Abstract

We introduce a modification of the Russian standardized AEAD MGM mode – an MGM2 mode, for
which a nonce is not encrypted anymore before using it as an initial counter value. For the new mode
we provide security bounds regarding security notions in the nonce-misuse setting (MRAE-integrity
and CPA-resilience). The obtained bounds are even better than the bounds obtained for the original
MGM mode regarding standard security notions.

Keywords: MGM, AEAD mode, security notion, security bounds, nonce-misuse, misuse-resistant

1 Introduction

Authenticated Encryption with Associated Data (AEAD) schemes, which aim at pro-
viding both integrity and confidentiality of data, are recently considered to be among
the most widely used cryptographic schemes. Therefore, the security of such schemes is
crucial. Security analysis of AEAD-schemes is usually carried out in the provable security
paradigm regarding standard notions, introduced in [6], they are IND-CCA and IND-CPA
for confidentiality and INT-CTXT for integrity.

One of the examples of such schemes is an MGM block cipher mode of operation, that
was adopted in Russia as a standard AEAD-mode [13]. The MGM plaintext encryption
procedure is quite similar to encryption in the CTR2 [18] mode. The main element of the
MGM authentication procedure is a multilinear function with secret coefficients produced
in the same way as the secret masking blocks used for plaintext encryption. Integrity and
confidentiality of MGM were analysed in [1] regarding standard security notions. Since
MGM was not developed with provable security in mind, security proofs turned out to be
cumbersome and, hence, difficult to verify.

Even though analysis of AEAD-schemes in the standard models is mandatory and
enough for use in many applications, some environments require other unusual crypto-
graphic properties, e.g. leakage resilience [5], RUP («Release of Unverified Plaintext»)
security [3], KDM («Key Dependent Message») security [10], misuse-resistance [19], etc.
In the current paper we focus on misuse-resistance or nonce misuse property [19]. A
nonce (number used only once) is an input to encryption or decryption algorithms of
AEAD-schemes that has to be unique (within a fixed key), but in some applications such
requirement is hard to obtain, not to mention implementation faults. Misuse-resistant
schemes aim to ensure the best possible security when faulty nonce is provided.

Security notions for misuse-resistant authenticated encryption were originally pro-
posed by Rogaway and Shrimpton in [19] and further developed in [4]. Strong variant of
misuse-resistant notions, called MRAE («Misuse-Resistant AE»), was introduced in [19].
This notion is the extensions of the IND-CCA and INT-CTXT notions by allowing an
adversary to repeat nonces in all of its’ queries. The MRAE notion is similar to a DAE notion [19] («Deterministic AE») where confidentiality is formalised as follows: ciphertext of each new query (not only new nonce) has to be indistinguishable from a random string. Providing such confidentiality is rather strong, and trying to achieve it seems to lead to loss in performance. All MRAE-secure modes, known to the authors, demand sufficiently larger amount of block cipher calls [16] or lose online property [15, 22]. For the reasons above, weak notions for confidentiality called CPA-res and CCA-res («Chosen Plaintext/Ciphertext Attack-resilience») were introduced in [4], where the confidentiality should be achieved only for messages that were encrypted correctly using unique nonces.

In nonce-misuse setting the MGM mode is obviously insecure in the MRAE model regarding confidentiality since counter-based encryption is actually used. MRAE-integrity of the MGM mode was analysed in [17]: the birthday type attack was proposed. However, no lower bounds for MGM were proven. So, there is a «hope» to provide non-trivial security bounds for MGM in the MRAE-integrity and CCA-res models.

Motivating by expectation that the security proof for the MGM mode in non-standard models will be even more complex, than in standard ones, we introduce modification of MGM mode – MGM2. The main difference between two modes lies in the way how secret masking blocks and secret coefficients of the multilinear function are produced – for the MGM2 mode this process is carried out in the CTR [18] style (without preliminary nonce encryption). Note, that the main cryptographic core of the construction, namely multilinear function, is not changed. We provide the security bounds for MGM2 in the MRAE-integrity and CPA-res models that turned out to be even better, than the bounds for the original MGM mode in the standard models. The corresponding security proofs are relatively short and, we hope, easier to verify. Among other advantages, the design of the MGM2 mode also allows to transparently integrate internal re-keying without a master key [2] (in the same way as for CTR-ACPKM [14] done) to achieve new security properties like leakage-resilience and increase key lifetime.

2 Preliminaries

By \{0,1\}^s we denote the set of s-component bit strings and by \{0,1\}^* we denote the set of all bit strings of finite length including the empty string. Let \(|a|\) be the bit length of the string \(a \in \{0,1\}^*\). For a bit string \(a\) we denote by \(|a|_n = \lceil |a|/n \rceil\) the length of the string \(a\) in \(n\)-bit blocks. By \(\{0,1\}^s\) we denote the set of bit strings which length is less or equal to \(s\).

For a string \(a \in \{0,1\}^*\) and a positive integer \(l \leq |a|\) let msb\(_l\)(\(a\)) be the string, consisting of the leftmost \(l\) bits of \(a\). For nonnegative integers \(l\) and \(i\) let str\(_l\)(\(i\)) be \(l\)-bit representation of \(i\) with the least significant bit on the right. For bit strings \(a \in \{0,1\}^n\) and \(b \in \{0,1\}^n\) we denote by \(a \otimes b\) a string which is the result of their multiplication in \(GF(2^n)\) (here strings encode polynomials in the standard way). If the value \(s\) is chosen from a set \(S\) uniformly at random, then we denote \(s \leftarrow^U S\). We define a function \(\text{Set}_{l,r} : \{0,1\}^n \rightarrow \{0,1\}^n\), \(\text{Set}_{l,r}(x) = x \oplus (\underbrace{0 \ldots 0}_{r \text{-component bit strings}} 1 \ldots 0), \ 0 \leq r < n\).

For any set \(S\), define \(\text{Perm}(S)\) as the set of all bijective mappings on \(S\) (permutations on \(S\)), and \(\text{Func}(S)\) as the set of all mappings from \(S\) to \(S\). A block cipher \(E\) (or just a cipher) with a block size \(n\) and a key size \(k\) is the permutation family \(\{E_K \in \text{Perm}(\{0,1\}^n) \mid K \in \{0,1\}^k\}\), where \(K\) is a key.
3 Security models

This section introduces models for an adversary that may repeat nonces in its queries.

We define security model using the notion of «experiment» or «game» played between a challenger and an adversary. The adversary and challenger are modelled using consistent interactive probabilistic algorithms. The challenger simulates the functioning of the analysed cryptographic scheme for the adversary and may provide him access to one or more oracles (for details see [8]).

We describe challengers and adversaries using pseudocodes with the following notations. If a variable \( x \) gets a value \( val \) then we denote \( x \leftarrow val \). Similarly, if a variable \( x \) gets the value of a variable \( y \) then we denote \( x \leftarrow y \). If the variable \( x \) gets the result of a probabilistic algorithm \( A \) we denote \( x \leftarrow A \). If we need to emphasize that \( A \) is deterministic than we denote it by \( x \leftarrow ^{d} A \). The event when \( A \) returned value \( val \) as a result is denoted by \( A \rightarrow val \) (or \( A \leftarrow A \rightarrow val \) if \( A \) is probabilistic).

Firstly, we introduce the general definition of an AEAD-scheme.

**Definition 1.** Let \( K \) be a set of keys, \( P \) be a set of plaintexts, \( A \) be a set of associated data, \( C \) be a set of ciphertexts, and \( T \) be a set of tags. An AEAD-scheme with nonce is a set of algorithms \( \Pi = \{ \Pi.\text{Gen}, \Pi.\text{Enc}, \Pi.\text{Dec} \} \), where

- \( \Pi.\text{Gen}( ) \rightarrow K \): A probabilistic key generation algorithm outputting a key \( K \in K \).
- \( \Pi.\text{Enc}(K, N, A, P) \rightarrow (C, T) \): A deterministic algorithm of authenticated encryption taking a key \( K \in K \), a nonce \( N \in N \), associated data \( A \in A \), a plaintext \( P \in P \). An output of the algorithm is a ciphertext \( C \in C \) and a tag \( T \in T \).
- \( \Pi.\text{Dec}(K, N, A, C, T) \rightarrow P \): A deterministic algorithm of authenticated decryption taking a key \( K \in K \), a nonce \( N \in N \), associated data \( A \in A \), a ciphertext \( C \in C \) and a tag \( T \in T \). An output of the algorithm is a plaintext \( P \in P \) or error symbol \( \perp \).

Let define a MRAE-int («Misuse-Resistant Authenticated Encryption - integrity») security notion for integrity (the integrity part of MRAE [19]).

**Definition 2 (MRAE-int).** For an AEAD-scheme \( \Pi \) the advantage of a MRAE-int-adversary \( A \) is defined as follows:

\[
\text{Adv}_{\Pi}^{\text{MRAE-int}}(A) = \Pr[\exp_{\Pi}^{\text{MRAE-int}}(A) \rightarrow 1],
\]

where experiment \( \exp_{\Pi}^{\text{MRAE-int}} \) is defined below:

\[
\begin{array}{llll}
\text{Exp}_{\Pi}^{\text{MRAE-int}}(A) & \text{Oracle Encrypt}(N, A, P) & \text{Oracle Decrypt}(N, A, C, T) \\
K \leftarrow \Pi.\text{Gen}( ) & (C, T) \leftarrow \Pi.\text{Enc}(K, N, A, P) & P \leftarrow \Pi.\text{Dec}(K, N, A, C, T) \\
\text{sent} \leftarrow \emptyset & \text{sent} \leftarrow \text{sent} \cup \{(N, A, C, T)\} & \text{if } (P \neq \perp) \land ((N, A, C, T) \notin \text{sent}) : \\
\text{win} \leftarrow \text{false} & \text{return } (C, T) & \text{win} \leftarrow \text{true} \\
\text{return win} & & \text{return } P
\end{array}
\]

Let introduce the CPA-res («Chosen Plaintext Attack - resilience») security notion for confidentiality, defined in [4].

**Definition 3 (CPA-res).** For an AEAD-scheme \( \Pi \) with the tag length \( s \) the advantage of a CPA-res-adversary \( A \) is defined as follows:

\[
\text{Adv}_{\Pi}^{\text{CPA-res}}(A) = \Pr[\exp_{\Pi}^{\text{CPA-res-1}}(A) \rightarrow 1] - \Pr[\exp_{\Pi}^{\text{CPA-res-0}}(A) \rightarrow 1],
\]

where experiments \( \exp_{\Pi}^{\text{CPA-res-b}}, b \in \{0, 1\} \), are defined below:
In [4] the CCA-res («Chosen Ciphertext Attack - resilience») security notion is also defined. This notion differs from CPA-res in that an adversary is provided with additional access to a decryption oracle. By the technique similar to one described in [20] it is easy to show that MRAE-int-security and CPA-res-security jointly imply CCA-res-security. Therefore, further we consider the CPA-res security notion only.

4 MGM2 mode

\[
\begin{array}{ll}
\text{MGM2.Gen()} & \text{MGM2.Dec}(K, N, A, C, T) \\
K \leftarrow \mathcal{U}(\{0,1\}^k) & h \leftarrow \{A\}_n, q \leftarrow \{C\}_n \\
\text{return } K & \ell \leftarrow h + q + 1 \\
\end{array}
\]

\[
\begin{array}{ll}
\text{MGM2.Enc}(K, N, A, P) & \text{Padding} \\
h \leftarrow \{A\}_n, q \leftarrow \{P\}_n & a \leftarrow n\{A\}_n - \|A\| \\
\ell \leftarrow h + q + 1 & c \leftarrow n\{C\}_n - \|C\| \\
\text{Encryption} & \text{len} \leftarrow \text{str}_{n/2}(\|A\|) \text{||} \text{str}_{n/2}(\|C\|) \\
\text{for } i = 1 \ldots q \text{ do:} & M_i \leftarrow \text{str}_{n/2}(\|A\|) \text{||} \text{str}_{n/2}(\|C\|) \\
\Gamma_i \leftarrow E_K(N||00||\text{str}_{n/2}(i-1)) & \text{len} \leftarrow M_1 \ldots || M_q \leftarrow A||0^n||C||0^n\text{||len} \\
C \leftarrow P \oplus \text{msb}_q(\Gamma_1 \ldots \Gamma_q) & \text{Tag verification} \\
\text{Padding} & \text{for } i = 1 \ldots \ell \text{ do:} \\
\text{for } i = 1 \ldots q \text{ do:} & H_i \leftarrow E_K(N||01||\text{str}_{n/2}(i-1)) \\
\tau \leftarrow \text{Set1}_r \left( \bigoplus_{i=1}^{\ell} M_i \oplus H_i \right) & \tau \leftarrow \text{Set1}_r \left( \bigoplus_{i=1}^{\ell} M_i \oplus H_i \right) \\
T \leftarrow \text{msb}_q(E_K(\tau)) & T' \leftarrow \text{msb}_q(E_K(\tau)) \\
\text{return } (C, T) & \text{if } T' \neq T: \text{ return } \bot \\
\end{array}
\]

Figure 1: AEAD mode MGM2

In this section we describe a new AEAD mode called MGM2 which is a slight modification of the MGM mode. By MGM2[\text{E, r, s}] we will denote the parametrized MGM2 mode
with a block cipher $E$ (with block size $n$ and key size $k$), a nonce length $r$, $\frac{n}{2} \leq r \leq \frac{3n}{4}$ and a tag length $s$, $1 \leq s \leq n$.

For $\text{MGM2}[E, r, s]$ the corresponding sets are as follows: $K = \{0,1\}^k$, $N = \{0,1\}^r$, $A = P = C = \{0,1\}^{\leq n(2^{n-r-2} - 1)}$, $T = \{0,1\}^s$. Moreover, the following condition should be satisfied: $0 < |A| + |P| \leq n(2^{n-r-2} - 1)$. The key generation, encryption and decryption algorithms are defined in Figure 1.

**Difference from MGM.** The main difference of the new MGM2 mode from the original MGM mode is in the modification of the way to produce the mask values for encryption ($\Gamma_i$), the coefficients of the multilinear function ($H_i$), and the tag values $T$. In MGM2 block cipher inputs, used to generate values for different use cases (we have three use cases: $\Gamma_i, H_i, T$), are separated by fixing the certain bits of inputs. Such a modification allows to obtain better security bounds, since the collision among block cipher inputs may occur only among values $\tau$.

## 5 Security analysis

The security of block cipher modes of operation is commonly analyzed under assumption that the underlying block cipher is PRP-CPA-secure (see [7]), i.e. $E_K$ for a random key is computationally indistinguishable from a random permutation. We follow this approach and provide security bounds directly for the mode with a random permutation.

We write $\text{MGM2}[\text{Perm}(n), r, s]$ for MGM2 that uses a random permutation $\pi$ as $E_K$ and we write $\text{MGM2}[\text{Func}(n), r, s]$ for MGM2 that uses a random function $\rho$.

### 5.1 Integrity

**Theorem 1.** For any MRAE-int-adversary $A$, making at most $Q_E$ queries to the Encrypt oracle and at most $Q_D$ queries to the Decrypt oracle, where the total block-length of associated data in all queries is at most $\sigma_A$ and the total block-length of plaintexts and ciphertexts in all queries is at most $\sigma_P$,

$$\text{Adv}_{\text{MGM2}[\text{Perm}(n), r, s]}^{\text{MRAE-int}}(A) \leq \left( \frac{Q(Q - 1)}{2^n} + \frac{Q_D}{2^s} \right) \left( 1 - \frac{\sigma - 1}{2^n} \right)^{-\sigma/2},$$

where $Q = Q_E + Q_D$ and $\sigma = 2\sigma_P + \sigma_A + 2Q$.

Note that for $n \geq 128$ and $\sigma \leq 2^{n/2}$, the bound (1) can be converted as follows:

$$\text{Adv}_{\text{MGM2}[\text{Perm}(n), r, s]}^{\text{MRAE-int}}(A) \leq 1.7 \left( \frac{Q(Q - 1)}{2^n} + \frac{Q_D}{2^s} \right).$$

Thus, the MGM2 mode provides integrity beyond birthday bound. Note that for the original MGM mode, if total amount of processed data achieves $2^{n/2}$, the bound, presented in [1], becomes trivial. This result also allows to use MGM2 as a MAC function (and even as a PRF, see further) by fixing $N$ with the constant value. Further we provide proof of the Theorem 1.

**Proof.** The proof is carried out in two steps. In the first step we introduce an auxiliary MAC-scheme with nonce called $\text{MGM2-MAC}[r,s]$ and estimate its security (see Section 5.1.1).

In the second step we show that the UF-CMA-security of the $\text{MGM2-MAC}[r,s]$ scheme tightly implies the MRAE-int-security of the $\text{MGM2}[\text{Func}(n), r, s]$ scheme (see Section 5.1.2).
Theorem 2.3. Due to that theorem for any distinguisher $D^*$ with oracle $f: \{0,1\}^n \to \{0,1\}^n$, making at most $q$ queries, the following inequality holds:

$$\Pr[D^* \to 1] \leq \Pr[D^\rho \to 1] \cdot \left(1 - \frac{q - 1}{2^n}\right)^{-q/2},$$

where $\pi \leftarrow \text{Perm}(n)$ and $\rho \leftarrow \text{Func}(n)$.

If we let $D$ be the algorithm $\text{Exp}_\text{MGM2[\cdot, \cdot, \cdot]}^{\text{MAE-int}}(A)$, where it makes queries to the oracle instead of calling the underlying function (block cipher), we obtain the target bound. \hfill $\square$

5.1.1 Security of MGM2-MAC

We introduce an auxiliary MAC-scheme with nonce called MGM2-MAC[$r, s$] based on the scheme MGM2[Func($n$), $r$, $s$]. Usually MAC-scheme is defined as a set of algorithms $\text{MAC} = \{\text{MAC.Gen}, \text{MAC.Tag}, \text{MAC.Verify}\}$, for MGM2-MAC[$r, s$] these algorithms are defined in Figure 2. This scheme is defined for the message set $\{M = M_1 || \ldots || M_\ell: M_i \in \{0,1\}^n, M_\ell \neq 0^n, 1 \leq \ell \leq 2^{n-r-2}\}$ (the message length is divisible by $n$, the last block is non-zero).

Figure 2: The scheme MGM2-MAC

Firstly, we introduce the standard PRF security notion (in nonce-misuse setting) for nonce-based MAC-schemes and obtain the PRF-security bound for the MGM2-MAC scheme.

Definition 4 (PRF). For a MAC-scheme MAC the advantage of a PRF-adversary $A$ is defined as follows:

$$\text{Adv}_{\text{MAC}}^{\text{PRF}}(A) = \Pr[\text{Exp}_\text{MAC}^{\text{PRF}-1}(A) \to 1] - \Pr[\text{Exp}_\text{MAC}^{\text{PRF}-0}(A) \to 1],$$

where experiments $\text{Exp}_{\text{MAC}}^{\text{PRF}-b}(A)$, $b \in \{0,1\}$ are defined below:

<table>
<thead>
<tr>
<th>$\text{Exp}_{\text{MAC}}^{\text{PRF}-b}(A)$</th>
<th>$\text{Oracle Tag}^1(N, M)$</th>
<th>$\text{Oracle Tag}^0(N, M)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>if $b = 1$: $K \leftarrow \text{MAC.Gen}()$</td>
<td>if $(N, M) \in \text{sent}$: return $\bot$</td>
<td>if $(N, M) \in \text{sent}$: return $\bot$</td>
</tr>
<tr>
<td>sent $\leftarrow \emptyset$</td>
<td>$T \leftarrow \text{MAC.Tag}(K, N, M)$</td>
<td>sent $\leftarrow \text{sent} \cup {(N, M)}$</td>
</tr>
<tr>
<td>$b' \leftarrow A^{\text{Tag}^b}(\ )$</td>
<td>$\text{sent} \leftarrow \text{sent} \cup {(N, M)}$</td>
<td>return $T$</td>
</tr>
<tr>
<td>return $b'$</td>
<td>return $T$</td>
<td></td>
</tr>
</tbody>
</table>
Lemma 1. For any PRF-adversary $\mathcal{A}$, making at most $Q$ queries to the Tag oracle:

$$\text{Adv}_{\text{MGM2-MAC}[r,s]}^{\text{PRF}}(\mathcal{A}) \leq \frac{Q(Q-1)}{2^n}.$$  

Proof. Let define auxiliary experiments $\text{Exp}^0$ and $\text{Exp}^1$ (see Figure 3), which differ from the experiment $\text{Exp}^{\text{PRF-1}}_{\text{MGM2-MAC}[r,s]}$ as follows. During the setup phase, the $\tau$ set is additionally initialized to empty value, and the flag $\text{bad}$ is set to false. During the experiment execution, the values $\tau$ are put in the $\tau$ set, and the flag $\text{bad}$ is set to true iff collision among the $\tau$ values occurs. Also, in the $\text{Exp}^0$ experiment the tag value is chosen from $\{0,1\}^s$ uniformly at random in the case of collision (see line in box).

<table>
<thead>
<tr>
<th>$\text{Exp}^0(A)$, $b \in {0,1}$</th>
<th>Oracle $Tag^b(N,M)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(\rho, \rho') \leftarrow \text{MGM2-MAC.Gen}(\cdot)$</td>
<td>if $(N,M) \in \text{sent}$:</td>
</tr>
<tr>
<td>$\text{bad} \leftarrow \text{false}$</td>
<td>return $\bot$</td>
</tr>
<tr>
<td>$\tau, \text{sent} \leftarrow \emptyset$</td>
<td>$\tau \leftarrow \text{PreTag}(\rho', N, M)$</td>
</tr>
<tr>
<td>$b' \leftarrow \mathcal{A}^{\tau=b'}(\cdot)$</td>
<td>$T \leftarrow \text{msb}_b(\rho(\tau))$</td>
</tr>
<tr>
<td>$\text{return } b'$</td>
<td>if $\tau \in \tau$:</td>
</tr>
<tr>
<td></td>
<td>$\text{bad} \leftarrow \text{true}$</td>
</tr>
<tr>
<td></td>
<td>if $b = 0$: $T \leftarrow {0,1}$</td>
</tr>
<tr>
<td></td>
<td>$\tau \leftarrow \tau \cup {\tau}$</td>
</tr>
<tr>
<td></td>
<td>$\text{sent} \leftarrow \text{sent} \cup {(N,M)}$</td>
</tr>
<tr>
<td></td>
<td>$\text{return } T$</td>
</tr>
</tbody>
</table>

Figure 3: Experiments $\text{Exp}^0$ and $\text{Exp}^1$

It is easy to see that $\text{Exp}^1$ is exactly the $\text{Exp}^{\text{PRF-1}}_{\text{MGM2-MAC}[r,s]}$ experiment. Moreover, for any $\mathcal{A}$ the value $\Pr[\text{Exp}^0(\mathcal{A}) \Rightarrow 1]$ is exactly the value $\Pr[\text{Exp}^{\text{PRF-0}}_{\text{MGM2-MAC}[r,s]}(\mathcal{A}) \Rightarrow 1]$. Indeed, in the $\text{Exp}^0$ experiment all tag values $T$ are produced according to the uniform distribution as in $\text{Exp}^{\text{PRF-0}}_{\text{MGM2-MAC}[r,s]}$ for the following reasons. For the queries, whose corresponding $\tau$ value is new (not in the current set $\tau$), the uniform random function $\rho$ is applied to the new input and, therefore, returns uniform output. For the other queries the $T$ value is directly sampled uniformly at random (see the line in box, Figure 3). Therefore,

$$\text{Adv}_{\text{MGM2-MAC}[r,s]}^{\text{PRF}}(\mathcal{A}) = \Pr[\text{Exp}^1(\mathcal{A}) \Rightarrow 1] - \Pr[\text{Exp}^0(\mathcal{A}) \Rightarrow 1].$$

Note that before the $\text{bad}$ flag is set to true (denote this event as $\text{bad} = \text{true}$) the $\text{Exp}^0$ and $\text{Exp}^1$ experiments are functioning identically, therefore (due to Lemma 2, [8]) the following inequality holds:

$$\Pr[\text{Exp}^1(\mathcal{A}) \Rightarrow 1] - \Pr[\text{Exp}^0(\mathcal{A}) \Rightarrow 1] \leq \Pr[\text{bad} = \text{true}].$$

Let estimate $\Pr[\text{bad} = \text{true}]$. Without loss of generality, we assume the adversary to be deterministic and making $Q$ pairwise different queries $(N_i, M'_i), \ i = 1, \ldots, Q$. Denote by $\bar{\rho}$ and $\bar{\rho}'$ the uniform random variables with sample space $\text{Func}(n)$. We will also use notation $\text{coll}^i, \ i = 2, \ldots, Q$, to denote the event that the $\text{bad}$ flag is set to true during the first $i$ queries processing. Thus,

$$\Pr[\text{bad} = \text{true}] = \sum_{i=2}^{Q} \Pr[\text{coll}^i \cap \overline{\text{coll}^{i-1}}],$$

7
where the probability is defined by the random variables \( \tilde{\rho} \) and \( \tilde{\rho}' \). Let estimate the value 
\[
\Pr[\text{coll}^i \cap \text{coll}^i'^{-1}] \quad \text{for any } i = 2, \ldots, Q.
\]

Note, that each \( i \)-th query – the pair \((N_i, M^i)\), where \( M^i = M^i_1 || \ldots || M^i_n \), \( M^i_j \in \{0, 1\}^n \) – is determined by the tag values \( T_1, \ldots, T_{i-1} \) previously obtained from the oracle. Without loss of generality, we assume \( l_1 = \ldots = l_i \). Indeed, otherwise we can pad the messages with zero blocks to the length \( l := \max(l_1, \ldots, l_i) \). This does not change the tag value, and the padded messages will stay pairwise different because of \( M^i_j \neq 0^n \). Therefore, the \( T_1, \ldots, T_{i-1} \) values fully determine \( l \) and \((N_1, M^1), \ldots, (N_i, M^i)\).

For fixed \( N_j \) we denote by \( H^i_k \), \( j = 1, \ldots, i; \ k = 1, \ldots, l \), the random variable 
\[
\tilde{\rho}'(N_j\|01||\text{str}_{n\rightarrow r}(k-1)).
\]
Notice that \( \Pr[H^i_k = B] = \frac{1}{2^n} \) for any \( B \in \{0, 1\}^n \). Note that the random variables \( \tilde{H}^j_k \) and \( \tilde{H}^j_k \) for some \( j \neq t \) and any \( k \) are dependent, namely 
\[
\Pr[\tilde{H}^j_k = \tilde{H}^j_k] = 1, \text{ iff } N_k = N_j.
\]

For short we denote by \( \tilde{H}^j \) the random variable \((\tilde{H}^j_1, \ldots, \tilde{H}^j_l)\). Also for set \( H = (H_1, \ldots, H_l) \) and message \( M = M_1 || \ldots || M_l \) let \( \tau(H, M) \) be the function 
\[
\text{Set}_1, \left( \bigoplus_{k=1}^l H_k \otimes M_k \right).
\]
So, we have 
\[
\Pr[\text{coll}^i \cap \text{coll}^i'^{-1}] = \sum_{T_1, \ldots, T_{i-1}} \Pr[\text{coll}^i \cap \text{coll}^i'^{-1} \cap \{\tilde{T}_j = T_j\}_{j=1}^{i-1}],
\]
where we write \( \tilde{T}_j \) for random variable \( \text{msb}_s(\tilde{\rho}(\tau(H^j, M^j))) \), and sum is taken over all \((T_1, \ldots, T_{i-1}) \in \{(0, 1)^s)^{i-1}\).

For fixed \((N_1, M^1), \ldots, (N_i, M^i)\) introduce the following conditions on set \( H^1, \ldots, H^i \), 
\( H^j := (H^j_1, \ldots, H^j_l), j = 1, \ldots, i \):

Condition \( E_1 \): \( \forall j, t, 1 \leq j < t \leq i - 1 : \tau(H^j, M^j) \neq \tau(H^t, M^t) \).

Condition \( E_2 \): \( \exists j, 1 \leq j \leq i - 1 : \tau(H^j, M^j) = \tau(H^j, M^j) \).

For any fixed \( T_1, \ldots, T_{i-1} \), and hence fixed \((N_1, M^1), \ldots, (N_i, M^i)\), the event \( \text{coll}^i \cap \text{coll}^i'^{-1} \) occurs iff random variables \( \tilde{H}^1, \ldots, \tilde{H}^i \) take such values \( H^1, \ldots, H^i \) that the conditions \( E_1 \) and \( E_2 \) are satisfied. For short we will denote the events that these conditions are satisfied by the same way, namely, by \( E_1 \) and \( E_2 \) correspondingly.

Note that fixing values \( H^j, j = 1, \ldots, i \), leads to fixing values \( \tau_j := \tau(H^j, M^j) \). Therefore, 
\[
\Pr[\text{coll}^i \cap \text{coll}^i'^{-1}] = \sum_{T_1, \ldots, T_{i-1}} \Pr[E_1 \cap E_2 \cap \{\tilde{T}_j = T_j\}_{j=1}^{i-1}
\]
\[
= \sum_{T_1, \ldots, T_{i-1}} \sum_{H^1, \ldots, H^i : E_1 \cap E_2} \Pr[\{\tilde{H}^j = H^j\}_{j=1}^{i-1} \cap \{\text{msb}_s(\tilde{\rho}(\tau_j)) = T_j\}_{j=1}^{i-1}]
\]
\[
= \sum_{T_1, \ldots, T_{i-1}} \sum_{H^1, \ldots, H^i : E_1 \cap E_2} \Pr[\{\tilde{H}^j = H^j\}_{j=1}^{i-1}] \cdot \Pr[\{\text{msb}_s(\tilde{\rho}(\tau_j)) = T_j\}_{j=1}^{i-1}].
\]

Here, sum is taken over all \( H^1, \ldots, H^i \), \( H^j \in \{(0, 1)^n)^i\), for which the \( E_1 \) and \( E_2 \) conditions are satisfied. The last transition is due to the fact that \( \tilde{\rho} \) and \( \tilde{H}^j \), \( j = 1, \ldots, i \), are independent.
Consider the value \( \Pr\{\{\text{msb}(\rho(T_j)) = T_j\}_{j=1}^{i-1}\} \). For any \( T_1, \ldots, T_{i-1} \) and \( H^1, \ldots, H^{i-1} \) for which the condition \( E_1 \) is satisfied, this probability is exactly the probability to sample function \( \rho \), such that \( i - 1 \) fixed inputs correspond to outputs with fixed \( s \) bits, i.e. \( \frac{1}{2^{s(i-1)}} \).

Thus:

\[
\Pr[\text{coll}^i \cap \text{coll}^{i-1}] = \sum_{T_1, \ldots, T_{i-1}} \sum_{H^1, \ldots, H^{i-1}} \Pr[\{\widetilde{H}^j = H^j\}_{j=1}^{i-1}] \cdot \frac{1}{2^{s(i-1)}} = \\
= \frac{1}{2^{s(i-1)}} \sum_{T_1, \ldots, T_{i-1}} \Pr[E_1 \cap E_2] \leq \frac{1}{2^{s(i-1)}} \sum_{T_1, \ldots, T_{i-1}} \Pr[E_2].
\]

Now consider \( \Pr[E_2] \) for any fixed \( T_1, \ldots, T_{i-1} \), and, hence, any fixed \( (N_1, M^1), \ldots, (N_i, M^i) \).

\[
\Pr[E_2] = \Pr[\exists j, 1 \leq j \leq i - 1: \tau(\widetilde{H}^i, M^i) = \tau(\widetilde{H}^j, M^j)] = \\
= \Pr\left[\bigcup_{j=1}^{i-1} \left\{ \tau(\widetilde{H}^i, M^i) = \tau(\widetilde{H}^j, M^j) \right\} \right] \leq \sum_{j=1}^{i-1} \Pr[\tau(\widetilde{H}^i, M^i) = \tau(\widetilde{H}^j, M^j)].
\]

Let estimate \( p := \Pr[\tau(\widetilde{H}^i, M^i) = \tau(\widetilde{H}^j, M^j)] \) for any \( j = 1, \ldots, i - 1 \). We consider two cases:

1. \( N_i \neq N_j \) (in this case \( \widetilde{H}^i_k \) and \( \widetilde{H}^j_k \) are independent).
2. \( N_i = N_j \) (in this case \( \widetilde{H}^i_k \) and \( \widetilde{H}^j_k \) are dependent).

**The first case:** \( p = \frac{\#\{H^i, H^j: \tau(H^i, M^i) = \tau(H^j, M^j)\}}{2^{2ni}} \).

\[
\#\{H^i, H^j: \tau(H^i, M^i) = \tau(H^j, M^j)\} = \\
= \#\left\{ H^i, H^j: \bigoplus_{k=1}^l H^i_k \otimes M^i_k = \bigoplus_{k=1}^l H^j_k \otimes M^j_k \right\} + \\
+ \#\left\{ H^i, H^j: \bigoplus_{k=1}^l H^i_k \otimes M^i_k = \bigoplus_{k=1}^l H^j_k \otimes M^j_k \oplus \text{Set}_r(0^n) \right\}.
\]

Since \( M^i_k \neq 0^n \) for any \( i \), the cardinality is \( 2 \cdot 2^{n(2i-1)} \). And, \( p = \frac{2}{2^n} \).

**The second case:** \( p = \frac{\#\{H^i: \tau(H^i, M^i) = \tau(H^i, M^i)\}}{2^{2ni}} \).

\[
\#\{H^i: \tau(H^i, M^i) = \tau(H^i, M^i)\} = \\
= \#\left\{ H^i: \bigoplus_{k=1}^l H^i_k \otimes (M^i_k \oplus M^j_k) = 0^n \right\} + \\
+ \#\left\{ H^i: \bigoplus_{k=1}^l H^i_k \otimes (M^i_k \oplus M^j_k) = \text{Set}_r(0^n) \right\}.
\]

Since for the same nonce the messages \( M^i \) and \( M^j \) should be different, there exists \( k \) such that \( M^i_k \oplus M^j_k \neq 0^n \). Therefore, the cardinality is \( 2 \cdot 2^{n(i-1)} \). And, \( p = \frac{2}{2^n} \).
Summing up, we have:
\[ \Pr[\text{bad} = \text{true}] = \sum_{i=2}^{Q} \frac{1}{2^{s(i-1)}} \sum_{T_1, \ldots, T_n} \sum_{j=1}^{i-1} \frac{2^{s}}{2^{n}} = \sum_{i=2}^{Q} \frac{i-1}{2^{n-1}} = \frac{Q(Q-1)}{2^{n}}. \]

\[ \square \]

Now we introduce the standard UF-CMA security notion for nonce-based MAC-schemes and obtain the UF-CMA-security bound for the MGM2-MAC scheme.

**Definition 5.** For a MAC-scheme MAC the advantage of a UF-CMA-adversary \( \mathcal{A} \) is defined as follows:
\[
\text{Adv}_{\text{MAC}}^{\text{UF-CMA}}(\mathcal{A}) = \Pr[\text{Exp}_{\text{MAC}}^{\text{UF-CMA}}(\mathcal{A}) \rightarrow 1],
\]
where experiment \( \text{Exp}_{\text{MAC}}^{\text{UF-CMA}}(\mathcal{A}) \) is defined below:

\[
\text{Exp}_{\text{MAC}}^{\text{UF-CMA}}(\mathcal{A}) \leftarrow \text{MAC.Gen}(\cdot) \\
\quad \text{sent} \leftarrow \emptyset \\
\quad \text{win} \leftarrow \text{false} \\
\quad \mathcal{A}^{\text{Tag,Verify}}(\cdot) \\
\quad \text{return } \text{win}
\]

<table>
<thead>
<tr>
<th>( K \leftarrow \text{MAC.Gen}(\cdot) )</th>
<th>( \text{Oracle Tag}(N, M) )</th>
<th>( \text{Oracle Verify}(N, M, T) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>\text{if } (N, M) \in \text{sent}:</td>
<td>\text{return } \bot \text{.}</td>
<td>\text{res} \leftarrow \text{MAC.Vf}(K, N, M, T)</td>
</tr>
<tr>
<td>\text{sent} \leftarrow \text{sent} \cup {(N, M)}</td>
<td></td>
<td>\text{if } \text{res} \land ((N, M) \notin \text{sent}):</td>
</tr>
<tr>
<td></td>
<td>\text{return } T \text{.}</td>
<td>\text{win} \leftarrow \text{true}</td>
</tr>
</tbody>
</table>

Using Proposition 7.3 [7] and Lemma 1 we obtain the following result.

**Corollary 1.** For any UF-CMA-adversary \( \mathcal{A} \), making at most \( Q_T \) queries to the Tag oracle and at most \( Q_V \) queries to the Verify oracle:
\[
\text{Adv}_{\text{MGM2-MAC}[r,s]}^{\text{UF-CMA}}(\mathcal{A}) \leq \frac{Q(Q-1)}{2^n} + \frac{Q_V}{2^s},
\]
where \( Q = Q_T + Q_V \).

### 5.1.2 Security of MGM2 with random function

**Lemma 2.** For any MRAE-int-adversary \( \mathcal{B} \), making at most \( Q_E \) queries to the Encrypt oracle and at most \( Q_D \) queries to the Decrypt oracle, there exists a UF-CMA-adversary \( \mathcal{A} \), making at most \( Q_E \) queries to the Tag oracle and at most \( Q_D \) queries to the Verify oracle, such that
\[
\text{Adv}_{\text{MGM2}[\text{Func}(n), r, s]}^{\text{MRAE-int}}(\mathcal{A}) \leq \text{Adv}_{\text{MGM2-MAC}[r,s]}^{\text{UF-CMA}}(\mathcal{B}).
\]

**Proof.** Let construct an adversary \( \mathcal{B} \), that uses the adversary \( \mathcal{A} \) as a black box. The adversary \( \mathcal{B} \) (see Figure 4) intercepts the queries of the adversary \( \mathcal{A} \) and process them by itself using its own oracles. For encryption/decryption \( \mathcal{B} \) implements lazy sampling for \( \rho'' \). For tag generation/tag verification the adversary \( \mathcal{B} \) implements the padding procedure and send the appropriate query to its oracles.

Note that the adversary \( \mathcal{B} \) simulates for \( \mathcal{A} \) exactly the experiment \( \text{Exp}_{\text{MGM2}[\text{Func}(n), r, s]}^{\text{MRAE-int}} \). Indeed, since for \( \text{MGM2}[\text{Func}(n), r, s] \) the inputs to the random function in case of 1) tag generation, 2) computing values \( H_i \) and 3) computing values \( \Gamma_i \) are different (because of fixed bits in inputs), using one random function is indistinguishable from using three independent random functions \( \rho, \rho', \rho'' \) for these three cases. Also, note that messages \( M \), formed by \( \mathcal{B} \), satisfy conditions for message set of \( \text{MGM2-MAC}[r, s] \).

If the adversary \( \mathcal{A} \) forges, then the adversary \( \mathcal{B} \) also forges in \( \text{Exp}_{\text{MGM2-MAC}[r,s]}^{\text{UF-CMA}} \). Indeed, if \( \mathcal{A} \) makes non-trivial valid query \( (N, A, C, T) \) to the Decrypt oracle, then the adversary makes \( \mathcal{B} \) corresponding non-trivial query \( (N, M = A||0^s||C||0^s||len, T) \) to the Verify oracle.

\[ \square \]
\[\mathcal{B}_A^{Tag, Verify}\]
\[\rho'' \xleftarrow{\$} Func(n) \text{ } / \text{ lazy sampling}\]
\[\text{return } A^{SEncrypt, SDecrypt}\]

\[SEncrypt(N, A, P)\]
\[h \leftarrow |A|_n, q \leftarrow |P|_n\]
\[\text{ ............ Encryption ............} \]
\[\text{for } i = 1 \ldots q \text{ do:}\]
\[\Gamma_i \leftarrow \rho''(N||00||\text{str}_{n/2} - 2(i - 1))\]
\[C \leftarrow P \oplus \text{msb}_P(\Gamma_1 \ldots \Gamma_q)\]
\[\text{ ............ Padding ............} \]
\[a \leftarrow n|A|_n - |A|\]
\[c \leftarrow n|C|_n - |C|\]
\[\text{len} \leftarrow \text{str}_{n/2}(|A|) \parallel \text{str}_{n/2}(|C|)\]
\[M \leftarrow A||0^n||C||0^n||\text{len}\]
\[\text{ ............ Tag Genetation ............} \]
\[T \leftarrow \text{Tag}(N, M)\]
\[\text{return } (C, T)\]

Oracle \(SDecrypt(N, A, C, T)\)
\[h \leftarrow |A|_n, q \leftarrow |C|_n\]
\[\text{ ............ Padding ............} \]
\[a \leftarrow n|A|_n - |A|\]
\[c \leftarrow n|C|_n - |C|\]
\[\text{len} \leftarrow \text{str}_{n/2}(|A|) \parallel \text{str}_{n/2}(|C|)\]
\[M \leftarrow A||0^n||C||0^n||\text{len}\]
\[\text{ ............ Tag Verification ............} \]
\[\text{if } Verify(N, M, T) = \text{false}:\]
\[\text{return } \bot\]
\[\text{ ............ Decryption ............} \]
\[\text{for } i = 1 \ldots q \text{ do:}\]
\[\Gamma_i \leftarrow \rho''(N||00||\text{str}_{n/2} - 2(i - 1))\]
\[P \leftarrow C \oplus \text{msb}_C(\Gamma_1 \ldots \Gamma_q)\]
\[\text{return } P\]

Figure 4: Adversary \(\mathcal{B}\)

### 5.2 Confidentiality

**Theorem 2.** For any CPA-res-adversary \(A\), making at most \(Q_1\) queries to the \(O_1\) oracle and at most \(Q_2\) queries to the \(O_2\) oracle, where the total block-length of associated data in all queries is at most \(\sigma_A\) and the total block-length of plaintext and ciphertexts in all queries is at most \(\sigma_P\),

\[
\text{Adv}^{\text{CPA-res}}_{\text{MGM2}[\text{Perm}(n), r, s]}(\mathcal{A}) \leq \frac{\sigma^2}{2^n} + \frac{Q(Q - 1)}{2^{n-1}},
\]

where \(Q = Q_1 + Q_2\) and \(\sigma = 2\sigma_P + \sigma_A + 2Q\).

**Proof.** Firstly, we apply PRP-PRF switching lemma [12] to replace \(\text{Perm}(n)\) by \(Func(n)\) (this gives us the term \(\frac{\sigma^2}{2^n}\) in the bound), and then we obtain the CPA-res-security bound for \(\text{MGM2}[\text{Func}(n), r, s]\).

The security bound for \(\text{MGM2}[\text{Func}(n), r, s]\) is obtained in the same way as in the proof of Theorem 1. Indeed, ciphertexts \(C\), received from the \(O_1\) oracle, are absolutely indistinguishable from uniform random strings since the inputs to the uniform random function \(\rho\) used to produce \(\Gamma_i\) are unique. The indistinguishability of the tags \(T\), received from the \(O_1\) oracle, from uniform random strings is estimated by constructing two PRF-adversaries for \(\text{MGM2-MAC}\) that uses CPA-res-adversary as a black box. Therefore, \(\text{Adv}^{\text{CPA-res}}_{\text{MGM2}[\text{Func}(n), r, s]}(\mathcal{A}) \leq \frac{Q(Q - 1)}{2^{n-1}}\).

\[\square\]

### 6 Conclusion

In the current paper we introduce the modification of the MGM mode — the MGM2 mode. For this mode we obtain the security bounds for non-standard notions MRAE-int
and CPA-res, allowing the adversary to repeat nonces. In comparison with the original mode, the security proof appears to be rather simple and short.

In the future work we are going to develop a SIV-construction (see [15]) of the MGM2 mode to achieve MRAE-conf-security. Also we are going to incorporate re-keying mechanisms in the MGM2 mode to achieve new security properties like leakage-resilience and increase key lifetime.

References


