# Lifting the Security of NI-MAC Beyond Birthday Bound 

Avijit Dutta and Goutam Paul<br>Cryptology and Security Research Unit (CSRU), R. C. Bose Centre for Cryptology \& Security, Indian Statistical Institute, Kolkata 700 108, India. avirocks.dutta13@gmail.com, goutam.paul@isical.ac.in


#### Abstract

In CRYPTO 1999, J. An and M. Bellare proposed a MerkleDamgård iteration based MAC construction called NI-MAC in order to avoid constant re-keying on multiblock messages in NMAC and to ease the security proof. In CRYPTO 2014, Gazi et al. revisited the proof of NI-MAC in the view of structure graph introduced by Bellare et al. in CRYPTO 2005 and gave a tight bound of order $\frac{l q^{2}}{2^{n}}$, which is an improvement over the trivial bound of order $\frac{l^{2} q^{2}}{2^{n}}$, for $q$ queries, each of length at most $\ell$ blocks. But this is again restricted to the birthday security. In order to prove the security of NI-MAC, Gazi et al. (CRYPTO 2014) introduced a variant of NI-MAC, called NI2-MAC and analyzed the advantage of NI2 MAC. Then he showed that the same proof technique will be applied to the security analysis of NI-MAC. In this paper, we lift the birthday bound of NI2-MAC construction beyond birthday $O\left(q^{2} l^{4} / 2^{2 n}\right)$ by a small change in the existing construction with one extra invocation of a independent keyed function. Finally, we argue how to lift the security of NI-MAC beyond birthday using the security proof for NI2-MAC.


Keywords: Beyond Birthday, MAC, NI, NI2, Structure-Graph.

## 1 Introduction

In symmetric key paradigm, MAC (Message Authentication Code) is used for preserving message integrity and message origin authentication. The design of a MAC should not only consider achieving security, but also target attaining efficiency. In the literature, three different approaches of designing a MAC exists: (a) universal hash function based MAC, a popular example of which is UMAC [6], (b) a compression function based MAC, like NMAC [2], HMAC [2], NI [1] etc.
(c) Block cipher based MAC, such as CBC MAC [4], PMAC [7], OMAC [12]. etc.

Most of the popular MACs are block cipher based MACs, but each one of them suffers from the same problem - security is guaranteed up to the birthday bound. When the block length of the underlying block cipher is 128-bit, then
birthday bound does not seem to be a problem, as we are guaranteed to have 64 bits of security which is well acceptable for many practical applications. But when we deal with 64 -bit block cipher as used in many light weight crypto devices, then birthday bound problem becomes the main bottleneck. Throughout the paper, we use $n$ to denote the block-length and $q$ to denote the number of MAC-queries by the adversary.
Related Work on Beyond birthday Secure MAC. In recent researches, many MAC constructions have been proposed with security beyond the birthday barrier without degrading the performance. The first attempt was made in ISO 9797-1 [3] without security proof. But Algorithm 4 of ISO 9797-1 was attacked by Joux et al. [15] that falsified the security bound. Algorithm 6 of ISO 97971 was proven to be secure against $O\left(2^{2 n / 3}\right)$ queries with restrictions on the message length [20]. In [20] Yasuda proved that the sum of two independent ECBC has beyond birthday bound. However, it requires four keys and it is rate $1 / 2$ construction as it requires two block cipher calls for processing each message block. In 2011, he proposed PMAC_Plus Construction [21] that achieves beyond birthday security. In 2012, Zhang et al. [22] proposed a 3key version of f9 MAC that achieves BBB.

There is also another deterministic MAC mode provides security beyond the birthday bound. Given an $n$-bit to $n$-bit fixed-key blockcipher with MAC security $\epsilon$ against $q$ queries, Dodis et al. [9] have designed a variable-length MAC achieving $O(\epsilon q \operatorname{poly}(n))$ MAC security. However, this design requires even longer keys and more block cipher invocations. By parity method, Bellare et al. present MACRX [3] with BBB security, conditioned on the input parameters are random and distinct. In [13], Jaulmes et al. proposed a randomized MAC that provides BBB security based on the ideal model (or possibly based on tweakable block cipher). Another BBB secure randomized construction called generic enhanced hash then MAC has been proposed in [18] by Minematsu. Recently Datta et al. in [8] unify PMAC_Plus and 3kf9 in one key setting with beyond birthday security.

In CRYPTO 1999, J. An and M. Bellare [1] proposed a Merkle-Damgård iteration based MAC construction called NI-MAC. The construction of NI-MAC is similar to that of NMAC [2], the only difference is that in NI-MAC the compression function $f$ takes an additional input key $k$ at each invocation. The motivation of designing NI was to avoid constant re-keying on multi-block messages in NMAC and to allow for a security proof starting by the standard switch from a PRF to a random function, followed by information-theoretic analysis.

In CRYPTO 2014, Gazi et al. [10] revisited the proof of NI-MAC in the view of structure graph introduced by Bellare et al. in CRYPTO 2005 [5] and gave a tight bound of order $\frac{l q^{2}}{2^{n}}$, which is an improvement over trivial bound of order $\frac{l^{2} q^{2}}{2^{n}}$, for $q$ queries, each of length at most $\ell$ blocks. But this is again restricted to the birthday security. In order to prove the security of NI-MAC, Gazi et al. [10] introduced a variant of NI-MAC, called NI2-MAC, and then derived the security of NI-MAC from the security analysis of NI2-MAC. In this paper, we propose an extension of NI2-MAC with a single invocation of an additional
pseudo-random function and prove (Section 4) that it achieves beyond-birthday security. Furthermore, we make a remark at the end that if we extend the NI MAC in the same way as we did for NI2 MAC, then also we achieve beyond birthday bound security.
Organization: Section 2 revisits the definition of prf, mac, structure graph. Section 3 contains the construction of $\mathrm{NI}^{+}$. Security analysis of NI2 ${ }^{+}$is shown from Section 4 to Section 6. We conclude the paper in Section 7.

## 2 Preliminaries

In this section, we briefly discuss the notations and definitions used in this paper. We also state some existing basic results.

### 2.1 PRF and Secure MAC

We denote $|S|$ as the cardinality of set $S$ and $S^{c}$ as the complement set of $S$. Let $x \stackrel{\$}{\leftarrow} S$ denote that $x$ is chosen uniformly at random from $S$. Let $\operatorname{Func}(A, B)$ denote the set of all functions from $A$ to $B$ A function $\rho: A \rightarrow B$ is said to be a random function, if $\rho$ is chosen uniformly at random from the $F u n c(A, B)$.

We will specify a random function by performing lazy sampling. In lazy sampling initially the function $\rho$ is undefined at every point of its domain. We maintain two sets that grows dynamically. One is domain, $\operatorname{Dom}(\rho)$ and another is Range, $\operatorname{Ran}(\rho)$, both initialized to be empty. $\operatorname{Dom}(\rho), \operatorname{Ran}(\rho)$ keeps the record of already defined domain points and range points of function $\rho$ respectively. Therefore, if $x \notin \operatorname{Dom}(\rho)$ then we will choose $y \stackrel{\$}{\leftarrow} B \backslash \operatorname{Ran}(\pi)$ and add $y$ in $\operatorname{Ran}(\rho)$ and $x$ in $\operatorname{Dom}(\rho)$. In this regard, $x$ is said to be fresh.

We consider that an adversary $\mathcal{A}$ is an oracle machine with access to its oracle $\mathcal{O}(\cdot)$ and outputs either 1 or 0 . Accordingly, we write $\mathcal{A}^{\mathcal{O}(\cdot)}=1$ or 0 . The resource of $\mathcal{A}$ is measured in terms of the time complexity $T(n)$ that it takes to interacts with its oracle $\mathcal{O}(\cdot)$ and the query complexity $q(n)$ which says the number of queries and replies exchanged between the adversary and its oracle. For practical purpose, we restrict to probabilistic polynomial time (PPT) adversaries only.

The PRF-advantage of a function $F_{k}: A \rightarrow B$ is defined as

$$
\operatorname{Adv}_{\mathbf{F}_{\mathbf{k}}}^{\left.\mathbf{P R F}^{(\mathcal{A}}\right)}\left(\mathcal{P r}\left[\mathcal{A}^{F_{k}(\cdot)}=1: k \stackrel{\$}{\leftarrow} \mathcal{K}\right]-\operatorname{Pr}\left[\mathcal{A}^{f(\cdot)}=1: f \stackrel{\$}{\leftarrow} \operatorname{Func}(A, B)\right]\right.
$$

If this advantage is negligible in the length of the input for all PPT adversaries, $F$ is said to be a secure PRF. Note that the first probability is calculated over
 probability is calculated over the randomness of $f \stackrel{\$}{\leftarrow} \operatorname{Func}(A, B)$.

The length of $M$ in bits is denoted by $\operatorname{len}(M)$. When it is not a multiple of $n$, we append $10^{n-1-l e n(M)} \bmod n$ to $M$ to make $\operatorname{len}(M)$ a multiple of $n$. We denote the maximum number of block in a query by $l$. We denote the partition
of a message $M$ as $M=M_{1}\left\|M_{2}\right\| \ldots \| M_{l}$ where each $M_{i}$ is an $n$-bit block and the number of blocks of $M$ is denoted by $l$.

An adversary attacking a MAC with $q$ queries obtains $q$ tags for $q$ distinct messages and produces a valid tag of a fresh message that he has not queried earlier. It is known [11] that any secure PRF is a secure MAC. Thus, to show that a MAC construction is secure, one needs to show that the PRF-advantage (which is a function of $q, l$ and $n$ ) of an adversary for the construction is negligible.

### 2.2 Structure Graphs

In this section, we briefly revisit the structure graph analysis of CBC-MAC [5] by Bellare et al. and that of NI-MAC [10] by Gazi et al.

Consider an iterated/cascaded construction with a function $f$, where $f$ could be a random permutation or a random function, that works on a message $M=$ $M_{1}\left\|M_{2}\right\| \ldots \| M_{l}$ of length $l$ blocks as follows:

$$
Y_{0}=\mathbf{0}, \text { and } Y_{i}=f\left(Y_{i-1}, M_{i}\right) \text { for } i=1, \ldots, l
$$

Note that for CBC-MAC analysis, $f(\alpha, \beta)$ is taken as $\pi(\alpha \oplus \beta)$ and for the NIMAC analysis, $f(\alpha, \beta)$ is taken as $\rho(\alpha \| \beta)$, where $\pi$ is a random permutation over $n$ bits and $\rho$ is a random function from $b+n$ bits to $n$ bits, where $b$ is the message block-length and $n$ is the length of the chaining variable as well as the tag.

For a set of any two fixed distinct messages $\mathcal{M}=\left\{M^{(1)}, M^{(2)}\right\}$ and a function $f$, we construct the structure graph $\mathcal{G}^{f}(\mathcal{M})$ with $\{0,1\}^{n}$ as the set of nodes as follows. We follow the computations for $M^{(1)}$ followed by those of $M^{(2)}$ by creating nodes labelled by the values $y_{i}$ of the intermediate chaining variables $Y_{i}$ with the edge $\left(Y_{i}, Y_{i+1}\right)$ labelled by the block $M_{i+1}$. In this process, if we arrive at a vertex already labelled, while not following an existing edge, we call this event an $f$-collision. An accident is an $f$-collision that does not close a cycle with alternating edge-directions such that the XOR of the labels of the cycle becomes 0.

More formally, let for two distinct messages $M^{(1)}$ and $M^{(2)}$ of $l_{1}$ and $l_{2}$ blocks respectively, where

$$
M^{(1)}=M_{1}^{(1)}\left\|M_{2}^{(1)}\right\| \ldots \| M_{l_{1}}^{(1)} \text { and } M^{(2)}=M_{1}^{(2)}\left\|M_{2}^{(2)}\right\| \ldots \| M_{l_{2}}^{(2)},
$$

the corresponding $Y$-values be given by

$$
Y_{0}^{(1)}, Y_{1}^{(1)}, Y_{2}^{(1)}, \ldots, Y_{l_{1}}^{(1)} \text { and } Y_{0}^{(2)}, Y_{1}^{(2)}, Y_{2}^{(2)}, \ldots, Y_{l_{2}}^{(2)}
$$

respectively. Let $\sigma=l_{1}+l_{2}$. We use the notation $M_{i}$ to refer to the block $M_{i}^{(1)}$, when $i<l_{1}$, otherwise to refer to the block $M_{i-l_{1}}^{(2)}$. Similarly, let $Y_{i}$ to refer to $\mathbf{0}$ when $i=0 ; Y_{i}^{(1)}$, when $1 \leq i \leq l_{1}$; and $Y_{i-l_{1}}^{(2)}$, when $l_{1}+1 \leq i \leq \sigma$. Now, consider the mappings

$$
[[\cdot]] \text { and }\left[[\cdot]^{\prime} \text { on }\{0, \ldots, \sigma\}\right.
$$

so that $[[i]]=\min \left\{j: Y_{i}=Y_{j}\right\}$ and $\left[\left[i^{\prime}\right]\right]=[[i]]$ for $i \neq l_{1}$ except that $\left[\left[l_{1}\right]\right]^{\prime}=0$.
For any fixed $f$ and any two distinct messages $\mathcal{M}=\left\{M^{(1)}, M^{(2)}\right\}$, we define the structure graph $\mathcal{G}^{f}(\mathcal{M})$ to be the triple $\mathcal{G}^{f}(\mathcal{M})=(V, E, L)$, where

$$
V=\{[[i]]: 0 \leq i \leq \sigma\}, \quad E=\left\{\left([[i-1]]^{\prime},[[i]]\right): 1 \leq i \leq \sigma\right\}
$$

and $L=E \rightarrow\{0,1\}^{n}$ is an edge-labeling function defined as

$$
L((u, v))=\left\{M_{i}:[[i-1]]^{\prime}=u \text { and }[[i]]=v\right\}
$$

Let $\left(V_{i}, E_{i}, L_{i}\right)$ be the graph obtained after processing only the first $i$ out of $\sigma$ blocks of $\mathcal{M}$. We say that $(i,[[i]])$ is an $f$-collision if $[[i]]<i$ and $M_{i} \notin$ $L_{i-1}\left([[i-1]]^{\prime},[[i]]\right)$. Note that the last condition on $M_{i}$ implies that collision occurred due to parallel edges with the same message label is not considered.

In [5], a general collision is called a true collision (except the collision that occurs due to parallel edges with same label on the edges). Further, a true collision is called an accident if it is not followed from a cycle $C$ with alternating edges with the sum of the labels of the edges involved in $C$ to $\mathbf{0}$, otherwise it is called an induced collision. However, for NI2-MAC, all $f$-collisions are accidents. In our work, we need to consider the accidents in $\mathcal{G}^{f}(\mathcal{M})$. Let $\mathcal{G}(\mathcal{M})$ denote the set of all structure graphs corresponding to the set of messages $\mathcal{M}$ (by varying $f$ over a function family). For a fixed $\operatorname{graph} G$, let $\operatorname{Acc}(G)$ denote the set of all accidents in $G$. We state the following known results.

Proposition 1. [10, Lemma 2] For a fixed graph $G, \operatorname{Pr}_{f}\left[\mathcal{G}^{f}(\mathcal{M})=G\right] \leq$ $2^{-n|A c c(G)|}$.

Proposition 2. [5, Lemma 7] $\operatorname{Pr}[G \stackrel{\$}{\leftarrow} \mathcal{G}(\mathcal{M}):|A c c(G)| \geq 2] \leq \frac{8 l^{4}}{2^{2 n}}$.

## 3 Proposed Construction of NI2 ${ }^{+}$for Beyond-Birthday Secure MAC

We present the schematic diagram of $\mathrm{NI}^{+}$in Fig. 3.1 followed by the description in Algorithm 1. $\quad f_{K_{1}}, f_{K_{2}}$ and $f_{K_{3}}$ are three independently chosen keyed


Fig. 3.1: Construction of NI2 ${ }^{+}$MAC

```
Input: \(f_{K_{1}}, f_{K_{2}}, f_{K_{3}}: K_{1}, K_{2}, K_{3} \stackrel{\&}{\leftarrow} \mathcal{K}, M \leftarrow\{0,1\}^{*}\)
Output: \(T \in\{0,1\}^{n}\)
\(M_{1}\left\|M_{2}\right\| \ldots M_{l} \leftarrow M \| 10^{*} ; / / l\) is the number of message blocks in \(M\)
\(Z \leftarrow 0^{n}\);
\(Y \leftarrow 0^{n}\);
    for \(i=1\) to \(l\) do
    \(Y \leftarrow f_{K_{1}}\left(M_{i}, Y\right) ;\)
    \(Z \leftarrow 2 \cdot(Z \oplus Y) ;\)
end
\(\theta \leftarrow Z ;\)
\(\Sigma \leftarrow f_{K_{2}}\left(0^{b}, Y\right) ;\)
\(T \leftarrow f_{K_{3}}(\Sigma, \Theta)\);
Return \(T\);
```

Algorithm 1: Algorithm for $\mathrm{NI}^{+}$MAC
functions such that $f_{K_{1}}, f_{K_{2}}:\{0,1\}^{n+b} \rightarrow\{0,1\}^{n}$ and $f_{K_{3}}:\{0,1\}^{2 n} \rightarrow\{0,1\}^{n}$. We denote

$$
\operatorname{Casc}^{f_{K_{1}}}(M):=f_{K_{1}}\left(\ldots\left(f_{K_{1}}\left(f_{K_{1}}\left(f_{K_{1}}\left(0, M_{1}\right), M_{2}\right), M_{3}\right), \ldots,\right) M_{l}\right)
$$

to be the output of the last message block in the upper lane of the construction depicted in Fig.3.1.

For any message $M \in\{0,1\}^{*}$, NI2 ${ }^{+}$MAC (after suitably padding with $10^{*}$ if the message length is not a multiple of the block length $b$ ) partitions $M$ into $l$ many blocks each of which is $b$ bits long. Then the blocks are iteratively processed as depicted in Fig.3.1. Final output $Y_{l}$ of $\operatorname{Casc}^{f_{K_{1}}}(M)$ as depicted in Fig.3.1 and $0^{b}$ becomes the input of $f_{K_{2}}(\cdot, \cdot)$ and the output of $f_{K_{2}}(\cdot, \cdot)$ is denoted as $\Sigma$. This is the so-called NI2 construction which we extend as follows. A linear combination of the intermediate chaining value of $\operatorname{Casc}^{f_{K_{t}}}(M)$ is denoted as $\Theta$. The symbol ' 2 ' in the construction is the root of an irreducible polynomial of degree $n . \Sigma$ and $\Theta$ are then fed into $f_{K_{3}}(\cdot, \cdot)$ and the output is returned as tag $T$.

Remark 1 NI-MAC, as originally proposed by An and Bellare in [1] replaces the 0 block at the input of $f_{K_{2}}$ with the bit length $|M|$ of the message $M$. We extend NI-MAC in the same way as we do for NI2-MAC and obtain NI ${ }^{+}$-MAC.

Note: In subsequent sections all the security proofs are done for NI2 ${ }^{+}$MAC.

## 4 Security Analysis of $\mathrm{NI}^{+}-\mathrm{MAC}$

Gazi et. al in [10] have shown that the advantage of distinguishing the output of NI-MAC from random output is bounded above by $\frac{q^{2}}{2^{n}}\left(l+\frac{64 l^{4}}{2^{n}}\right)$ and that for NI2-MAC is $\frac{q^{2}}{2^{n}}\left(l d^{\prime}(l)+\frac{64 l^{4}}{2^{n}}\right)$ where $d^{\prime}(l)=\max _{l^{\prime} \in\{1, \ldots, l\}}\left|\left\{d \in \mathbb{N}: d \mid l^{\prime}\right\}\right|$. In
this section we analyze the advantage of our construction $\mathrm{NI}^{+}{ }^{+}$-MAC and show that the advantage of our construction achieves beyond birthday bound security; better than that of NI-MAC or NI2-MAC. Thus we have the following theorem.

Theorem 1. Let $f:\{0,1\}^{n} \times\{0,1\}^{n} \times\{0,1\}^{b} \rightarrow\{0,1\}^{n}$ be a $\left(\epsilon_{1}, t, q\right)$ secure PRF and $\left(\epsilon_{2}, t, l q\right)$ secure PRF. Let $h:\{0,1\}^{n} \times\{0,1\}^{n} \times\{0,1\}^{n} \rightarrow\{0,1\}^{n}$ be a $\left(\epsilon_{3}, t, q\right)$ secure $P R F$. Then NI2 $^{+}$be a $\left(\epsilon^{\prime}, t^{\prime}, q, l\right)$ secure $P R F$, where

$$
\epsilon^{\prime} \leq \epsilon_{1}+\epsilon_{2}+\epsilon_{3}+\frac{11 q^{2} l^{4}}{2^{2 n}}
$$

such that $t=t^{\prime}+\tilde{O}(l q)$.
Proof. We give the sketch of the proof of Theorem 1 below. Let $\mathcal{A}$ be a adaptive PRF-adversary against $\mathrm{NI}^{+}$running in time $t$ and asking at most $q$ queries, each of length at most $l$ blocks. NI2 ${ }^{+}$uses three independent keyed functions $f_{1}, f_{2}$ and $h_{3}$. Now if we replace $f_{1}, f_{2}$ and $h_{3}$ by three different random functions $r_{1}, r_{2}$ and $r_{3}$ respectively such that $r_{1}, r_{2} \stackrel{\$}{\leftarrow} \operatorname{Func}\left(\{0,1\}^{n} \times\{0,1\}^{n} \times\{0,1\}^{b},\{0,1\}^{n}\right)$ and $r_{3} \stackrel{\$}{\leftarrow} \operatorname{Func}\left(\{0,1\}^{n} \times\{0,1\}^{n} \times\{0,1\}^{n},\{0,1\}^{n}\right)$ and call the resulting construction $\mathrm{NI} 2_{r}^{+}$, then we have

$$
\Delta^{\mathcal{A}}\left(\mathrm{NI}^{+}, R\right) \leq \epsilon_{1}+\epsilon_{2}+\epsilon_{3}+\Delta^{\mathcal{A}}\left(\mathrm{NI}_{r}^{+}, R\right)
$$

where $\epsilon_{i}$ is the PRF-advantage of $f_{i}, i=1,2$ and $\epsilon_{3}$ is the PRF-advantage of $h_{3}$ and $R:\{0,1\}^{*} \rightarrow\{0,1\}^{n}$ be a uniform random function.

Therefore to prove Theorem 1, we only need to prove

$$
\Delta^{\mathcal{A}}\left(\mathrm{NI} 2_{r}^{+}, R\right) \leq \frac{11 q^{2} l^{4}}{2^{2 n}}
$$

In the experiment where $\mathcal{A}$ interacts with $\mathrm{NI}_{r}^{+}$, let $C_{i}$ denotes the event that during the first $i$ queries, the inputs to $r_{3}$, i.e., $(\Sigma, \Theta)$ for any two distinct queries $M^{(j)}$ and $M^{(k)}$ are also distinct. That means $\left(\Sigma^{(j)}, \Theta^{(j)}\right) \neq\left(\Sigma^{(k)}, \Theta^{(k)}\right), \quad \forall 1 \leq$ $j, k \leq i$. Therefore, as long as the monotone condition [17] $C=C_{0}, C_{1}, \ldots$ remains satisfied, the distribution of the responses of $\mathrm{NI} 2_{r}^{+}$to distinct queries will be exactly identical to the distribution of the outputs of $r_{3}$ on distinct inputs and thus to independent uniform random values. In other words, we have

$$
\mathrm{NI} 2_{r}^{+} \mid C \equiv R
$$

Thus, using Lemma 1 in [10] we have, $\Delta^{\mathcal{A}}\left(\mathrm{NI}_{r}^{+}, R\right)$ is upper-bounded by the probability that a distinguisher $\mathcal{A}$ issuing $q$ queries to $\mathrm{NI}_{r}^{+}$makes the monotone condition $C$ fail. This probability is denoted by $\operatorname{Pr}_{\mathcal{A}}\left[\mathrm{NI}_{r}^{+} ; C\right]$. Thus,

$$
\begin{equation*}
\Delta^{\mathcal{A}}\left(\mathrm{NI}_{r}^{+}, R\right) \leq \operatorname{Pr}_{\mathcal{A}}\left[\mathrm{NI} 2_{r}^{+} ; \bar{C}\right] \tag{1}
\end{equation*}
$$

Now we explain how to construct a non-adaptive PRF adversary $\mathcal{A}_{n a}$ from the above adaptive PRF adversary $\mathcal{A}$.

Construction of Non-adaptive PRF Adversary. Let $\mathcal{A}_{n a}$ be the non adaptive PRF adversary that we want to construct from the adaptive PRF adversary $\mathcal{A}$. $\mathcal{A}_{n a}$ will simulate the adaptive PRF adversary $\mathcal{A}$ in the following way. At the time of $i^{t h}$ query, $M^{(i)}$, where $1 \leq i \leq q$, asked by adversary $\mathcal{A}, \mathcal{A}_{n a}$ will return random string in response of $i^{t h}$ query to $\mathcal{A}$. After all the $q$ queries are over, $\mathcal{A}_{n a}$ will (non-adaptively) ask all the queries that $\mathcal{A}$ asked during simulated interaction.

Therefore, we have the following

$$
\begin{equation*}
\operatorname{Pr}_{\mathcal{A}}\left[\mathrm{NI} 2_{r}^{+} ; \bar{C}\right]=\operatorname{Pr}_{\mathcal{A}_{n a}}\left[\mathrm{NI} 2_{r}^{+} ; \bar{C}\right] . \tag{2}
\end{equation*}
$$

The maximum probability over all such non-adaptive distinguishers $\mathcal{A}_{n a}$ is given by

$$
\begin{equation*}
\operatorname{Pr}\left[\mathrm{NI}_{r}^{+} ; \bar{C}\right]=\max _{\mathcal{A}_{n a}} \operatorname{Pr}_{\mathcal{A}_{n a}}\left[\mathrm{NI} 2_{r}^{+} ; \bar{C}\right] \tag{3}
\end{equation*}
$$

With respect to the $\mathrm{NI} 2_{r}^{+}$construction, let $\operatorname{Coll}(l)$ denotes the probability that for random choice of the compression function $f_{1}$ and $f_{2}$, results in a collision in $\Sigma$ and $\Theta$ maximized over the choice of two distinct inputs $M^{(i)}, M^{(j)}$, each of which is at most $l$ blocks long.

More formally, for $f_{1}, f_{2} \stackrel{\$}{\leftarrow} \operatorname{Func}\left(\{0,1\}^{n} \times\{0,1\}^{n} \times\{0,1\}^{b} \rightarrow\{0,1\}^{n}\right)$ we define,

$$
\operatorname{Coll}(l):=\max _{M^{(i)} \neq M^{(j)}\left\|M^{(i)}\right\|,\left\|M^{(j)}\right\| \leq l} \operatorname{Pr}^{f_{1}, f_{2}}\left[\left(\Sigma^{(i)}, \Theta^{(i)}\right)=\left(\Sigma^{(j)}, \Theta^{(j)}\right)\right]
$$

Note that, $\left(\Sigma^{(i)}, \Theta^{(i)}\right)=\left(\Sigma^{(j)}, \Theta^{(j)}\right)$ implies $\Sigma^{(i)}=\Sigma^{(j)}$ and $\Theta^{(i)}=\Theta^{(j)}$. Therefore, to bound the probability of occurrence of a collision in the input of $r_{3}$ necessarily implies to bound the probability of occurrence of a collision in $\Sigma$ and a collision in $\Theta$. That means

$$
\begin{equation*}
\operatorname{Pr}^{f_{1}, f_{2}}\left[\left(\Sigma^{(i)}, \Theta^{(i)}\right)=\left(\Sigma^{(j)}, \Theta^{(j)}\right)\right]=\operatorname{Pr}^{f_{1}, f_{2}}\left[\Sigma^{(i)}=\Sigma^{(j)} \wedge \Theta^{(i)}=\Theta^{(j)}\right] \tag{4}
\end{equation*}
$$

Note that, $\mathcal{A}_{n a}$ violates the monotone condition $C$ only when the collision occurs at the input of $r_{3}$. Therefore from Equation (1), (2) and (3), and using union bound we obtain,

$$
\begin{equation*}
\Delta^{\mathcal{A}}\left(\mathrm{NI}_{r}^{+}, R\right) \leq \operatorname{Pr}\left[\mathrm{NI}_{r}^{+} ; \bar{C}\right] \leq \frac{q^{2}}{2} \operatorname{Coll}(l) . \tag{5}
\end{equation*}
$$

In Lemma 1 of Section 4.1, we show that $\operatorname{Coll}(l) \leq \frac{22 l^{4}}{2^{2 n}}$. Therefore, plugging in the bound of $\operatorname{Coll}(l)$ into Equation (5), we get the result.

### 4.1 Computation of $\operatorname{Coll}(l)$

Recall that, $\operatorname{Coll}(l)$ was defined as $\operatorname{Pr}\left[\Sigma^{(i)}=\Sigma^{(j)} \wedge \Theta^{(i)}=\Theta^{(j)}\right]$ maximized over the choice of pair of distinct inputs $M^{(i)}$ and $M^{(j)}$, each of length at most $l$ blocks. Therefore, to establish the bound on $\operatorname{Coll}(l)$, we derive the bound on $\operatorname{Pr}\left[\Sigma^{(i)}=\Sigma^{(j)} \wedge \Theta^{(i)}=\Theta^{(j)}\right]$

Lemma 1. Given two fixed distinct messages $M^{(i)}, M^{(j)}$, each of length is at most l blocks, we have

$$
\operatorname{Pr}\left[\Sigma^{(i)}=\Sigma^{(j)} \wedge \Theta^{(i)}=\Theta^{(j)}\right] \leq \frac{22 l^{4}}{2^{2 n}}
$$

Proof. Let $Z^{(i)}=Y_{l_{i}}^{(i)}$ denote the input to the function $r_{2}$ for message $M^{(i)}$ (refer to Fig.3.1). Similarly, we set $Z^{(j)}=Y_{l_{j}}^{(j)}$. So, we have,

$$
\begin{align*}
& \operatorname{Pr}\left[\Sigma^{(i)}=\Sigma^{(j)} \wedge \Theta^{(i)}=\Theta^{(j)}\right] \\
= & \operatorname{Pr}\left[\Sigma^{(i)}=\Sigma^{(j)} \wedge \Theta^{(i)}=\Theta^{(j)} \wedge Z^{(i)}=Z^{(j)}\right]+ \\
& \operatorname{Pr}\left[\Sigma^{(i)}=\Sigma^{(j)} \wedge \Theta^{(i)}=\Theta^{(j)} \wedge Z^{(i)} \neq Z^{(j)}\right]  \tag{6}\\
\leq & \operatorname{Pr}\left[Z^{(i)}=Z^{(j)} \wedge \Theta^{(i)}=\Theta^{(j)}\right]+ \\
& \operatorname{Pr}\left[\Sigma^{(i)}=\Sigma^{(j)} \wedge \Theta^{(i)}=\Theta^{(j)} \wedge Z^{(i)} \neq Z^{(j)}\right]  \tag{7}\\
\leq & \operatorname{Pr}\left[Z^{(i)}=Z^{(j)} \wedge \Theta^{(i)}=\Theta^{(j)}\right]+ \\
& \left(\sum_{k=0}^{1} \operatorname{Pr}\left[\Sigma^{(i)}=\Sigma^{(j)} \wedge \Theta^{(i)}=\Theta^{(j)} \wedge N C O L=k \mid Z^{(i)} \neq Z^{(j)}\right] \cdot \operatorname{Pr}\left[Z^{(i)} \neq Z^{(j)}\right]\right) \\
& +\operatorname{Pr}[N C O L \geq 2]  \tag{8}\\
\leq & \operatorname{Pr}\left[Z^{(i)}=Z^{(j)} \wedge \Theta^{(i)}=\Theta^{(j)}\right]+ \\
& \sum_{k=0}^{1} \operatorname{Pr}\left[\Sigma^{(i)}=\Sigma^{(j)} \wedge \Theta^{(i)}=\Theta^{(j)} \wedge N C O L=k \mid Z^{(i)} \neq Z^{(j)}\right]+\operatorname{Pr}[N C O L \geq 2]
\end{align*}
$$

Since the event $Z^{(i)}=Z^{(j)}$ is a subset of the event $\Sigma^{(i)}=\Sigma^{(j)}$, the first term of Equation (6) is equal to $\operatorname{Pr}\left[Z^{(i)}=Z^{(j)} \wedge \Theta^{(i)}=\Theta^{(j)}\right]$.

According to Claim 1, we have $\operatorname{Pr}\left[Z^{(i)}=Z^{(j)} \wedge \Theta^{(i)}=\Theta^{(j)}\right] \leq \frac{l d^{\prime}(l)}{2^{2 n}}+\frac{8 l^{4}}{2^{2 n}}$. From Proposition 2 we have, $\operatorname{Pr}[N C O L \geq 2] \leq \frac{8 l^{4}}{2^{2 n}}$. From Claim 2, we have, $\sum_{k=0}^{1} \operatorname{Pr}\left[\Sigma^{(i)}=\Sigma^{(j)} \wedge \Theta^{(i)}=\Theta^{(j)} \wedge N C O L=k \mid Z^{(i)} \neq Z^{(j)}\right] \leq \frac{4 l^{2}+1}{2^{2 n}}$. Therefore,

$$
\begin{aligned}
\operatorname{Pr}\left[\Sigma^{(i)}=\Sigma^{(j)} \wedge \Theta^{(i)}=\Theta^{(j)}\right] & \leq \frac{l d^{\prime}(l)}{2^{2 n}}+\frac{8 l^{4}}{2^{2 n}}+\frac{4 l^{2}+1}{2^{2 n}}+\frac{8 l^{4}}{2^{2 n}} \\
& \leq \frac{l d^{\prime}(l)}{2^{2 n}}+\frac{16 l^{4}}{2^{2 n}}+\frac{4 l^{2}+1}{2^{2 n}} \\
& \leq \frac{22 l^{4}}{2^{2 n}}
\end{aligned}
$$

In the next two sections, we state and prove the two claims above.

## 5 Details of the Proof of Claim 1

Claim 1 Fix two distinct messages $M^{(i)}, M^{(j)}$ each of length at most $l$ blocks. Then,

$$
\operatorname{Pr}\left[Z^{(i)}=Z^{(j)} \wedge \Theta^{(i)}=\Theta^{(j)}\right] \leq \frac{l d^{\prime}(l)}{2^{2 n}}+\frac{8 l^{4}}{2^{2 n}}
$$

where $Z^{(i)}=Y_{l_{i}}^{(i)}, Z^{(j)}=Y_{l_{j}}^{(j)}$, and $l_{i}, l_{j}$ are the number of blocks of $M^{(i)}, M^{(j)}$ respectively.

Proof. We prove the claim using the structure graph. After fixing two messages $M^{(i)}$ and $M^{(j)}$ and choosing a function $f$ uniformly at random from the set of all functions over $\{0,1\}^{b} \times\{0,1\}^{n} \rightarrow\{0,1\}^{n}$, we analyze the structure graph $G:=G^{f}\left(M^{(i)}, M^{(j)}\right)$. In particular, we analyze the probability of the event $Z^{(i)}=Z^{(j)} \wedge \Theta^{(i)}=\Theta^{(j)}$ in view of number of collisions (say, NCOL) occurred in the corresponding structure graph $G$. Therefore, we have,

$$
\begin{aligned}
\operatorname{Pr}\left[Z^{(i)}=Z^{(j)} \wedge \Theta^{(i)}=\Theta^{(j)}\right]= & \operatorname{Pr}\left[Z^{(i)}=Z^{(j)} \wedge \Theta^{(i)}=\Theta^{(j)} \wedge N C O L=1\right] \\
& +\operatorname{Pr}\left[Z^{(i)}=Z^{(j)} \wedge \Theta^{(i)}=\Theta^{(j)} \wedge N C O L \geq 2\right]
\end{aligned}
$$

In Section 5.1, we show that

$$
\begin{equation*}
\operatorname{Pr}\left[Z^{(i)}=Z^{(j)} \wedge \Theta^{(i)}=\Theta^{(j)} \wedge N C O L=1\right] \leq \frac{l d^{\prime}(l)}{2^{2 n}} \tag{9}
\end{equation*}
$$

where $d^{\prime}(l)$ is the maximum number of positive divisors of the integer $l^{\prime}$ from $[1, l]$.

When NCOL in the graph is at least 2, then using Proposition 2 we have,

$$
\begin{equation*}
\operatorname{Pr}\left[Z^{(i)}=Z^{(j)} \wedge \Theta^{(i)}=\Theta^{(j)} \wedge N C O L \geq 2\right] \leq \operatorname{Pr}[N C O L \geq 2] \leq \frac{8 l^{4}}{2^{2 n}} \tag{10}
\end{equation*}
$$

Therefore, combining Equations (9) and (10), we get the result.
Now the only part of the proof that remains is to prove Equation (9).

### 5.1 Proof of Equation (9)

We can write

$$
\begin{align*}
& \operatorname{Pr}\left[Z^{(i)}=Z^{(j)} \wedge \Theta^{(i)}=\Theta^{(j)} \wedge N C O L=1\right] \\
= & \operatorname{Pr}\left[Z^{(i)}=Z^{(j)} \wedge N C O L=1\right] \cdot \operatorname{Pr}\left[\Theta^{(i)}=\Theta^{(j)} \mid Z^{(i)}=Z^{(j)} \wedge N C O L=1\right] . \tag{11}
\end{align*}
$$

In Equation (11), there are two probabilities that need to be computed. First, we compute $\operatorname{Pr}\left[Z^{(i)}=Z^{(j)} \wedge N C O L=1\right]$ by considering different structure graphs with $N C O L=1$, corresponding to the construction NI2 ${ }_{r}^{+}$. Let $G$ denote the set of all structure graphs with $N C O L=1$ and $Z^{(i)}=Z^{(j)}$. Without loss of generality, let $l_{i}$ and $l_{j}$ be the lengths of the messages $M^{(i)}$ and $M^{(j)}$ respectively, with $l_{i} \geq l_{j}$. Let $G_{1} \subset G$ be the set of all structure graphs such that the $M^{(i)}$ path does not contain any loop. The $G_{2}=G \backslash G_{1}$ is the set of the remaining structure graphs. For the ease of understanding blue colored path represents the $M^{(i)}$ path and red colored path represents the $M^{(j)}$ path.

Analysis of $\boldsymbol{G}_{\boldsymbol{1}}$. If $M^{(j)}$ is a proper prefix of $M^{(i)}$, then $\left|G_{1}\right|=0$, since in that case $Z^{(i)}$ won't be equal to $Z^{(j)}$. So without loss of generality, lets assume that $M^{(j)}$ is not a prefix of $M^{(i)}$. Suppose the first $p$ blocks constitute the common prefix. Define $t^{*}=\min \left\{t>l_{i}+p:[[t]] \leq l_{i}\right\}$. Thus, the edge $\left(\left[\left[t^{*}-1\right]\right]^{\prime},\left[\left[t^{*}\right]\right]\right)$ in $G$ creates the collision and from that point onwards, $M^{(j)}$ path will follow the rest of $M^{(i)}$ path which is nothing but the common suffix part of $M^{(i)}$ and $M^{(j)}$.


Fig. 5.1: Structure Graph of type $G_{1}$

The scenario is explained in Fig. 5.1. Since there are $\leq l$ choices for $t^{*}$, we have $\left|G_{1}\right| \leq l$.

Analysis of $\boldsymbol{G}_{\mathbf{2}}$. In graph $G_{2}, M^{(i)}$ path creates a collision by creating a self loop. We define $t^{*}=\min \{t:[[t]] \leq t\}$ and let $p^{*}=\left[\left[t^{*}\right]\right]$. Therefore, $\left(t^{*}, p^{*}\right)$ denotes the collision in $M^{(i)}$ path. Now we can split $M^{(i)}$ into three mutual disjoint strings $x, y, z$ such that $x:=M_{1}^{(i)}\|\ldots\| M_{p^{*}}^{(i)}, y:=M_{p^{*}+1}^{(i)} \| \ldots M_{t^{*}}^{(i)}$ and some $z$ chosen to be the smallest string so that we can write $M^{(i)}=x\left\|y^{a}\right\| z$ for some $a \geq 1$.

Note that to have $Z^{(i)}=Z^{(j)}$ and one collision has already been occurred in the loop, therefore, $M^{(j)}$-path must be a subpath of $M^{(i)}$-path and it cannot bifurcate from $M^{(i)}$ path and then collide with the last output block of $M^{(i)}$ as that would increases the number of collisions to 2 . Thus, the $M^{(j)}$-path must be of the form $x\left\|y^{b}\right\| z$, where $b<a$ (since $l_{i}>l_{j}$ in this case). Hence, the number of blocks in $y$, i.e., $t^{*}-p^{*}$, in the diagram must divide $l_{i}-l_{j}$. This


Fig. 5.2: Structure Graph of type $G_{2}$
scenario is explained in Fig. 5.2. There are at most $l$ choices for such a $t^{*}$ and $d^{\prime}(l)$ choices for such a $p^{*}$. Hence, $\left|G_{2}\right| \leq l d^{\prime}(l)$. In the special case, when $l_{i}=l_{j}$, then obviously, $\left|G_{2}\right|=0$.

Therefore, considering $G_{1}$ and $G_{2}$ together, by Proposition 1, we have

$$
\begin{equation*}
\operatorname{Pr}\left[Z^{(i)}=Z^{(j)} \wedge N C O L=1\right] \leq \frac{l d^{\prime}(l)}{2^{n}} \tag{12}
\end{equation*}
$$

Now, we compute the second probability of Equation 11, i.e., $\operatorname{Pr}\left[\Theta^{(i)}=\right.$ $\left.\Theta^{(j)} \mid Z^{(i)}=Z^{(j)} \wedge N C O L=1\right]$. Note that $\Theta^{(i)}=\Theta^{(j)}$ gives an equation of the form

$$
\begin{equation*}
2^{l_{i}} Y_{1}^{(i)}+2^{l_{i}-1} Y_{2}^{(i)}+\cdots+2 Y_{l_{i}}^{(i)}=2^{l_{j}} Y_{1}^{(j)}+2^{l_{i}-1} Y_{2}^{(i)}+\cdots+2 Y_{l_{j}}^{(j)} \tag{13}
\end{equation*}
$$

The condition $Z^{(i)}=Z^{(j)}$ and $N C O L=1$ is equivalent to the condition $Y_{l_{i}}^{(i)}=$ $Y_{l_{j}}^{(j)}$ and $Y_{a}^{(i)} \neq Y_{b}^{(j)}$, whenever either $a<l_{i}$ or $b<l_{j}$. With this condition, Equation 13 becomes

$$
\begin{equation*}
2^{l_{i}} Y_{1}^{(i)}+2^{l_{i}-1} Y_{2}^{(i)}+\cdots+2^{2} Y_{l_{i}-1}^{(i)}=2^{l_{j}} Y_{1}^{(j)}+2^{l_{i}-1} Y_{2}^{(i)}+\cdots+2^{2} Y_{l_{j}-1}^{(j)} \tag{14}
\end{equation*}
$$

Now, for both the graphs $G_{1}$ and $G_{2}$, we will be able to find at least one $Y$ variable belonging to the part between $p$ and $t^{*}$, such that Equation (14) becomes non-trivial for such variable $Y$, giving a probability of $\frac{1}{2^{n}}$ for the second term of Equation (11). When this along with Equation (12) is plugged in Equation (11), the probability in Equation (11), i.e., in Equation (9), becomes bounded by $\frac{l d^{\prime}(l)}{2^{2 n}}$.

## 6 Details of the Proof of Claim 2

Claim 2 Fix two distinct messages $M^{(i)}, M^{(j)}$ each of length at most $l$ blocks. Then,

$$
\sum_{k=0}^{1} \operatorname{Pr}\left[\Sigma^{(i)}=\Sigma^{(j)} \wedge \Theta^{(i)}=\Theta^{(j)} \wedge N C O L=k \mid Z^{(i)} \neq Z^{(j)}\right] \leq \frac{4 l^{2}+1}{2^{2 n}}
$$

where $Z^{(i)}=Y_{l_{i}}^{(i)}, Z^{(j)}=Y_{l_{j}}^{(j)}, l_{i}, l_{j}$ is the number of blocks of $M^{(i)}, M^{(j)}$ respectively.

Proof. It is to be noted that, under the condition $Z^{(i)} \neq Z^{(j)}, \Sigma^{(i)}=\Sigma^{(j)}$ is independent on $\Theta^{(i)}=\Theta^{(j)} \& N C O L=k$ for $k=0,1$. Therefore, we can write

$$
\begin{aligned}
& \sum_{k=0}^{1} \operatorname{Pr}\left[\Sigma^{(i)}=\Sigma^{(j)} \wedge \Theta^{(i)}=\Theta^{(j)} \wedge N C O L=k \mid Z^{(i)} \neq Z^{(j)}\right] \\
= & \operatorname{Pr}\left[\Sigma^{(i)}=\Sigma^{(j)} \mid Z^{(i)} \neq Z^{(j)}\right]\left(\sum_{k=0}^{1} \operatorname{Pr}\left[\Theta^{(i)}=\Theta^{(j)} \wedge N C O L=k \mid Z^{(i)} \neq Z^{(j)}\right]\right) .
\end{aligned}
$$

Now, $\operatorname{Pr}\left[\Sigma^{(i)}=\Sigma^{(j)} \mid Z^{(i)} \neq Z^{(j)}\right] \leq \frac{1}{2^{n}}$ as $f_{3}$ is independent from $f_{1}$ and $f_{2} ;$ collision probability of a random function. Again from Claim 2 we have,

$$
\begin{equation*}
\sum_{k=0}^{1} \operatorname{Pr}\left[\Theta^{(i)}=\Theta^{(j)} \wedge N C O L=k \mid Z^{(i)} \neq Z^{(j)}\right] \leq \frac{4 l^{2}+1}{2^{n}} \tag{15}
\end{equation*}
$$

Combining the collision probability of a random function and Equation (15), we get the result.

Therefore, we are only left with the proof of Equation (15).

### 6.1 Proof of Equation (15)

To prove the equation, we separately bound the following $\operatorname{Pr}\left[\Theta^{(i)}=\Theta^{(j)} \wedge\right.$ $\left.N C O L=0 \mid Z^{(i)} \neq Z^{(j)}\right]$ and $\operatorname{Pr}\left[\Theta^{(i)}=\Theta^{(j)} \wedge N C O L=1 \mid Z^{(i)} \neq Z^{(j)}\right]$ separately.

Again we consider two distinct messages $M^{(i)}$ and $M^{(j)}$ with lengths $l_{i}$ and $l_{j}$ respectively, with $l_{i} \geq l_{j}$. Since we are given the condition $Z^{(i)} \neq Z^{(j)}$, the structure graphs will have the common feature that the end-point $Y_{l_{i}}^{(i)}$ of $M^{(i)}$ path and the end-point $Y_{l_{j}}^{(j)}$ of $M^{(j)}$-path must be different, i.e., from Equation (13), we have $Y_{l_{i}}^{(i)} \oplus Y_{l_{j}}^{(j)}=c \neq 0$. Thus, Equation (13) becomes non-trivial, with probability $\frac{1}{2^{n}}$.

Now, we need to count the number of distinct structure graphs for each of the cases $N C O L=0$ and $N C O L=1$.

Clearly, when $N C O L=0$, only such structure graph is possible, as shown in Fig. 6.1. Thus, we have


Fig. 6.1: Structure Graph of accident 0

$$
\begin{equation*}
\operatorname{Pr}\left[\Theta^{(i)}=\Theta^{(j)} \wedge N C O L=0 \mid Z^{(i)} \neq Z^{(j)}\right] \leq \frac{1}{2^{n}} \tag{16}
\end{equation*}
$$

Now, let us consider the case $N C O L=1$. Let $G$ be the set of all structure graphs with $N C O L=1$ with $Z^{(i)} \neq Z^{(j)}$. Let $G_{1} \subset G$ be the set of all structure graphs such that the $M^{(i)}$-path does not contain any loop. The $G_{2}=G \backslash G_{1}$ is the set of remaining structure graphs.

Analysis of $\boldsymbol{G}_{\mathbf{1}}$. For $G_{1}$, the $M^{(j)}$ path can either intersect with $M^{(i)}$ exactly once or $M^{(j)}$ path does not intersect with $M^{(i)}$ but it creates a loop with itself. In the first case, $M^{(j)}$-path cannot have any loop as shown in Fig. 6.2 as that would increase the number of collision to 2 , and in the second case, the $M^{(j)}$ path cannot intersect $M^{(i)}$-path at all as that would again increase the number of collision to 2 as shown in Fig. 6.3. In either case, the number of such graphs is at most $l^{2}$.


Fig. 6.2: Structure Graph of type $G_{1} ; M^{(i)}$ path has no loop


Fig. 6.3: Structure Graph of type $G_{1} ; M^{(i)}$ path has no loop, $M^{(j)}$ path has loop

Analysis of $\boldsymbol{G}_{\mathbf{2}}$. For $G_{2}$, note that $M^{(i)}$ path contains a loop. Now the $M^{(j)}$ path may or may not intersects $M^{(i)}$ path. If it does, then it must follow the same loop as $M^{(i)}$ and then exit either from the loop or afterwards, as shown in Fig. 6.4. $M^{(j)}$ path may also bifurcate from $M^{(i)}$ path before the loop and then it should not intersect with $M^{(i)}$ path again or it should not make any self loop with itself as both of the cases would increases the number of collision to 2 . Note that $M^{(j)}$ path cannot intersect $M^{(i)}$ path before the loop as that would increase the number of collision to 2 .

If $M^{(j)}$ path does not intersect $M^{(i)}$ path, then $M^{(j)}$ path cannot make a loop with itself as that would increase the number of collision to 2 . Therefore, again the case is similar to Fig. 6.3 where the blue colored path will then represent the $M^{(j)}$ path and red colored path will represent $M^{(i)}$ path. In either case, the number of such graphs is at most $l^{2}$.


Fig. 6.4: Structure Graph of type $G_{2} ; M^{(i)}, M^{(j)}$ both path contain a loop

Thus, for the above $4 l^{2}$ graphs (combined $G_{1}$ and $G_{2}$ ),

$$
\begin{equation*}
\operatorname{Pr}\left[\Theta^{(i)}=\Theta^{(j)} \wedge N C O L=1 \mid Z^{(i)} \neq Z^{(j)}\right] \leq \frac{4 l^{2}}{2^{n}} \tag{17}
\end{equation*}
$$

Therefore, from Equation (16) and (17), we get

$$
\sum_{k=0}^{1} \operatorname{Pr}\left[\Theta^{(i)}=\Theta^{(j)} \wedge N C O L=k \mid Z^{(i)} \neq Z^{(j)}\right] \leq \frac{4 l^{2}+1}{2^{n}}
$$

Remark 2 We have achieved $B B B$ security for the $N_{2} 2^{+}$MAC which is the extended version of NI2 MAC. Note that NI2 MAC is a variant of NI MAC. One can easily show that same modification on NI MAC gives BBB security. It is to be noted that in case of $\mathrm{NI}^{+}$MAC when we calculate

$$
\operatorname{Pr}\left[\Sigma^{(i)}=\Sigma^{(j)} \wedge \Theta^{(i)}=\Theta^{(j)} \wedge Z^{(i)}=Z^{(j)}\right]
$$

then we should consider only the structure graph that does not contain any loop as we need to consider the $i^{\text {th }}$ and $j^{\text {th }}$ message having same length.

## 7 Conclusion and Future Work

Rcently, NI2-MAC was introduced in order to prove the security of NI-MAC. In this paper, we show a modified construction of NI2-MAC and prove its security to be beyond birthday. While we use we use an extra keyed function $\left(f_{K_{3}}\right)$ in $\mathrm{NI} 2^{+}$, an interesting research problem would be to avoid the usage of this extra keyed function and achieves beyond birthday security.

## References

1. Jee Hea An and Mihir Bellare. Constructing vil-macsfrom fil-macs: Message authentication under weakened assumptions. In Wiener [19], pages 252-269.
2. Mihir Bellare, Ran Canetti, and Hugo Krawczyk. Keying hash functions for message authentication. In Neal Koblitz, editor, Advances in Cryptology - CRYPTO '96, 16th Annual International Cryptology Conference, Santa Barbara, California, USA, August 18-22, 1996, Proceedings, volume 1109 of Lecture Notes in Computer Science, pages 1-15. Springer, 1996.
3. Mihir Bellare, Oded Goldreich, and Hugo Krawczyk. Stateless evaluation of pseudorandom functions: Security beyond the birthday barrier. In Wiener [19], pages 270-287.
4. Mihir Bellare, Joe Kilian, and Phillip Rogaway. The security of cipher block chaining. In Yvo Desmedt, editor, Advances in Cryptology - CRYPTO '94, 14 th Annual International Cryptology Conference, Santa Barbara, California, USA, August 2125, 1994, Proceedings, volume 839 of Lecture Notes in Computer Science, pages 341-358. Springer, 1994.
5. Mihir Bellare, Krzysztof Pietrzak, and Phillip Rogaway. Improved security analyses for CBC macs. In Victor Shoup, editor, Advances in Cryptology - CRYPTO 2005: 25th Annual International Cryptology Conference, Santa Barbara, California, USA, August 14-18, 2005, Proceedings, volume 3621 of Lecture Notes in Computer Science, pages 527-545. Springer, 2005.
6. John Black, Shai Halevi, Hugo Krawczyk, Ted Krovetz, and Phillip Rogaway. UMAC: fast and secure message authentication. In Wiener [19], pages 216-233.
7. John Black and Phillip Rogaway. A block-cipher mode of operation for parallelizable message authentication. In Knudsen [16], pages 384-397.
8. Nilanjan Datta, Avijit Dutta, Mridul Nandi, Goutam Paul, and Liting Zhang. Onekey double-sum mac with beyond-birthday security. Cryptology ePrint Archive, Report 2015/958, 2015. http://eprint.iacr.org/.
9. Yevgeniy Dodis and John P. Steinberger. Domain extension for macs beyond the birthday barrier. In Kenneth G. Paterson, editor, EUROCRYPT 2011, volume 6632 of Lecture Notes in Computer Science, pages 323-342. Springer, 2011.
10. Peter Gazi, Krzysztof Pietrzak, and Michal Rybár. The exact prf-security of NMAC and HMAC. In Juan A. Garay and Rosario Gennaro, editors, Advances in Cryptology - CRYPTO 2014-34th Annual Cryptology Conference, Santa Barbara, CA, USA, August 17-21, 2014, Proceedings, Part I, volume 8616 of Lecture Notes in Computer Science, pages 113-130. Springer, 2014.
11. Oded Goldreich, Shafi Goldwasser, and Silvio Micali. On the cryptographic applications of random functions. In G. R. Blakley and David Chaum, editors, Advances in Cryptology, Proceedings of CRYPTO '84, Santa Barbara, California, USA, August 19-22, 1984, Proceedings, volume 196 of Lecture Notes in Computer Science, pages 276-288. Springer, 1984.
12. Tetsu Iwata and Kaoru Kurosawa. OMAC: one-key CBC MAC. In Johansson [14], pages 129-153.
13. Éliane Jaulmes, Antoine Joux, and Frédéric Valette. On the security of randomized CBC-MAC beyond the birthday paradox limit: A new construction. In Fast Software Encryption, 2002, volume 2365 of Lecture Notes in Computer Science, pages 237-251. Springer, 2002.
14. Thomas Johansson, editor. Fast Software Encryption, 10th International Workshop, FSE 2003, Lund, Sweden, February 24-26, 2003, Revised Papers, volume 2887 of Lecture Notes in Computer Science. Springer, 2003.
15. Antoine Joux, Guillaume Poupard, and Jacques Stern. New attacks against standardized macs. In Johansson [14], pages 170-181.
16. Lars R. Knudsen, editor. Advances in Cryptology - EUROCRYPT 2002, International Conference on the Theory and Applications of Cryptographic Techniques, Amsterdam, The Netherlands, April 28-May 2, 2002, Proceedings, volume 2332 of Lecture Notes in Computer Science. Springer, 2002.
17. Ueli M. Maurer. Indistinguishability of random systems. In Knudsen [16], pages 110-132.
18. Kazuhiko Minematsu. How to thwart birthday attacks against macs via small randomness. In Seokhie Hong and Tetsu Iwata, editors, Fast Software Encryption, 2010, volume 6147 of Lecture Notes in Computer Science, pages 230-249. Springer, 2010.
19. Michael J. Wiener, editor. Advances in Cryptology - CRYPTO '99, 19th Annual International Cryptology Conference, Santa Barbara, California, USA, August 1519, 1999, Proceedings, volume 1666 of Lecture Notes in Computer Science. Springer, 1999.
20. Kan Yasuda. The sum of CBC macs is a secure PRF. In Josef Pieprzyk, editor, Topics in Cryptology - CT-RSA 2010, The Cryptographers' Track at the RSA Conference 2010, San Francisco, CA, USA, March 1-5, 2010. Proceedings, volume 5985 of Lecture Notes in Computer Science, pages 366-381. Springer, 2010.
21. Kan Yasuda. A new variant of PMAC: beyond the birthday bound. In Phillip Rogaway, editor, Advances in Cryptology - CRYPTO 2011-31st Annual Cryptology Conference, Santa Barbara, CA, USA, August 14-18, 2011. Proceedings, volume 6841 of Lecture Notes in Computer Science, pages 596-609. Springer, 2011.
22. Liting Zhang, Wenling Wu, Han Sui, and Peng Wang. 3kf9: Enhancing 3gpp-mac beyond the birthday bound. In Xiaoyun Wang and Kazue Sako, editors, Advances in Cryptology - ASIACRYPT 2012 - 18th International Conference on the Theory and Application of Cryptology and Information Security, Beijing, China, December 2-6, 2012. Proceedings, volume 7658 of Lecture Notes in Computer Science, pages 296-312. Springer, 2012.
