1, \ldots, p\}$ as well as a session $s \in_R \{1, \ldots, S\}$ – thus F is selecting $\Pi^s_{i,j}$ as its guess for the initiator oracle against which E will succeed. F has a MAC oracle, per definition 4, that runs on a key K chosen randomly from $\{0,1\}^k$ which it will use to calculate the tag for messages between i and j. For all principals in the set $\{1, \ldots, p\} \setminus \{i, j\}$, F also selects keys k_l for each pair of principals; these keys will be used to calculate MAC_{k_l} and $\mathrm{MAC.ver}_{k_l}$ on message flows between all principals other than i and j.

F answers all of E's `Send` and `Corrupt` queries, according to the protocol. However, should E ask a `Corrupt` query on the principals corresponding to either of the instances i or j, F will give up. If E asks a `Send` query, for $\Pi^l_{i,j}$ or $\Pi^m_{j,i}$ for any l or m, F will compute the response with its MAC generation oracle and, if necessary, F checks incoming MACs using its MAC verification oracle

Against Initiator: Suppose that E succeeds against at least one initiator oracle with non-negligible probability of success. From E an adversary F will be constructed against the MAC.

Note: if E never calls on i to initiate a protocol run, F gives up.

Now suppose that E does call on i to initiate a protocol run. Then at some time τ_0, $\Pi^s_{i,j}$ will send out a flow R_i. For some time $\tau_2 > \tau_0$, $\Pi^s_{i,j}$ must receive a flow of the form $\mathcal{E}_K(R_j, R_i, I_i)$, else F gives up. If F has already received $\mathrm{MAC}_K(R_j, R_i, I_i)$ from its oracle, then it gives up; else it returns $\mathrm{MAC}_K(R_j, R_i, I_i) = \big((R_j, R_i, I_i),\ Tag_1\big)$ as its guess for $\mathcal{E}_K(R_j, R_i, I_i)$.

Game of F: If we assume that E succeeds on the instance guessed by F, then the oracle $\Pi^s_{i,j}$ accepts. Given also that there are no collisions, F cannot have previously obtained the flow $\mathcal{E}_K(R_j, R_i, I_i)$. Therefore F outputs a valid forgery for $\mathcal{E}_K(R_j, R_i,\ I_i)$ (i.e. a valid forgery for the MAC per Definition 4).

Ergo (assuming that E always succeeds against at least one initiator oracle),

$$\mathbf{Adv}^{\mathrm{MAC}}_{\Pi}(F) \geq \frac{1}{p^2 S} \cdot \Pr[\neg \text{ Match.Conv } | \neg \text{ NC}] \tag{2}$$

Against Responder: Suppose that E succeeds against at least one responder oracle but no initiator oracle with non-negligible probability of success. Similarly to the initiator case above, an adversary F of the MAC will be constructed.

Note: if E never calls on j as a responder to a protocol run, or if E succeeds against some initiator oracle, F gives up.

Now suppose that E does call on j as a responder oracle. Then at some time τ_1, $\Pi^t_{j,i}$ must receive a flow R_i and respond with a flow $\mathcal{E}_K(R_j, R_i, I_i)$. At time $\tau_3 > \tau_1$, $\Pi^t_{j,i}$ must receive a flow $\mathcal{E}_K(R_i, R_j)$, else F gives up. If F has already calculated $\mathrm{MAC}_K(R_i, R_j)$, then it gives up; else it computes $\mathrm{MAC}_K(R_i, R_j) = \big((R_i, R_j),\ Tag_2\big)$ and returns this as its guess for $\mathcal{E}_K(R_i, R_j)$.

Game of F: As above, if we assume that the probability that E succeeds on the instance guessed by F then the oracle $\Pi^t_{j,i}$ accepts. Given also that there are no collisions, F cannot have previously obtained the flow $\mathcal{E}_K(R_i, R_j)$. Therefore F outputs a valid forgery for $\mathcal{E}_K(R_i, R_j)$ (i.e. a valid forgery for the MAC).

Therefore (under the assumption that E always succeeds against at least one responder oracle but no initiator oracle),

$$\mathbf{Adv}^{\mathrm{MAC}}_{\Pi}(F) \geq \frac{1}{p^2 S} \cdot \Pr[\neg \text{ Match.Conv } | \neg \text{ NC}]. \tag{3}$$

Combining equations (2) and (3), and taking into account the two mutually exclusive cases, we have:

$$\mathbf{Adv}^{\mathrm{MAC}}_{\Pi}(F) \geq \frac{1}{2}\left(\frac{1}{p^2 S} \cdot \Pr[\neg \text{ Match.Conv } | \neg \text{ NC}]\right)$$

or

$$2p^2 S \cdot \mathbf{Adv}^{\mathrm{MAC}}_{\Pi}(F) \geq \Pr[\neg \text{ Match.Conv } | \neg \text{ NC}]. \tag{4}$$

Negligible probability of success

By equations (1) and (4), the probability that E secures its goal of oracle acceptance, while maintaining the absence of another oracle in matching conversation, is negligible. Particularly,

$$\begin{aligned}\mathbf{Adv}^{\mathrm{MA}}(E) &\leq \Pr[\neg \text{ Match.Conv }] \\ &\leq \Pr[\mathrm{NC}] + \Pr[\neg \text{ Match.Conv } | \neg \text{ NC}] \\ &\leq q^2/2^{k+1} + 2p^2S \cdot \mathbf{Adv}_{\Pi}^{\mathrm{MAC}}(F).\end{aligned}$$

Moreover, if E runs in time t and asks q queries, then F runs in time $t_F \approx t$ and asks $q_F = q$ queries. Thus, the protocol is secure with $\mathcal{E}_K$ implemented as a MAC function, where $\mathcal{E}_K(M) = \mathrm{MAC}(M) = (M, T)$, for a message M.

5 Analysis with Associated Data – RS Model

While the analysis above demonstrates security of the ISO 9798–2.4 protocol core, it omits an important aspect of the original protocol: optional text fields. As with most protocols, these fields allow additional data to be sent – sometimes authenticated – during the mutual authentication process. However, the addition of this data would nullify the security statement in §4 since the inclusion of additional fields was not considered.

Rogaway and Stegers [13] introduced a model that addresses this issue by splitting the protocol into two parts: the partially specified protocol core (PSP) and the protocol details (PD). In essence, the protocol details selects the optional text fields to be added. Yet, since there is no restriction on the data that is sent in these fields, it is necessary to maintain the perspective that data choice could weaken the protocol. Fundamentally, this weakness is modeled by allowing the adversary itself to choose the optional text fields; thus, not only does the adversary call the security game but the game also calls the adversary.

Data fields in the model fall into two categories: associated data (AD) which are authenticated by the protocol and, by protocol security, should be guaranteed to be mutually held by all parties, and ancillary but unauthenticated data. While the RS model addresses both categories of data, the former is of salient concern. Even though the unauthenticated data fields are relevant to the security of the protocol, as they may influence the selection of the AD fields, they are also subject to being changed by an adversary en route. Consequently, no authenticity claims can be made on the non-authenticated fields.

Succinctly, the ISO 9798–2.4 protocol has text fields Text_l for $l \in \{1, \ldots, 5\}$. Data fields Text_1, Text_3, and Text_5 are not authenticated and can therefore be modified en route later in the protocol. Hence they cannot be classified as AD. Likewise, since Text_4 is sent in the last message by the initiator, there is no guarantee that it will be received by the responder and is consequently also not AD, although it is authenticated. Ultimately, this leaves field Text_2 as the only AD. Applying the Rogaway–Stegers (RS) framework, it is our goal to

demonstrate that ISO 9798–2.4 is still secure even when the AD selection is under adversarial control.

Rogaway and Stegers combine their AD framework with a particular mutual authentication model, using session IDs, and apply it to a variant of the Needham-Schroeder-Lowe protocol [13, p. 7]. For application of the Rogaway–Stegers AD framework, we use the BR mutual authentication model with matching conversations, as in §4. To avoid trivial breaks of the matching conversations by an adversary, conversation transcripts will not include unauthenticated text fields – i.e. $Text_1$, $Text_3$, and $Text_5$

Notably, the RS model is a public-key mutual authentication model, whereas ISO 9798–2.4 is a symmetric-key protocol. As a result of these details, slight model adaptations are required. Nonetheless, security in RS framework for associated data can be summarized as shown below.

Definition 3. RS Framework for Associated Data with BR Mutual Authentication Model [4,13]. *Let* $\mathtt{No-Matching}^E(k)$ *be the event that there exists an uncorrupted oracle* $\Pi^s_{i,j}$ *which accepted and there is no uncorrupted oracle* $\Pi^t_{j,i}$ *which engaged in matching conversation with* $\Pi^s_{i,j}$*. The protocol* Π *is a* secure mutual authentication *protocol if for any polynomial time adversary* $\mathcal{A}$*,*

1. *Matching conversations* $\Rightarrow$ *acceptance.*
   If oracles $\Pi^s_{i,j}$ *and* $\Pi^t_{j,i}$ *have matching conversations, then both oracles accept.*
2. *Acceptance* $\Rightarrow$ *matching conversations.*
   The probability of $\mathtt{No-Matching}^E(k)$ *is negligible.*
3. *Matching Conversations* $\Rightarrow$ *Matching AD.*
   If oracles $\Pi^s_{i,j}$ *and* $\Pi^t_{j,i}$ *have matching conversations, then the associated data in the protocol is guaranteed to be mutually held.*

Theorem 51 *Let* Π *be the ISO 9798–2.4 protocol implemented with a strongly unforgeable MAC algorithm* $\mathcal{E}_K(M) = (m, \mathrm{MAC}_K(M))$*, including the optional text fields* $Text_l$ *and the associated data* $Text_2$*. Then advantage of an polynomial-time adversary against the mutual authentication scheme can be reduced to the adversarial advantage against the MAC:*

$$\mathbf{Adv}^{MA}(\mathcal{A}) \leq 2p^2S \cdot \mathbf{Adv}^{\mathrm{MAC}}_{\Pi}(F) + q^2/2^{k+1}.$$

Proof. Succinctly, this proof will build on that of §4 and the following previously used notation will continue, with the addendum of matching AD. For conciseness, text fields $Text_l$ will be denoted T_l.

- NC: Two different instances accept with the same nonce pair.
- Match.Conv: Two different instances are in matching conversation.
- Match.AD: The AD held by both parties at the end of the protocol matches.
- q: number of calls allowed to the adversary.
- S: number of sessions.
- p: number of principals.

– 1^k: security parameter.

Simply, the first requirement for mutual authentication in definition 3 follows from the protocol description. It remains to be shown that acceptance still implies matching conversations even with the optional text fields included and that this in turn guarantees the associated data is mutually held by both parties at termination. Adversarial advantage against the mutual authentication scheme thus complies with the following inequalities. The final reduction will serve as a triad proof infrastructure.

$$\begin{aligned}\mathbf{Adv}^{\mathrm{MA}}(\mathcal{A}) &\leq \Pr[(\neg \text{ Match.Conv}) \vee (\neg \text{ Match.AD} \mid \text{Match.Conv})] \\ &\leq \Pr[\neg \text{ Match.Conv}] + \Pr[\neg \text{ Match.AD} \mid \text{Match.Conv}] \\ &\leq \Pr[\mathrm{NC}] + \Pr[\neg \text{ Match.Conv} \mid \neg \text{ NC}] \\ &+ \Pr[\neg \text{ Match.AD} \mid \text{Match.Conv} \ \wedge \neg \text{ NC}]\end{aligned} \tag{5}$$

Nonce Collision

As q calls are allowed to $\mathcal{A}$ and nonces are selected independently, the birthday bound yields

$$\Pr[\mathrm{NC}] \leq q^2/2^{k+1}. \tag{6}$$

Acceptance Implies Matching Conversations

Immediately, this proof is in parallel to that in §4. Furthermore, nonce collision has already been accounted for. Correlatively, the following are addenda to the proof and reduction statement presented in §4:

Case 1: (continued from §4.) Let F be an adversary against the MAC, having a MAC oracle that runs on a key K chosen randomly from $\{0,1\}^k$. Suppose that the probability that $\mathcal{A}$ succeeds in having an initiator oracle accept without being in matching-conversation is non-negligible.

When the PSP requires a choice of text fields, it calls on the PD, answered by $\mathcal{A}$, to select T_l. If the responder exists, at time τ_1, the PSP calls the PD which responds with its selection for all text fields in the second message flow while also setting AD $= T_2$ for the responder. Regardless of the responder's view, when the initiator receives the flow $T_3||\mathcal{E}_K(R_j, R_i, I_i, T_2)$ at time τ_2, the PD sets AD $= T_2$ for the initiator.

Even though $\mathcal{A}$ has power over the PD and is therefore able to choose the AD, it is deterministic in its selection. Essentially, $\mathcal{A}$ is not allowed to simply change T_2 at a later date; once chosen, $\mathcal{A}$ may not attempt to insure that principals i and j hold different AD values by simply reselecting T_2 via the PD when setting the AD from the initiator's view. Thus, any attempt by $\mathcal{A}$ to ensure that conversations do not match by changing T_2, and therefore the AD, must be made by exchanging the flow with a previous one or by a valid forgery.

Consequently, as in §4, F has previously calculated the flow $\mathcal{E}_K(R_j, R_i, I_i, T_2)$ or F outputs a valid forgery for it (i.e. a valid forgery for the MAC). Since there are no nonce collisions, F has not previously calculated the flow. Therefore it must output a valid forgery for the MAC.

Case 2: (continued from §4.) Let F be an adversary against the MAC, having a MAC oracle that runs on a key K chosen randomly from $\{0,1\}^k$. Suppose that the probability that $\mathcal{A}$ succeeds against a responder oracle but no initiator oracles is non-negligible.

When the PSP requires a choice of text fields, it calls on the PD, answered by $\mathcal{A}$, to select T_l. At time τ_1, the PD sets AD $= T_2$ for the responder. When it receives the flow $T_5||\mathcal{E}_K(R_j, R_i, I_i, T_2)$, the PD again sets AD $= T_2$ from the initiator's view at time τ_2.

As in Case 1, the AD is chosen deterministically in the PD call and may not simply be changed later by the PD to ensure non-matching conversations. Consequently, in order to get the responder to accept at time τ_3, F has either previously calculated the flow $\mathcal{E}_K(R_i, R_j, T_4)$ or F outputs a valid forgery for it (i.e. a valid forgery for the MAC). Since there are no nonce collisions, F has not previously calculated the flow. Therefore it must output a valid forgery for the MAC.

Combining Case 1 and 2, the reduction is summarized as

$$\mathbf{Adv}_{\Pi}^{\mathrm{MAC}}(F) \geq \tfrac{1}{2}\left(\tfrac{1}{p^2 S} \cdot \Pr[\neg \text{ Match.Conv } | \neg \text{ NC}]\right).$$

Hence,

$$\Pr[\neg \text{ Match.Conv } | \neg \text{ NC}] \leq 2p^2 S \cdot \mathbf{Adv}_{\Pi}^{\mathrm{MAC}}(F). \tag{7}$$

Associated Data Agreement

Compactly, it can be assumed that an instance and its partner are in matching conversations and that there are no nonce collisions. It remains to show that the same AD, T_2, is equally held by both sides.

Since i and j are in matching conversation, at some time τ_1 the responder sent the message $T_{3_j}||\mathcal{E}_K(R_j, R_i, I_i, T_{2_j})$ and at some time τ_2 the initiator received a message

$$T_{3_i}||\mathcal{E}_K(R_j, R_i, I_i, T_{2_i}),$$

where T_{3_i} and T_{3_j} may or may not be equal. Moreover, this was authenticated under the symmetric key K, $T_{2_i} = T_{2_j}$. Therefore $\mathrm{AD}_i = \mathrm{AD}_j$.

Ergo,

$$\Pr[\neg \text{ Match.AD } | \neg \text{NC} \wedge \text{ Match.Conv}] = 0. \tag{8}$$

Combining the reductions from equations 6–8 with equation 5 yields the full reduction of security for ISO 9798–2.4 with inclusion of associated data.

Remark: As previously observed, there is no guarantee that the final message flow is received, which limits T_4 from inclusion in the AD. However, assuming matching conversations, a responder that receives a flow $T_5||\mathcal{E}_K(R_i, R_j, T_4)$ can be assured that T_4 has been authenticated by the sender. Namely, this follows from the assumptions above – if $T_{5_j}||\mathcal{E}_K(R_i, R_j, T_4)$ is received by an instance $\Pi^t_{j,i}$ under a symmetric key, then the flow $T_{5_i}||\mathcal{E}_K(R_i, R_j, T_4)$ was sent by a partner instance $\Pi^s_{i,j}$ of principal i and $\mathcal{E}_K(R_i, R_j, T_4)$ must also form part of the conversation transcript of $\Pi^s_{i,j}$.

6 Using Authenticated Encryption

As previously stated, the ISO 9798–2 standard currently does not specify the primitive to be used as the encipherment function $\mathcal{E}_K$. Likewise, while the standard concurs that its integrity requirements "can be achieved in many ways" [6, p. 4], authenticated encryption per the ISO/IEC 19772 standard is recommended. Consequently, it is desirable to check that a protocol implemented with an AE primitive will have security traceable to that of a protocol under a MAC primitive, proven in §4 and §5.

For the lemmas and theorem below, *strongly-unforgeable authenticated encryption* (SUF-AE) is defined as in Definition 5. The reader should note that this definition may be more commonly referred to as INT-CTXT$_{\mathcal{AE}}$ [3] – however, the term SUF-AE is used here for the sake of clarity with its relationship to SUF-CMA.

Lemma 1. *Suppose that $(\mathcal{K}, \mathcal{E}, \mathcal{D})$ is a strongly-unforgeable authenticated encryption algorithm according to Definition 5. Then the MAC algorithm with $\mathrm{MAC}_K(M) = (M, \mathcal{E}(K, M))$ is SUF-CMA secure, according to Definition 4.*

Proof. Suppose that E is an adversary that succeeds against the MAC with advantage $\mathbf{Adv}^{\text{SUF-CMA}}_{\text{MAC}}(E)$, which is non-negligible; from E an adversary F will be constructed against the authenticated encryption algorithm.

Let MAC be a strongly-unforgeable message authentication code. F starts the game, chooses $K \xleftarrow{\$} \mathcal{K}$ and initiates adversary E on 1^n.

Provide F with an oracle for $\mathrm{MAC}_K(\cdot) = (\cdot, \mathcal{E}_K(\cdot))$ which it will use to answer E's MAC queries. Since E forges the MAC under key K, it can output a valid pair (M, C) such that $\mathcal{D}_K(C) = M$ and E did not ask a query $\mathrm{MAC}_K(M)$ such that $C = \mathcal{E}_K(M)$.

Now, if F wishes to forge an authenticated encryption on message $M_l \in$ `Message`, $l \in \{0, \ldots, w\}$, it will call E on M_l. Respectively, E will output the message-tag pair (M_l, C_l) where $\mathcal{D}_K(C_l) = M_l$ and C_l was not previously the answer to an oracle call $\mathrm{MAC}_K(M_l)$. Since E succeeds with non-negligible probability $\mathbf{Adv}^{\text{SUF-CMA}}_{\text{MAC}}(E)$, F will also succeed in forging the authenticated encryption with $\mathcal{E}_K(M_l) = C_l$, where C_l has not previously been produced as a ciphertext on M_l, with non-negligible probability. Thus,

$$\mathbf{Adv}^{\text{SUF-CMA}}_{\text{MAC}}(E) \leq \mathbf{Adv}^{\text{SUF-AE}}_{(\mathcal{K},\mathcal{E},\mathcal{D})}(F).$$

Lemma 2. *Let Π be the 9798–2.4 protocol implemented with a strongly unforgeable authenticated encryption algorithm $(\mathcal{K},\mathcal{E},\mathcal{D})$. Let Π' be the 9798–2.4 protocol implemented with the MAC as in Lemma 1. An efficient adversary against Π can be efficiently converted into an adversary against Π'.*

Proof. Let $\mathcal{A}$ be an efficient adversary against the mutual authentication protocol Π operating with advantage $\mathbf{Adv}^{\text{MA-AE}}_{\Pi}(\mathcal{A})$; from $\mathcal{A}$, an adversary $\mathcal{B}$ will be constructed against Π'. Let the advantage of $\mathcal{B}$ be denoted $\mathbf{Adv}^{\text{MA-MAC}}_{\Pi'}(\mathcal{B})$.

Starting the protocol game, $\mathcal{B}$ chooses $K \xleftarrow{\$} \mathcal{K}$ and initiates adversary $\mathcal{A}$ on 1^k. Let n be a polynomial bound on the number of queries allowed to $\mathcal{A}$.

$\mathcal{B}$ will answer $\mathcal{A}$'s first `Send` query in the open. Thereafter, all `Send` queries will be answered by submitting them to $\text{MAC}_K(M)$, $\mathcal{B}$'s oracle for the MAC, and only passing on the tag part $\mathcal{E}(K,M)$ of the answer pair $(M,\mathcal{E}(K,M)) \leftarrow \text{MAC}_K(M)$ to $\mathcal{A}$.

When $\mathcal{B}$ wishes to succeed against Π' it must convince an instance to accept without matching conversations. To do this, $\mathcal{B}$ will pass all messages to $\mathcal{A}$, which will answer each `Send` query, outputting encryptions C_i on message queries $i = 2, 3$ (second and third protocol flows). $\mathcal{B}$ will then relay the pair (M_i, C_i) to its respective instances in the second and third protocol flows as required.

Since $\mathcal{A}$ can succeed against Π and therefore get one of its instances to accept without matching conversations with non-negligible probability, relaying the message-tag pair will also ensure that one of the instances in Π' will accept without matching conversations, so long as $\mathcal{A}$ has not made a forgery on the AE. Should $\mathcal{A}$ forge the AE, then the message added by $\mathcal{B}$ to create the message-tag pair will not match the decryption of the ciphertext output by $\mathcal{A}$. However, with n queries allowed to $\mathcal{A}$ and a probability of forgery $\mathbf{Adv}^{\text{SUF-AE}}_{(\mathcal{K},\mathcal{E},\mathcal{D})}(F)$, this only occurs with negligible probability. Thus,

$$\begin{aligned}\mathbf{Adv}^{\text{MA-AE}}_{\Pi}(\mathcal{A}) &\leq \mathbf{Adv}^{\text{MA-AE}}_{\Pi}(\mathcal{A} \text{ succeeds } |\mathcal{A} \text{ does not forge}) + \Pr(\mathcal{A} \text{ forges}) \\ &\leq \mathbf{Adv}^{\text{MA-MAC}}_{\Pi'}(\mathcal{B}) + n \cdot \mathbf{Adv}^{\text{SUF-AE}}_{(\mathcal{K},\mathcal{E},\mathcal{D})}(F).\end{aligned}$$

Theorem 61 *Let Π be the 9798–2.4 protocol implemented with a strongly unforgeable authenticated encryption algorithm $(\mathcal{K},\mathcal{E},\mathcal{D})$. Let Π' be the 9798–2.4 protocol implemented with the MAC as in Lemma 1. In the Rogaway–Stegers framework with associated data considered, an efficient adversary against Π can be efficiently converted into an adversary against Π'.*

Proof. Applying the reductions from Lemma 6, Lemma 2, and §5 Game 5, it follows that

$$\begin{aligned}\mathbf{Adv}^{\text{MA-AE}}_{\Pi}(\mathcal{A}) &\leq \mathbf{Adv}^{\text{MA-MAC}}_{\Pi'}(\mathcal{B}) + n \cdot \mathbf{Adv}^{\text{SUF-AE}}_{(\mathcal{K},\mathcal{E},\mathcal{D})}(F) \\ &\leq 2p^2S \cdot \mathbf{Adv}^{\text{SUF-CMA}}_{\text{MAC}}(E) + q^2/2^{k+1} + n \cdot \mathbf{Adv}^{\text{SUF-AE}}_{(\mathcal{K},\mathcal{E},\mathcal{D})}(F) \\ &\leq 2p^2S \cdot \mathbf{Adv}^{\text{SUF-AE}}_{(\mathcal{K},\mathcal{E},\mathcal{D})}(F) + q^2/2^{k+1} + n \cdot \mathbf{Adv}^{\text{SUF-AE}}_{(\mathcal{K},\mathcal{E},\mathcal{D})}(F) \\ &= (2p^2S + n) \cdot \mathbf{Adv}^{\text{SUF-AE}}_{(\mathcal{K},\mathcal{E},\mathcal{D})}(\mathcal{F}) + q^2/2^{k+1}.\end{aligned}$$

7 Conclusion

Ultimately, these results underscore the security of the ISO 9798–2.4, a real-world mutual authentication standard. Basing security on matching conversations, the protocol core was first analyzed in the Bellare–Rogaway model. Being more efficient than an authenticated encryption scheme, a MAC function was used in the security assessment of the protocol and shown to be sufficient. While this no longer yields privacy, it attests to the security of the mutual authentication scheme in the most fundamental cases – when merely integrity and authenticity are required.

Integrated into the proof of security for the protocol core is a polynomial-time reduction to the security of the MAC. Furthermore, we have shown that while a strongly-unforgeable MAC is sufficient for security, the current recommendation of ISO 9798–2 of authenticated encryption will also result in a secure protocol, albeit a less efficient one. Strong unforgeability is required for both the MAC and the AE since an adversary's ability to produce a different, yet valid, encipherment for a message flow would trivially result in principals accepting without matching conversations.

Subsequently, the full protocol, inclusive of associated data, was analyzed in the RS model. With additional data fields included and adversarial selection of the data for those fields allowed, the protocol was again demonstrated to be secure under the MAC implementation.

Ad interim, a parallel symbolic analysis of the protocol for juxtaposition was also performed using Scyther, albeit omitted from this paper. In comparison to the symbolic analysis by Basin *et al.* [2] which applied symmetric encryption and checked for aliveness and weak-agreement, we reconnoitered security while implementing a MAC for the encipherment function as well as demanding non-injective agreement and non-injective synchronization; thereby twinning restrictions with those used in the computational analysis. Results from the separate symbolic analysis affirmed those in this research.

Collectively, these results demonstrate a clarification and efficiency improvement to the ISO standard's current requirements, while also validating the protocol's security in the computational model. Even though authenticated encryption is recommended for implementation of the standard, it is confirmed in this research that security is achievable for the mutual authentication scheme using only a MAC. If authenticated encryption is applied, our security analysis demands that it be strong unforgeable.

While this research addresses a specific ISO 9798 protocol, the standard covers several variants for use in authentication. Of these, some highlight particular aspects that would need to be taken under consideration, should similar analyses be performed. Notably, the BR model would require adjustment in the case of one-pass authentication, as this is not addressed in the traditional model. Likewise, some ISO 9798 protocols utilize timestamps instead of nonces for freshness and care would be required in modeling these if an analysis is to be performed in the manner of our work.

References

1. Abdalla, M., Bellare, M., Rogaway, P.: The Oracle Diffie-Hellman Assumptions and an Analysis of DHIES. In: Topics in Cryptology – CT-RSA. Lecture Notes in Computer Science, vol 2020, Springer-Verlag (2001)
2. Basin, D., Cremers, C., Meier, S.: Provably repairing the ISO/IEC 9798 standard for entity authentication. Journal of Computer Security 21(6), 817–846 (2013)
3. Bellare, M., Namprempre, C.: Authenticated Encryption: Relations among notions and analysis of the generic composition paradigm. In: Advances in Cryptology Proc. ASIACRYPT 2000, volume 1976 of Lecture Notes in Computer Science. pp. 531–545 (2000)
4. Bellare, M., Rogaway, P.: Entity Authentication and Key Distribution. In: Advances in Cryptology - Proc. CRYPTO 93, volume 773 of Lecture Notes in Computer Science. pp. 232–249. Springer (1993)
5. Blake-Wilson, S., Johnson, D., Menezes, A.: Key agreement protocols and their security analysis. In: Proceedings of sixth IMA International Conference on Cryptography and Coding, LNCS 1355. pp. 30–45. Springer-Verlag (1997)
6. ISO: Information technology – security techniques – entity authentication – part 2: Mechanisms using symmetric encipherment algorithms. ISO ISO/IEC 9798-2:2008, International Organization for Standardization, Geneva, Switzerland (2008)
7. ISO: Information technology – security techniques – entity authentication – part 2: Mechanisms using symmetric encipherment algorithms. ISO ISO/IEC 9798-2:2008/Cor 1:2010, International Organization for Standardization, Geneva, Switzerland (Technical Corrigendum 1, 2010)
8. Jutla, C.: Encryption Modes with Almost Free Message Integrity. In: Advances in Cryptology – EUROCRYPT 2001. Lecture Notes in Computer Science, vol 2045, Springer-Verlag (2001)
9. LaMacchia, B., Lauter, K., Mityagin, A.: Stronger Security of Authenticated Key Exchange. In: ProvSec 2007. pp. 1–16. LNCS vol. 4784, Springer (2007)
10. R., C., Krawczyk, H.: Analysis of Key–Exchange protocols and Their Use for Buliding Secure Channels. In: Advances in Cryptology – EUROCRYPT 2001. pp. 453–474. Springer-Verlag (2001)
11. Rogaway, P.: Authenticated-Encryption with Associated-Data. In: Ninth ACM Conference on Computer and Communications Security (CCS-9). ACM Press (2002)
12. Rogaway, P., Bellare, M., Black, J.: OCB: A Block-Cipher Mode of Operation for Efficient Authenticated Encryption. In: Eighth ACM Conference on Computer and Communications Security (CCS–8). pp. 365–403. ACM Press (2003), `www.cs.ucdavis.edu/~rogaway`
13. Rogaway, P., Stegers, T.: Authentication without Elision: Partially Specified Protocols, Associated Data, and Cryptographic Models Described by Code. In: Proceedings of the 2009 22nd IEEE Computer Security Foundations Symposium. pp. 26–39. IEEE Computer Society (2009)

A Definitions for security of MAC and authenticated encryption

Definition 4. *([5,1], modified.) A strongly-unforgeable message authentication code is a probabilistic polynomial-time algorithm* $\mathrm{MAC}_{(\cdot)}(\cdot)$. *Let* `Message` =

$\{0,1\}^*$, $\texttt{mKey} = \{0,1\}^k$ *for some number* k, *and* $\texttt{Tag} = \{0,1\}^{tLen}$ *for some number* $tLen$.

A message authentication code is defined by a pair of algorithms $\big(\mathrm{MAC}_{(\cdot)}(\cdot), \mathrm{MAC}.ver_{(\cdot)}(\cdot,\cdot)\big)$. *To compute the MAC,* $\mathrm{MAC}_{(\cdot)}(\cdot)$ *takes a key* $K \in mKey$ *and a message* $M \in \texttt{Message}$ *and computes:*

$$(M, T) = \mathrm{MAC}_K(M).$$

The authenticated message is the pair (M, T); T *is called the tag on* M.

To verify a purported message-tag pair (M, τ), *any entity with key* K *computes*

$$\mathrm{MAC}.ver_K(M, \tau),$$

which returns either 0 *(message unauthentic) or* 1 *(message authentic). It is required for all* $K \in mKey$ *and* $M \in \texttt{Message}$, $\mathrm{MAC}.ver_K(MAC_K(M)) = 1$.

An adversary F *(of the MAC) is a probabilistic polynomial-time algorithm which has access to an oracle that computes MACs under a randomly chosen key* K'. *The output of* F *is a pair* (M, T) *such that* (M, T) *was not previously output by the MACing oracle.*

A MAC is secure if, for every adversary F *of the MAC, the function* $\epsilon(k)$ *defined by*

$$\epsilon(k) = P[K' \leftarrow \{0,1\}^k; (M, T) \leftarrow F : (M, T) = MAC_{K'}(M)]$$

is negligible. Note that F *wins even if it outputs a different tag on a previously queried message.*

Definition 5. *([11], modified.) A strongly-unforgeable authenticated encryption scheme (SUF-AE scheme) is a three-tuple* $(\mathcal{K}, \mathcal{E}, \mathcal{D})$, *with associated set* $\texttt{Message} \subseteq \{0,1\}^*$, *satisfying* $M \in \texttt{Message} \Rightarrow M' \in \texttt{Message}$ *for any* M' *of equal length to* M.

Algorithm $\mathcal{E}$ *is probabilistic, returning a string* $C \stackrel{\$}{=} \mathcal{E}_K(M)$ *on input of a string* $K \in \mathcal{K}$ *and* $M \in \texttt{Message}$.

Algorithm $\mathcal{D}$ *is deterministic, taking in a string* $K \in \mathcal{K}$ *and* $C \in \{0,1\}^*$, *and returning* $\mathcal{D}_K(C)$ *which is either a string in* $\texttt{Message}$ *or a symbol* $\perp$. *Moreover, it is required that* $\mathcal{D}_K(\mathcal{E}_K(M)) = M$ *for all* $K \in \mathcal{K}$ *and* $M \in \texttt{Message}$.

Let $\$(\cdot)$ *be an oracle that, on input of* M, *returns a random string of length* $l(|M|)$ *where* l *is the length function of* $(\mathcal{K}, \mathcal{E}, \mathcal{D})$. *Let* $\mathcal{A}$ *be an adversary. Define* $\mathbf{Adv}^{IND\$\text{-}CPA}_{(\mathcal{K},\mathcal{E},\mathcal{D})}(\mathcal{A}) = P[K \stackrel{\$}{\rightarrow} \mathcal{K} : \mathcal{A}^{\mathcal{E}_K(\cdot)} = 1] - P[\mathcal{A}^{\$(\cdot)} = 1]$. *In this instance, IND$-CPA is the indistinguishability from random bits under a chosen-plaintext attack.*

Let $(\mathcal{K}, \mathcal{E}, \mathcal{D})$ *be a SUF-AE scheme. Choose* $K \stackrel{\$}{\leftarrow} \mathcal{K}$ *and run the adversary* $\mathcal{A}$, *providing it with an oracle for* $\mathcal{E}_K(\cdot)$. *We say that adversary* $\mathcal{A}$ *forges, under*

key K for the particular run, if $\mathcal{A}$ outputs a ciphertext C where $\mathcal{D}_K(C) \neq \perp$ and $\mathcal{A}$ did not ask a query $\mathcal{E}_K(M)$ such that $C = \mathcal{E}_K(M)$. However, $\mathcal{A}$ is allowed to have previously queried $\mathcal{E}_K(M)$, such that $C_1 = \mathcal{E}_K(M)$, as long as $C \neq C_1$. Let $\mathbf{Adv}^{SUF\text{-}AE}_{(\mathcal{K},\mathcal{E},\mathcal{D})}(\mathcal{A})$ be the probability that $\mathcal{A}$ forges against the authentication. The probability is over the random choice of K.

B Security Revisited without Unique Identifier

Per the ISO 9798–2.4 protocol, the unique identifier of the initiator I_B may be excluded if either uni-directional keys are used or the protocol environment precludes reflection attacks. These two cases will be discussed below for the security proofs of §2 and §5.

B.1 Analysis of Core Security Proof Revisited

The security argument for the core proof of §4 remains largely unchanged for an environment where the unique identifier I_B is excluded. In particular, in the case of uni-directional keys, F has two MAC oracles, per definition 4, that run on keys K_i and K_j chosen randomly from $\{0,1\}^k$ which it will use to calculate the tags for messages sent to instances of i and j, respectively. Thus the tag in the case of $\mathcal{A}$'s success against an initiator is calculated as $\mathrm{MAC}_{K_i}(R_j, R_i)$ and the tag in the case against a responder is likewise changed to $\mathrm{MAC}_{K_j}(R_i, R_j)$. Thus, as in the original proof, F must win by producing a forgery

Furthermore, in an environment without reflection attacks, the case of F against an initiator remains unchanged. Against a responder oracle, $\mathcal{A}$ cannot win by reflecting back to $\Pi^t_{j,i}$ the flow $\mathcal{E}_K(R_j, R_i)$. Consequently, F must again win as in §4 by producing a forgery.

B.2 Analysis with Associated Data Revisited

Theorem B1 *Let Π be the ISO 9798–2.4 protocol implemented with a strongly unforgeable MAC algorithm $\mathcal{E}_K(M) = (m, \mathrm{MAC}_K(M))$, including the optional text fields $Text_l$ and the associated data $Text_2$, uni-directional keys in an environment that precludes reflection attacks, and no unique identifier I_B. Then advantage of an polynomial-time adversary against the mutual authentication scheme can be reduced to the adversarial advantage against the MAC:*

$$\mathbf{Adv}^{MA}(\mathcal{A}) \leq 2p^2S \cdot \mathbf{Adv}^{\mathrm{MAC}}_{\Pi}(F) + q^2/2^{k+1}.$$

Proof. Essentially, the theorem follows from the proof in §5 with some alterations as noted below.

Nonce Collision

Proof as shown in §5.

$$\Pr[\mathrm{NC}] \leq q^2/2^{k+1}. \tag{9}$$

Acceptance Implies Matching Conversations

Proof as in §5 with the added notes of §B.1.

$$\Pr[\neg \text{ Match.Conv } | \neg \text{ NC}] \leq 2p^2 S \cdot \mathbf{Adv}_{\Pi}^{\text{MAC}}(F). \tag{10}$$

Associated Data Agreement

Compactly, it can be assumed that an instance and its partner are in matching conversations and that there are no nonce collisions. It remains to show that the same AD, T_2, is equally held by both sides.

Uni-Directional Keys Since i and j are in matching conversation, at some time τ_1 a responder sent the message $T_{3_j} || \mathcal{E}_{K_i}(R_j, R_i, T_{2_j})$ and at some time τ_2 an initiator received a message $T_{3_i} || \mathcal{E}_{K_i}(R_j, R_i, T_{2_i})$, where T_{3_i} and T_{3_j} may or may not be equal. Then the responder must be an instance $\Pi_{j,i}^t$ of j, and the initiator must be an instance $\Pi_{i,j}^s$ of i. Moreover, the data T_2 has been authenticated under the key K_i, so $\Pi_{i,j}^s$ can be assured of the integrity of T_2. Therefore $\text{AD}_i = \text{AD}_j$.

Ergo,

$$\Pr[\neg \text{ Match.AD } | \neg \text{NC} \wedge \text{ Match.Conv}] = 0. \tag{11}$$

Reflection Attacks Disallowed Proof as in §3, with unique identifier I_B removed.

Combining the reductions from equations 9–11 with equation 5 yields the full reduction of security for ISO 9798–2.4 with inclusion of associated data.