A Novel Encryption Technique to Protect Patient Health Information Electronically Using Playfair Cipher 15 by 14 Matrix

Nisarga Chand, Subhajit Bhattacharyya, and Angsuman Sarkar

Abstract Healthcare data such as patient name, DOB, address, phone number, fax number, E-mail address, etc., need to be kept private to protect patient privacy. It has been found that unauthorized person always try to get this information. With the help of proper encryption technique, we can protect the information. Play fair cipher is a very much well-known encryption technique for text messages. It is an example of a digraph substitution cipher. In this technique, pair of alphabets (digraphs) was encrypted instead of a single alphabet. That is why, it becomes so strong and not easy to break by conventional method. Play fair is fairly fast to apply and calls for no unique tools. It converts a plain text to cipher text message mathematically with the help of a secret key. Likewise, while restoring the original message from encrypted message, the same secret key has been used. However, the traditional play fair cipher supports only 25 uppercase alphabets. Here, in our work, we use 210 characters instead of 25 characters earlier, which make the cipher text stronger. These 210 characters include all the alphanumeric characters as well as 148 special characters supported by MATLAB. Here, we carried out cryptanalysis and found that the cipher is a strong one also it is thus almost impossible to be broken by any cryptanalytic attack. Finally, we have represented the final result with the help of MATLAB with two iteration steps.

Keywords Play fair cipher · Encryption · Iteration steps · Special characters · Cryptanalysis

N. Chand (✉)
Adamas University, Kolkata, India

S. Bhattacharyya
Mallabhum Institute of Technology, Bishnupur, Bankura, India

A. Sarkar
Kalyani Government Engineering College, Kalyani, Nadia, India
e-mail: angsumansarkar@ieee.org

M. Mukherjee et al. (eds.), *Advances in Medical Physics and Healthcare Engineering*,
Lecture Notes in Bioengineering, https://doi.org/10.1007/978-981-33-6915-3_42

1 Introduction

In the year 2000 surprisingly, it has been found that many patients got free samples of medication recommendations for the disease they were just diagnosed a few months ago. The case amazed the patients that how the pharmaceutical companied came to know about their disease. A high level and prolonged investigation were initiated which proved that a physician, a pharmaceutical company, and a well-known pharmacy chain were involved in this case of the breach of confidentiality.

So, the government needed to come forward and create guidelines to protect patient privacy. Then, Health Insurance Portability and Accountability Act (HIPPA) comes into picture in the year 1996 (Cohen and Mello 2018). According to HIPPA, any information that can recognize a patient is regarded protected health information (PHI). Anything related to health, treatment or billing that could identify a patient is PHI. This includes (a) Name, (b) DOB, (c) Address, (d) Phone number, Fax number, (e) e-mail address etc.

Healthcare data generally means protected health information (PHI) which must be kept confidential, so that no unauthorized person can access it (Kayaalp 2018). The most common method to guard information is the technique of encryption and decryption. In encryption sender, conceal the message with a secret key following some sort of rules. These rules and the secret key are only known to the person who is supposed to receive the message. With the proper set of rules, the secret key receiver restores the original message that is referred to as symmetric encryption (Stallings 2006). The encryption method to protect the health information must be very strong, so that if somehow this information goes into wrong hand, it cannot be decoded very easily. The traditional encryption structure be made up of original message which needs to be encrypted, a set of rules called encryption algorithm, confidential key, cipher text, again a set of rules called decryption algorithm. We require a sturdy encryption algorithm with a view to encrypt the original message content into cipher text (Kahate 2007). The one who encrypt and transmit the message and the one who collect and decrypt the message have to acquire the confidential key, so that any third party cannot access it. Amidst each and every prevailing encryption algorithm, play fair cipher dominates.

The basic building block of play fair cipher algorithm is a 5 by 5 matrix and a secret keyword (Kahate 2007). Each element of this 5 by 5 matrix is filled by secret key first and then with alphabets which are not present in the keyword sequentially. Twenty-five English uppercase alphabets are only allowed in this algorithm, that is why we have enhanced the previous matrix size from 5 × 5 to 15 × 14, so that we can incorporate sixty-two alphanumeric as well as one forty-eight special characters in it.

Table 1 5 × 5 traditional play fair cipher matrix with keyword AMPHE

A	M	P	H	E
B	C	D	F	G
I/J	K	L	N	O
Q	R	S	T	U
V	W	X	Y	Z

2 Traditional Play Fair Cipher

Conventional play fair cipher uses a 5 × 5 matrix, and each element of this matrix is filled up by twenty-five distinct English alphabets (Khan 2015a). First the confidential keyword is placed without any repetition of letters while traversing row wise sequentially. Remaining elements of the matrix is being filled up by rest of the letters in alphabetical order. Here, we choose confidential keyword as "AMPHE," and the resultant conventional matrix is shown in Table 1.

Play fair cipher algorithm always consider I & J as one letter. It is visible that the policies of encryption practice a couple of plaintext characters. If there are even numbers of characters in the original message, then the algorithm works fine. But if somehow the original message consists of odd number of characters, then an extra letter X is appended after the termination of original message (Khan 2015b). In addition, if same letter found in any pair, then an extra letter X is inserted between them, and again new pair of letters are reconstructed.

The protocols are mentioned below:

(i) Original digraphs which is appeared in the same row are changed by the immediate right contents, also the first content of the row circularly attached with the last content. As an example, UV is converted to VW during the encryption process.

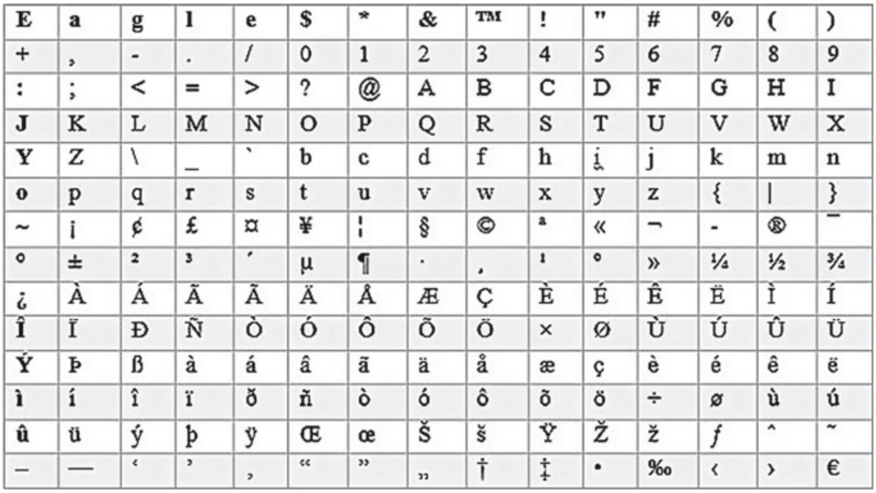
(ii) Original digraphs which is appeared in the same column are changed by the below contents in the matrix, also the top content of the column circularly attached with the last content. As an example, HF is converted to FP during the encryption process.

(iii) Most of the cases, if the digraphs are not mentioned in the same row or column, then choose other two intersecting points as a content of the rectangular shape. So that, SU is converted to LX, and CA becomes BM (Bhattacharyya et al. 2014).

3 Modified 15 × 14 Play Fair Cipher Algorithm

This revised play fair cipher algorithm uses a 15 by 14 matrix, 210 characters instead of 25 earlier, and a keyword. Out of 210 characters, 148 are special characters which are supported by MATLAB. The elements of the matrix are filled up by characters

Table 2 15 × 14 play fair cipher matrix with keyword Eagle$*&™



E	a	g	l	e	$	*	&	™	!	"	#	%	(	)
+	,	-	.	/	0	1	2	3	4	5	6	7	8	9
:	;	<	=	>	?	@	A	B	C	D	F	G	H	I
J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
Y	Z	\	_	`	b	c	d	f	h	i	j	k	m	n
o	p	q	r	s	t	u	v	w	x	y	z	{	|	}
~	¡	¢	£	¤	¥	¦	§	©	ª	«	¬	-	®	¯
°	±	²	³	´	µ	¶	·	¸	¹	º	»	¼	½	¾
¿	À	Á	Â	Ã	Ä	Å	Æ	Ç	È	É	Ê	Ë	Ì	Í
Î	Ï	Ð	Ñ	Ò	Ó	Ô	Õ	Ö	×	Ø	Ù	Ú	Û	Ü
Ý	Þ	ß	à	á	â	ã	ä	å	æ	ç	è	é	ê	ë
ì	í	î	ï	ð	ñ	ò	ó	ô	õ	ö	÷	ø	ù	ú
û	ü	ý	þ	ÿ	Œ	œ	Š	š	Ÿ	Ž	ž	ƒ	ˆ	˜
–	—	‘	’	‚	“	”	„	†	‡	•	‰	‹	›	€

of keyword first from left hand side to right hand side and from upper to lower, and the remaining elements of the matrix are filled up by rest characters.

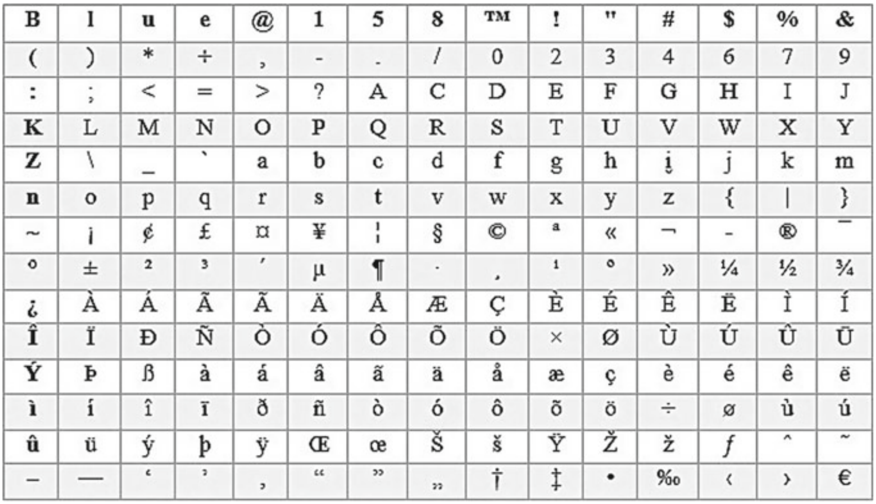
3.1 Assumption

In this algorithm implementation, two keywords are taken. First one is Eagle$*&™, and the second one is Blue@158™. With the assist of those two keywords, two new 15 by 14 matrices have been formed which are depicted in Tables 2 and 3.

3.2 Algorithm

- At first, user enter the message which needs to be encrypted.
- In the second stage, any space between the words of the entered message is removed.
- In the third stage, symbol "X" is inserted in the middle of two characters which are same.
- In the fourth stage, the output of the second stage is encrypted with the help of the keyword "Eagle$*&™".
- The output of the fourth stage is again encrypted with the help of the keyword "Blue@158™" in the fifth stage.
- For same row pair character if any one of the characters situated at last column, then during encryption, it becomes foremost column character. And for same

Table 3 15 × 14 play fair cipher matrix with keyword Blue@158™



B	l	u	e	@	1	5	8	™	!	"	#	$	%	&
(	)	*	+	,	-	.	/	0	2	3	4	6	7	9
:	;	<	=	>	?	A	C	D	E	F	G	H	I	J
K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y
Z	\	_	`	a	b	c	d	f	g	h	i	j	k	m
n	o	p	q	r	s	t	v	w	x	y	z	{	|	}
~	¡	¢	£	¤	¥	¦	§	©	ª	«	¬	-	®	¯
°	±	²	³	´	µ	¶	·	¸	¹	º	»	¼	½	¾
¿	À	Á	Ã	Ã	Ä	Å	Æ	Ç	È	É	Ê	Ë	Ì	Í
Î	Ï	Ð	Ñ	Ò	Ó	Ô	Õ	Ö	×	Ø	Ù	Ú	Û	Ū
Ý	Þ	ß	à	á	â	ã	ä	å	æ	ç	è	é	ê	ë
ì	í	î	ī	ð	ñ	ò	ó	ô	õ	ö	÷	ø	ù	ú
û	ü	ý	þ	ÿ	Œ	œ	Š	š	Ÿ	Ž	ž	ƒ	ˆ	˜
–	—	‘	’	‚	“	”	„	†	‡	•	‰	‹	›	€

column pair character if any one of the characters situated at first column, then during encryption, it becomes rearmost row character. This is a very vital rule for encryption.

- Next, in the sixth stage, we decrypt the output of fifth stage with Keyword “Blue@158™” and then again decrypt the output of fourth stage with the keyword “Eagle$*&™”.
- The above-mentioned vital rule for encryption is also applied in case of decryption process.
- From the last stage, we get the original message back which is the output of step number two.

3.3 Simulation Flowchart

See Fig. 1.

3.4 Matrix Formation with Keyword

See Tables 2 and 3.

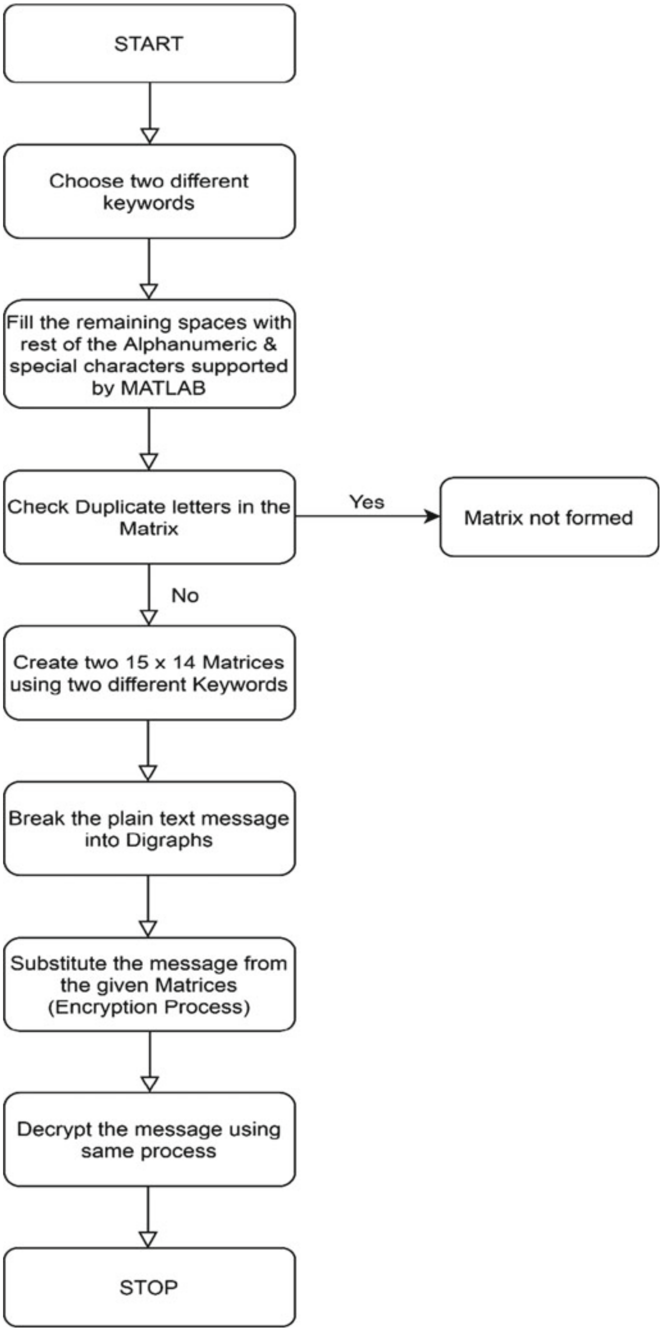

START
Choose two different keywords
Fill the remaining spaces with rest of the Alphanumeric & special characters supported by MATLAB
Check Duplicate letters in the Matrix
Yes
Matrix not formed
No
Create two 15 x 14 Matrices using two different Keywords
Break the plain text message into Digraphs
Substitute the message from the given Matrices (Encryption Process)
Decrypt the message using same process
STOP

Fig. 1 Flowchart of encryption & decryption process using 15 × 14 play fair cipher matrix

4 Cryptanalysis of Modified Cipher

Encryption is used to protect the information from attackers who are always tried to decrypt the original message from encrypted text. Different forms of cryptanalytic attacks are discussed below.

- As the matrix size is 15 × 14, we can choose the secret key in 210! (Factorial 210) different ways. So, it can be easily concluded that it is immune to brute force attack.
- If frequency analysis cipher text attack is used to decrypt the message, then also number of combinations to be searches would be 210 * 210 = 44,100.
- But if somehow attacker knows the encrypted message and original message behind this encrypted message, then he can easily get the key and decrypt all the message encrypted using this key.

5 Advantages of 15 × 14 Play Fair Cipher Algorithm

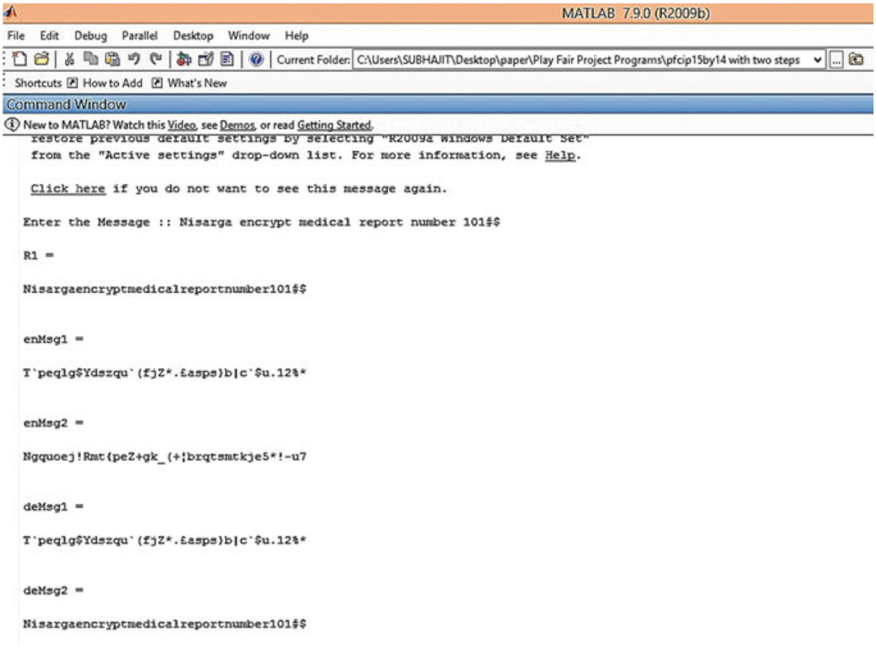
The primary downside of conventional play fair cipher is that original message can be constructed using only 25 English uppercase alphabets. One cannot include lowercase English alphabets, special characters, and decimal numbers in it. To triumph over the disadvantages, we put into effect a changed cipher which employs a 15 × 14 matrix to be able to incorporate nearly all of the printable characters supported through MATLAB. So, any secret message can be encrypted with the help of this methodology.

6 Experimental Result

To implement the modified play fair cipher 15 by 14 matrix, we have used MATLAB. In Fig. 2, we have presented the result that we have obtain in MATLAB. From the figure, one can easily identify original text message, two strong encrypted message and finally decrypted message which is same as original message.

7 Conclusion

The main objective of this paper was to secure healthcare data of patient which is at risk now a days according to survey conducted by different agencies. For this purpose, a new strong encryption technique has been demonstrated here. In our work, we have used the original play fair cipher concept and enriched it by adding more features into it. Also, the limitations of earlier play fair cipher matrices like 5 × 5, 6


MATLAB 7.9.0 (R2009b)
File Edit Debug Parallel Desktop Window Help
Current Folder: C:\Users\SUBHAJIT\Desktop\paper\Play Fair Project Programs\pfcip15by14 with two steps
Shortcuts How to Add What's New
Command Window
New to MATLAB? Watch this Video, see Demos, or read Getting Started.

```
  restore previous default settings by selecting "R2009a Windows Default Set"
  from the "Active settings" drop-down list. For more information, see Help.

  Click here if you do not want to see this message again.

Enter the Message :: Nisarga encrypt medical report number 101#$

R1 =

Nisargaencryptmedicalreportnumber101#$


enMsg1 =

T`peqlg$Ydszqu`(fjZ*.£asps)b|c`$u.12%*


enMsg2 =

Ngquoej!Rmt(peZ+gk_(+¦brqtsmtkje5*!-u7


deMsg1 =

T`peqlg$Ydszqu`(fjZ*.£asps)b|c`$u.12%*


deMsg2 =

Nisargaencryptmedicalreportnumber101#$
```

Fig. 2 Output of 15 × 14 play fair cipher algorithm

× 6 and 10 × 9 have been eliminated in this work. Here, we have enriched the basic play fair cipher by adding two additional features. First one is enhanced matrix size which is of size 15 × 14. As the matrix size is increased, we can easily incorporate more characters into our matrix. So, the second additional feature is incorporation of almost all characters supported by MATLAB into our new enhanced matrix. So naturally, it has been found that our encrypted message looks very strong than earlier one. For further enhancement of our work, two key can be used one for encryption of message and other for decryption of message. Also, modern encryption method such as steganography can be included to enrich the security of information. So, at last, we can conclude that this work can help the people in coming future so secure their medical privacy.

References

Bhattacharyya S, Chand N, Chakraborty S (2014) A modified encryption technique using playfair cipher 10 by 9 matrix with six iteration steps. Int J Adv Res Comput Eng Technol 3:307–312

Cohen IG, Mello MM (2018) HIPAA and protecting health information in the 21st century. JAMA 320(3):231–232. https://doi.org/10.1001/jama.2018.5630

Kahate A (2007) Cryptography and network security, 2nd edn. Tata McGraw-Hill Publishing Company Limited, New Delhi

Kayaalp M (2018) Patient privacy in the era of big data. Balkan Med J. 35(1):8–17. https://doi.org/10.4274/balkanmedj.2017.0966

Khan SA (2015a) Design and analysis of playfair ciphers with different matrix sizes. Int J Comput Netw Technol 3:117–122

Khan SA (2015b) Design and analysis of playfair ciphers with different matrix sizes. Int J Comput Netw Technol 3(3):117–122

Stallings (2006) Cryptography and network security: principles and practice, 4th edn. Prentice Hall