In this paper, we present new generic multiset attacks against generalized Feistel networks, by which we can recover all the unknown round functions completely instead of deciding whether an unknown encryption oracle is such network or a random permutation. With one $r$-round multiset distinguisher, we can recover the outermost round functions for $r+1$-round block cipher. Next we propose the dummy-round technique, which allows us to make a full-round decomposition if the outermost round is recovered. Moreover, the dummy-round technique barely increases the complexity of our attack. Using this generic method, we propose attacks on 7-round RC6-like and 7-round CLEFIA-like structures. Our attacks can recover all the secret round functions, requiring only $O(10 \times 2^{0.7n})$ time complexity and $O(5 \times 2^{n/2})$ chosen plaintexts, where $n$ indicates the block size of the cipher. For 64-bit ciphers of these two structures, our results will lead to a practical attack.

1. Introduction

The architecture is a fundamental part of a block cipher. It plays an important role in both security aspects and implementation performances of the cipher. Two of the most frequently used architectural structures nowadays are the Substitution-Permutation Networks and the generalized Feistel Networks; the latter contains the standard Feistel Network and its variants. Typical examples of the variant Feistel Networks include CLEFIA [1], RC6 [2], and CAST256 [3].

Among all the attacks against block ciphers, structural attack is an interesting branch which studies the security of the architectures. In this cryptanalysis, all of the internal functions are unknown or key-dependent. The only information available to the attacker is the type of the general structure of the block cipher and size parameters of its components. The aim of the attack is to recover all the internal functions. Since this cannot exploit particular weaknesses (such as bad differential properties or weak avalanche effects) of concrete functions, structural recovery attacks are often weaker than traditional differential attacks or linear attacks on given cryptosystems. The advantage of these attacks is that they are applicable to large classes of cryptosystems, which is thus very useful in establishing general design rules for strong cryptosystems and in dealing with the algorithms with unknown design criteria.

Related Works. The structural attack is far from being new. In 2001, Biryukov and Shamir [4] investigated the recovery problem of iterated SPN ciphers, in which the substitutions and permutations are all secret and key-dependent. In ASIACRYPT2014, Biryukov et al. proposed a recovery on ASASA scheme, which was designed by claiming that it could resist traditional attacks [5]. Soon this result was improved by Dinur et al. in [6], and a more efficient recovery algorithm was proposed. In [7], Tiessen et al. proposed a structural attack on a variant of AES (in which the S-boxes are kept unknown). Their attack was indeed an improved integral attack and could recover all the secret information up to 6 rounds.

The structural attack against generalized Feistel Networks was first studied in [8]: it was presented in their work that if the Feistel functions were completely unknown, the yoyo game could attack up to 5 rounds. The use of small Feistel Networks for lightweight S-Box design was investigated in [9], and an efficient decomposition was discovered for the secret S-Box of recent Russian standards [10] using reverse-engineering.

Our Contribution. This paper mainly concentrates on the recovery attacks against generalized Feistel ciphers with
bijective round functions. The main results of this paper are as follows:

(1) We propose a new and special integral distinguisher for structural attack. If an \( r \)-round distinguisher is detected, we can always launch an efficient \( r+1 \)-round structural attack of the outermost round functions.

(2) We put forward the dummy-round technique. This technique shows that if the decomposition of the outermost round function in an \( r \)-round iterative cipher is with (data/time) complexity \( \mathcal{N} \), then the complexity to decompose all the internal round functions is at most with (data/time) complexity \( r \times \mathcal{N} \); more precisely, given an algorithm which can recover the round functions at the last round of an \( r \)-round cipher, we add a dummy round at the beginning, so that the number of rounds remains the same; consequently we can use the same algorithm to recover the round functions from the second last round, and so on.

(3) Therefore, for an \( r \)-round iterative cipher, we can use our integral distinguisher to make a full-round decomposition without a significantly increased complexity. Applying our results, we propose recovery attacks on 7-round CLEFIA-like and 7-round RC6-like ciphers. To the best of our knowledge, these are the first and best structural attack results against these structures.

Organization. The rest of this paper is organized as follows. Section 2 introduces several basic definitions that will be used throughout this paper. Section 3 elaborates the generic multiset attack against generalized Feistel ciphers. Section 4 applies our attack on 7-round CLEFIA/RC6-like ciphers, respectively. Section 5 concludes the paper.

2. Preliminaries

2.1. Generalized Feistel Ciphers. First, we will clarify what generalized Feistel cipher means in this paper.

Let \( k \) be an even integer and a single round (denoted by \( RF \)) of \( k \)-cell generalized Feistel cipher is defined as

\[
RF(x_0, \ldots, x_{k-1}) = \pi(x_0, F_0(x_0) \oplus x_1, \ldots, F_{(k-2)/2}(x_{k-2}) \oplus x_{k-1})
\]

(1)

in which \( F_i : \{0,1\}^m \rightarrow \{0,1\}^m \) is a keyed function called a round function and \( \pi : \{0,1\}^{mk} \rightarrow \{0,1\}^{mk} \) is a cell-wise permutation.

The encryption of a generalized Feistel cipher is defined as

\[
(y_0, \ldots, y_{k-1}) = \bigoplus_{i=1}^r RF_i(x_0, \ldots, x_{k-1})
\]

(2)

where the first input \((x_0, \ldots, x_{k-1})\) is the plaintext and the \( r \)-round output \((y_0, \ldots, y_{k-1})\) is the ciphertext.

2.2. Multiset Properties on Generalized Feistel Ciphers. Multiset attack [4] is a generic class of attacks which appeared in the literature under three different names: the square attack [11, 12], the saturation attack [13, 14], and the integral cryptanalysis [15], which can also be treated as a special variant of the higher-order differential attack [16], cube attack [17], and also of the division property [18]. This attack generally uses a set of chosen plaintexts that contain all possible values for some bits and has a constant value for the other bits. Corresponding ciphertexts are calculated from plaintexts in the set by using an encryption oracle. If ciphertexts just add up to zero in certain bits, we say that this cipher has the multiset distinguisher.

In [4], Biryukov and Shamir defined multiset, which can be represented as a list of (value, multiplicity) pairs, and the size of the multiset is the sum of all its multiplicities.

Example 1. The multiset \( \mathcal{M} = \{1, 2, 2, 2, 3, 3, 3, 3\} \) can be represented as \( \mathcal{M} = \{(1, 2), (2, 3), (3, 4)\} \); the size of \( \mathcal{M} \) is 9.

We define five multiset properties as follows:

**Multiset Properties**

- \( \mathcal{A} \) (All): Every possible value appears exactly once in the multiset.
- \( \mathcal{B} \) (Balance): The XOR of all values in the multiset is 0.
- \( \mathcal{E} \) (Even): Each value occurs even times in the multiset.
- \( \mathcal{C} \) (Constant): The value is fixed to a constant for all texts in the multiset.
- \( \mathcal{U} \) (Unknown): The output multiset is unknown.

Note that the definitions of \( \mathcal{B}, \mathcal{E}, \mathcal{C} \) are the same as in [4], and the property \( \mathcal{A} \) is equal to the property \( \mathcal{P} \) as defined in [4].

Generalized Feistel ciphers make use of three basic operations: XOR-operation, branching operation, and secret round functions \( F_i \). Multiset properties over these operations comprise the multiset property of the ciphers and obey the major rules (see Table 1). For showing these rules, it is crucial to require the round functions \( F_i \) of the generalized Feistel cipher to be invertible (or bijective).

In this paper, we will use multiset distinguisher in our attack, which is of the form \( (\alpha \rightarrow \beta) \), where \( \alpha \in \{\mathcal{C}, \mathcal{A}\}^k, \beta \in \{\mathcal{C}, \mathcal{A}, \mathcal{B}, \mathcal{U}\}^k \), and the input state \( \alpha \) contains at least one cell equal to \( \mathcal{A} \), and the output cells in state \( \beta \) are not all equal to \( \mathcal{U} \).

3. Generic Multiset Attack against Generalized Feistel Ciphers

By applying Table 1, one can build the multiset propagation system for any fixed generalized Feistel ciphers. In this attack, we are only interested in the multiset distinguishers with at least one cell of the output state satisfying property \( \mathcal{B} \) but neither \( \mathcal{E} \) nor \( \mathcal{A} \), which we denoted as \( G\mathcal{B} \).
3.1. Recover the Outermost Round Function. For \( r + 1 \) rounds generalized Feistel cipher
\[
c = RF_r \circ \cdots \circ RF_0 (p),
\]
we will start by decomposing the outermost round.

The recovered result of the last round will be given by a look-up table. More precisely, we identify the secret function by fixing all the entries of the last round function. In order to achieve this goal, we build linear equations related to the entries with the help of the multiset property and then apply the Gaussian elimination algorithm to get all the entries.

Our attack uses a set of chosen plaintexts that contain all possible values at \( t \) cells \( 0 < t < k \) in positions \( d_1, d_2, \ldots, d_t \) and constant value \( v_0 \) \( (k - t) \) cells for the rest (denoted as \( \mathcal{A}_{v_0 \prec \{d_1, d_2, \ldots, d_t\}} \)). After \( r \)-round encryption, if some cell of the output state satisfies \( G \mathcal{B} \)-property, we can recover the \( r + 1 \)-th round function as follows.

First, we denote the set \( \{c_1, c_2, \ldots, c_{2^m}\} \), which is the ciphertexts encrypted by plaintexts in the set \( \mathcal{A}_{v_0 \prec \{d_1, d_2, \ldots, d_t\}} \). Then, we consider a part of the last round which corresponds to the cell of the state after \( r \)-round encryption, where we know that the multiset of all values has property \( G \mathcal{B} \), and denote \( f_{r+1} \) as the inverse of that part of the last-round \( RF_r \). Finally, we get the equation, which is only related to the final round, satisfying
\[
\bigoplus_{j=1}^{2^m} f_{r+1} (c_j) = 0.
\]

Then we assume that the multiset of the values \( f_{r+1} (c_j) \) has property \( G \mathcal{B} \). By the definition of the \( G \mathcal{B} \)-property, the left part of the equation above does not fall into the case in which each value occurs even times, thus being nontrivial (see Figure 1).

Remark. We need to mention that, for 16-bit/32-bit size block ciphers, we have checked the multiset property by simple experiments. The results indicate that the randomly chosen round functions present a good chance to lead \( G \mathcal{B} \)-properties for both CLEFIA-like and RC6-like structures.

Next, we change the constant \( v_0 \) to obtain new sets \( \mathcal{A}_{v_0 \prec \{d_1, d_2, \ldots, d_t\}} \) to get more linear equations with randomly-looking subsets of variables. When sufficient linear equations are collected, we can solve the system by Gaussian elimination to recover \( f_{r+1} \), and the solving process of each system requires \( 2^{2.81m} \) steps with Strassen’s Algorithm.

In the process of collecting linear equations, most generalized Feistel ciphers have the rank deficiency problem, which means we can never get a system of equations with a full rank of \( 2^m \). Similar problems also appear in the decompositions of SPN [4–6] and standard Feistel cipher [8]. This is due to the fact that, for any of these ciphers, there exist several equivalent ciphers. Picking up any one from the equivalents, the encryption mapping keeps correct. Therefore, decompositions of such cipher structures are not unique. We will show it in our practical decompositions later.

3.2. Dummy-Round Technique for the Inner Round Functions. After finding the outermost round function, we can just repeat our attack in the reverse direction by using chosen ciphertexts and recover the first round; then we are left with the rest inner round functions. If there exists an attack with much lower complexity, the complexity of recovery of all rounds is dominated by that of recovering the outer round. However, in some literatures now available, people still have to find new ways to recover the inner round functions [4–6, 8], mainly because the technique they used to attack the outer round cannot be applied in attacking the inner rounds.
A straightforward way is to transfer the inner round recovery problems into the outer round decomposition. We next provide a general technique called the dummy-round technique.

**Dummy-Round Technique.** Let \( RF_1 \) be the \( i \)-th round of the generalized Feistel cipher; then an \( r+1 \)-round generalized Feistel cipher could be represented by \( RF_r \circ RF_{r-1} \circ \cdots \circ RF_0 \). Using the multiset distinguisher and a linear equations system solver, we are able to recover the last round \( RF_r \). In order to recover the rest of the \( r \)-round functions, we transfer it into the known \( r+1 \)-round issues. We randomly choose \( k/2 \) round functions \( G_0, G_1, \ldots, G_{(k-2)/2} \) and construct a new round \( RF_{-1} \) (called dummy-round), i.e.,

\[
RF_{-1} = \pi(x_0, G_0(x_0) \oplus x_1, \ldots, G_{(k-2)/2}(x_{k-2}) \oplus x_{k-1}),
\]

and then we get a new cipher \( RF_{-1} \circ \cdots \circ RF_0 \circ RF_1 \).

Let \( c = RF_r \circ RF_{r-1} \circ \cdots \circ RF_0(p) \); then we get

\[
RF_{-1}^{-1}(c) = RF_{-1} \circ \cdots \circ RF_1 \circ RF_0^{-1}(p),
\]

since both \( RF_r \) and \( RF_{-1} \) are known to us, then \( C = RF_{-1}^{-1}(c) \) and \( P = RF_0^{-1}(p) \) are available. Then the equation above could be rewritten as

\[
C = RF_{-1} \circ \cdots \circ RF_0 \circ RF_{-1}(P),
\]

which is exactly the same as the original structure. So for the original structure, if we are able to recover the outermost round \( RF_r \), we can use the exact same method to find \( RF_{-1} \) by introducing the dummy-round \( RF_{-1} \) (see Figure 2).

Therefore, the complexity of recovery of each round is dominated by the complexity of recovery of the final round. The dummy-round technique allows reusing the final-round attack for all rounds. When several rounds are attacked, it is very likely that there exists an attack with much lower complexity for the inner rounds. Generally, in this case the total attack complexity is at most multiplied by the number of rounds.

### 4. Recovery on CLEFIA-Like and RC6-Like Structures

In this section, we describe two existing generalized Feistel structures, named the CLEFIA-like and RC6-like structures. Single rounds of these two structures are listed as follows (see Figure 3).

The \( i \)-th round of CLEFIA-like structure:

\[
(y_0, y_1, y_2, y_3) = (F_{2i}(x_0) \oplus x_1, x_2, F_{2i+1}(x_2) \oplus x_3, x_0).
\]

The \( i \)-th round of RC6-like structure:

\[
(y_0, y_1, y_2, y_3) = (F_{2i+1}(x_2) \oplus x_3, x_0, F_{2i}(x_0) \oplus x_1, x_2).
\]

#### 4.1. Multiset Distinguishers of CLEFIA-Like and RC6-Like Structures

Choosing a fixed plaintext \( (a_0, a_1, a_2, a_3) \), we fulfill the following set of \( 2^m \) plaintexts which will help to find multiset distinguishers for CLEFIA/RC6-like structures.

\[
\mathcal{A} = \{(a_0, x \oplus a_1, a_2, a_3) \}: x \text{ ranges from } 0 \text{ to } 2^m - 1.
\]

The traces of these integrals through CLEFIA/RC6-like structures are depicted in Figure 4. Thus we build a multiset distinguisher \( ((\mathcal{C}, \mathcal{D}, \mathcal{E}, \mathcal{F}) \rightarrow (\mathcal{U}, \mathcal{B}, \mathcal{E}, \mathcal{F})) \) for CLEFIA/RC6-like structures. Consequently, we collect one equation

\[
\bigoplus_{x \in \{0,1\}^m} (F_{13} [c_3(x)] \oplus c_4(x)) = 0.
\]

Similarly, we can also prove that \( ((\mathcal{C}, \mathcal{D}, \mathcal{E}, \mathcal{F}) \rightarrow (\mathcal{U}, \mathcal{B}, \mathcal{E}, \mathcal{F})) \) is also legal for both of these two structures, which tells

\[
\bigoplus_{x \in \{0,1\}^m} (F_{12} [c_1(x)] \oplus c_2(x)) = 0.
\]

Next we can change the value of \( a_0, a_1, a_3 \) in the chosen plaintexts set and generate sufficiently linear equations. When enough linear equations are obtained, we can solve the linear system by Gaussian elimination to recover \( F_{12} \) and \( F_{13} \).

#### 4.2. Equivalent Structure and Rank Deficiency

In the equation-collection phase of CLEFIA/RC6-like structures, we cannot get a system of equations with a full rank of \( 2^m \). This rank deficiency phenomenon is caused by the existence of equivalent structures, more precisely, due to the fact that a given structure instance is not uniquely determined by its round functions.
Proposition 3. Let \((F_0, F_1, F_2, F_3, F_4, F_5)\) be a decomposition solution of 3-round CLEFIA-like (RC6-like, resp.) mapping; then for any constants \(a, b\), let

\[
\begin{align*}
  f_0 (x) &= F_0 (x) \oplus a, \\
  f_1 (x) &= F_1 (x) \oplus b, \\
  f_2 (x) &= F_2 (x \oplus a), \\
  f_3 (x) &= F_3 (x \oplus b), \\
  g_2 (x) &= F_2 (x \oplus b), \\
  g_3 (x) &= F_3 (x \oplus a), \\
  f_4 (x) &= F_4 (x) \oplus b, \\
  f_5 (x) &= F_5 (x) \oplus a,
\end{align*}
\]

then \((f_0, f_1, f_2, f_3, f_4, f_5) ((f_0, f_1, g_2, g_3, f_4, f_5), \text{resp.})\) is also a decomposition solution of CLEFIA-like (RC6-like, resp.) mapping.

Proof. By computing the encryption details of each cell in Figure 5, we can verify the correctness of this proposition directly.

Proposition 3 provides 3-round equivalents for each structure. Combining 3-round equivalents, we can get \(r\)-round \((r \geq 3)\) equivalent structures for these two ciphers: if \(F_2\) and \(F_{2+1}\) are replaced by \(F_2 \oplus a\) and \(F_{2+1} \oplus b\), respectively, then we can still keep the correctness of the whole structure by adding constants on rounds \(r-1\) and \(r-2\).

A natural question is, except for this type of equivalents, if there still exists any other equivalent structure, we have to be faced with the rank deficiency problem again. Since proving the nonexistence of equivalents is quite difficult, we tested this issue in an actual implementation of the attack for \(m = 4\). Fortunately, we always got a linear system of rank 15 in 16 variables, which indicates that the constant addition type is the unique equivalent structure of these two ciphers. Since the arbitrarily chosen \(a\) can be used to be added to the output of the “real” \(F_{13}\), the various solutions are simply equivalent keys which represent the same plaintext/ciphertext mapping.

4.3. Recover the Round Functions

Recover the Outermost Round \(F_{12}\) and \(F_{13}\). For the original structure, the outermost round consists of \(F_{12}\) and \(F_{13}\). In order to get \(2^m\) equations for \(F_{12}\) and \(F_{13}\), respectively, we
should use $2^{2m}$ chosen plaintexts of the form $(a_0, x, a_2, y)$, in which $a_0, a_2$ are constants. For each fixed $x$, we get a single equation of $F_{13}$ by varying $y$ through all the possible $2^n$ values. Also, we can get an additional equation of $F_{13}$ by fixing $y$ and varying $x$ through all the $2^n$ possible values.

Solving each system of linear equations by Gaussian elimination requires $2^{2.81m}$ steps, and thus we need $2 \times 2^{2.81m} = 2^{2.81m+1}$ steps to recover $F_{13}$ and $F_{15}$.

Recover $F_i$ (for $i = 1, 2, 3, 4$). Since we have found a way to recover the outermost round, i.e., $F_{12}$ and $F_{15}$, of these two 7-round structures, then for the inner rounds, we can use the dummy-round technique introduced in the last section to recover the rest round functions of CLEFIA/RC6-like structures.

According to the basic principle of dummy-round technique, we peel off the last round and add a dummy-round before the first round, and then we apply the “outermost-round recovery algorithm” to recover $F_{10}$ and $F_{11}$; then we repeat this process again and again, until all the internal round functions from rounds 6 to 3 are recovered.

It should be noticed that the shortest round numbers of equivalent structures of the two target structures are both 3. In decomposing these structures from rounds 6 to 3, we can still ignore the influence of the rank deficiency for $F_i$ (for $i = 1, 2, 3, 4$). Therefore, the total complexity of this procedure can be obtained by multiplying the number of rounds by the complexity of the outermost round decomposition, i.e., $4 \times 2^{2m}$ chosen plaintexts and $4 \times 2^{2.81m+1}$ steps.

Recover $F_0 \sim F_3$. For the remaining last two rounds, we are able to get $F_0, F_1, F_2, F_3$ by the plaintext-ciphertext comparison.

For 2-round CLEFIA structure, the encryption satisfies the system of equations (see Figure 6).

$$
F_0 (p_0) = p_1 \oplus c_3 \\
F_1 (p_2) = p_3 \oplus c_1 \\
F_2 (c_5) = p_2 \oplus c_0 \\
F_3 (c_1) = p_0 \oplus c_2
$$

(14)

And for 2-round RC6 structure, the similar system of equations can be obtained.

We need about $4 \times 2^n$ calls of the codebook to recover these 4 round functions.

If we use $n$ to denote the block size of the structure, i.e., $n = 4 \times m$, then the total time complexity is about $10 \times 2^{2.81m} = 10 \times 2^{2.7m}$ and the data complexity is about $5 \times 2^m = 5 \times 2^{0.5n}$. Our result will lead to a practical decomposition for the case of $n = 64$.

5. Summary

Structural attack is now a generic attack against secret-component based block ciphers. In this paper we propose an efficient decomposition algorithm for the generalized Feistel structure with bijective round functions. We use the integral property to find the outermost round and introduce the dummy-round technique to find the rest. This technique allows the final-round attack to be used on all the rounds left and does not depend on how the final round is recovered. Our attack provides a practical threat for 7-round CLEFIA-like and 7-round RC6-like ciphers with data length up to 64 bits. We believe that the new progress of the integral attacks, such as the division property [18] and cube attack [17], will lead to more efficient decompositions. Future work will concentrate on discovering more efficient decomposition algorithms.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.
Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant Nos. 61772547, 61402523, and 61272488).

References


