Perspectives in Communication, Embedded-Systems and Signal-Processing (PiCES) – An International Journal
ISSN: 2566-932X, Special Issue 1

Cryptography Based on DNA Analysis

Anand M
Associate professor, ECE, East West Institute of Technology, Bangalore, India, anandm@ewit.edu

Anusha G
ECE, East West Institute of Technology, Bangalore, India, anushabharadwaj88@gmail.com

Aravind Kumar MJ
ECE, East West Institute of Technology, Bangalore, India, aravindmj005@gmail.com

Bharath Kumar R
ECE, East West Institute of Technology, Bangalore, India, bharathgowda531@gmail.com

Hema Varna M
ECE, East West Institute of Technology, Bangalore, India, varnahema30@gmail.com

Abstract: Secure data transmission is a challenging and fundamental aspect nowadays. Secure and nonvulnerable transmission of authenticated data is ensured by using cryptography and steganography methods. Advancement in technology leads to a new cryptographic algorithm each day. Recent advancement has led to a new technique called DNA cryptography. In this paper, the RSA algorithm, DNA cryptography, and steganography methods are followed to encrypt the plain text by assuring triple security. This combination will produce a more robust, more secure algorithm that is hard for the intruder to decrypt the message.

Keywords: DNA cryptography; RSA algorithm; Steganography; Secure data transmission; Triple security

I. INTRODUCTION

As we know in the current technological world, where most people deal with the internet, online business, and online transactions comprise confidential data. These data play an important role in one's life. If this data is not protected, then there is always a chance for hackers to corrupt, use, manipulate, and transmit it.

The primary mission of information technology is to ensure that systems and their contents remain secure. However, attacks on information systems are on a daily occurrence and the need for data security grows along with the trailblazing of such attacks. So, for the well management of network security most effective and strongest concept must be applied. This paper suggests ways of enhancing and improving security requirements. The main idea proposed here is that to provide multilayer security to important data so only intended users can retrieve or access it.

The proposed paper is on DNA cryptography. DNA cryptography is a new and very promising direction in cryptography research. It provides secure communication in the presence of malicious third parties and brings forward a new hope for unbreakable algorithms. In simple this paper is a combination of mathematics and security engineering using DNA cryptography and steganography along with RSA providing multi-layered security to the data. The RSA algorithm is widely used for securing sensitive data which will use a secure key to communicate. It is one of the best-known public-key cryptosystems. RSA is used in most digital data, information, and telephone security applications because it promises the privacy and authenticity of the data. The main purpose of using the RSA algorithm is that it solves the problem of distributing the key for encryption and decryption.

DNA cryptography can be defined as the method to hide the data in terms of the DNA sequence. DNA (Deoxy Ribonucleic Acid) is a long strand of polymer containing nitrogen bases named Adenine(A), Cytosine(C), Guanine(G), and Thymine(T). These four bases play an important role in encoding and decoding the plain text. The text data which is in the form of a DNA sequence is intended to be secret inside the image which is not secret and is visible to the audience. Using steganography, it is possible to communicate and considerably reduce information leakage. The steganography technique proposed in this paper offers immunity of secret data against modification, removal, compression, etc. Steganography has become fascinating and gruelling in the field of research. So, in this steganography, the hidden data can be distributed more evenly over the whole image to make it stronger and the scope for adding lots of data is allowed. The proposed algorithm in this paper is fast and robust than other cryptography algorithms.

The paper is structured as follows. In Section II the proposed methodology of the project is described in brief. Section III gives detailed explanation of the method applied and implemented using a combination of cryptography and steganography. The Experimental results are discussed in Section IV. Finally, the conclusion and the future scope is described in section V.

II. METHODOLOGY

The security of the data is a major concern with an increase in the use of the internet nowadays. With an intension to transfer the data securely an encryption model is proposed. This section describes the method followed i.e. DNA cryptography and chaotic neural network and then the proposed encryption method is explained in brief.

A. DNA cryptography

DNA (Deoxy Ribonucleic Acid) is a long strand of polymer containing nitrogen bases named Adenine(A), Cytosine(C), Guanine(G), and Thymine(T). These four bases play an important role, as they can be utilized for encoding the data being stream of 1's and 0's. The purpose of using DNA is that it can store and transmit a massive amount of data. In the proposed technique the plaintext is first converted to its respective ASCII value. Each ASCII value is then converted to a corresponding binary format. Then the binary sequence is transformed into a DNA sequence by mapping ATGC bases to the binary pairs.

B. Chaotic neural network

The main reason to add a chaotic neural network in this paper is because of its captivating communication system. It includes high-speed information transmission, tolerance to noise, and broadband power spectrum. In general, cryptography is easy to frame up with a chaotic system. One main concept in a chaotic system is that it generates a secret key which is pseudorandom in nature. It is considered to be more protected since it is troublesome to integrate the chaotic neural network with time varying delay.

C. RSA Algorithm

RSA was developed in 1977 by Ron Rivest, Adi Shamir, and Leonard Adleman. RSA is an asymmetric cryptosystem and consists of two different keys Public and Private keys. The public key is to encrypt the message and can be shared with everyone, whereas the private key is to decrypt the message and must be kept secret. The plain text is encrypted using the public key whereas it can only be decrypted using the private key [1]. The steps involved in the algorithm are:

a) Key generation

It involves generation of two key pairs, public and private key. To generate the key following steps are to be followed.

- Choose two large distinct prime no's p and q • Compute n=p*q.
- Compute phi=(p-1)*(q-1).
- Choose public exponent e(co-prime) between the interval [1, phi].
- The public key pair (n, e) is generated.
- The private key d is calculated using the e value by computing d*e= 1 (mod phi). The private key pair (n, d) is generated.

b) Encryption

The encryption step involves the generation of the ciphertext by using the public key value e. Any user with the public key can encrypt the plaintext.

$$C = M^e mod n \quad (1)$$

c) Decryption

The end user only with the private key value d can decode the ciphertext.

$$M = C^e mod n \quad (2)$$

D. Encryption process

The encryption process involves the RSA algorithm, chaotic neural network, DNA encoding, and steganography by guarantying triple layer security. The plain text is converted to a DNA sequence, which is wrapped by an image, upon which the RSA algorithm protects the data. The basic block diagram of the proposed method is shown in Fig 1. below.

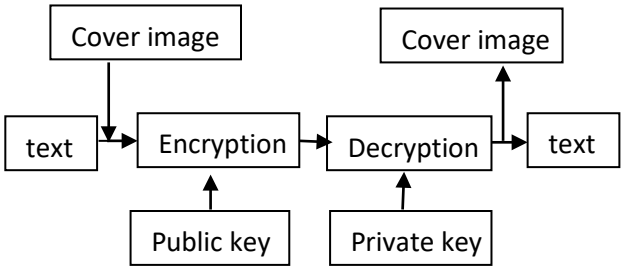


Fig 1. Basic block diagram of proposed method.

Firstly, the plaintext is converted to its corresponding ASCII value. The binary equivalent (Bj) of the respective ASCII values is obtained. A random binary sequence is generated using a chaotic neural network. The Hopfield delay is first calculated by solving the differential equation of the delay function using the fourth-order Range-Kutta method by considering n=2 and step function h=0.01. After the calculation of the delay a random binary sequence generator, generates a random sequence of length 42 bit. Suppose say, the random binary sequence of length 42 bit generated is considered as 111011000111001110011000110011101011100011. The sequence is then split into three subsequences. First 32 bits (Rj1) are used as a random key which is mainly used in XOR operations, the next 9 bits (Rj2) are used to set the number of shifts the binary sequences must undergo to create confusion to the intruder and the last bit is used to select the trajectory of the chaotic map. The sequence Rj1 and the binary equivalent of the ASCII value are subjected to left cyclic shifting by a certain bit specified by the sender.

The resultant sequences are then XORed, producing a 32bit stream of 1's and 0's (Ej).

$$E_i = R_i' \oplus B_i' \quad (3)$$

The resultant sequence (Ej) is then paired and mapped to DNA bases A, T, G and C respectively [4] by following Table 1, thus the plaintext is completely masked by DNA bases.

A	T	G	C
00	11	10	01

Table 1. DNA mapping

The DNA sequence obtained is to be hidden in an image using a steganographic process. The proposed work is based on the LSB steganography method, in which the cover image is captured in real-time with the aid of a web camera. The cover image captured is an RGB image and has to be converted to a greyscale image for the ease of computation. The greyscale is chosen over an RGB image because the greyscale image consists of an 8-bit plane and is easy to extract them whereas the RGB image consists of a 24-bit plane and it may require more computational step to extract the bit plane. The DNA sequence is then mapped back to the binary. The bit plane in which the sequence to be hidden is selected. The pixel value of that bit plane is extracted to which the binary sequence is appended at the least significant bit place. The DNA sequence hidden in the image is referred to as stegno-image, which is transmitted to the end-user.

E. Decryption process

The Steganographic image can only be decrypted by an end-user with access to private key. Firstly, the steganographic image is obtained by providing an accurate private key d. The next step is to decrypt the steganographic image in order to retrieve the DNA sequence. At first, the bit plane in which the sequence is hidden is extracted from the greyscale image. The pixels of the image are then obtained is converted to binary values, the least significant bits are recognized and are grouped. In the proposed method three least significant bits are considered. If the bit range increases a noise or distortion in the steganographic image can be noticed hence choosing three least significant bits is efficient. Thus, the DNA sequence is retrieved back from the image [6].

The DNA sequence is decoded using Table 1 resulting a binary sequence (Ej). The sequence Ej is then XORed with the random sequence (Rj'), (Bj') is obtained as a result. Right cyclic shift operation is performed on Bj', which yields (Bj) the binary equivalent of the ASCII code. The plain text is retrieved by considering the generated ASCII values.

III. IMPLEMENTATION

In this section, the encryption and decryption process of the proposed algorithm based on chaotic neural network, DNA cryptography and steganography are briefly described.

A. Binary Sequence Generator

A random binary sequence is generated based on chaotic neural network which is identical in nature. The binary sequence generated randomly depends on the trajectories of chaotic map [7]. Consider a chaotic map x(.) generating a binary sequence in the interval k=[s,t] can be defined as,

$$\frac{x-s}{t-s} = 0.a1(x).a2(x).....ai(x), x \in [s,t] \quad (4)$$

where ai(x) ϵ {0,1} and

$$ai(x) = -\sum_{j=1}^{2^i-1}(-1)^{j-1}\emptyset(t-2)(j \div 2^i) + s^y \quad (5)$$

where, ∅(t) is the threshold function. The random binary sequence is generated by nth iteration.

B. Encryption process

The popular RSA algorithm followed by DNA coding is used to encrypt the message. The public key pair and private key pair are generated by using the RSA algorithm. The plain text is then encrypted by using the public key (e), which yields the ciphertext. The DNA encoding is achieved by the means of a chaotic neural network [8], the steps followed in the encryption process is shown in Fig 2.

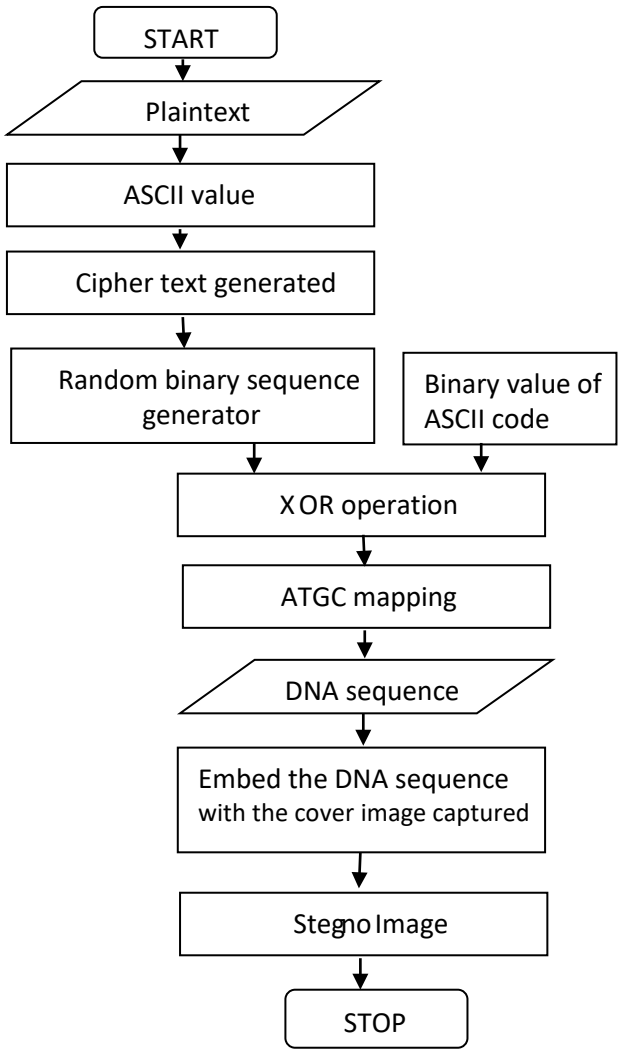


Fig 2. Flowchart of the encryption process

Step1: A random binary sequence of length 42-bit is generated. Which is then split into three sections. The first 32 bit can be used as key; the next 9 bits are used by the sender to set the decimal value for the shifting operations and the last bit is used as a trajectory.

Step2: The binary equivalent value of the ASCII code of the plain text is then XORed with the random binary sequence generated by a chaotic neural network (here the sequence are subjected to left shift by j times before XOR operation, j is the decimal value generated by the random binary sequence generator), which yields a stream of binary sequence [9].

Step3: The resultant binary sequence after performing XOR operations is then grouped into pairs. The pairs are then mapped into DNA sequence by simple substitution for 00 as A, 01 as C, 10 as G and 11 as T as shown in Table 1.

Step4: The DNA sequence formed, is then embedded into an image [10] by following the Steganographic method.

C. Steganographic process

A real-time captured image is used as a cover image to hide the DNA sequence. The least significant bits (LSB) method is used to hide the DNA sequence in the cover image. The method involves replacing the least significant bits of the cover image pixels of a given byte. LSB method is significant and easier, but if the bits exceed more than three bits place a small distortion noticed in the steganographic image [11], Fig 2 explains the steps involved in steganography. The steps followed in embedding the DNA sequence with the cover image are :

Step1: Capture the cover image by the means of a web camera. The captured image will be a color image.

Step2: Convert the cover image to greyscale image by using an image conversion unit by selecting any channel, in this paper a green channel is selected.

Step3: Convert the DNA sequence to the stream of bytes by mapping back to binary sequence referring the Table1.

Step4: Select a bit plane (from 1 to 8-bit planes) to hide the DNA sequence.

Step5: Compute the segment of the cover image.

D. Decryption process

The DNA sequence is retrieved from the pixels [12] of the cover image by considering the bit plane in which the sequence is hidden. The steps followed in decoding the DNA sequence is shown in Fig 3 and is explained below

Step1: The DNA sequence is mapped back to the binary sequence by following ATGC mapping by assigning a pair of binary values for the respective DNA bases as shown in Table 1.

Step2: The streams of 1's and 0's obtained after DNA mapping is then XORed with the random binary sequence.

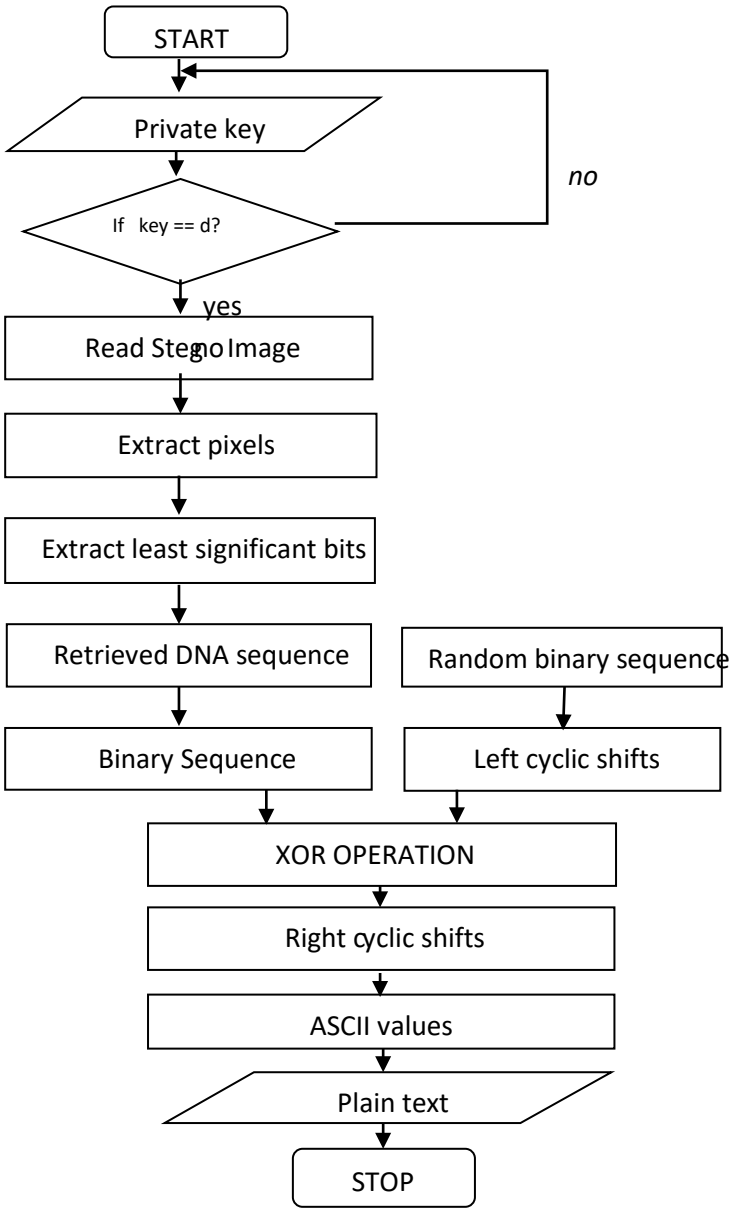


Fig 3. Flowchart of the decryption process

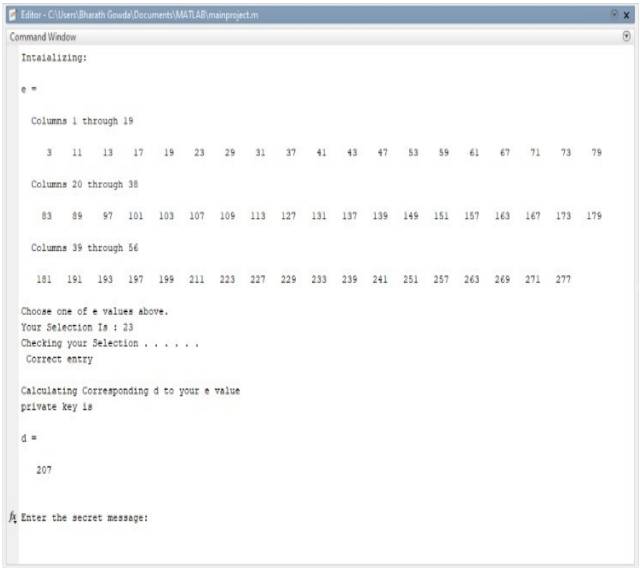
Step3: Right shift is performed j times on the stream of a binary sequence. The ASCII values are obtained using the plain text is retrieved.

IV. RESULTS AND DISCUSSION

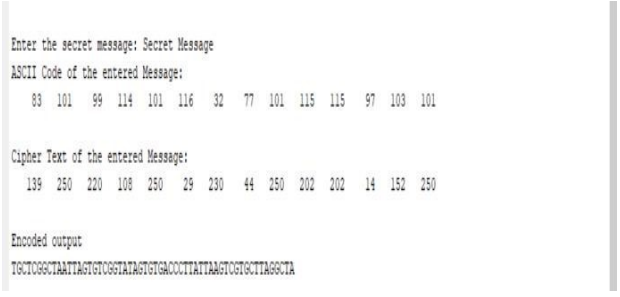
The results of the proposed method are shown above in Fig 4 and Fig 5. Fig 4 describes the complete steps involved in the encryption process, firstly a menu presenting all the steps involved in the proposed work is displayed. The process begins with key generation using the RSA algorithm followed by DNA encoding the plaintext. The DNA sequence is then hidden in an image captured by a web camera generating a steganographic image as shown in Fig 4 d). Fig 5 describes the results obtained from each step of decryption. The inference made from the above results is that the proposed method

guarantees triple layer security to the data so that the sender and the receiver remain contented.

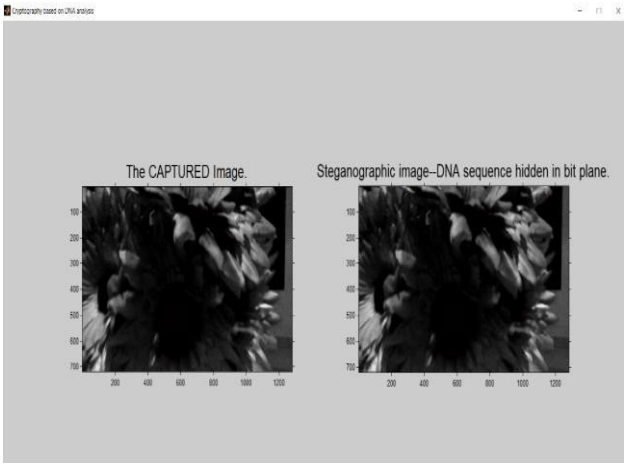
(a)



(b)



(c)



(d)

Fig 4. (a) menu displaying all the steps involved in the program (b) RSA key generation (c) Plaintext masked into the DNA sequence. (d) Generation of Steganographic image by hiding the DNA sequence in any one of the bit planes of the greyscale image.

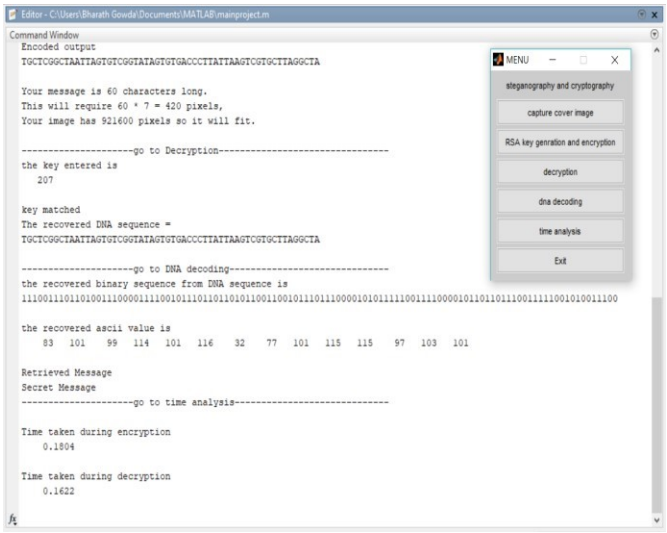


Fig 5. Decryption of the plaintext

V. CONCLUSION

The proposed paper is based on the union of cryptography and steganography, the combination produces more robust and secure algorithm. Using DNA and RSA algorithm we could provide multi-layer reliability of the confidential information. So, this technique can be applied and used for various applications to keep the data safe. Further DNA cryptography methods can be enhanced using polymerase chain reaction (PCR) and text can also be embedded within a video or audio media.

REFERENCES

[1] Taki, A. E., Deen, E., & Gobran, S. N. (2014). Digital Image Encryption Based on RSA Algorithm, 9(1), 69–73.

[2] Yu, Wenwu, and J. Cao. "Cryptography based on delayed chaotic neural networks," Physics Letters, vol. A 356.4, pp. 333-338, August 2006.

[3] C. Fu, Z. Zhu, "A chaotic image encryption scheme based on circular bit shift method.", in The 9th International Conference for Young Computer Scientists. pp. 3057–3061, 2008.

[4] Qiang Zhang, Xianglian Xue, and XiaopengWei, "A Novel Image Encryption Algorithm Based on DNA Subsequence Operation", The Scientific World Journal, 2012, Volume 2012, Article ID 286741

[5] Ms. Pallavi Hemant Dixit," Arm Implementation of LSB Algorithm of Steganography" Submission, Shivaji University, Kolhapur, Maharashtra, India

[6] Sajisha K S, Dr. Sheena Mathew, "An Encryption based on DNA cryptography and Steganography", IEEE, 2017.

[7] T. Kohda, A. Tsuneda, "Statistics of chaotic binary sequences," IEEE Transactions on information theory, vol. A 43.1, pp. 104-112, January 1997

[8] A. Khare A, B. Shukla P, C. Silakari S. Secure and Fast Chaos based Encryption System using Digital Logic Circuit. Int J Comput Netw Inf Secur [Internet]. 2014; 6(6): 25–33.

[9] Tian Tian Zhang, Shan Jun Yan, Cheng Yan Gu, Ran Ren and Kai Xin Liao, "Research on Image Encryption Based on DNA Sequence and Chaos Theory", IOP Conf. Series: Journal of Physics: Conf. Series 1004 (2018)