Swarm Intelligence based Key Generation for Text Encryption in Cellular Networks

Sreelaja.N.K,* G.A.Vijayalakshmi Pai #
* Research Scholar
PSG College of Technology,
Coimbatore, India.
*sreelajank@gmail.com, #paigav@mca.psgtech.ac.in

Abstract— **Encryption of data traffic in cellular network is essential since it is vulnerable to eavesdropping. This paper focuses on encrypting the data sent between the mobile stations and base stations using a stream cipher method. However, the keys for encryption are generated using a swarm intelligence approach. Swarm intelligence is the emergent collective intelligence of groups of simple autonomous agents. The novel technique termed Ant Colony Key Generation Algorithm (AKGA) employs a character code table for encoding the keys. The advantage of this approach is that it reduces the number of keys to be stored and distributed. Experimental results demonstrating AKGA's encrypting text of different lengths and the comparison of its performance with other stream cipher methods are presented.**

Keywords-Stream Cipher, Character Code Table, Key Generation, Swarm Intelligence, Cellular Networks

I. INTRODUCTION

Cellular communication is seeing an explosive growth due to increased usage. However, it is vulnerable to eavesdropping which poses a threat to security and privacy of the user. It is therefore essential that the data traffic across the cellular communication network is encrypted. A Cellular network consists of mobile stations attached to a base station (BS).A cluster of BS's which is fixed, is attached to a mobile telephone switching office which is connected to the public switched telephone network (PSTN). Cryptographic schemes are developed for protecting alphanumeric data since the emerging wired and wireless IP networks are vulnerable to eavesdropping. Thus in the case of the cellular network, the messages sent between the mobile station and the base station is encrypted using a stream cipher method.

Stream cipher is a symmetric key encryption where each bit of data is encrypted with each bit of key. The Crypto key used to encrypt the plain text is randomly changed so that the cipher text produced is mathematically impossible to break. The changing of random keys will not allow any pattern to be repeated which would give a clue to the cracker to break the cipher text. Some of the techniques such as RC4 algorithm and one time pad cipher are examples of stream cipher method.

The RC4 algorithm, a stream cipher method is vulnerable to analytic attacks of the state table. The drawback is that one in every 256 keys can be a weak key. These keys are identified by cryptanalysis that is able to find circumstances under which one of more generated bytes are strongly correlated with a few bytes of the key [5]. Also the same sequence of keys are repeated which would enable the hacker to break the cipher text. Also the first three words of the secret key can be found and by iteration each word of the key used in RC4 can be obtained [4].

The Vernam cipher is a type of one-time pad considered to be a perfect cipher. In this method a large non repeating set of keys are glued together in a pad. The receiver needs a pad identical to the sender to decrypt the message. The encryption involves a long non repeating sequence of random numbers that are combined with the plain text. Each letter in the plain text has a numeric equivalent. The encrypted text is the XOR operation of the characters of the plain text with the corresponding stream of random numbers. The drawback is the need for the unlimited number of keys. The distribution and storage of large number of random keys becomes a problem. Also if the random number sequence is found then the key used for encryption can be traced easily [1].

Wu et al [2] has proposed an encryption system in which a standard stream cipher is combined with a Multiple Huffman Table (MHT) encryption scheme. A standard stream cipher uses a key stream generator to produce a pseudo-random binary sequence having the same length as the plain text. A bit wise XOR operation is performed between the key stream and the plain text. Each segment of the plain text is encoded using MHT encryption and the corresponding segment in the key stream is used as the segment key. The disadvantage is that more number of keys has to be stored and the keys can be found if the keystream generator is cracked.

Swarm Intelligence [8], is an algorithm that models the collective behavior of social insects. Ant system is an evolution from the swarm intelligence forming an evolutionary algorithm to solve optimization problems. Artificial Ants [6] have some characteristics which do not find counterparts with real ants. They live in a discrete world and the moves consist of transitions from discrete state to discrete states. They have an internal state. This private state contains the memory of the ant agent's past action. They deposit a particular amount of pheromone, which is a function of the quality of the solution found. An artificial ant's timing in

The authors express their sincere thanks to the All India Council for Technical Education, New Delhi for supporting this research under the Research Promotion Scheme (F.No 8023/BOR/RPS-104/2006-07) .

pheromone deposition is problem dependent and often does not reflect real ant's behaviour.

This paper proposes a stream cipher method to encrypt data messages sent between the mobile stations and the base station. A Swarm intelligence based technique is used to generate the keys for encryption. A novel approach called Ant colony Key Generator Algorithm (AKGA) is proposed to generate the keystream. The Keystream is a group of characters denoting the keys for encryption. The length of the keystream should be lesser than or equal to the length of the plain text. The keystream generation is done based upon the distribution of characters in the plain text.

Depending on the distribution of characters in the plain text, a tree is formed to generate the code for each character. The decimal equivalent of the code gives the value for the character forming the character code table. Mutation is done at different positions of the tree to form multiple character code tables. The characters in the plain text are encoded using the values from the character code table chosen for encryption. Each key in the keystream is given a value. The keys in the keystream occurring in the plaintext are encoded using values in the character code table. If the length of the keystream is lesser than the length of the plain text then the values of the keys in the keystream are added to a predetermined value to generate the keys for the characters in the plain text which is at a position greater than the length of the keystream. An XOR operation is performed with the keys and the plain text to obtain the encrypted text.

The proposed method uses a swarm intelligence approach of choosing a keystream based on the distribution of the characters in the plain text so that the keys in the keystream is encoded using the character code table. The keys in the keystream are used to generate the keys for the portion of the plain text which exceeds the length of the keystream. This reduces the storage and distribution of keys while comparing with Vernam cipher method. Also it makes it difficult to break the cipher since the keys in the keystream and the plain text is encoded using the character code table. The key generator cannot be cracked in AKGA algorithm where as in Vernam cipher and MHT encryption scheme the keys can be found if the random number generator used to generate the keys is cracked. Also the same sequence of keys will not be repeated as in the case of RC4 algorithm.

The rest of the paper is organized as follows. Section II, describes the Cellular Network System. Section III, describes the encryption technique using AKGA algorithm. Section IV describes the swarm intelligence based key generation in AKGA algorithm. Section V discusses multiple character code table encryption. The experimental results are discussed in Section VI. Section VII shows a performance comparison with other stream cipher methods. The Security analysis is discussed in Section VIII.Conclusion is presented in Section IX.

II. CELLULAR NETWORK SYSTEM

Cellular communication is vulnerable to fraud and eavesdropping since it is carried out over the air interface. Thus it becomes essential to have cell phone users authenticated and voice or data traffic to be encrypted. Encryption is to use a key to scramble a message such that the message cannot be read by anyone but the holder of the correct key[3].

A cellular network includes mobile stations (MS's), with wireless access to the public switched telephone network (PSTN). The Mobile stations in the cellular network are served by a base station (BS). The BS is fixed, and it is connected to the mobile telephone switching office (MTSO), also known as the mobile switching center. A cluster of BS's is attached to an MTSO which is connected to the PSTN. The data and voice messages sent between the BS and the Mobile station is encrypted. Fig. 1 illustrates a typical cellular network.

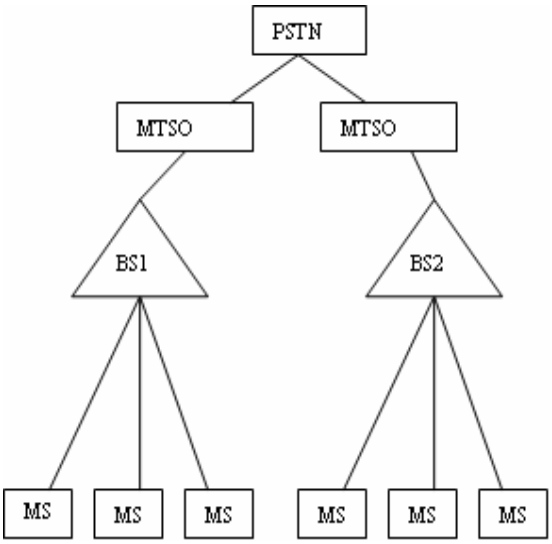

Figure 1. Model of a Cellular Network System

III. ENCRYPTION TECHNIQUE USING AKGA ALGORITHM

Fig. 2 shows a model of an encryption system of a stream cipher. The Plain text is taken and the keystream for encryption is generated by AKGA algorithm. To ensure security the same sequence of keys is never used more than once [7]. The characters in the keystream that are present in the plain text are replaced by the values in the character code table chosen for encrypting the plain text.

The characters in the keystream denote the keys to be used for encryption. If the length of the keystream is lesser than the length of the plain text to be encoded, the value of each character in the keystream is added by some predetermined value and the key for the remaining portion of the text is generated. This reduces the computational load and the key storage.

The characters in the plain text are also encoded using the values in the character code table chosen for encryption. The keys used for encryption occurring in the chosen character code table are also encoded using the values corresponding to it in the table. An XOR operation is done between the keys and the plain text to obtain the cipher text.

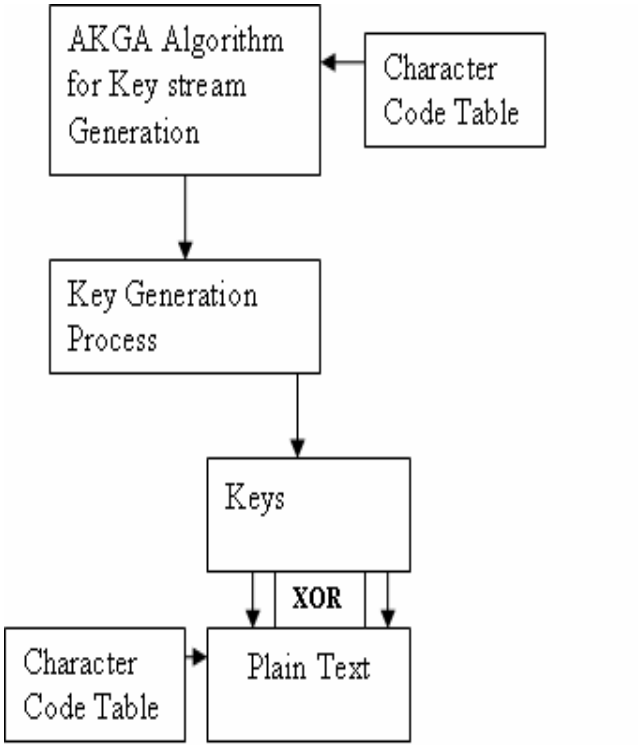

Figure 2. Model of an Encryption System

IV. SWARM INTELLIGENCE BASED KEY GENERATION IN AKGA ALGORITHM

The Ant Colony Key Generator algorithm is a novel approach to generate keystream for text encryption. Each ant agent has a pheromone deposition comprising of characters representing the keystream. The number of characters in the ant agent pheromone deposition occurring in the plain text is counted. The energy level of the ant agent is a count of the characters in the key stream occurring in the plain text divided by the length of the keystream. The ant agent with a maximum energy level greater than the specified threshold value is chosen as the keystream for text encryption.

A. Model of an Ant System

An Ant System model is used to choose the key for text encryption. The ants in the system work together to obtain an optimal solution. The pheromones on the paths traveled by the ants serve as a means of communication between other ants. The initial sets of ants have characters representing the pheromone deposition. The energy of each ant agent is computed as the number of characters in the key stream occurring in the plain text divided by the length of the keystream. The ant agent with an energy level lesser than the specified threshold changes its pheromone deposition until a value greater than the specified threshold value is obtained.

A Tabu-list is used to list the energy of each ant agent depending on the pheromone deposition making the ant agent not to choose the same keystream while updating the pheromone deposition. Once the energy level of the ant agent reaches the specified threshold limit the solution is obtained. The pheromone deposition, tabu-list, and energy monitoring help this novel ant system (AS) to obtain a solution and the pheromone deposition of the ant agent with a maximum energy value in a trial is chosen as the key stream for encryption. Fig. 3 shows a model of an Ant system for AKGA algorithm.

B. Key Stream Representation using Ant Agent

Each Keystream is represented as an ant agent having a pheromone deposition consisting of combinations of characters. A total of 94 characters are taken and a permutation of these characters is done to get groups of characters of all possible orderings without any repetition. These groups of characters form the keystream. Fig. 4 shows the generation process of the keystream which is a combination of characters representing the pheromone deposition of the ant agent. For a set of 4 characters (A, Q, T, S) the possible ways of obtaining the keystream of length 3 and 4 by permuting the characters is shown in Fig. 5. Fig. 6 illustrates the ant agent representing the keystream containing 6 characters denoting the pheromone deposition. Each character in the keystream denotes a key. The total number of possible keystreams generated from a group of 94 characters is given in (1).

$$\sum_{r=1}^{94} 94!/(94-r)! \qquad (1)$$

$$\approx 94!(e)$$

C. Assignment of Energy Value

Energy value is a measure of how well a desired behaviour is performed by an ant agent. Each ant agent has a combination of characters representing the keystream. The energy value of the ant agent is computed by taking the number of characters in the keystream occurring in the plain text divided by the keystream length. The pheromone deposition of the ant agent with a maximum energy value greater than a specified threshold value is the solution and the keystream is chosen for encryption. Let A_i be the ant agent and Energy value is calculated using (2).

$$Energy(Ai) = count(C_j^i \in P)/length(i) \qquad (2)$$

where A_i represents the ant agent, *i =1, 2……No: of Ant Agents*, C_j^i represents the *j* th character of the *i* th key stream, *j=1, 2…Length of the key stream*,*P* represents the plain text.

D. Selection of Pheromone Deposition by the Ant Agent

Each ant agent has a tabu list which maintains the energy value and the keystream representing the pheromone deposition of the earlier trial. The agent updates the pheromone deposition by changing its keystream and the energy value is computed. A comparison of the keystream with the previous trials of the ant agent is made and the antagent do not take the same keystream present in the tabu list.

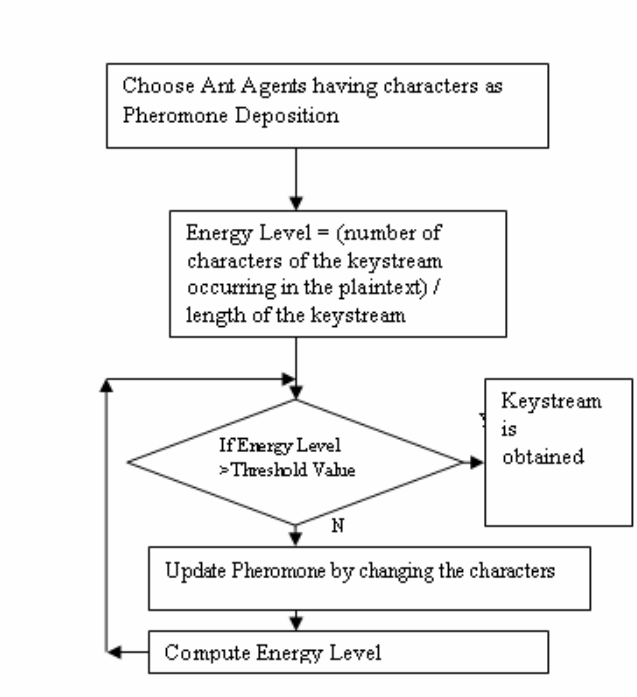

Figure 3. Model of an Ant System for AKGA algorithm

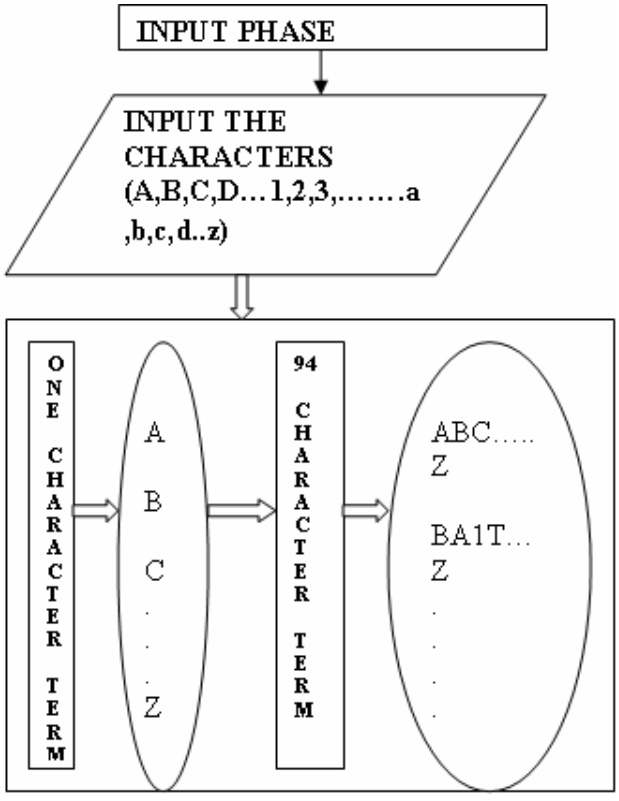

Figure 4. Generation Process of Keystream

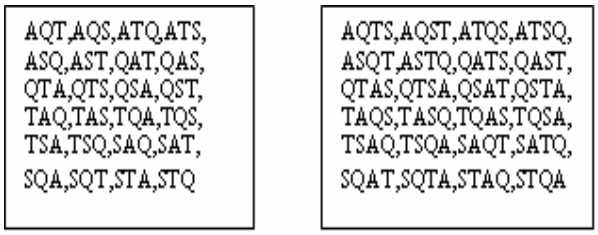

Figure 5. Keystream of length 3 and 4 generated from 4 characters

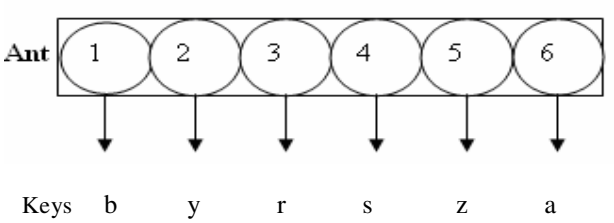

Figure 6. Representation of Keys in a Keystream as an Ant agent

E. Algorithm: Ant Colony Key Generation Algorithm

Fig. 7 illustrates the AKGA algorithm in pseudocode.

```
Procedure AKGA ()
K=0;
Choose Ant agents with characters
denoting the keystream representing
the Pheromone deposition;
Evaluate the energy level of each ant
agent Ai according to the energy
function
        Energy(Ai) = count(C_j^i ∈ P) / length(i)
If (energy (Ai)>threshold value) then
return (ant agent with energy
     value=maximum energy value);
else
repeat
Update the Pheromone deposition by
changing the keystream and compute the
energy of the ant agent;
Select the ant agents with
energy (ant agent in current trial)>
threshold value;
return (ant agent with energy
       value=maximum energy value);
until (energy value in a trial>
      threshold value);
end AKGA;
```

Figure 7. Pseudocode for AKGA Algorithm

V. MULTIPLE CHARACTER CODE TABLE ENCRYPTION

Entropy coding has some properties linked with cryptography. The characters in the plain text are replaced by the values in the predefined character code table by simple lookup operations. Multiple tables are used to increase the model space keeping the structure of the tree representing the characters unchanged. Encryption should be combined with entropy coding using multiple statistical tables in a secret order. The advantage of using multiple tables is that encryption can be done at a reasonable high level of security [2].

A. Code Generation and Value Assignment for the characters in the plain text

The characters in the plain text are counted and their probability of occurrence in the plain text is found. The characters are represented in the form of a tree according to the probability of occurrence. The left hand side branch of the tree is labeled as '0' and the right hand side branch is labeled as '1'.Fig. 8 shows a tree representing the codes for the characters in the plain text 'text'. The character't' occurs twice and 'e' and 'x' occurs once. The code for each character in the plain text is found by traversing the tree. The value of the character is found by taking the decimal equivalent of the code. The code and the value of the character form the character code table. Table I shows the codes and the values generated for the characters in the plain text.

B. Generation of Multiple Tables

The original tree is generated and the label pairs are mutated to get a new tree according to a mutation process. In the tree an inner node is connecting each labeled pair. A tree with t-1 inner nodes and labeled pairs would generate 2^{t-1} trees using mutation process. Fig. 9a, Fig. 9b, Fig. 9c shows the mutation process. Table II shows the different code and values for the characters in the plain text after mutation. The tree is traversed after mutation and the characters are assigned the codes. The values are assigned to the characters by taking the decimal equivalent of the code. If the value of the characters is the same in a table, the code length is added to the value of the corresponding character in order to get unique values for characters present in a table.

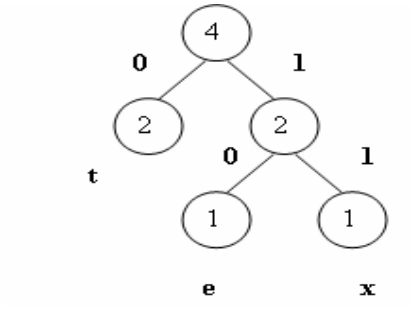

Figure 8. Tree Representing the Character Code

TABLE I. CODES AND VALUES FOR CHARACTERS IN THE PLAIN TEXT

Character	Code	Value
t	0	0
e	10	2
x	11	3

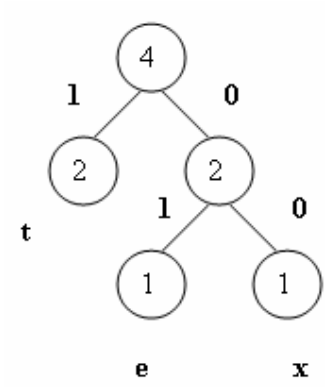

Figure 9(a). Tree after Mutation at Positions1,2 and 3,4

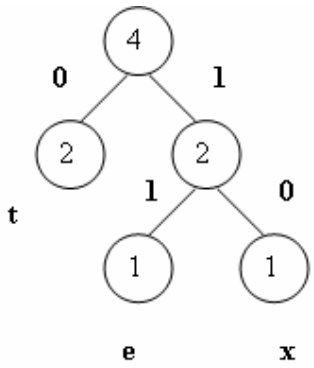

Figure 9(b). Tree after Mutation at Positions 3,4

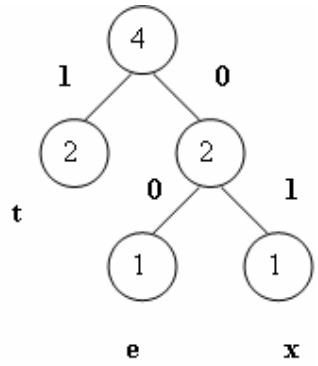

Figure 9(c). Tree after Mutation at Positions 1, 2

TABLE II. VALUES FOR CHARACTERS IN THE PLAIN TEXT AFTER MUTATION

Plain Text	Code	Value	Code	Value	Code	Value
t	1	1	1	1	0	0
e	00	0	01	3	11	3
x	01	3	00	0	10	2

VI. EXPERIMENTAL RESULTS

Consider the process of encrypting a text in which the keystream for encryption is generated using AKGA algorithm. Consider the text to be encrypted is "theskyisblue". The threshold value is assumed to be 0.75. Each ant agent has a pheromone deposition comprising of characters representing the keystream. The energy level of the ant agent is a count of the characters in the key stream occurring in the plain text divided by the length of the keystream. The ant agent with a maximum energy level greater than the specified threshold value is chosen as the keystream for text encryption. Table. III shows the pheromone deposition of ant agents denoting the keystream and their corresponding energy value.

Since the second ant agent in Trial II has the maximum energy value 0.88 which is greater than the threshold value the keystream **kbelsuthz** corresponding to that ant agent is chosen for encryption. Each character in the keystream is chosen as the key for encryption.

A Character code table is generated for the characters in the plain text depending upon their probability of occurrence in the plain text. In the given plain text to be encoded the characters t, h, k, y, i, b, l, u occurs once and s, e occurs twice. A character code tree is generated and the tree is traversed and the code is generated for the characters in the plain text. The values for the characters are found by taking the decimal equivalent of the code. The length of the character code is added to the value of the character for the characters with the same value in order to keep the values unique in a table. To generate multiple character code tables, mutation is done at various positions of the character code tree. In this example a mutation operation is done at first and third position of the tree and a character code table can be generated as follows. Table IV shows the values for the characters in the plain text before and after mutation.

The mutated table is chosen as the character code table to encode the plain text and the keystream. The plain text to be encoded is replaced with the corresponding values in the mutated character code table. Also the characters in the keystream occurring in the plain text are replaced with the values in the mutated character code table. The characters in the keystream not occurring in the plain text are replaced with their ASCII values. Since the keystream is smaller than the length of the plain text to be encoded, the values of the keys of the keystream are added to a predetermined value to generate the keys for the remaining portion of the plain text. The predetermined value can be generated by dividing the length of the plain text by half of its length. Here the value is chosen as 2.Thus the keys for the portion of the plain text exceeding the length of the keystream is generated by adding the values of the keys in the keystream with the value 2. An XOR operation is performed with the plain text and the keys to obtain the encrypted text. Table V shows the encryption process.

The keys used for encryption looks like a series of random numbers but not exactly a set of random numbers. The number of keys to be stored and distributed is 9 which is lesser than the length of the plain text. Also the same pattern of keys will not be repeated. Fig. 10 shows the time taken to encrypt messages of different lengths.

TABLE III. ANT AGENT FINDING THE KEYSTREAM

Trial-I	Energy	Trial-II	Energy
aghiotrsv	0.44	asturvskl	0.66
aegtxazyr	0.22	kbelsuthz	0.88
lkaytusxz	0.66	irsyztuad	0.55
yagktlmsb	0.66	erasyltuk	0.77
Maximum Energy Value	0.66		0.88

TABLE IV. VALUES FOR CHARACTERS AFTER MUTATION

Character	Value Using Original Tree	Value After Mutation At Positions 1,2 And 5,6
s	0	1
e	2	0
t	6	3
h	14	4
k	30	10
y	62	22
i	126	46
b	254	94
l	510	190
u	511	191

TABLE V. ENCRYPTION PROCESS

	Keys	Plain Text	Value Of Plain Text	Cipher Text Values
k	6	t	3	5
b	94	h	4	90
e	0	e	0	0
l	190	s	1	191
s	1	k	6	7
u	191	y	22	169
t	3	i	46	45
h	4	s	1	5
z	122	b	9	36
k	8	l	4	182
b	96	u	190	223
e	2	e	191	2

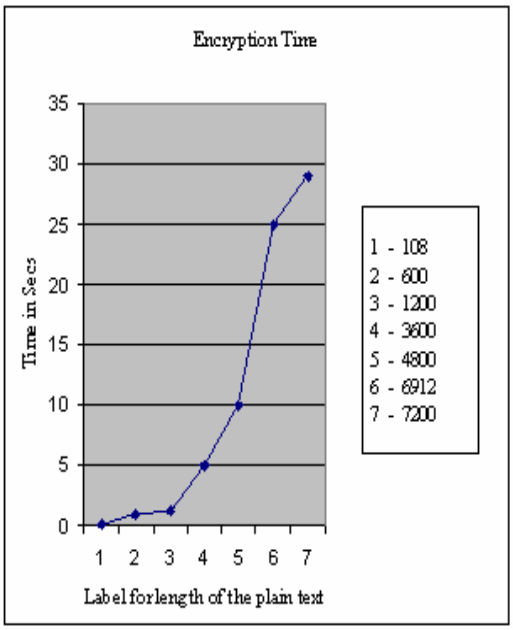

Figure 10. Time Taken for Encryption

VII. COMPARISON WITH OTHER STREAM CIPHER METHODS

The AKGA algorithm method for key generation for encrypting text is compared with other well known stream cipher methods such as RC4 and VernamCipher. Table VI gives a comparison between the various stream cipher methods. Fig. 11 shows a comparison between the number of keys to be stored for AKGA algorithm and Vernam cipher methods.

TABLE VI. COMPARISON BETWEEN RC4, VERNAM CIPHER AND AKGA ALGORITHM

	Key Generation	**Key Storage**	**Cracking Of Keys**
RC4	A state table is initialized with 1 to 256 bytes to produce a Pseudo random stream of keys by swapping the elements in the 256 byte state table. This set of keys is XORed with the plain text to produce the cipher text.	The number of keys to be stored is lesser compared to vernam cipher.	This stream cipher method is vulnerable to analytic attacks of the state table. 1 out of every 256 keys is a weak key. These keys can be identified by cryptanalysis which can find whether the generated bytes are strongly correlated with the bytes of the key.
Vernam Cipher	Keys are randomly generated using Random stream generator	Large number of keys has to be stored depending on the length of the plain text	If the random number generator is cracked the keys can be found.
AKGA Algorithm	Keystream is generated based on the distribution of characters in the plain text	A comparatively smaller number of keys has to be stored since the keys for the remaining portion of the text are generated using the keys in the keystream.	Even if the keys are found the values assigned for the key occurring in the plain text has to be found from the tables. Since there are 2^{t-1} tables the hacker should search all the tables to know which table is used to encode the key. Also the value of the keys not present in the character code table has to be found.

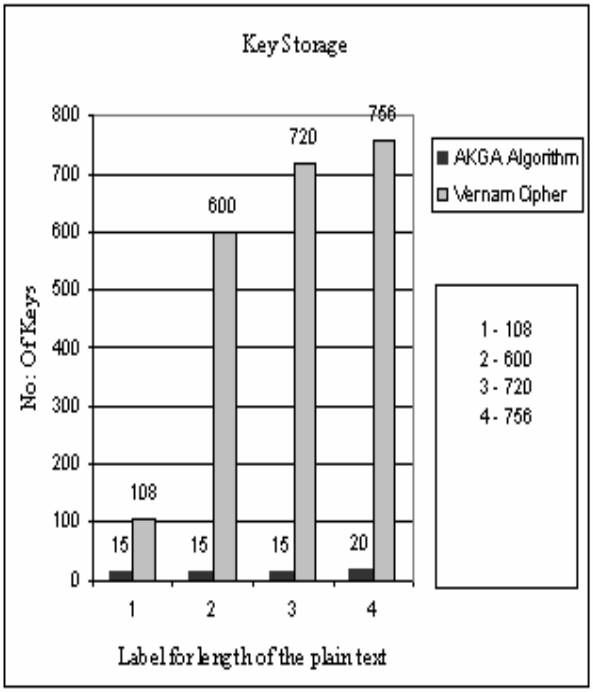

Figure 11. Comparison of Key Storage in AKGA and Vernam Cipher

VIII. SECURITY ANALYSIS

The Key is changed for each character of the plain text to produce a cipher text that is mathematically impossible to break. Since 94 characters are chosen the total number of keystreams will be 94!*2.718.Thus a hacker has to try all such keystreams to find an appropriate one. This method makes it difficult for the hacker to find out the keystream used for encryption.

The characters of the keystream occurring in the plain text are replaced by the values from the character code table. This would increase the security in such a manner that it is difficult to know the values assigned for the characters in the

keystream.This is because there are 2^{t-1} character code tables and the hacker has to search those tables for the values. Even if the plaintext value is found, it is encoded using character code table making it difficult for the hackers to find the value.

IX. CONCLUSION

Encryption is an important issue in wireless communication since it is carried out over the air interface, and is more vulnerable to fraud and eavesdropping. A cellular network integrates cell phones into the public switched telephone network. This paper deals with a stream cipher method used for encrypting the data messages transferred between the mobile station and the fixed base stations in a cellular network. A Swarm Intelligence based approach termed AKGA algorithm is used for generating keystream for the plain text to be encrypted. Also the keystream is used to generate the keys for the portion of the plain text exceeding the length of the keystream. This method of encryption using a stream cipher reduces the number of keys to be stored and distributed. Also the keys and the plain text are encoded using the values in the multiple character code tables. This makes it difficult for the hacker to trace the character code table used for encoding the keys and the plain text. AKGA demonstrates the potential to encrypt messages of different lengths. Also it overcomes the drawbacks of Vernam cipher and RC4 stream cipher method.

ACKNOWLEDGMENT

The first author expresses her sincere thanks to the Management and Principal of Sri Krishna College of Engineering and Technology, Coimbatore for the extended support for the research work.

REFERENCES

[1] Charles Pfleeger, Shari Lawrence Pfleeger, "Security in computing", Third Edition, Prentice Hall of India ,2003.

[2] Chung-Ping Wu, C.C. Jay Kuo, "Design of Integrated Multimedia Compression and Encryption Systems", IEEE Transactions on Multimedia, Volume 7, Issue 5, Oct. 2005 Page(s): 828 – 839.

[3] Jingyuan Zhang, Ivan Stojmenovic, "Cellular Networks", University of Alabama, University of Ottawa, Canada.

[4] Mantin and A. Shamir. , "Weaknesses in the key scheduling algorithm of RC4", Lecture Notes in Computer Science, Vol. 2259, Revised Papers from the 8th Annual International Workshop on Selected Areas in Cryptography Pages: 1 - 24, Year of Publication: 2001.

[5] RC4 Encryption Algorithm, " http://www.vocal.com"

[6] N.P.Padhy,"Artificial Intelligence and Intelligent Systems", Oxford University press, 2005.

[7] B.Schneier, " Applied Cryptography Second Edition: protocols, algorithms and source code in c",Wiley, 1996.

[8] Yang Liu and Kevin M. Passino, "Swarm Intelligence", Literature Overview, Dept. of Electrical Engineering, The Ohio State University, March 30, 2000.