Mathematical Backdoors in Symmetric Encryption Systems
Proposal for a Backdoored AES-like Block Cipher

Arnaud Bannier and Éric Filiol
Operational Cryptology and Virology Lab, ESIEA, 38 rue des Drs Calmette et Guérin, 53000 Laval, France

Keywords: Cryptography, Encryption Algorithms, Backdoor, Trapdoor, Cryptanalysis, Block Cipher, AES.

Abstract: Recent years have shown that more than ever governments and intelligence agencies try to control and bypass the cryptographic means used for the protection of data. Backdoor encryption algorithms is considered as the best way to enforce cryptographic control. Until now, only implementation backdoors (at the protocol/implementation/management level) are generally considered. In this paper we propose to address the most critical issue of backdoors: mathematical backdoors or by-design backdoors, which are put directly at the mathematical design of the encryption algorithm. While the algorithm may be totally public, proving that there is a backdoor, identifying it and exploiting it, may be an intractable problem. We intend to explain that it is probably possible to design and put such backdoors. Considering a particular family (among all the possible ones), we present BEA-1, a block cipher algorithm which is similar to the AES and which contains a mathematical backdoor enabling an operational and effective cryptanalysis. The BEA-1 algorithm (80-bit block size, 120-bit key, 11 rounds) is designed to resist to linear and differential cryptanalyses. A challenge will be proposed to the cryptography community soon. Its aim is to assess whether our backdoor is easily detectable and exploitable or not.

1 INTRODUCTION

Despite the fact that in the late 90s/early 2000s, citizens have partially obtained the freedom for using cryptography, the recent years have shown that more than ever, governments and intelligence agencies still try to control and bypass the cryptographic means used for the protection of data and of private life. Snowden’s leaks were a first upheaval. A tremendous number of secret projects (from NSA, GCHQ) have been revealed to the public opinion which proves this situation.

While the need for the security of everyday life activities (for companies, for citizens) requires more and more cryptography, recent bothering initiatives by political decision-makers ask for an even stronger control over cryptography not to say preparing the simple prohibition or ban of cryptographic application such as telegram. At the same time, the EU as well as a number of security agencies (such as French ANSSI, German BSI) confirmed that it was nonsense and militate for the mandatory use of end-to-end encryption.

The recurring approaches and attempts consist in making the implementation of backdoors mandatory. The simplest and naïve approach consists in enforcing key escrowing at the operators’ level. But point-to-point encryption solutions (which are not equal to end-to-end encryption) like Telegram or Proton mail enable to prevent it. A number of different backdoor techniques are regularly mentioned or proposed.

The most critical aspect in implementation backdoors lies on the fact that hackers or analysts may find them more or less easily and worse may exploit them. This is the reason why it is likely that IT operators or developers are very reluctant to accept backdoors until now. In case of leak, they will inevitably lose users’ confidence and favor the development of trusted services abroad. In fact, the backdoor issue arises due to the fact that only implementation backdoors (at the protocol/implementation/management level) are generally considered.

In this paper we address the most critical issue of backdoors: mathematical or by-design backdoors. In other words, the backdoor is put directly at the mathematical design of the encryption algorithm. While the algorithm may be totally public, proving that there is a backdoor, identifying it and exploiting it, may be an intractable problem, unless you know the backdoor. To some extent, the RSA’s Dual_EC_DRBG standard case falls within this category (Shumow and Fergu-
son, 2007). Other non-public examples are known within the military cryptanalysis community, and partially revealed to the public thanks to the 1995 Hans Buehler case (Strehle, 1994). This kind of backdoor is the most difficult one to address and there is quite no public work on that topic. It is generally the technical realm of a few among the most eminent intelligence agencies (namely NSA, GCHQ, SVR/GRU) which moreover have the ability and power to step in and to influence the international standardization processes.

We intend to explain that it is probably possible to design and put such backdoors. Considering a particular case of mathematical backdoors (among all the possible ones) based on our previous work (Bannier et al., 2016), we present a block cipher algorithm which is similar to the AES and which contains a mathematical backdoor enabling an operational and effective cryptanalysis (in other words in a limited time on a modern desktop computer and with a limited number of plaintext/ciphertext pairs). This block cipher algorithm (80-bit block, 120-bit key size, 11 rounds) is designed to resist to linear and differential cryptanalyses.

This paper is organized as follows. In Section 2 we explore the concept of backdoors and trapdoors and we identify two main categories, each containing itself subcategories depending on the nature of the cipher (stream or block ciphers). This observation is backed by the personal experience of the second author as a military cryptanalyst. We also present the state-of-the-art, history and previous work regarding backdoors, mostly in symmetric cryptography. In Section 3, we present our backdoored block cipher algorithm BEA-1 (standing for Backdoored Encryption Algorithm 1), based on our work (Bannier et al., 2016). This is a particular family of trapdoors using a suitable partition of the plaintext and ciphertext spaces. In Section 4, we discuss the cryptographic security of this cipher, with respect to linear and differential cryptanalyses. We also propose a cryptographic challenge to the cryptography community\(^1\), regarding the backdoor identification and exploitation. We suppose that this backdoor is likely to be detected. Such a challenge should enable to prove or disprove this claim. Lastly we conclude in Section 5 and explore future work.

\(^1\)This challenge will be launched on the Arxiv repository right after the conference, at the end of February.

2 THE CONCEPT OF BACKDOOR

2.1 Definition and Classification

Proposal

Trapdoors are a two-face, key concept in modern cryptography. It is primarily related to the concept of “trapdoor function” — a function that is easy to compute in one direction, yet difficult to compute in the opposite direction without special information, called the “trapdoor”. This first “face” relates most of the time to asymmetric cryptography algorithms. It is a necessary condition to get reversibility between the sender/receiver (encryption) or the signer/verifier (digital signature). The trapdoor mechanism is always fully public and detailed. The security and the core principle is based on the existence of a secret information (the private key) which is essentially part of the trapdoor. In other words, the private key can be seen as the trapdoor.

The second “face” of the concept of trapdoor relates to the more subtle and perverse concept of “mathematical backdoor” and is a key issue in symmetric cryptography (even if the issue of backdoors may be extended to asymmetric cryptography; see for example the case of the \( \text{DUAL_BECDBG} \) (Shumow and Ferguson, 2007), or the case of trapdoored primes addresses recently in (Fried et al., 2016)).

In this case, the aim is to insert hidden mathematical weaknesses which enable one who knows them to break the cipher. If possible, these weaknesses should be independent of the secret key. Somehow, it consists to create a hidden asymmetry to the detriment of the legitimate users of the communication and to the benefit of the eavesdropper. In this context, the existence of a backdoor is a strongly undesirable property.

In the rest of the present section, we will oppose the term of trapdoor (desirable property) to that of backdoor (undesirable property). While the term of trapdoor has been already used in the very few literature covering this second face of this problem, we suggest however to use the term of backdoor to describe the issue of hidden mathematical weaknesses. This would avoid ambiguity and maybe would favor the research work around a topic which is nowadays mostly addressed by governmental entities in the context of cryptography control and regulations.

Inserting backdoors in encryption algorithms underlies quite systematically the choice of cryptographic standards (DES, AES…). The reason is that the testing, validation and selection process is always conducted by governmental entities (NIST or equivalent) with the technical support of secret entities (NSA or equivalent). So an interesting and critical
research area is: “how easy and feasible is it to design and to insert backdoors (at the mathematical level) in encryption algorithms?”. In this paper, we intend to address one very particular case of this question. It is important to keep in mind that a backdoor may be itself defined in the following two ways.

- As a “natural weakness” known — but non disclosed — only by the tester/validator/final decision-maker (e.g. the NSA as it could have been the case for the AES challenge). The best historic example is that of the differential cryptanalysis. Following Biham and Shamir’s seminal work in 1991 (Biham and Shamir, 1993), NSA acknowledged that it was aware of that cryptanalysis years ago (Schneier, 2000). Most of experts estimate that it was nearly 20 years ahead. However a number of non public, commercial block ciphers in the early 90s may be weak with respect to differential cryptanalysis.

- As an intended design weakness put by the author of the algorithm. To the authors knowledge, there is no known cases for public algorithms yet.

As far as symmetric cryptography is concerned, there are two major families of cipher systems for which the issue of backdoor must be considered differently.

- **Stream ciphers.** Their design complexity is rather low since they mostly rely on algebraic primitives (LFSRs and Boolean functions which have intensely been studied in the open literature). Until the late 70s, backdoors relied on the fact that quite all algorithms were proprietary and hence secret. It was then easy to hide non primitive polynomials, weak combining Boolean functions. The Hans Buehler case in 1995 (Strehle, 1994) shed light on that particular case.

- **Block ciphers.** This class of encryption algorithms is rather recent (end of the 70s for the public part). They exhibit so a huge combinatorial complexity that it is reasonable to think to backdoors. As described in (Daemen and Rijmen, 2002) for a $k$-bit secret key and a $m$-bit input/output block cipher there are $(2^m)!/2^k$ possible such block ciphers. For such an algorithm, the number of possible internal states is so huge that we are condemned to have only a local view of the system, that is, the round function or the basic cryptographic primitives. We cannot be sure that there is no degeneration effect at a higher level. This point has been addressed in (Daemen and Rijmen, 2002) when considering linear cryptanalysis. Therefore, it seems reasonable to think that this combinatorial richness of block cipher may be used to hide backdoors.

Since block ciphers are the most widely used encryption algorithms nowadays by the general public and the industry, we will focus on them in the rest of the paper. Backdoors in stream ciphers have quite never been exposed to the public.

### 2.2 Previous Work

One of the first trapdoor cipher was proposed in 1997 by Rijmen and Preneel (Rijmen and Preneel, 1997). The S-boxes are selected randomly and then modified to be weak to the linear cryptanalysis. They are finally applied to a Feistel cipher such as CAST or LOKI91. But because of the big size of the S-boxes, the linear table of such an S-box cannot be computed. However the knowledge of the trapdoor gives a good linear approximation of the S-boxes which is then used in a linear cryptanalysis. As an example, the authors created a 64-bit block cipher based on CAST cipher, and four $4 \times 32$ S-boxes. If the parameters of the trapdoors are known, a probabilistic algorithm allows to recover the key easily. Such a family of trapdoor ciphers leads to recover only a part of the key, and the authors claim that the trapdoor is undetectable. But in (Wu et al., 1998), Wu and al. discovered a way to recover the trapdoor if the attacker knows its global design but not the parameters. They also showed that there exists no parameter allowing to hide the trapdoor. Nevertheless, it is worthwhile to mention that in practice, if a real cipher containing a trapdoor is given, the presence of the trapdoor will certainly not be revealed.

In (Paterson, 1999), a DES-like trapdoor cipher exploiting a weakness induced by the round functions is presented. The group generated by the round functions acts imprimitively on the message space to allow the design of the trapdoor. In other words, this group preserves a partition of the message space between input and output of the round function. Such a construction leads to the design of a trapdoor cipher composed of 32 rounds and using a 80 bits key. The knowledge of the trapdoor allows to recover the key using $2^{13}$ operations and $2^{25}$ plaintexts. Even if the mathematical material to build the trapdoor is given, no general algorithm is detailed to construct such S-boxes. Furthermore, as the author says, S-boxes using these principles are incomplete (half of the cipher text bits are independent of half of the plaintext bits). Finally, the security against the differential attack is said not as high as one might expect.

More recently in (Angelova and Borissov, 2013), the authors created non-surjective S-boxes embedding a parity check to create a trapdoor cipher. The mes-
sage space is thus divided into cosets and leads to create an attack on this DES-like cipher in less than $2^{23}$ operations. The security of the whole algorithm, particularly against linear and differential cryptanalyses is not given and the authors admit that their attack is dependent on the first and last permutation of the cipher. Finally, the non-surjective S-boxes may lead to detect easily the trapdoor by simply calculating the image of each input vector. This problem is naturally avoided in a Substitution-Permutation Network (SPN) in which S-boxes are bijective by definition.

In a slightly different context, Caranti and al. answer to Patterson’s question by the affirmative in (Caranti et al., 2006), by proving that the imprimitivity of the group generated by round functions is actually related to the cosets of a linear subspace. They also give some conditions to create such a primitive group to design a secure cipher that cannot contain such trapdoor, and finally show that the AES respects these conditions. They add in (Caranti et al., 2009) an algorithm to verify this last condition simply and show that AES and Serpent S-boxes verify this property.

3 DESCRIPTION OF BEA-1

The algorithm BEA-1 (standing for Backdoored Encryption Algorithm version 1) is based on our research work on partition-based trapdoors (Bannier et al., 2016). This section is intended to describe this cipher precisely. The cipher operates on 80-bit data blocks using a 120-bit master key. Our algorithm is directly inspired by Rijndael (Daemen and Rijmen, 2002), the block cipher designed by Joan Daemen and Vincent Rijmen which is now known as the AES (the encryption standard proposed by the USA, under the auspices of NIST and NSA (NIST, 2000)). Consequently, our cipher is a Substitution-Permutation Network.

The encryption consists in applying eleven times a simple keyed operation called *round function* to the data block. A different 80-bit round key is used for each iteration of the round function. Since the last round is slightly different and uses two round keys, the encryption requires twelve 80-bit round keys. These round keys are derived from the 120-bit master key using an algorithm called *key schedule* (depicted in Figure 1).

The round function in Figure 2 is made up of three distinct stages: a *key addition*, a *substitution layer* and a *diffusion layer*. The key addition is just a bitwise XOR between the data block and the round key. The substitution layer consists in the parallel evaluation of four different S-boxes and is the only part of the cipher which is not linear or affine. Following the design principles of the AES, the diffusion layer comes in two parts: the *ShiftRows* and the *MixColumns* operations.

The decryption is straightforward from the encryption since all the components are bijective. Thus, to decrypt, we just have to apply the inverse operations in the reverse order. Remark that the key addition and the *ShiftRows* are involutions, therefore the same operations are used in the decryption process. In contrast to the AES, the algorithm works with bundles of 10 bits instead of 8 bits. Let $\mathbb{F}_2$ denote the Galois Field of order 2. Any 10n-bit block $x$ is seen as n-tuple of 10-bit bundles $(x_0, \ldots, x_{n-1})$, thus as an element of $(\mathbb{F}_2^{10})^n$. The hexadecimal notation is used to denote any 10-bit bundle. For example, $37A$ stands for 1101111010 in $\mathbb{F}_2^{10}$.

The S-boxes $S_0$, $S_1$, $S_2$ and $S_3$ are four permutations of $\mathbb{F}_2^{10}$ given in Appendix. The linear map $M : (\mathbb{F}_2^{10})^4 \rightarrow (\mathbb{F}_2^{10})^4$ processes four 10-bit bundles. Because of the linearity of this map, $M$ is only defined on the standard basis of $(\mathbb{F}_2^{10})^4$. For convenience, its inverse $M^{-1}$ is also given in Appendix.

A pseudo-code for the key schedule is given in Algorithm 1. To provide an overview of its structure, the first step is represented in Figure 1. This representation also emphasizes the similarities between our key schedule and the AES one. The pseudo-code for the encryption and decryption functions are respectively given in Algorithms 2 and 3. The notation $a \parallel b$ denotes the concatenation of the vectors $a$ and $b$. Again, an overview of the round function is given in Figure 2.

4 CRYPTOGRAPHIC SECURITY

ANALYSIS OF BEA-1

4.1 Differential and Linear Cryptanalyses

In (Daemen and Rijmen, 2002), Daemen and Rijmen introduced the differential and the linear branch numbers of a linear transformation. With an exhaustive search, it can be checked that the differential and linear branch numbers of $M$ are both equal to 5, which is the maximum. This implies that any 2-round trail has at least 5 active S-boxes. Thus, a 10-round trail involves at least 25 active S-boxes.

Note that all the S-boxes are (at most) differentially 40-uniform and linearly 128-uniform. Therefore, the probability of any 10-round differential trail is upper bounded by $\frac{40}{128}^{25} \approx 2^{-116.9}$ and the absolute bias of a 10-round linear trail is upper bounded
Algorithm 1 - ExpandKey
Input. The 120-bit master key \( K = (K_0, \ldots, K_{11}) \in (\mathbb{F}_{12}^5)^{12} \).
Output. The twelve 80-bit round keys \( K^0, \ldots, K^{11} \in (\mathbb{F}_{12}^5)^{80} \).

1. \((k_0, \ldots, k_{11}) \leftarrow (K_0, \ldots, K_{11}) \)
2. For \( i \) from 0 to 11 do
   3. \( x \leftarrow \text{M} \left( k_i \right) \oplus k_{i-1} \)
   4. \( x \leftarrow (k_i \oplus 2^8) \times x \)
   5. \( (k_{2i}, \ldots, k_{2i+1}) \leftarrow \left( k_{2i+1}, \ldots, k_{2i} \right) \)
   6. \( (k_{i+1}, \ldots, k_{i+10}) \leftarrow \left( k_{i+10}, \ldots, k_{i+1} \right) \)
   7. \( (k_{i+11}, \ldots, k_{i+11}) \leftarrow \left( k_{i+11}, \ldots, k_{i+10} \right) \)
   8. \( (k_{i+11}, \ldots, k_{i+11}) \leftarrow \left( k_{i+11}, \ldots, k_{i+10} \right) \)
9. For \( i \) from 0 to 11 do
10. \( \text{k} \leftarrow (k_0, \ldots, k_{11}) \)

Algorithm 2 - Encrypt
Input. The 120-bit master key \( K \in (\mathbb{F}_{12}^5)^{12} \) and the 80-bit plaintext block \( p \in (\mathbb{F}_{12}^5)^{80} \).
Output. The 80-bit ciphertext block \( c \in (\mathbb{F}_{12}^5)^{80} \).

1. \( \text{k} \leftarrow c \)
2. For \( i \) from 0 to 11 do
3. \( x \leftarrow \text{S} \left( k_i \right) \oplus x \)
4. \( x \leftarrow (k_i \mod 4) \times x \)
5. \( x \leftarrow M \left( k_i \times x \right) \times x \)
6. \( x \leftarrow M \left( k_i \right) \times x \)
7. \( x \leftarrow (k_i \mod 4) \times x \)
8. \( x \leftarrow (k_i \mod 4) \times x \)
9. \( x \leftarrow (k_i \mod 4) \times x \)
10. \( x \leftarrow (k_i \mod 4) \times x \)
11. \( x \leftarrow (k_i \mod 4) \times x \)
12. \( x \leftarrow (k_i \mod 4) \times x \)

Algorithm 3 - Decrypt
Input. The 120-bit master key \( K \in (\mathbb{F}_{12}^5)^{12} \) and the 80-bit ciphertext block \( c \in (\mathbb{F}_{12}^5)^{80} \).
Output. The 80-bit plaintext block \( p \in (\mathbb{F}_{12}^5)^{80} \).

1. \( \text{k} \leftarrow c \)
2. For \( i \) from 0 to 11 do
3. \( x \leftarrow \text{S} \left( k_i \right) \oplus x \)
4. \( x \leftarrow (k_i \mod 4) \times x \)
5. \( x \leftarrow M \left( k_i \times x \right) \times x \)
6. \( x \leftarrow M \left( k_i \right) \times x \)
7. \( x \leftarrow (k_i \mod 4) \times x \)
8. \( x \leftarrow (k_i \mod 4) \times x \)
9. \( x \leftarrow (k_i \mod 4) \times x \)
10. \( x \leftarrow (k_i \mod 4) \times x \)
11. \( x \leftarrow (k_i \mod 4) \times x \)
12. \( x \leftarrow (k_i \mod 4) \times x \)

by \((128/57)^{25} = 2^{-50}\). Consequently, a differential cryptanalysis of the 10-round version of our cipher would require at least \(2^{17}\) chosen plaintext/ciphertext pairs and a linear cryptanalysis would require \(2^{20}\) known plaintext/ciphertext pairs.

Even if this is a rough approximation since it does not take into account the inter-column diffusion provided by the ShiftRows operation, it suffices to prove the cipher practical resistance against classical differential and linear cryptanalyses. In fact, there is only \(2^{30}\) different plaintext/ciphertext pairs for a fixed master key.

4.2 Statistical Analysis of BEA-1

Any cryptographic algorithm must behave as a random generator or at least must exhibit enough randomness properties. Therefore, its outputs for different classes of inputs must pass all the reference statistical tests. The most widely used is the NIST’s Statistical Test Suite (STS) (Rukhin et al., 2001).

We have performed the statistical analysis for BEA-1 with respect to all the tests which are implemented in STS. Our encryption has passed all the tests successfully. This result is of rather high importance since

- STS is the tool recommended by the US government to evaluate statistical properties of any secure encryption algorithm. It is explicitly mandatory to consider it in the industry.

- The presence of our backdoor remains statistically undetectable which proves that if statistical properties are a necessary condition for cryptographic security it is absolutely not a sufficient property. It may be bypassed by considering statistical simulation techniques (Filiol and Josse, 2007). Alge-
4.3 Cryptographic Challenge

We propose a cryptographic challenge whose aims is twofold:

- identifying and explaining what our backdoor consists in,
- exploiting this backdoor in the most efficient way (in terms of computing time, memory requirements, the number of required plaintext/ciphertext pairs).

We have run our own full cryptanalysis implementation several times. Each time, we retrieve the 120-bit key successfully.

This challenge will be officially launched right after the presentation of the present paper, on the arxiv.org. To take part, participants must send the following data to both authors (prior to any publication):

- the description of the backdoor,
- the description of the attack to exploit the backdoor successfully.
- the relevant source codes. They will help us to sort the different proposals with respect to, first the number of required plaintext/ciphertext pairs, second the computing time on our reference computer.

Incentive (non monetary) awards will be awarded to the three best attacks. Our attack as the reference solution will be presented in at the RusCrypto 2017 con-
ference around end of March 2017. Consequently, the challenge holds until this date. Moreover the best attack will be considered for publication in the Journal in Computer Virology and Hacking Techniques.

5 CONCLUSION AND FUTURE WORK

In this paper, we have proposed an AES-like encryption algorithm which contains a backdoor at its design level. This algorithm, named BEA-1, exhibits many of the desirable properties that any secure algorithm should. However, it is absolutely unsuitable for actually protection information. Indeed, we manage to break it with a rather limited amount of resources successfully.

While it is a humble, first step in a larger research work, it illustrates the issue of using foreign encryption algorithms which may contain such hidden weaknesses. The very final aim of our work is to prove that it is feasible to embed such undetectable intended weaknesses. It is consequently a critical issue to have a broader work conducted in this research area and we hope that other people will also consider it as such.

The next step will be to consider more sophisticated combinatorial structures.

REFERENCES


APPENDIX

The present appendix contains the different tables for the S-boxes (Figures 5, 6, 7 and 8), the linear map $M$ (Figure 3) and its inverse $M^{-1}$ (Figure 4). They can be copied and pasted for a practical implementation of the encryption algorithm.
x \mapsto M(x)

(001,000,000,000) \mapsto (112,1BC,36C,0C5)
(002,000,000,000) \mapsto (544,394,342,165)
(004,000,000,000) \mapsto (23F,15B,0C7,0A7)
(008,000,000,000) \mapsto (21F,11F,1E0,2B7)
(010,000,000,000) \mapsto (2D9,1A0,0C4,095)
(020,000,000,000) \mapsto (231,120,322,016)
(040,000,000,000) \mapsto (3C6,010,0EC,261)
(080,000,000,000) \mapsto (32C,199,2C5,07A)
(100,000,000,000) \mapsto (35C,13E,212,110)
(200,000,000,000) \mapsto (13E,20F,253,0BC)

x \mapsto M^{-1}(x)

(001,000,000,000) \mapsto (10B,221,09D,398)
(002,000,000,000) \mapsto (1AE,1E9,2CB,245)
(004,000,000,000) \mapsto (1E8,11E,05F,34A)
(008,000,000,000) \mapsto (080,04D,016,34C)
(010,000,000,000) \mapsto (0A0,337,3C5,2D4)
(020,000,000,000) \mapsto (322,3FD,3D5,0E5)
(040,000,000,000) \mapsto (002,246,282,380)
(080,000,000,000) \mapsto (1E9,3FE,238,329)
(100,000,000,000) \mapsto (0F5,1BD,210,210)
(200,000,000,000) \mapsto (2D8,209,353,243)

Figure 3: Specification of \( M \).

Figure 4: Specification of \( M^{-1} \).
Figure 5: Specification of the S-box $S_0$.

Figure 6: Specification of the S-box $S_1$. 

ForSE 2017 - 1st International Workshop on FORmal methods for Security Engineering
Figure 7: Specification of the S-box $S_3$.  

Figure 8: Specification of the S-box $S_3$.  

Mathematical Backdoors in Symmetric Encryption Systems - Proposal for a Backdoored AES-like Block Cipher