Linear Cryptanalysis of Baby Rijndael

Josef Kokeš
Czech Technical University in Prague,
Faculty of Information Technology,
Thákurova 9, 16000 Prague 6, Czech Republic
Email: josef.kokes@fit.cvut.cz

Róbert Lórencz
Czech Technical University in Prague,
Faculty of Information Technology,
Thákurova 9, 16000 Prague 6, Czech Republic
Email: robert.lorencz@fit.cvut.cz

***Abstract*—We present results of linear cryptanalysis of Baby Rijndael, a reduced-size model of Rijndael. The results were obtained using exhaustive search of all approximations and all keys and show some curious properties of both linear cryptanalysis and Baby Rijndael, particularly the existence of different classes of linear approximations with significantly different success rates of recovery of the cipher's key.**

***Index Terms*—Linear cryptanalysis, linear approximations, Baby Rijndael, key recovery, success rate**

I. Introduction

Since their introduction to the public cryptanalytic research, linear and differential cryptanalysis have been accepted by the cryptological community and are now the standard techniques to be used to analyze the security of existing and proposed ciphers alike. However, it seems that differential cryptanalysis [2] is the more successful – not only in itself, but also in its various modifications such as impossible differential cryptanalysis [4][3][15], boomerang attacks [17], related key attacks [5][6] or biclique attacks [7], which often represent the state-of-the-art in cryptanalysis of current ciphers.

Linear cryptanalysis, in comparison, can't boast such successes. When Matsui originally published it [14], it gained a significant critical acclaim and was successful in breaking a number of then-current cryptosystems (particularly DES [14], but also PRESENT [9] or GOST [16]), and some interesting extensions have been found such as zero-correlation linear cryptanalysis [8], but it seems that the success rate is somewhat smaller than that of its older companion technique.

We feel that the current research into linear cryptanalysis can be improved in several aspects. Particularly, we focus on actual experimental evaluation of Matsui's Algorithm 2 [14] as applied to Baby Rijndael [1], a reduced Rijndael-family cipher, where the technique exhibits several interesting properties: our experimental results show that the success rate of various linear approximations is highly dependent on their structure, not only on their probability bias, that certain key bits are more easily recoverable than other key bits, and that these key bits are easier to recover for some keys than for other keys.

In this paper, we will briefly describe Baby Rijndael, Matsui's Algorithm 2 and our modifications to get a better performance out of them. We show that for Baby Rindael, there exist many different optimal linear approximations, and present the experimental results which demonstrate the different success rate of these approximations under Matsui's Algorithm 2. We then modify the algorithm to recover only selected key bits and show experimental data which demonstrate how the success rate of recovery of different bits differs. Finally, we adapt multiple approximation analysis [11] to our data and present experimental results demonstrating the varying difficulty of recovering key bits for different keys.

II. Our Approach

In our study of the properties of linear cryptanalysis, we wanted to evaluate our subject cipher by experimental verification, to verify whether there are any unexpected parameters to the success rate of the technique used. Unfortunately, this is particularly difficult for modern ciphers which use key and block sizes much to large to allow for exhaustive tests of all possibilities. For this reason we decided to use a cipher which is not realistic in the sense that it allows for exhaustive processing, but at the same time relates to real-world ciphers in a significant way, so that the results obtained could potentially be extended to these ciphers. From among the available options, we chose Baby Rijndael by Cliff Bergman.

A. Baby Rijndael

Baby Rijndael is a block cipher proposed by Cliff Bergman [1] as an educational block cipher. It is modelled after Rijndael (AES), but with reduced key- and block-space: it uses 16-bit blocks and 16-bit keys. Its design, however, follows the design of the full Rijndael, respecting the requirements, implementations and design decisions set by Daemen and Rijmen in Rijndael proposal [10].

The state of the cipher is represented by a column-major 2×2 matrix A, where each element a_{ij} is a four-bit number. The state is initially filled by the plaintext and xor-ed by the key; then it undergoes three rounds of transformations consisting of a sequence of SubBytes, ShiftRows, MixColumns and AddRoundKey; the last, fourth round, omits the MixColumns transformation. The individual transformations, as well as the key schedule, are defined much like the same-named transformations of Rijndael, except for the size of the state and individual elements.

A detailed description of Baby Rijndael can be found in [12] (in Czech), a more complete version along with the discussion of the mapping of Baby Rijndael to Rijndael is a part of article [13] (currently under review). Here it suffices to say that Baby Rijndael follows all relevant requirements and design decisions

ISBN: 978-1-4799-8451-0/15/$31.00 ©2015 IEEE

of Rijndael's authors, if we apply them to a smaller cipher state.

B. Linear Cryptanalysis

Our approach to linear cryptanalysis follows Matsui's specifications for Algorithm 2 [14]:

We select such a linear approximation of Baby Rijndael which has the highest probability bias. Using this approximation, we describe the first three rounds of the cipher in the form

$$P[i_1, i_2, ..., i_a] \oplus D(C, K_4)[j_1, j_2, ..., j_b] = K[k_1, k_2, ..., k_k] \quad (1)$$

where $P[i_1, i_2, ..., i_a]$ is the xor of some bits of the plaintext, C is the ciphertext, K_r is the key in the last (r-th) round of the cipher, $D(C, K_r)[j_1, j_2, ..., j_b]$ is the xor of some bits of the cipher's inner state which is generated by performing one decryption round of the ciphertext using key K_r and $K[k_1, k_2, ..., k_k]$ is the xor of some bits of the (master) key. This approximation is satisfied with probability $P = 0.5 + \epsilon$, where ϵ is the linear probability bias of the approximation.

Then we can apply all candidates for key K_r^i against a given set of N plaintext-ciphertext samples encrypted by a fixed key K and for each candidate calculate the number of times T_i the equation 1 was satisfied. Let T_{max} be the maximum of all T_i and T_{min} the minimum. Then:

1) If $|T_{max} - \frac{N}{2}| > |T_{min} - \frac{N}{2}|$, then the recovered key is the key corresponding to T_{max} and furthermore guess that $K[k_1, k_2, ..., k_k]$ is 0 if $\epsilon > 0$ or 1 if $\epsilon < 0$.
2) If $|T_{max} - \frac{N}{2}| < |T_{min} - \frac{N}{2}|$, then the recovered key is the key corresponding to T_{min} and furthermore guess that $K[k_1, k_2, ..., k_k]$ is 1 if $\epsilon > 0$ or 0 if $\epsilon < 0$.

C. Adaptation of Baby Rijndael to Matsui's Algorithm 2

Early in our research when we adapted Baby Rijndael to Matsui's Algorithm 2, we noticed that can significantly reduce the number of non-linear operations used if we modify the description of the cipher so that 1) we switch the order of SubBytes and ShiftRows, and 2) combine two SubBytes and one MixColumns operation into one larger transformation "BigSub":

- We can perform this modification because SubBytes and ShiftRows are completely independent of each other: Each element of the state matrix is processed by SubBytes the same way regardless of its position, and ShiftRows only changes the positions of the whole elements, never of their parts, and never mixes the content of two or more elements.
- Each transformation block of a cipher can be considered a non-linear transformation, even though this is usually undesirable as it increases the number of non-linear elements and thus decreases bias.
- A sequence of transformations can be replaced by a composite transformation, as long as all input and all output bits of all member transformations are included in the composite transformation.

When we performed the linear analysis of the "BigSub" transformation, we discovered that the maximum reachable bias for it is $\pm\frac{1}{4}$, same as for the SubBytes transformation; that means that we can replace two non-linear transformations with one non-linear transformation (which improves the overall bias) while maintaining the bias of the components of the transformation (which doesn't change the overall bias), thus improving the overall bias.

Furthermore, we noticed in our earlier work that the success ratio of the recovery of the last-round key – that is, the probability that a key selected by the Algorithm 2 is the key actually used to encrypt the set of plaintexts – is relatively low. For that reason we decided not to use the guess for $K[k_1, k_2, ..., k_k]$ in our current work; that will be used only when the success rate becomes significantly high.

III. Results

During our research, we achieved several relevant results. We created a list of all linear approximations for the Baby Rijndael cipher. We evaluated all of them as to their ability to successfully recover the correct key. We then chose a set of approximations with the highest success rate and tried to improve the probability of success, finding several interesting facts along the way.

A. Success Rate of Baby Rijndael's Linear Approximations

We created a program which could generate all possible linear approximations of Baby Rijndael using an exhaustive search, and list those approximations with the highest achievable bias – we will call these the *optimal approximations*. We did, however, limit the search in these aspects:

- We only generated approximations for the first three rounds of the Baby Rijndael, in accordance with the requirements of Algorithm 2, where the last round is run "backwards" with all possible candidate keys.
- We only focused on approximation whose active bits end in two SubBytes blocks – we will call these the *active SubBytes*. This optimization is also done to facilitate Algorithm 2, as it requires an exhaustive search of all candidate keys and then an exhaustive search of the remaining bits of the full key. With key N bits long, the optimum separation for one-pass linear cryptanalysis is into $N/2$ bits for the candidate keys and $N/2$ bits for the remainder of the key, leading up to the overall complexity of $2 \cdot 2^{N/2}$. It is possible to perform multiple passes of linear cryptanalysis in sequence with fewer active SubBytes in each, but we left this approach for future study.

For Baby Rijndael with four existing SubBytes in each round, there are 6 different classes of linear approximations if we distinguish them by the active SubBytes: counting from left, first two SubBytes are active (we will denote this as the "1100" class), the first and the third is active ("1010") etc. The number of approximations belonging to each class is surprisingly regular, probably owing to the way the SubBytes transformation is constructed.

Then we generated 65536 sets of plaintext-ciphertext samples, one for each existing key, which we will call the *correct key*. In each sample set, there were 65536 plaintext-ciphertext pairs, i.e. our sample sets contained all possible plaintexts and respective ciphertexts for a given key.

Finally, we applied Algorithm 2 to every sample set to recover a last round key, which we transformed (using a look-up table) to an actual *recovered master key*. We then compared the recovered master key to the correct key, and calculated the *rank of the key* as the number of recovered master keys we would have to try before we could reach the correct key; that is, if recovered master key equals the correct key, then the rank of the key is 1; if the keys do not match, the rank of the key would be 2, 3 and so on, up to 256 (the correct key was the last key suggested by Algorithm 2, in other words, the algorithm completely failed).

This calculation was done for all approximations of the "1010", "1001", "0110" and "0101" classes and to 200 randomly selected[1] approximations of the "1100" and "0011" class. The results are shown in table I.

We can see that the success rate of the approximations, expressed as the rank of the correct key, significantly depends on the active SubBytes of the approximation: While approximations of the "0011", "0110", "1001" and "1100" class are only a little better than a random guess (which would yield the rank of 128), approximations of the "0101" and "1010" class achieve a much better rank. This is contrary to the expectation that all linear approximations with the same bias should be interchangeable as far as their success rate is concerned. It is as yet unclear why Baby Rijndael's approximations with alternating active and passive SubBytes should prove so much better than the other approximation.

For our subsequent tests, we selected the class of approximation "0101", which performed equally well as the "1010" class in the average case and marginally better in the median case.

B. Success Rate of the Approximations in the "0101" Class

For all linear approximations in the "0101" class, we evaluated the success rate over all existing keys. The best approximation achieved the average rank of 40.27, significantly better than the 57.40 rank of the worst approximation. However, this difference only translates to 0.52 bits of complexity saved by the best approximation over the worst approximation. You can see the best and the worst approximations in table II.

It is interesting to note that if we consider all 48 approximations, the standard deviation of the rank tends to grow with the average value of the rank. The correlation between the two statistics is not perfect, but it is very high, with Pearson's correlation coefficient of 0.9865.

However, when we consider the average number of correctly recovered bits of the key, that is, the number of bits in the recovered master key which match their respective bits in the correct key, the results are rather discouraging, the number of recovered bits varies between 4.896 (best) and 4.398 (worst) – in other words, on average we are only able to recover a little over one half of the bits correctly, which compares poorly to random guessing. See table III for details.

We observed that the rank of the correct key often is not 1, as expected by the linear cryptanalysis theory, but rather a different value. It is not unreasonable to expect, though, that the candidate keys with better ranks would have more of their bits correctly guessed than candidate keys with worse ranks. In order to verify this expectation, we tried to calculate a weighted average of up to 10 highest ranking candidate keys, with weights assigned according to the distance of each particular T_i from N (notation as per Matsui's Algorithm 2 above). The number of bits correctly guessed using this method is shown in table IV.

The results rather convincingly show that the expectation above is probably not correct and averaging top-ranking candidate keys does not lead to improved accuracy of key recovery.

C. Recovery of individual key bits

We tried another approach to improve the accuracy of key recovery: We modified the Algorithm 2 to recover individual bits rather than the whole candidate key. The modification is straightforward: We perform Algorithm 2 as usual, but only consider a subset of the bits of the recovered master key. We then evaluate how these bits match against the respective bits of the correct key, hopefully improving the accuracy.

We performed exhaustive testing of this idea on all the "0101" approximations, all keys and a complete plaintext-ciphertext sample set for each combination of an approximation and a key. We measured the probability that a given single bit of the recovered last-round key is correct, i.e. equal to the respective bit of the correct last-round key, across all possible keys. Unlike the previous experiments, here we only consider the match between the bits of the recovered last-round key and the correct last-round key; the rank of the correct last-round key is without meaning in this approach.

Table V shows that this approach is promising, in several important aspects:

- There are significant variations between individual approximations, further supporting the expectation that there are other factors to an approximation's power than just the bias.
- Some key bits are easier to recover than other bits, which suggests that there may exist a set of keys in the Rijndael family of ciphers which are more susceptible to the linear cryptanalysis. When taken as a whole, bits 3 and 11 are the easiest to recover - among the 20 most successful approximations, 8 recovered bit 11, 7 recovered bit 3, 3 recovered bit 0 and 2 recovered bit 8. The most successful approximation for recovery of bit 10 was 29th, etc.
- With the best approximations, we can achieve a significantly higher success rate than when recovering the entire candidate key. This means that recovery of individual bits is indeed possible, and these recovered bits can then be

[1] Using a cryptographically secure PRNG

TABLE I

SUCCESS RATE OF BABY RIJNDAEL'S LINEAR APPROXIMATIONS. THE LOWER THE AVERAGE RANK, THE BETTER CAN ALGORITHM 2 RECOVER THE CORRECT KEY. THE VALUE OF ONE-HALF OF THE NUMBER OF CANDIDATE KEYS IS THE WORST CASE, THE LINEAR APPROXIMATIONS CAN'T DETERMINE THE CORRECT KEY BETTER THAN A RANDOM GUESS.

	Active SubBytes					
	0011	0101	0110	1001	1010	1100
Optimal approximation's probability bias	$\pm\frac{1}{256}$	$\pm\frac{1}{256}$	$\pm\frac{1}{256}$	$\pm\frac{1}{256}$	$\pm\frac{1}{256}$	$\pm\frac{1}{256}$
Nr. of opt. approximations	3840	48	48	48	48	3840
Nr. of candidate keys	256	256	256	256	256	256
Average rank of the correct key	114.75	49.58	111.91	111.90	49.58	114.72
Median rank of the correct key	114.91	49.85	111.77	111.65	49.88	115.08
Std. deviation of the correct key	2.89	6.03	2.23	2.32	6.03	2.77

TABLE II

SUCCESS RATE OF BABY RIJNDAEL'S LINEAR APPROXIMATIONS OF THE "0101" CLASS. THE LOWER THE AVERAGE RANK, THE BETTER CAN ALGORITHM 2 RECOVER THE CORRECT KEY. "INNER STATE" REPRESENTS THE BITS AT THE BEGINNING OF THE LAST ROUND, AFTER PERFORMING THE SHIFTROWS TRANSFORMATION. THE ACTIVE BITS ARE COUNTED FROM THE RIGHT, THAT IS, THE LEFT-MOST BIT IS NUMBER 15, THE RIGHT-MOST BIT IS NUMBER 0.

Active bits		Average	Std. dev.
Plaintext	Inner state	rank	of rank
Best approximations			
0, 2, 3	1, 2, 9, 10, 11	40.27	46.72
12, 14, 15	1, 2, 9, 10, 11	40.37	46.82
8, 10, 11	1, 2, 3, 9, 10	40.38	46.84
4, 6, 7	1, 2, 3, 9, 10	40.43	46.85
Worst approximations			
4, 6, 7	0, 1, 8, 11	57.20	69.50
4, 6	0, 3, 9, 11	57.20	70.50
12, 14, 15	0, 3, 8, 9	57.25	69.55
0, 2, 3	0, 3, 8, 9	57.40	69.65
Average over all 48 approximations			
		49.58	61.04

TABLE III

THE AVERAGE NUMBER OF BITS OF THE CORRECT KEY RECOVERED BY BABY RIJNDAEL'S LINEAR APPROXIMATIONS OF THE "0101" CLASS. THE HIGHER THE NUMBER THE BETTER. "INNER STATE" AND "ACTIVE BITS" ARE DEFINED AS IN TABLE II.

Active bits		Average
Plaintext	Inner state	rank
Best approximations		
0	0, 3, 8, 9	4.895
0, 2	1, 3, 8, 11	4.877
8, 10	0, 3, 9, 11	4.876
8	0, 1, 8, 11	4.875
Worst approximations		
12, 14	3, 9, 10	4.421
4, 6	1, 2, 11	4.414
8, 10	1, 2, 11	4.399
0, 2	3, 9, 10	4.398
Average over all 48 approximations		
		4.632

TABLE IV

THE AVERAGE NUMBER OF BITS OF THE CORRECT KEY RECOVERED BY BABY RIJNDAEL'S LINEAR APPROXIMATIONS OF THE "0101" CLASS, IF WE CALCULATE A WEIGHTED AVERAGE OF BITS OF A GIVEN NUMBER OF THE TOP-RANKING CANDIDATE KEYS. THE HIGHER THE NUMBER THE BETTER.

Number of candidate keys used	Average number of bits recovered		
	Best case	Worst case	Average case
1	4.895	4.398	4.632
2	3.345	2.834	3.070
3	4.806	4.385	4.566
4	3.855	3.367	3.604
5	4.111	3.683	3.882
6	2.918	2.485	2.691
7	2.506	2.143	2.319
8	1.610	1.302	1.449
9	1.268	1.022	1.142
10	0.768	0.589	0.677

ISBN: 978-1-4799-8451-0/15/$31.00 ©2015 IEEE

TABLE V
THE PROBABILITY OF RECOVERY OF INDIVIDUAL BITS OF THE KEY, WHEN CALCULATED AS THE PROBABILITY THAT THE GIVEN BIT IN A RECOVERED LAST-ROUND KEY IS CORRECT ACROSS ALL POSSIBLE KEYS.

Active bits		Recovered	Probability
Plaintext	Inner state	bit	of recovery
Best approximations			
0, 2	1, 3, 8, 11	3	0.703
8, 10	0, 3, 9, 11	11	0.701
4	0, 1, 8, 11	3	0.683
12	0, 3, 8, 9	11	0.682
0	0, 3, 8, 9	11	0.682
8	0, 1, 8, 11	3	0.679
12, 14, 15	0, 3, 8, 9	11	0.677
10, 11	1, 2, 3, 9, 10	0	0.676
Worst approximations			
12, 14, 15	1, 2, 9, 10, 11	11	0.511
4, 6, 7	1, 2, 3, 9, 10	3	0.510
8, 10	1, 2, 11	11	0.493
0, 2	3, 9, 10	3	0.491

used to facilitate recovery of other bits, possibly using different cryptanalytic techniques.

IV. DISCUSSION AND CONCLUSION

Thanks to the small size of Baby Rijndael, we were able to perform an exhaustive analysis of a number of its aspects. We were able to find all optimal linear approximations of the cipher and discovered that there is a great number of these optimal approximations. This will be useful further on, when we consider the combined effect of multiple linear approximations on the same sample set – as Kaliski and Robshaw suggest [11], we can use multiple approximations to generate a new statistic for our set of candidate keys, one which would reduce variance of the result and thus decrease the size of the required sample set. We didn't use this approach in this paper, but we do have preliminary results which seem quite promising; we will definitely proceed with research in this direction.

An interesting aspect of the optimal linear approximation is their varying ability to recover the master key. We can divide the approximations into several classes, and through exhaustive testing of all possible keys we showed that approximations from some of these classes are, on average, significantly more successful than approximations from other classes. This suggests that, when we consider linear cryptanalysis, we need to pay attention not only to the probability bias, but also the choice among several possible approximations. We don't as yet know why the classes "0101" and "1010" are so much better than the others, but we intend to find out.

Even within a class of approximations, the variations in key recovery abilities are quite surprising. We would definitely like to precisely measure, what makes the "good" approximations different from the "bad" ones. If we could discover some metric which would let us choose the best approximation in advance, without having to perform intensive calculations, it would certainly be a useful result in its own.

It came as a distinct surprise to us that individual key bits show quite a large difference in their ability of being successfully recovered. We weren't able to formulate the reasons behind this behavior so far, but this is also one of our research targets. We are particularly eager to continue our research in this area, because we would like to explore our idea of using information revealed by one cryptanalytic technique (e.g. linear cryptanalysis in this case) to enhance the power of another technique (e.g. algebraic cryptanalysis) – and then inject the results back to the original technique. It seems possible to achieve synergistic effects here.

The ultimate goal, of course, isn't cryptanalysis of Baby Rijndael, however interesting it may be. We hope, though, that the principles we discover will eventually allow us to attack even the full Rijndael itself – or, failing that, at least give us some idea of which approaches do have a hope of succeeding and which do not.

REFERENCES

[1] Bergman, C.: A Description of Baby Rijndael. Iowa State University, 2005.
[2] Biham, E., Shamir, E.: Differential Cryptanalysis of DES-like Cryptosystems. Lecture Notes in Computer Science Volume 537, 1991, pp 2-21.
[3] Biham, E.: Impossible cryptanalysis of Skipjack. CRYPTO '98, 1998.
[4] Biryukov, A.: Miss-in-the-middle attacks on IDEA. CRYPTO '98, 1998.
[5] Biryukov, A., Khovratovich, D.: Related-key Cryptanalysis of the Full AES-192 and AES-256. Lecture Notes in Computer Science Volume 5912, 2009, pp 1-18.
[6] Biryukov, A., Dunkelman, O., Keller, N., Khovratovich, D., Shamir, A.: Key Recovery Attacks of Practical Complexity on AES Variants With Up To 10 Rounds. Cryptology ePrint Archive, Report 2009/374, 2009.
[7] Bogdanov, A., Khovratovich, D., Rechberger, C.: Biclique Cryptanalysis of the Full AES. Advances in Cryptology – ASIACRYPT 2011.
[8] Bogdanov, A., Wang, M.: Zero Correlation Linear Cryptanalysis with Reduced Data Complexity. Lecture Notes in Computer Science Volume 7549, 2012, pp 29-48.
[9] Cho, J. T.: Linear Cryptanalysis of Reduced-Round PRESENT. Lecture Notes in Computer Science Volume 5985, 2010, pp 302-317.
[10] Daemen, J., Rijmen, V.: The design of Rijndael: AES – the Advanced Encryption Standard. Springer-Verlag, 2002, ISBN 3-540-42580-2.
[11] Kaliski, B., Robshaw, M.: Linear Cryptanalysis Using Multiple Approximations. Advances in Cryptology – CRYPTO '94. Lecture Notes in Computer Science Volume 839, 1994, pp 26-39.
[12] Kokeš, J.: Cryptanalysis of Baby Rijndael. Diploma thesis, Faculty of Information Technology, Czech Technical University in Prague, 2013.
[13] Kokeš, J., Lórencz, R.: Properties of Baby Rijndael Under Linear Cryptanalysis. 2015 (Under review.)
[14] Matsui, M.: Linear Cryptanalysis Method for DES Cipher. Lecture Notes in Computer Science 765, 1994, ISBN 978-3-540-57600-6, pp 386-397.

[15] Shamir, A.: Impossible differential attacks. CRYPTO '98, 1998.
[16] Shorin, V., Jelezniakov, V., Gabidulin, E.: Linear and Differential Cryptanalysis of Russian GOST. Electronic Notes in Discrete Mathematics Volume 6, April 2001, pp 538547.
[17] Wagner, D.: The Boomerang Attack. 6th International Workshop on Fast Software Encryption (FSE '99). Rome: Springer-Verlag. pp. 156170.

ISBN: 978-1-4799-8451-0/15/$31.00 ©2015 IEEE