Secure Symmetric-Key Block Cipher Based on Generalized Finite Fields

Czesław Kościelny [1] 2005

Abstract

Taking into account the need for personal communications impervious to eavesdroppers, and commercial communications resistant to industrial espionage, the author shows how to build very strong and flexible, yet easily implemented symmetric-key block cipher, using an algebraic system, which can be a little "defected", and which is named a generalized finite field. The cipher presented works in CBC mode, may practically have any block length, any key length, and can be used for encrypting messages of any kind and of any size.

Keywords: Block ciphers, generalized finite fields, Maple.

1 Introduction

At the beginning of the silicon era technological applications of semiconductors in the form of pure crystalline germanium or silicon were very limited. The meaningful development of semiconductor electronics has begun only when the trace amounts of dopants, causing defects of the crystal's structure, to the silicon or germanium crystals have been added. It is possible to perceive some analogy between contemporary cryptography and the pre-semiconductor era in electronics: generally in all currently proposed and used cryptographic systems encrypting/decrypting procedures compute cryptograms corresponding to given plaintexts, and vice versa, using pure algebraic structures such as groups, rings and fields. It has been verified by the author that application in cryptographic operations of algebraic structures with small "defects" can have positive influence on the properties of ciphers. This approach makes the interrelation between the cryptogram and the secret key very complicated, brings about smooth dispersion of the statistical parameters of the message across the cryptogram and makes cryptanalysis remarkably difficult, even if cryptographic algorithms are very simple. Deficient-looking algebraic systems may also be successfully applied to build good random number generators.

[1] Academy of Management in Legnica, Poland, Faculty of Computer Science,
e mail: c.koscielny@wsm.edu.pl

Encrypting/decrypting procedures described in the paper employ operations in the system, named a generalized finite field and denoted as $gff(n^m)$, in which the multiplication table of non-zero elements may not be a Latin square.

The paper is organized as follows: Section 2 contains the definition of a generalized finite field $gff(n^m)$, Section 3 presents a method of constructing a block cipher based on the algebraic system discussed in Section 2 and Section 4 comprises some important remarks on the properties and the implementation of the cipher. The author feels obliged to notice that although the cryptographic procedures described seem to be very simple, the problem of practical verification of properties of the presented cipher consisted in carrying out difficult programming task, because to construct a strong cipher, one should use a generalized finite field extremely large in size, e.g. having the number of elements of order form 10^{100} to 10^{10000}, and even more. The one of many possible solutions of this task has been shown in detail in the Maple worksheet, which accompanies this paper.

2 Generalized Finite Fields

Let n be any integer ≥ 2, $m-$ an arbitrary integer $\geq\ 1$, $f(x)-$ an arbitrary polynomial of degree m over the ring Z_n. Next let

$$gff(n^m) = \langle \mathsf{F}[x],\ +,\ \cdot \rangle, \tag{1}$$

be an algebraic system consisting of the set $\mathsf{F}[x]$ of all n^m polynomials of degree d, $0 \leq d \leq m-1$, 0 included, over the ring $\mathrm{Z_n}$ and of operations of addition and multiplication of these polynomials. Operations on elements from $gff(n^m)$ are performed nearly exactly alike as in $GF(p^m)$: addition over the ring $\mathrm{Z_n}$, multiplication over the same ring modulo the polynomial $f(x)$.

It is easy to observe that $gff(n^m)$ fulfills the following set of axioms:

A1. The system $\langle \mathsf{F}[x],\ + \rangle$, is an abelian group.

A2. The system $\langle \mathsf{F}[x]^*,\ \cdot \rangle$ is an abelian pseudo-semigroup,
$\mathsf{F}[x]^* = \mathsf{F}[x]^* \setminus \{0\}$, where 0 is an additive identity element.

A3. $\forall a(x),\ b(x),\ c(x)\ \in \mathsf{F}[x]$
$(a(x) \cdot (b(x) + c(x)) = a(x) \cdot b(x) + a(x) \cdot c(x)) \wedge$
$((a(x) + b(x)) \cdot c(x) = a(x) \cdot c(x) + b(x) \cdot c(x)).$

The author had a small problem with naming the system appearing in the axiom **A2**, since, for the time being, similar systems were not considered in the mathematical literature. The multiplicative system of $gff(n^m)$ has some properties of semigroups (and, more generally, of groupoids), then it has been named a pseudo-semigroup (because is not closed under multiplication: for some non-zero $a(x)$, $b(x) \in \mathsf{F}[x]^*$ the case $a(x) \cdot b(x) = 0$ may occur while 0 is not a member of the set $\mathsf{F}[x]^*$, and this set contains non-invertible elements).

Table 1: Multiplication table in $gff(n^m)$.

<table>
<tr><th>$\cdot$</th><th>ei_1</th><th>ei_2</th><th>$\cdots$</th><th>ei_{N_i}</th><th>en_1</th><th>en_2</th><th>$\cdots$</th><th>en_{N_n}</th></tr>
<tr><td>ei_1</td><td colspan="4" rowspan="4">A</td><td colspan="4" rowspan="4">B^T</td></tr>
<tr><td>ei_2</td></tr>
<tr><td>$\vdots$</td></tr>
<tr><td>ei_{N_i}</td></tr>
<tr><td>en_1</td><td colspan="4" rowspan="4">B</td><td colspan="4" rowspan="4">C</td></tr>
<tr><td>en_2</td></tr>
<tr><td>$\vdots$</td></tr>
<tr><td>en_{N_n}</td></tr>
</table>

The system considered has many interesting properties [1]. For example, the elements of the pseudo-semigroup belong to two disjoint sets: a set of invertible elements

$$ie = \{ei_1,\, ei_2,\, \ldots,\, ei_{N_i}\}, \tag{2}$$

and a set of non invertible elements

$$nie = \{en_1,\, en_2,\, \ldots,\, en_{N_n}\}, \tag{3}$$

where $N_i = |ie|$, $N_n = |nie|$. Any invertible element is a generator of a cyclic group, being a subgroup of the pseudo-semigroup. Therefore, if n is not a prime of if $f(x)$ is not irreducible, then the multiplication table of the non-zero elements of $gff(n^m)$ has the form shown in Table I.

The interior of the multiplication table consists of two square matrices $A_{N_i \times N_i}$, $C_{N_n \times N_n}$ and of two rectangular ones $B_{N_n \times N_i}$, $B^T_{N_i \times N_n}$. It should be observed that only the matrix A of the table is a Latin square.

Furthermore, one should know that if n is not a prime or if $f(x)$ is not irreducible then the multiplicative system of $gff(n^m)$ has the divisors of 0 and $gff(n^m)$ is not an integral domain.

The axioms **A1**, **A2** and **A3**, defining the system $gff(n^m)$, are satisfied if the operations on its elements are performed according to the way shown beneath.

Let

$$a(x) = \sum_{i=1}^{m} a_{m-i}\, x^{m-i}, \quad b(x) = \sum_{i=1}^{m} b_{m-i}\, x^{m-i} \tag{4}$$

be two elements of $\mathsf{F}[x]$. Then their sum will be

$$a(x) + b(x) = c(x) = \sum_{i=1}^{m} c_{m-i}\, x^{m-i}, \tag{5}$$

where

$$c_i \equiv a_i + b_i \pmod{n}, \;\; i = 0,\, 1,\, \ldots,\, m-1.$$

Similarly

$$a(x) - b(x) = d(x) = \sum_{i=1}^{m} d_{m-i}\, x^{m-i}, \tag{6}$$

where

$$d_i \equiv a_i - b_i \pmod{n}, \;\; i = 0,\, 1,\, \ldots,\, m-1.$$

The multiplication is more complicated. To calculate the product of two elements belonging to $gff(n^m)$ one ought to compute first

$$g(x) = a(x) \cdot b(x) =$$

$$= g_{2m-2}\, x^{2m-2} + g_{2m-3}\, x^{2m-3} + \cdots + g_2\, x^2 + g_1\, x + g_0$$

where

$g_0 \equiv a_0\, b_0 \pmod{n}$,
$g_1 \equiv a_1\, b_0 + a_0\, b_1 \pmod{n}$,
$g_2 \equiv a_2\, b_0 + a_0\, b_2 + a_1\, b_1 \pmod{n}$,
. . . .
$g_{2m-3} \equiv a_{m-1}\, b_{m-2} + a_{m-2}\, b_{m-1} \pmod{n}$,
$g_{2m-2} \equiv a_{m-1}\, b_{m-1} \pmod{n}$.

Next, to obtain finally the product $h(x)$ of two $gff(n^m)$ elements (4), we ought to represent $g(x)$ as

$$g(x) = u(x) \cdot f(x) + h(x) \tag{7}$$

using addition and multiplication modulo n, wherefrom

$$h(x) = a(x) \cdot b(x).$$

The operation of multiplication in $gff(n^m)$ may also be shortly written as

$$h(x) \equiv a(x) \cdot b(x) \pmod{f(x)}.$$

The multiplicative inverse $a^{-1}(x)$ of the element $a(x)$ can be determined the most quickly by means of the extended Euclidean algorithm for polynomials, which yields:

$$a(x) \cdot a^{-1}(x) + w(x) \cdot f(x) = 1,$$

that is

$$a(x) \cdot a^{-1}(x) \equiv 1 \pmod{f(x)}.$$

So

$$a(x)/b(x) = a(x) \cdot b^{-1}(x).$$

Thus, we can compute in $gff(n^m)$ as in any field, performing addition, subtraction, multiplication, division and the operation of rising to a power (by using, for example, square-and-multiply algorithm). The presented principles of computing in a generalized finite field may be suitably optimized or improved to be well adapted for hardware or software implementation, necessary for applications.

It's worth mentioning that elements of $gff(n^m)$ can be represented not only as polynomials or vectors over Z_n, but also as integers. The latter case is the most interesting for cryptography, therefore, we will continue the problem of computing in $gff(n^m)$ considering the system

$$gff(n^m) = \langle \mathsf{F}_{n^m},\ \mathsf{+},\ \bullet \rangle, \tag{8}$$

where $\mathsf{F}_{n^m} = \{0,\ 1,\ \ldots,\ n^m - 1\}$.

The system (8) is obtained from the system (1) in which a set of polynomials is replaced by a set of integers. To do the required transformations from polynomials to integers we may use the isomorphic mapping

$$\sigma : \mathsf{F}[x] \to \mathsf{F}_{n^m}, \tag{9}$$

defined by the function

$$\sigma(a(x)) = a(n) = \mathrm{A} \ \in \mathsf{F}_{n^m}, \tag{10}$$

converting a polynomial $a(x) \in \mathsf{F}[x]$ to an integer A from the set F_{n^m}. The mapping σ is an isomorphism, so the inverse mapping σ^{-1} exists and is described by means of the following two-step algorithm:

Step 1:

convert a base 10 integer $\mathrm{A} \in \mathsf{F}_{n^m}$ to base n, namely,

$$\mathrm{A} = a_{m-1} \ \cdots \ a_1 \, a_0, \ a_i \in \{0, 1, \ldots, n-1\},$$

Step 2:

write down the sequence of numerals

$$a_{m-1} \ \cdots \ a_1 \, a_0$$

as a polynomial, obtaining

$$\sigma^{-1}(\mathrm{A}) = a_{m-1} \, x^{m-1} + \cdots + a_1 \, x + a_0 \ \in \mathsf{F}[x].$$

Thus,

$$\begin{aligned} \forall \ \mathrm{A},\, \mathrm{B} \in \mathsf{F}_{n^m} \ (\mathrm{A} \bullet \mathrm{B} = \sigma(\sigma^{-1}(\mathrm{A}) \cdot \sigma^{-1}(\mathrm{B}))) \wedge \\ (\mathrm{A} + \mathrm{B} = \sigma(\sigma^{-1}(\mathrm{A}) \ + \ \sigma^{-1}(\mathrm{B}))). \end{aligned} \tag{11}$$

To the family of systems $gff(n^m)$ belongs a big class of algebraic structures. E.g. if $n = p$, p is a prime, and if $f(x)$ is not irreducible then the multiplicative structure of $gff(p^m)$ forms so-called spurious multiplicative group of $GF(p^m)$, applicable in the implementation of user-friendly ElGamal public-key encryption scheme [1]. If, in addition, $f(x)$ is irreducible, $gff(p^m)$ becomes $GF(p^m)$. Thus, $gff(n^m)$, is a generalization of finite field, and it fully deserves to be called a generalized finite field. To form $gff(n^m)$ we may use any monic polynomial of degree m over $\mathrm{Z_n}$, in other words, there exist n^m generalized finite fields $gff(n^m)$ with various multiplicative systems, having a different number of invertible elements. Although all properties of $gff(n^m)$ are not yet known, this algebraic structure will certainly be broadly applied, mainly in cryptography and coding.

Example 1.

Let $n = 6$, $m = 3$, $f(x) = x^3 + 5\,x + 1$, and let us construct the 216-element generalized finite field

$$gff(6^3) = gff(216) = \langle \mathsf{F}_{216},\ \mathsf{+},\ \bullet \rangle,$$

where $\mathsf{F}_{216} = ie \cup nie$, *ie* denoting the following set of 182 invertible elements:

ie ={(1), (5), (6, 181), (7, 186), (8, 195), (9, 202), (10, 207), (11, 210), (13, 171), (15, 149), (17, 159), (19, 130), (20, 113), (175, 200), (22, 109), (23, 128), (25, 99), (27, 73), (29, 87), (30, 41), (31, 42), (32, 51), (33, 56), (34, 63), (35, 66), (36, (211), (37, 80), (38, 141), (39, 106), (40, 191), (43, 208), (44, 93), (45, 124), (46, 79), (47, 180), (48, 143), (49, 92), (50, (215), (52, 203), (53, 104), (54, 103), (55, 206), (57, 196), (58, 83), (59, 120),(60, 139), (61, 82), (62, 201), (64, 187), (65, (90), (67, 182), (68, 107), (69, 122), (70, 95), (71, 194), (75, 173), (77, 157), (78, 133), (81, 118), (85, 147), (89, 161), (91, 114), (94, 135), (97, 169), (101, 145), (102, 137), (105, (116), (110, 131), (112, 127), (115, 204), (117, 184), (119, 192), (121, 150), (123, 164), (125, 174), (132, 199), (136, 189), (138, 167), (134, 213), (140, 177), (142, 153), (151, 190), (152, 183), (154, 205), (155, 198), (162, 193), (163, 188), (165, 214), (166, 209), (176, 197), (178, 185), (179, 212)}

and *nie* the set of 34 non-invertible elements:

nie = {0, 2, 3, 4, 12, 14, 16, 18, 21, 24, 26, 28, 72, 74, 76, 84, 86, 88, 96, 98, 100, 108, 111, 126, 129, 144, 146, 148, 156, 158, 160, 168, 170, 172}.

The invertible elements are listed in pairs - an element and its multiplicative inverse (with the exception of 1 and 5, since $1^{-1} = 1$ and $5^{-1} = 5$). For example, (6, 181) means that $6^{-1} = 181$, etc.

It is essential to mention in this place that the extended Euclidean algorithm for polynomials over Z_n fails if n is not a prime, and it cannot be able to determine all invertible elements in $gff(n^m)$. For example, using Maple routine implementing this algorithm, one can find in the considered $gff(216)$ only 50 invertible elements. Here are these invertible elements, determined by means of Maple:

iempl = {(1), (5), (6, 181), (7, 186), (8, 195), (9, 202), (10, 207), (11, 210), (30, 41), (31, 42), (32, 51), (33, 56), (34, 63), (35, 66), (36, 211), (40, 191), (43, 208), (47, 180), (50, 215), (52, 203), (55, 206), (57, 196), (62, 201), (64, 187), (67, 182), (71, 194)}.

Assuming that

$$\mathrm{A} = 102, \quad \mathrm{B} = 200,$$

and taking into account (10) and (11), we will follow the process of performing addition and multiplication in $gff(216)$:

$$
\begin{aligned}
\sigma^{-1}(\mathrm{A}) &= a(x) = 2\,x^2 + 5\,x,\\
\sigma^{-1}(\mathrm{B}) &= b(x) = 5\,x^2 + 3\,x + 2,\\
c(x) &= a(x) + b(x) = 7\,x^2 + 8\,x + 2,\\
d(x) &\equiv c(x) \pmod 6 = x^2 + 2\,x + 2,\\
\mathrm{A} + \mathrm{B} &= \sigma(d(x)) = d(6) = 50,\\
g(x) &= a(x) \cdot b(x) =\\
&= 10\,x^4 + 31\,x^3 + 19\,x^2 + 10\,x,\\
h(x) &\equiv g(x) \pmod 6 = 4\,x^4 + x^3 + x^2 + 4\,x,\\
r(x) &\equiv h(x) \pmod{x^3 + 5\,x + 1} =\\
&= -19\,x^2 - 5\,x - 1,\\
r(x) \pmod 6 &= 5\,x^2 + x + 5,\\
\mathrm{A} \bullet \mathrm{B} &= \sigma(r(x)) = r(6) = 191.
\end{aligned}
$$

Other operations in $gff(216)$ can be carrying out alike.

3 Symmetric-Key Block Cipher Based on $gff(n^m)$

It is known that any modern block cipher (e.g. AES) provides confusion (substitutions that make the relationship between the key and ciphertext as complex as possible) and diffusion (transformations that dissipate the statistical properties of the plaintext across the ciphertext) through several composite enciphering transformations over large blocks of data. In the discussed cipher the desired level of confusion and diffusion is achieved much more easily, only by using two operations in $gff(n^m)$ per enciphered block of plaintext. It is an advantageous feature, which is very needed in the implementation of fast ciphers.

The proposed cipher is flexible, which means that its block size (and simultaneously key space) can be of any size.

A block cipher encrypts message in fixed-size blocks. For messages exceeding the block size, the simplest approach is to partition the message into t blocks and encrypt each separately, using an appropriate mode of operation. The cipher considered works in CBC (cipher-block chaining) mode of operation [4], which may be specified in the following algorithm:

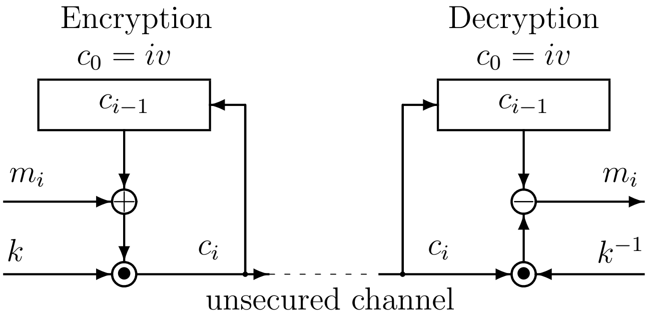

Figure 1: A $gff(n^m)-$based block cipher working in CBC mode.

Input:

k - secret key, iv - initial vector, n, $f(x)$ - data defining $gff(n^m)$ (the last three parameters may be common to all entities), t plaintext blocks $m_1,\ m_2,\ \ldots,\ m_t$.

Encryption:

$c_0 \leftarrow iv$, for $1 \leq i \leq t$, $c_i \leftarrow k \bullet (m_i + c_{i-1})$

(produces t ciphertext blocks $c_1,\ c_2,\ \ldots,\ c_t$).

Decryption:

$c_0 \leftarrow iv$, for $1 \leq i \leq t$, $m_i \leftarrow c_i \bullet k^{-1} - c_{i-1}$

(decrypts $c_1,\ c_2,\ \ldots,\ c_t$ to recover plaintext).

Thus, the idea of the cipher is illustrated in Fig. 1.

The secret key, initial vector, plaintext blocks and ciphertext blocks are elements of any given $gff(n^m)$ (it is evident that k must be an invertible element). The operations of addition, subtraction and multiplication on these data are performed in $gff(n^m)$. If the parameters iv, n and $f(x)$ are not common to all users, then they can be used either for secret sharing or to increase considerably the key space.

Example 2.

We will now construct a cipher using $gff(216)$ from Example 1, to verify the mode of encrypting and decrypting. With this end in view let us assume that

$$k = 102,\ iv = 13,$$

and let

$$M = [1, 4, 2, 8, 5, 7]$$

be six blocks of a message. Then we obtain the following six blocks of cryptogram

$$C_1 = (1{+}13) \bullet 102 = 86,$$

$$C_2 = (4{+}86) \bullet 102 = 100,$$

etc., and finally

$$C = [86, 100, 12, 60, 77, 120].$$

In the like manner we recover the message correctly from the cryptogram using

$$k^{-1} = 137$$

during decryption.

If we assume in this example that the polynomial $f(x)$ is secret, then the key space has 4536 elements and the key size is equal to 13 bits. The key size is not impressive, but this example makes it easy to verify encrypting/decrypting algorithms.

Example 3.

This example, concerning the cipher over $gff(n^m)$ constructed using

$$n \;=\; 3^{48} \cdot 11^{12} \cdot 29^{12} \cdot 418849^{12}, \;\; f(x) \;=\; x^2 + 1234567x, \;\; m \;=\; 2,$$

by means of the accompanying Maple worksheet has been computed. We see that now $gff(n^m)$ has approximately $0.6668012974 \cdot 10^{241}$ elements, block size bs is 100 bytes, the key length equals to 800 bits, n is not a prime and $f(x)$ is not irreducible. After carrying out several experiments the reader will observe and understand that the optimal confusion and diffusion are obtained if one of the following equations

$$n{=}\lceil e^{(8\ \ln(2)\, bs/m)}\rceil, \tag{12}$$

$$m{=}\lceil\frac{8\ \ln(2)\, bs}{\ln(n)}\rceil, \tag{13}$$

$$bs{=}\lceil\frac{\ln(n^m)}{8\ \ln(2)}\rceil, \tag{14}$$

is satisfied more or less. This condition has been taken into account in the worksheet.

Using the cipher, the three diverse files: `m.one`, `m.pdf`, and `m.wav`, containing a text document (the letter `k`, repeated 206553 times), a PDF document, and a short musical fragment, respectively, have been used as plaintexts. Looking at Figs. 2, 3, and 4 we may verify that the distribution of

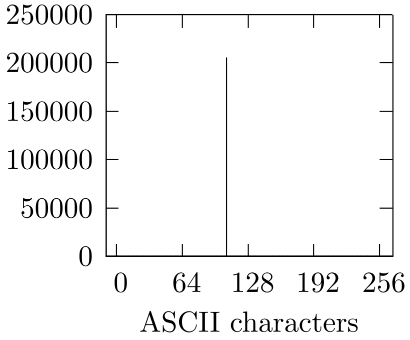

Figure 2: Frequency of occurrence of ASCII characters in the plaintext file `m.one` (size 206553 bytes).

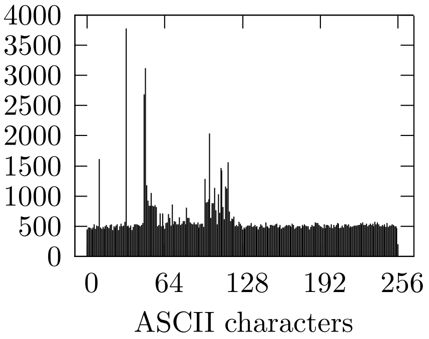

Figure 3: Frequency of occurrence of ASCII characters in the plaintext file `m.pdf` (size 157787 bytes).

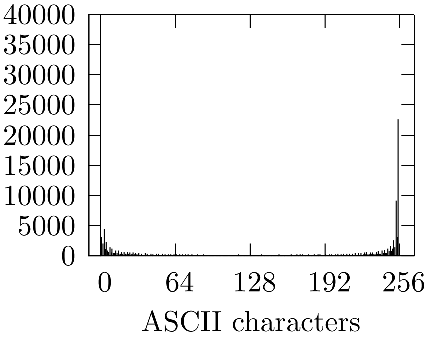

Figure 4: Frequency of occurrence of ASCII characters in the plaintext file `m.wav` (size 167376 bytes).

ASCII characters in these files are distinct.

Next, the files have been enciphered into cryptograms, and the frequency

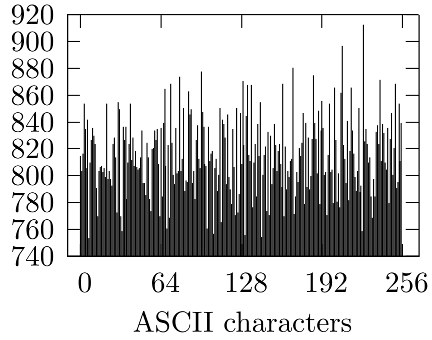

Figure 5: Frequency of occurrence of ASCII characters in the cryptogram of the file `m.one` (size 208700 bytes).

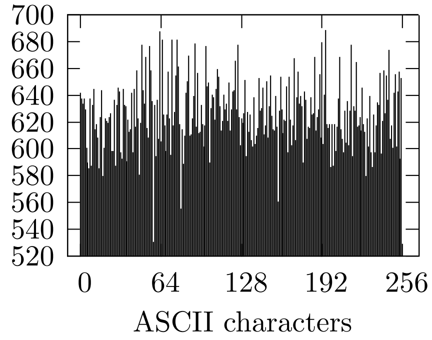

Figure 6: Frequency of occurrence of ASCII characters in the cryptogram of the file `m.pdf` (size 160400 bytes).

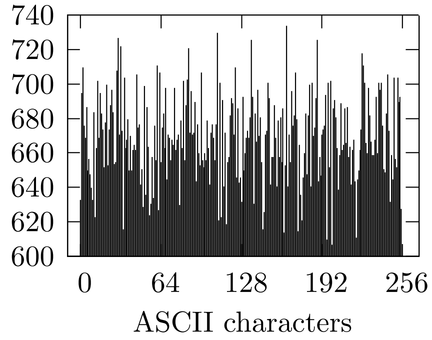

Figure 7: Frequency of occurrence of ASCII characters in the cryptogram of the file `m.wav` (size 171300 bytes).

of occurrence of ASCII characters in each cryptogram has been computed. As it was supposed, the distributions of ASCII characters in each cryptogram are

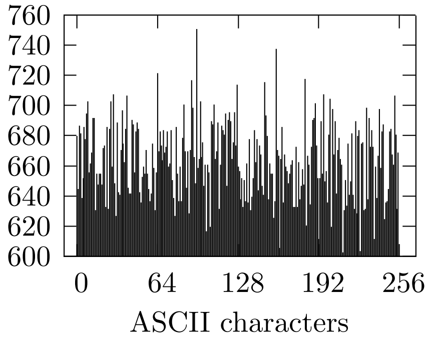

Figure 8: Frequency of occurrence of ASCII characters in the the file containing 170000 characters generated at random using Maple random number generator.

much the same, and are very similar to the output of good random number generator, although they represent very different messages (compare Figs. 5, 6, and 7 with Fig. 8).

4 Conclusion

An example of the application of a new, very simple, yet effective and promising algebraic tool for cryptography, named generalized finite fields, offering construction of fast, flexible and secure ciphers has been presented. To convince the reader that the cipher is really fast, flexible and secure, in the file `gffbskbc.m` are stored the Maple routines allowing to explore $gff(n^m)$ and to test the cipher considered. These routines may be used either immediately as elements of encrypting/decrypting programs in the Maple environment or they can be translated into any compiled programming language. In the latter case encryption/decryption can be performed about 150 times faster than in the Maple environment.

It is evident that using $gff(n^m)$ it is possible to construct many other ciphers and cryptographic protocols. The presented idea of $gff(n^m)$ is worthy of notice because it generates new interesting research problems for mathematicians and computer science specialists. For example, from the practical point of view especially crucial, attractive and absorbing is the task of finding all reversible elements of $gff(n^m)$.

References

[1] C. Kościelny, *Spurious Multiplicative Group of $GF(p^m)$: a New Tool for Cryptography*, Quasigroups and Related Systems, vol. 12, pp. 61 – 73, 2004.

[2] C. Kościelny, *A New Approach to the ElGamal Encryption Scheme*, International Journal of Applied Mathematics and Computer Science, vol. 12, no 2, 2004, pp. 265 – 267.

[3] C. Kościelny, *Computing in $GF(p^m)$ and in $gff(n^m)$ using Maple*, accepted for publication in Quasigroups and Related Systems, vol. 13, 2005.

[4] A. Menezes, P. van Oorschot, S. Vanstone, *Handbook of Applied Cryptography*, CRC Press, 1997, p. 230.