Parallel Bit-Stream Cipher with Cellular Automata Number 30 in Java

Antonio J. Tomeu, Alberto G. Salguero
Dpto. de Ingeniería Informática
Escuela Superior de Ingeniería, Universidad de Cádiz
C/Chile s/n, E-11002
Cádiz (España)
antonio.tomeu, Alberto.salguero}@uca.es

Manuel I. Capel
Dpto. de Lenguajes y Sistemas Informáticos
Escuela Técnica Superior de Ingeniería Informática y de Telecomunicaciones, Universidad de Granada
C/Periodista Daniel Saucedo s/n, E-18071
Granada, (España)
manuelcapel@ugr.es

Abstract—In previous works we have made the integration of one-dimensional linear cellular automata into bit-stream ciphers. These automata present a high degree of confusion-diffusion information and reasonable entropy levels. Our purpose is to obtain one-time pad ciphers to use in mobile communications. Cellular automaton number 30 is tested and validated to code the cipher. This work shows how to parallelize our encryption algorithm using Subramanian’s equation to distribute the computational effort over a multicore CPU. Several standard multicore platforms will be used to verify performance gain.

Keywords—cellular automata; bit-stream cipher; entropy; randomness; parallel processing, multicore programming, Vernam cipher, randomness test, complexity, self-organization.

I. INTRODUCTION

The encryption of data is currently a must and unavoidable necessity. Several available encryption paradigms, such as symmetric and asymmetric (or public key), exist to solve this problem. The symmetric one is still widely used due to its low cost and efficiency [10]. When it comes to sending relatively small amounts of information, the use of standard symmetric key encryption algorithms (DES, AES) becomes inappropriate ([3], [4] and [10]). The same occurs with public key algorithms, which require a lot of time for processing a relatively small amount of information. Commercial approaches such as Pretty Good Privacy (PGP) are useful to reduce this problem.

In these cases, a more appropriate, efficient and quick alternative is to resort to encrypted bitstream approach, with or without using one-time pad ([4], [10] and [11]), which uses the encryption protocol just once, and then for next timedicards and replaces it with a different one.

This is accomplished easily and efficiently by using the bitstream encryption cipher ([10] and [18]), also known as Vernam cipher, as the general framework to build our one-time pad.

A bit stream cipher encrypts the characters in a plain text message using a sequence of binary digits in the form

\[ \gamma_i = \mu_i \oplus \nu_i \]  

for \( i = 1, 2, 3... \) (1)

where \( \mu_i \) are the bits of the plain text and \( \nu_i \) are bits of the cipher key, obtained by the equation \( \gamma_i \), which ciphers the bits using the XOR function \( \oplus \) (bitwise addition modulo 2). A bit stream cipher is obtained when \( \nu_i \) bits are generated randomly and independently.

\[ \mu_i = \gamma_i \oplus \nu_i \]  

for \( i = 1, 2, 3... \) (2)

as the boolean function \( \oplus \) is an involution.

Shannon proved ([4], [10]) that if a particular condition on the entropy of the key \( H(K) \) and message \( H(M) \) holds, being

---

1 Credits of this image are given to Prof. Dr. Jorge Ramió Aguirre and his e-book “Seguridad Informática y Criptografía”, at free disposal in the author’s website.
M, C and K random variables that respectively denote the flat texts and the encryption key, then the encryption is unconditionally safe if the cipher bit sequence is randomly chosen and independent of the bit-sequences of the message. This condition is given by the inequality,

\[ H(K) \geq H(M) \]

This is because \( H(M|C) = H(M) \) and therefore the amount of information \( I(M;C) \) must be considered null. Of course, once the one-time pad has been established, multiple instances of it can be used and discarded for each connection. It is at that point where the one-dimensional cellular automata, properly used, provides a stream of pseudo-random bits that allows the use of a one-time Vernam cipher, or alternatively the implementation of a reusable cipher stream changing the key of the cipher.

II. ONE-DIMENSIONAL CELLULAR AUTOMATA

A. Cellular automata model definition

A one-dimensional cellular automaton [19] is a tuple \(<\mathbb{Z}^\infty, \kappa, \rho, \delta>\) where:

- \(\mathbb{Z}^\infty\) is a row (or configuration) of discrete components called cells with possible values within a finite Abelian ring or residual class modulo \(\kappa\). The evolution of the cellular automaton becomes a sorted and deterministic succession in \(\mathbb{Z}^\infty\).

- \(\rho\) is the radius of neighborhood, and is a natural number greater than or equal to one. It determines the range of surrounding cells with which a given cell exchanges information.

- \(\delta\) is the transition function that is used to recalculate the row of cells to obtain a succession of configurations.

The system evolves by discrete time steps, generating new configurations (or “generations” in bio-inspired terminology) of the cellular automaton by synchronously applying the transition function \(\delta\) to each of the cells.

This transition function exploits the principle of locality to communicate any cell with each other and exchanges information between them, so that when a particular cell changes its state takes into account the information provided by the surrounding \(\rho\) cells left and right. The form of the transition function adopts the scheme shown in equation number 3:

\[
\alpha^{(t+1)}_i = \delta(\alpha^{(t)}_{i-\rho}, \ldots, \alpha^{(t)}_i, \ldots, \alpha^{(t)}_{i+\rho}) \tag{3}
\]

The symmetry of the neighborhood range of a particular cell is often imposed as a model condition, although this is not strictly necessary.

When the model is used in real problems, it is necessary to impose boundary conditions, being the cylindrical boundary the most common and most suitable shape to our needs. It is also useful to represent the transition functions using the Wolfram numerical coding technique, which lets us to specify a particular cellular automaton by a single number ([19]).

In [19] a classification of cellular automata behavior has been carried out. The results obtained are the same whether empirical or analytical analysis are used. All one-dimensional cellular automata behave according to one of the following four possibilities:

- Class 1: evolution leads to a metastable configuration that is perpetuated indefinitely. This configuration destroys all the information encoded in the initial configuration.
- Class 2: evolution leads to a set of simple patterns that are repeated regularly or remain stable.
- Class 3: evolution leads to a pattern with presence of chaos.
- Class 4: evolution leads to isolated structures with highly complex behaviors associated with dendritic-like structures that hold information processing capabilities (comparable to the class of Turing-computable functions, [13]) and sometimes are capable of universal computation.

In figure 2, examples of the four kinds of behavior above described for cellular automata with \(\kappa=2\) states per cell are shown. Left column shows automata with a random initial configuration, whereas the right column shows automata with all their cells but the central one have a value of 0.

The boundary conditions used for computing the spatiotemporal automata shown in figure 2 are cylindrical.

Since it is intended to achieve a process that forces the confusion-diffusion of the information to be encrypted, it is necessary to choose cellular automata that exhibit chaotic behavior of class 3 ([2], [9]) as potential candidates for the generation of one-time pad reusable encryption flow.
B. Analysis methodology

Compliance with the Shannon criterion to ensure a strong encryption cannot be satisfied by simply identifying a pattern with chaotic behavior. It is necessary to analyze the pattern using an analytical modeling to establish that such a pattern is not spurious, but real ([2]).

For this, the method suggested in [19] and completely developed in [12] and [16] is applied. This methodology characterizes a one-dimensional cellular automaton with cylindrical border as chaotic and member of class 3 when the following three criteria are met:

- Passes five of a battery of six tests of randomness, at least: $\chi^2$, runs, series, permutations, intervals and overlapped triplets.
- Has a configuration entropy value close to the unit.
- Has a temporal entropy defined by the equation

$$S_t = -\sum_{i=0}^{k-1} p_{i}^{(j)} \log_k p_{i}^{(j)} \quad (4)$$

- Has a temporal entropy defined by the equation

$$S_t = -\sum_{i=0}^{k-1} p_{i}^{(j)} \log p_{i}^{(j)} \quad (5)$$

close to the unit.

With the above criteria, a powerful simulation package (©CaSim) has been designed in order to search for automata with good randomness properties that belong to class 3, which can be used to form one-time pads. This package is able to go over the space of the rules of one-dimensional cellular automata for any desired combination of pairs ($\kappa, \rho$), and then automatically generating a database of promising automata. It is also possible to search for cellular automata with the right properties in order to integrate them into a reusable stream cipher. In this work, we will follow the latter approach.

Figure 3 shows a snapshot of CaSim simulator running for a cellular automaton with three states per cell.

C. Cellular Automata Number 30

Given all of the binary one-dimensional cellular automata with neighborhood range equal to unit, the one with additive code number 30 has been chosen as the candidate to form the stream cipher. Its spatiotemporal evolution pattern is described in figure 4. Its transition function is given by

$$\alpha^{(i)} = (\alpha^{(i-1)}_{k-1} + \alpha^{(i-1)}_k + \alpha^{(i-1)}_{k+1} + (\alpha^{(i-1)}_k \alpha^{(i-1)}_{k+1})) \quad (6)$$

Additionally, its performance in terms of randomness filtering and parameters entropy is described in table I. It is also widely studied in the literature ([19]). Table I summarizes the means of temporal entropy and configuration parameters.

<table>
<thead>
<tr>
<th>Rule</th>
<th>$S$ (mean)</th>
<th>$S_t$ (mean)</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>0.96</td>
<td>0.98</td>
</tr>
</tbody>
</table>
The same rule, when using a random initial configuration behaves as figure 5 shows.

As shown, the behavior of the rule number 30 belongs to class 3 in terms of spatiotemporal evolutionary pattern. It is, therefore, a candidate to be used in the cipher. The next section verifies this assumption.

III. PROPOSED BIT-STREAM CIPHER

Once the two-dimensional cellular automaton that will act as a generator of pseudo-random cipher sequences has been chosen, it is necessary to consider how to get the sequences from the automaton. There are several proposals in the literature ([3], [4], [6], [7] and [8]), although the standard in these cases consists on taking the spatiotemporal evolution of the k-th central cell of the cellular automaton. From the above equation it is possible to write the pseudocode of the bitstream encryption algorithm as follows,

\[ \gamma_i = \mu_i \oplus a_k(i) \quad \text{for } i = 1, 2, \ldots \]

where the cipher text bits \( c_i \) are obtained from the XOR function applied to the plain text bits \( m_i \) and the cipher bits \( a_k(i) \). The latter is provided by the temporal evolution of the k-th central cell of the cellular automaton. From the above algorithm it is possible to write the pseudocode of the bitstream encryption algorithm as follows,

Bitstream cipher algorithm

INPUT: a message \( \{\mu_i\} \) of n bits and a key \( \chi^{(k)} \)

OUTPUT: cryptogram \( \{\gamma_i\} \) of n bits

METHOD:
1. Initialize the cellular automaton with key \( C^{(k)} \)
2. For all \( i = 1 \) to \( n \)
   2.1. \( \gamma_i = \mu_i \oplus a_k(i) \)
   2.2. \( a_k(i+1) = (a_k(i-1) + a_k(i) + a_k(i+1) + (a_k(i) \cdot a_k(i+1))) \mod 2 \)

The Java class below shows the implementation of the above algorithm to compute the cellular automaton Rule 30. The program used to implement stream ciphers is also shown.

```java
import java.util.Random;

public class aC30{
    int nCelulas;
    int[] estadoAC;
    int[] aux;
    int[] temp;

    public aC30(int nCelulas){
        this.nCelulas = nCelulas;
        Random p = new Random();
        estadoAC = new int[nCelulas];
        aux = new int[nCelulas];
        for(int i=0; i<estadoAC.length; i++)
            estadoAC[i] = p.nextInt(2);
    }

    public int fTransicion(int i){
        if(i==0)
            return((estadoAC[estadoAC.length-1]+estadoAC[i]
            +estadoAC[i+1]+(estadoAC[i]*estadoAC[i+1]))%2);
        else if(i==estadoAC.length-1)
            return((estadoAC[i-1]+estadoAC[i]+estadoAC[0]+
            estadoAC[i]*estadoAC[0])%2);
        else return((estadoAC[i-1]+estadoAC[i]+estadoAC[i]
            +estadoAC[i+1]+(estadoAC[i]*estadoAC[i+1]))%2);
    }

    public void nuevoEstadoAC(){
        for(int i=0; i<estadoAC.length; i++)
            aux[i] = fTransicion(i);
        temp = estadoAC;
        estadoAC = aux;
        aux = temp;
    }

    public int valCelula(int i){
        return(estadoAC[i]);
    }
}
```

Ann. Mult. GPU Prog. 21
Code 1 listing: Java class to model the cellular automaton 30.

The above class computes successive generations of cellular automata with rule number 30. It can be integrated into a bit stream cipher application as follows,

```java
import java.util.Random;

public class cifraVernam {
    public static int nCel = 1000;
    public static int longTexto;
    public static Random p = new Random();
    public static int[] textoLlano;
    public static int[] textoCifrado;
    public static aC30 miAC = new aC30(nCel);

    public static void main(String[] args) {
        longTexto = Integer.parseInt(args[0]);
        textoLlano = new int[longTexto];
        textoCifrado = new int[longTexto];
        for (int i = 0; i < textoLlano.length; i++)
            textoLlano[i] = p.nextInt(2);
        long inicTiempo = System.nanoTime();
        for (int i = 0; i < textoLlano.length; i++)
            textoCifrado[i] = (textoLlano[i] + miAC.valCelula((int)nCel/2)) % 2;
        miAC.nuevoEstadoAC();
        long tiempoTotal = (System.nanoTime() - inicTiempo) / (long)1.0e9;
        System.out.println("Calculo finalizado en "+tiempoTotal+" segundos");
    }
}
```

Source code 2: Java class to compute plain text encryption.

The bits of the plain text string have been simulated using a random generator. The bits of the encrypted text have been computed using the cellular automata in `ac30.java`.

**B. Encryption security**

Although it has been considered as a secondary aspect in this work, whose main objective has been the parallelization of the bit stream encryption process using a cellular automaton number 30, security aspects must be also analyzed. A fast parallel encryption is useless if it is not safe. An elemental analysis of the security of the cellular automaton number 30 cipher is important.

The analysis starts studying the randomness of this automaton with respect to the previous tests described above.

The following control parameters have been used:

- 1000 cells array.
- Random initial configuration.
- Cylindrical frontier condition.
- 1000 generations.

Table II below shows the results obtained:

<table>
<thead>
<tr>
<th>CA30</th>
<th>(\chi^2)</th>
<th>Runs</th>
<th>Perm.</th>
<th>Int.</th>
<th>Series</th>
<th>Trip.</th>
</tr>
</thead>
</table>

A comparison with other usual techniques for obtaining random sequences of digits is shown in table III.

**TABLA III: COMPARATIVA CON OTROS GENERADORES**

<table>
<thead>
<tr>
<th></th>
<th>(\chi^2)</th>
<th>Rachas</th>
<th>Perm.</th>
<th>Int.</th>
<th>Series</th>
<th>Trip.</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA60</td>
<td>(\sqrt{2})</td>
<td>(\sqrt{2})</td>
<td>(\sqrt{2})</td>
<td>(\sqrt{2})</td>
<td>(\sqrt{2})</td>
<td>(\sqrt{2})</td>
</tr>
<tr>
<td>LGC</td>
<td>(-)</td>
<td>(\sqrt{2})</td>
<td>(\sqrt{2})</td>
<td>(\sqrt{2})</td>
<td>(\sqrt{2})</td>
<td>(\sqrt{2})</td>
</tr>
<tr>
<td>LFSR</td>
<td>(\sqrt{2})</td>
<td>(\sqrt{2})</td>
<td>(\sqrt{2})</td>
<td>(\sqrt{2})</td>
<td>(\sqrt{2})</td>
<td>(\sqrt{2})</td>
</tr>
<tr>
<td>(\sqrt{2})</td>
<td>(\sqrt{2})</td>
<td>(\sqrt{2})</td>
<td>(\sqrt{2})</td>
<td>(\sqrt{2})</td>
<td>(\sqrt{2})</td>
<td>(\sqrt{2})</td>
</tr>
<tr>
<td>(\Pi)</td>
<td>(\sqrt{2})</td>
<td>(\sqrt{2})</td>
<td>(\sqrt{2})</td>
<td>(\sqrt{2})</td>
<td>(\sqrt{2})</td>
<td>(\sqrt{2})</td>
</tr>
</tbody>
</table>

The comparison between the chosen cellular automata and the other alternatives (some of them commercialized within many commonly used compilers, as LGC) is favorable to the former one. Additionally, the analysis of both the temporal and configuration entropies with rule number 30 automaton, in both cases, provides values close to the unit. The encryption bitstream is thus properly characterized regarding to its randomness.

Finally, it should be noted that a correlation analysis between the plain text and the equivalent cipher text sample shows differences less than 1%.

IV. BITSTREAM CIPHER PARALLELIZATION

The objective of this work is the parallelization of the encryption algorithm assuming be used in multi-core processors platforms. The encryption application should adapt the number of cryptographic tasks according to the number of available cores. The structure of the proposed parallelization strategy is shown in Figure 6.
Fig. 6: Parallel encryption architecture with multiple instances of the cellular automaton.

For the development of the cipher, plain text will be divided into equal length blocks of bits ([15]). There will be many instances of cellular automata as blocks. The final encrypted text will be obtained applying the cipher to each plain text block.

A. Number of tasks

To decide the number of encryption tasks the known Subramanians’ equation is used ([14]). This equation establishes the optimal number of parallel tasks taking into account two parameters:

- The number of available cores, \( N_{nd} \).
- Blocking coefficient of the tasks, \( C_b \).

From the above parameters, this equation establishes that the number of tasks is given by the following expression

\[
N_t = \frac{N_{nd}}{1 - C_b} \quad (8)
\]

where the blocking coefficient roughly quantifies the fraction of time that a particular task is waiting because of inputs/outputs, network delays, or other kind of latency. In this case, given that all tasks perform intensive computation, it is assumed that the blocking coefficient will have a value equal to zero. If a version dedicated to encrypt files is considered, the value of \( C_b \) will be in the range [0.5-0.99]. The actual value must be empirically estimated.

B. Parallel tasks structure

The development of a parallel version of the previous code requires making a partition of the plain text sequence into equal size segments. The number of segments is determined by the above equation. Each of the segments will be encrypted by a dedicated task. Those tasks are independent of each other. Source code 3 shows the class that has been implemented to model such a task.

```java
public class tareaVernam implements Runnable{
    private final int linf;
    private final int lsup;
    private final int[] tL;
    private final int[] tC;
    private final aC30Sinc miac;
    private final int nCel;

    public tareaVernam(int linf, int lsup, int[] tL, int[] tC, aC30Sinc miac, int nCel){
        this.linf = linf;
        this.lsup = lsup;
        this.tL = tL;
        this.tC = tC;
        this.miac = miac;
        this.nCel = nCel;
    }

    public void run(){
        for(int i=linf; i<=lsup;i++){
            tC[i]=(int)(tL[i]+miac.valCelula(nCel/2))%2;
            miac.nuevoEstadoAC();
        }
    }
}
```

Source code 3: Runnable task for the parallel encryption.

Each of the instances of the above class receives a segment of the plain text bits sequence to be encrypted. This sequence is initialized with the parameters \( l_{inf} \) and \( l_{sup} \), along with a reference to an instance of the class \( a30{Sinc} \), which has been already described in this work. In this case, all of the container structures are objects of the class `java.util.concurrent.atomic.AtomicIntegerArray` which makes the read/write operations of tasks safe when they run in parallel. This is not strictly necessary in our approach for the encryption parallelization. Not using secure data structures for parallelism, the encryption time values offered in the analysis section may be slightly improved ([14]).

The main program determines how many independent cipher tasks are going to run in parallel. For this, it starts by making a system call to determine how many logical cores are available, considering a \( C_b \) value equal to zero

```java
Runtime.getRuntime().availableProcessors();
```

and then processed using an executor service with fixed capacity, which can be easily created in Java just using the method `newFixedThreadPool(nTareas)` of the class `java.util.concurrent.Executors`. After that, the program just instantiates the tasks and transfer them to the executor in order to be processed:

```java
executor.execute(  
    new tareaVernam(  
        linf,  
        lsup-1,  
        textoLlano,  
        textoCifrado,  
        new aC30Sinc(Clave),  
        clave)  
    );
```

As shown when instantiating a new object of the synchronized cellular automata class, each of the instances that are going to be run in parallel needs to receive its own key. In order encryption be effective and cryptographically robust,
these keys should be different. This ensures that each parallel encryption task generates a different cipher bitstream. The generation of the different keys can be achieved in two ways:

- using different keys explicitly, depending on the number of tasks that are going to be used or
- obtaining the keys from a single key entered by the user, thus using a technique that rotates its bits.

The latter is the usual technique in these situations and it is the one that has been used in the design and implementation of the proposed encryption algorithm.

V. PERFORMANCE ANALYSIS

To carry out the performance analysis of the parallel encryption algorithm we have used the following platforms:

- **Intel Core i7-3632QM** processor with eight logical cores (using Hyper Threading\(^2\)), running at 2.2 GHz, eight GBytes of RAM memory and Windows 8 (64 bits) as the operative system (in figure 7).

- **Intel Core i3-2310M** processor with four logical cores (using Hyper Threading), running at 2.1 GHz, four GBytes of RAM memory, and Windows 8 (32 bits) and Linux Fedora 19 (*Schrödinger’s Cat*) as the operative systems (in figure 8).

The obtained results are shown in figures 7 and 8.

As shown in figure 7 and as expected, the encryption time in the parallel version of the algorithm linearly grows. The same occurs with the sequential version, which uses just one core of the processor. The usual improvement gap is present, resulting from the use of all available cores that the platform offers. The obtained improvement is still low and some improvements are further applicable.

These results show the need to always scale the parallel encryption technique depending on the number of available cores, even using the GPU architectures when possible.

Using the second of the platforms, the obtained results are those shown in figure 8. In this case, the same behavior in figure 7 is obtained. The explanation is also the same. The main difference is the performance reduction obtained in the parallel version. This is a consequence of using half of the cores that we previously had. Also, it can be seen how the threads in Linux have a better performance than the threads in Windows, once the JVM threads are mapped to system threads.

Regarding the load of processor cores, it has also been obtained the expected behavior. The parallel version uses the full computation power that the platforms can offer, while the sequential versions use just a fraction of this.

![Fig. 7: Encryption time depending on the size of the plain text message (in bits). Sequential and parallel versions of the algorithm running on an Intel i7 platform.](image)

Disk files oriented cipher versions have blocking coefficients values in the range \([0.5-0.99]\), with corresponding associated latencies plus delays derived from the safe writing of the parallel tasks in the encrypted file. In this situation, the usage of the processor by the threads in the proposed parallel algorithm is lower than those obtained here but always superior to sequential versions.

---

\(^2\) Hyper-Threading is the simultaneous multi-threading commercial implementation of the Intel Corp.
VI. CONCLUSIONS AND FUTURE WORK

As expected, the distribution of the workload of the sequential encryption algorithm to all the available cores in the local processor has reasonably improved the encryption time. By calculating subkeys using rotations of the original key, the security is also improved.

To use the new parallel encryption algorithm is possible for application design based on multi-core mobile technology, supported nowadays by Android. This is an interesting future line of research.

On the other hand, the literature provides standard methods for increasing the safety of bitstream cipher by deploying, among others, no linearization techniques ([17]), clock control encryption, linear filters encryption. A possible adaptation of the algorithm presented in this work is to implement these variations, it is also being necessary to test the security enhancement by means of the same techniques used here. To check how these solutions behave once parallelized is also a must.

Additionally, the other major paradigm currently used in the field of symmetric ciphers is the use of encryption algorithms in blocks of bits, using Feistel networks as their mathematical basis for the encryption [15]. The bit block cipher is a proposal that is well suited to the parallel algorithm proposed here by its nature. Thus it will open a complementary line of research to one presented here, i.e., the parallelization of oriented block ciphers using cellular automata of class 3.

The parallelization technique for the encryption that we have used here also adapts well to a scheme of massive parallelism, like GPU-based architectures. Another of the lines of future work will consists of developing the encryption algorithm proposed on this kind of platforms. Also, a test about the cipher of images using processor cluster is planned.

The integration of a fuzzy cellular automata model, under current analysis ([5], [12]) within the cryptographic protocols with parallelized algorithms is also an interesting direction of study that worths the effort.

ACKNOWLEDGMENTS

The authors want to thank the Computing Infrastructure Service at the University of Cadiz who made available the processor cluster.

REFERENCES


