ENCRYPTION TECHNOLOGY FOR TAMIL LANGUAGE (TAMILAN CIPHER)

First Author: P.Thamizhikkavi **Second Author: Dr.S.Magesh**

Abstract-Tamilan cipher is a new encryption tool with is used for the encryption for regional language (Tamil). In this cipher plain text is in Tamil language and the key is been taken in values. This encryption tool is been developed for the increase in the data security over the internet.

Keywords-cipher text, encryption, key value, mei, plain text, uir, uirmei.

I INTRODUCTION:

In modern technology there are many encryption technics such as RSA, DES, etc. All these encryption technics are been evolved for the international language English[5]. Since at the beginning only English is been used over the internet for the worldwide communication. Now days confidential data are been transfer over the internet in regional language also. We know that all the encryption technics can be cracked by the third person and once the plain text is been captured it can be easily translated into English with the help on online translator. To overcome this problem here evolves the new technic called as Tamilan cipher.

II EXISTING SYSTEM:

All the encryption techniques are been cracked by any third person and the data is been theft. How much ever the encryption level may increase since the methodology is been known worldwide decryption process can be done by any one. Even if the data is not in English it can be easily translated into any language using the online translator.

First Author:*P.Thamizhikkavi,MTech (ISCF),Department of IT,S.R.M University.*

Second Author: *Dr.S.Magesh, Asst.Professor (Sr.G), Department of IT, S.R.M University,*

III PROPOSED SYSTEM:

If the encryption method is been changed or modified or updated over the existing method it increases the level of security over the data. New method created to increase the level of security is called as Tamilan cipher. Tamilan cipher is a substitution technic for the Tamil characters.

IV PROCEDURE:

Encryption process for a substitution methods like ceaser cipher, play fair cipher, rail fence cipher vigenere cipher...etc. in all these methods a character is been substituted/changed to the other character mapped by some formula or relation. For Tamilan cipher will make use of the method followed n full vigenere cipher.

A. Full Vigenere Cipher:

In normal vigenere cipher there is a substitution box which is been created in sequence manner such the key is also an alphabet. The drawback of this type is if one word is been identified the with the sequence of alphabets we can easily guess or obtain the full decrypted text (also called as plain text). To overcome this problem they created a new cipher called full vigenere chipher. In full vigenere cipher the plain text and the key is same but the characters inside the substitution box is been randomly placed according to the combinations of character such that no same character is been repeated in the same row or same column [1].

B. Tamilan Cipher:

In English language there are only 26 characters available. Even though it is been classified into two types such as vowels and non-vowels when come for substitution techniques all the characters are been taken into account. Whereas in Tamil language there are totally 247 character available[2]. These 247 characters are future classified into four type's names uir, Mei, uirmei

ISSN: 2278 – 7798

and aiudham. There are 12 characters available in uir (i.e. '□'). 18 characters are available in mei (i.e. '□□'). In uirmei there are totally 216 characters (12 x 18 = 21) in other words it can be classified as uir x mei =uirmei. i.e. '□□ + □ = □'. Ayudha eluthu which is just a single character called as akku (i.e. '□'). Below figure illustrate the characters available in Tamil language. Few characters from other languages are also been used in Tamil language but not often.

The above is general character format on Tamil language whereas when comes to computer part character are classified into same four type but a small change in it. Mei is without the dot and in uirmei combination are been made with the help of additional characters such as '□ + ி = கி'[4].

∴	அ	ஆ	இ	ஈ	உ	ஊ	எ	ஏ	ஐ	ஒ	ஓ	ஔ
க்	க	கா	கி	கீ	கு	கூ	கெ	கே	கை	கொ	கோ	கௌ
ங்	ங	ஙா	ஙி	ஙீ	ஙு	ஙூ	ஙெ	ஙே	ஙை	ஙொ	ஙோ	ஙௌ
ச்	ச	சா	சி	சீ	சு	சூ	செ	சே	சை	சொ	சோ	சௌ
ஞ்	ஞ	ஞா	ஞி	ஞீ	ஞு	ஞூ	ஞெ	ஞே	ஞை	ஞொ	ஞோ	ஞௌ
ட்	ட	டா	டி	டீ	டு	டூ	டெ	டே	டை	டொ	டோ	டௌ
ண்	ண	ணா	ணி	ணீ	ணு	ணூ	ணெ	ணே	ணை	ணொ	ணோ	ணௌ
த்	த	தா	தி	தீ	து	தூ	தெ	தே	தை	தொ	தோ	தௌ
ந்	ந	நா	நி	நீ	நு	நூ	நெ	நே	நை	நொ	நோ	நௌ
ப்	ப	பா	பி	பீ	பு	பூ	பெ	பே	பை	பொ	போ	பௌ
ம்	ம	மா	மி	மீ	மு	மூ	மெ	மே	மை	மொ	மோ	மௌ
ய்	ய	யா	யி	யீ	யு	யூ	யெ	யே	யை	யொ	யோ	யௌ
ர்	ர	ரா	ரி	ரீ	ரு	ரூ	ரெ	ரே	ரை	ரொ	ரோ	ரௌ
ல்	ல	லா	லி	லீ	லு	லூ	லெ	லே	லை	லொ	லோ	லௌ
வ்	வ	வா	வி	வீ	வு	வூ	வெ	வே	வை	வொ	வோ	வௌ
ழ்	ழ	ழா	ழி	ழீ	ழு	ழூ	ழெ	ழே	ழை	ழொ	ழோ	ழௌ
ள்	ள	ளா	ளி	ளீ	ளு	ளூ	ளெ	ளே	ளை	ளொ	ளோ	ளௌ
ற்	ற	றா	றி	றீ	று	றூ	றெ	றே	றை	றொ	றோ	றௌ
ன்	ன	னா	னி	னீ	னு	னூ	னெ	னே	னை	னொ	னோ	னௌ

Fig 1. Total characters in Tamil language[2]

This Tamilan cipher is been created based on the computer format of Tamil language.

C. Flow of Tamilan Cipher:

In Tamilan cipher plain text will be in Tamil character and key will be in values. Once the plain text and key is obtained from the user system will select the first character in the plain text and process that character. Since this is a substitution cipher only single character will be taken into process. Once the first key is been selected it will check for the type of that character either it comes under uir or mei or uirmei or others. If it falls on any one type such as uir or mei or uirmei the following procedure is been followed and the following S box are been used for substitution. Below flowchart represent the main working of Tamilan cipher.

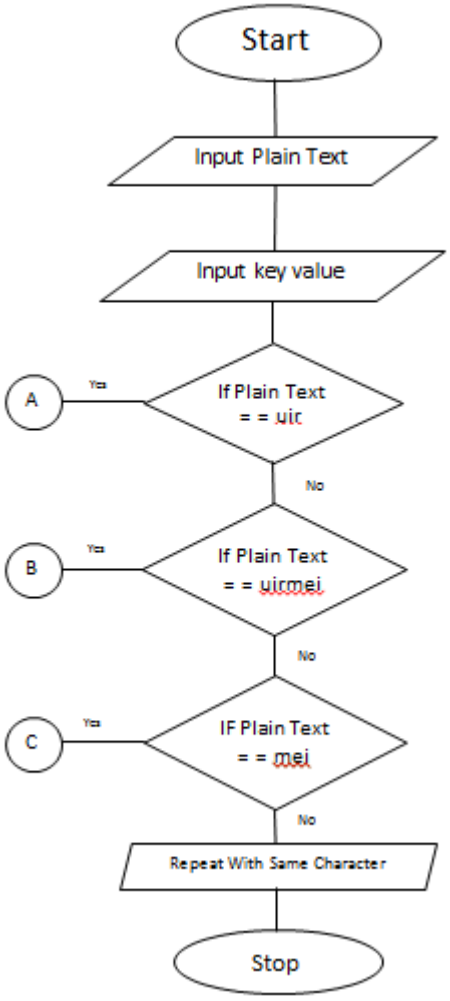

Fig 2. Flowchart of Tamilan cipher

D. List of S Box

As we discussed above there are three substitution box created for the working process of Tamilan cipher. Each box is been explained below

>> UIR

The first classification is called as uir eluthu. There are 12 characters in this classification. Those characters are '□□□□□□□□□□□□'. If the plain text is of this type then the following flow chat will be followed for the encryption process. Key value % 12 is been calculated so that the value ranges from 0 to 11. Since there are only 12 characters available.

ISSN: 2278 – 7798

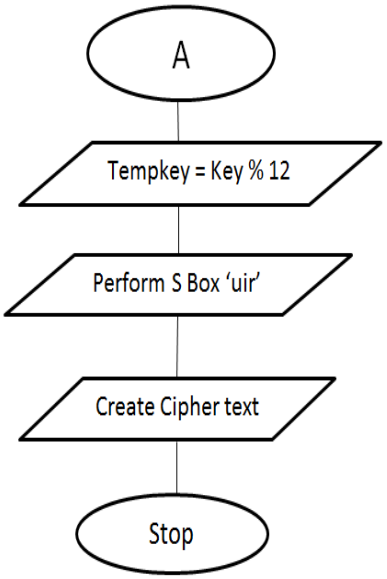

Fig 3. Flowchart of 'uir'

These characters are individual characters i.e. these characters will not exit with the combinations of any other characters. So these characters are needed to be replaced only with the same classification of characters. Thus the substitution box is been created in such manner. The below diagram shows the substitution box for uir eluthu.

	0	1	2	3	4	5	6	7	8	9	10	11
அ	ஃ	ஹ	உ	ஔ	ஒ	ஓ	அ	இ	எ	ஆ	ஏ	ஐ
ஆ	ஹ	அ	இ	எ	ஔ	ஐ	ஃ	ஆ	ஒ	உ	ஓ	ஏ
இ	இ	ஆ	ஹ	ஏ	ஐ	ஔ	உ	அ	ஓ	ஃ	ஒ	எ
ஃ	ஒ	ஔ	ஓ	ஹ	ஃ	உ	எ	ஐ	அ	ஏ	ஆ	இ
உ	ஐ	ஏ	ஔ	ஆ	இ	ஹ	ஓ	எ	உ	ஒ	ஃ	அ
ஹ	ஓ	ஐ	ஏ	இ	உ	ஆ	ஔ	ஒ	ஹ	எ	அ	ஃ
எ	ஏ	ஒ	எ	ஃ	ஆ	அ	ஐ	ஓ	இ	ஔ	ஹ	உ
ஏ	எ	ஓ	ஒ	உ	அ	ஃ	ஏ	ஔ	ஆ	ஐ	இ	ஹ
ஐ	அ	உ	ஃ	ஓ	எ	ஒ	ஆ	ஹ	ஏ	இ	ஐ	ஔ
ஒ	ஔ	எ	ஐ	அ	ஹ	இ	ஒ	ஏ	ஃ	ஓ	உ	ஆ
ஓ	உ	இ	ஆ	ஐ	ஓ	ஏ	ஹ	ஃ	ஔ	அ	எ	ஒ
ஔ	ஆ	ஃ	அ	ஒ	ஏ	எ	இ	உ	ஐ	ஹ	ஔ	ஓ

Fig 4. Substitution box for 'uir'

With the new modulated key and the plain text (character) a new character is been replaced with the help of the above box. I.e. for plain character '□' and the key value 7 the cipher character is '□'

>> MEI

The second classification is called as mie eluthu. There are 18 characters in this classification. Those characters are '□□□□□□□□□□□□□□□□□□'. If the plain text is of this type then the following flow chat will be followed for the encryption process. Key value % 18 is been calculated so that the value ranges from 0 to 17. Since there are only 18 characters available.

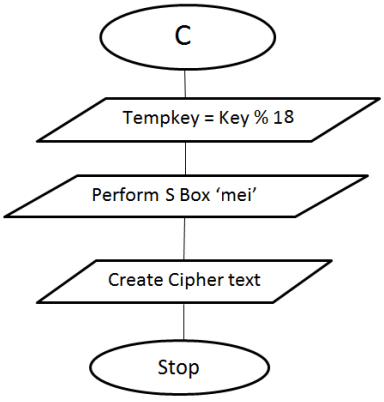

Fig 5. Flowchart for 'mei'

These characters are individual and combined characters. Either these characters are displayed individual or been combined with other classification called as uirmei there exit the meaning for that character where are they are not supposed to be replaced with other classification. Thus the substitution box is been created in such manner. The below diagram shows the substitution box for mei eluthu.

With the new modulated key and the plain text (character) a new character is been replaced with the help of the above box. I.e. for plain character '□' and the key value 7 the cipher character is '□'.

ISSN: 2278 – 7798

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
க	க	வ	ஞ	ல	த	ற	ப	ள	ந	ன	ண	ழ	ச	ர	ங	ம	ட	ய
ங	ம	க	ய	ஞ	ள	ண	ந	ட	ல	ச	ப	ங	வ	ழ	த	ற	ர	ன
ச	த	ங	ம	க	ல	ட	ச	ஞ	வ	ண	ர	ப	ழ	ன	ந	ள	ய	ற
ஞ	ய	ண	ந	ட	ற	ச	த	ங	ன	ஞ	ம	க	ல	ள	ர	வ	ப	ழ
ட	ந	ச	த	ங	வ	ஞ	ர	க	ழ	ட	ய	ண	ன	ற	ப	ல	ம	ள
ண	ர	ஞ	ப	ண	ன	க	ய	ச	ள	ங	ந	ட	ற	வ	ம	ழ	த	ல
த	ப	ட	ர	ச	ழ	ங	ம	ண	ற	க	த	ஞ	ள	ல	ய	ன	ந	வ
ந	ல	ந	ள	த	க	ய	ழ	ம	ங	ர	ன	ச	ப	ண	வ	ஞ	ற	ட
ப	ண	ள	ச	ன	ர	ல	ட	ழ	ம	வ	ங	ற	ய	ந	ஞ	ப	க	த
ம	ன	ம	ழ	ர	ண	த	ற	ப	ஞ	ந	வ	ய	ட	ங	ள	ச	ல	க
ய	ங	ழ	க	வ	ந	ள	ண	ல	ப	ற	ட	ன	ர	ய	ச	த	ஞ	ம
ர	ற	ர	வ	ய	ட	ப	ல	ந	ண	ம	ள	த	க	ஞ	ன	ங	ழ	ச
ல	ச	ற	ண	ழ	ப	வ	ஞ	ன	ய	ல	க	ள	ம	த	ட	ர	ங	ந
வ	ள	த	ற	ம	ஞ	ர	வ	ய	க	ப	ழ	ந	ங	ச	ல	ட	ன	ண
ழ	ஞ	ல	ட	ள	ம	ன	ங	ற	த	ழ	ச	வ	ந	ப	க	ய	ண	ர
ள	ழ	ய	ன	ப	ச	ந	ள	ர	ட	த	ல	ம	ஞ	க	ற	ண	வ	ங
ற	ட	ன	ங	ற	ய	ழ	க	வ	ர	ள	ஞ	ல	த	ம	ண	ந	ச	ப
ன	வ	ப	ல	ந	ங	ம	ன	த	ச	ய	ற	ர	ண	ட	ழ	க	ள	ஞ

Fig 6. Substitution box for 'mei'

>> UIRMEI

The last classification is called as uirmei eluthu. There are 12 characters in this classification. Those characters are '□□□□□□□□□□□□ □□'. If the plain text is of this type then the the following flow chat will be followed for the encryption process. Key value % 12 is been calculated so that the value ranges from 0 to 11. Since there are only 12 characters available.

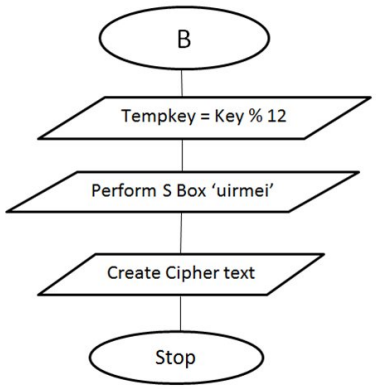

Fig 7. Flowchart for 'uirmei'

These characters are not individual characters i.e. these characters will exist only with combinations of mei eluthu. These characters are not supposed to be printed as such, which has no meaning. Thus the substitution box is been created in such manner. The below diagram shows the substitution box for uirmei eluthu.

	0	1	2	3	4	5	6	7	8	9	10	11
்	ீ	ை	ா	ெள	ொ	ோ	்	ி	ெ	ள	ே	ை
ா	ை	்	ி	ெ	ெள	ை	ீ	ள	ொ	ா	ோ	ே
ி	ி	ள	ை	ே	ை	ெள	ா	்	ோ	ீ	ொ	ெ
ீ	ொ	ெள	ோ	ை	ீ	ா	ெ	ை	்	ே	ள	ி
ா	ை	ே	ெள	ள	ி	ை	ோ	ெ	ா	ொ	ீ	்
ை	ோ	ை	ே	ி	ா	ள	ெள	ொ	ை	ெ	்	ீ
ெ	ே	ொ	ெ	ீ	ள	்	ை	ோ	ி	ெள	ை	ா
ே	ெ	ோ	ொ	ா	்	ீ	ே	ெள	ள	ை	ி	ை
ை	்	ா	ீ	ோ	ெ	ொ	ள	ை	ே	ி	ை	ெள
ொ	ெள	ெ	ை	்	ை	ி	ொ	ே	ீ	ோ	ா	ள
ோ	ா	ி	ள	ை	ோ	ே	ை	ீ	ெள	்	ெ	ொ
ெள	ள	ீ	்	ொ	ே	ெ	ி	ா	ை	ை	ெள	ோ

Fig 8. Substitution bor for 'uirmei'

With the new modulated key and the plain text (character) a new character is been replaced with the help of the above box. I.e. for plain character '□' and the key value 7 the cipher character is '□'.

For all uir, mei , uirmei S-box, box are been created in such a manner than no same character is been repeated in same row or same column. Even if the key value is 0 the cipher text will not be same as plain text.

E. Remaining Words:

There are other few characters available along with the ayudha eluthu "□" those characters are '□□□□'[3]. These characters are not used often but exist somewhere in Tamil language. Is this characters few are individual and few can come with the combination of other classification. Since the occurrence of these characters are less not of same types box cannot be created. If the s box is been created those characters the cipher characters do not exist in Tamil languages.

F. Decryption Process:

For the decryption process, same substitution box is been used. But process is little different form the encryption process. During decryption, key value is taken into account and the respective column is been checked. It will check for the cipher text in the key value column. After finding the cipher text in the column it will be replaced by the first character in that column which is called as the plain text. i.e. for decryption of cipher text '□' for the key value 7 then the plain text will be '□'. The procedure is been followed for all uir,mei and uirmei substitution box.

V CONCLUSION:

Once this Tamilan cipher is been used for the government sector or any other areas where critical data travels over online. Data is been

ISSN: 2278 – 7798

transmitted over internet in a secure manner. Even if the online data is been decrypted it cannot be translated into any language since it is been a Tamil cipher text where only this tool can replace/decrypt the cipher text.

VI REFERENCE:

[1]Introduction to cryptography with java applets, David Bishop. P – 14.
[2] Tamil language in context: A comprehensive approach to learning Tamil, Vasu Renganathan. P – 10.
[3]http://tamilcube.com/learn-tamil/tamil-alphabets -chart.aspx
[4] http://jrgraphix.net/r/Unicode/0B80-0BFF
[5] Enhancing DES Using Local Languages, C.P.Ronald Reagan, S.Selvi, Dr.S.Prasanna Devi, Dr.V.Natarajan.
[6] Holistic Recognition of Handwritten Tamil Words, 2012 Third international conference of emerging application of information technology(EAIT).
[7]Encrypted SMS application on Android with Combination of ceaser cipher and vigenere algorithm.
[8]Learning the ceaser and vigenere cipher by Hierarchical evolutionary Re-combination, 2013 IEEE congress on evolutionary computation

First Author:P.Thamizhikkavi,MTech (ISCF),Department of IT,S.R.M University.
Second Author: Dr.S.Magesh, Asst.Professor (Sr.G), Department of IT, S.R.M University.

ISSN: 2278 – 7798