A Data Confidentiality Approach to SMS on Android

Tun Myat Aung(✉), Kaung Htet Myint, and Ni Ni Hla

University of Computer Studies, Yangon, Myanmar
tma.mephi@gmail.com, kolynn.2013@gmail.com,
ni2hla@ucsy.edu.mm

Abstract. Short Message Service (SMS) is a text messaging service component of mobile communication systems. It uses standardized communications protocols to exchange short text between mobile devices. SMS does not have any built-in procedure to offer security for the text transmitted as data. Most of the applications for mobile devices are designed and developed without taking security into consideration. In practical use, SMS messages are not encrypted by default during transmission. Therefore, a data confidentiality approach to SMS on Android will be developed in the paper. It includes design, implementation and confidentiality measurement of RC4 stream cipher for SMS data confidentiality on mobile networks.

Keywords: SMS security · Data confidentiality · Mobile application
Cryptogarphy · RC4 stream cipher

1 Introduction

Data confidentiality is a protection of data from unauthorized disclosure. It is the most common aspect of information security. It not only applies to the storage of information, also applies to the transmission of information. We need to protect our sensitive information from malicious actions during transmission of Short Message Service (SMS). For data confidentiality security service, an encipherment security mechanism can be used.

Short Message Service (SMS) is a mechanism of delivery of short messages over the mobile networks. It is a store and forward way of transmitting messages to and from mobiles. The message (text only) from the sending mobile is stored in a central short message center (SMS) which then forwards it to the destination mobile. Global System for Mobiles (GSM), Code Division Multiple Access (CDMA) and Time-Division Multiple Access (TDMA) are supporting SMS transmission.

The primary purpose of SMS is to deliver text messages from one mobile device to another. It provides many benefits to our everyday life. But, it is now considered as a safe and secure tool when sensitive information is transmitted using the typical SMS services. Nowadays, there are many possible threats on SMS, therefore, it is important not only to prevent the SMS message from being illegally intercepted by illegal sources but also to ensure the origin of the message from the legitimate sender.

P. Vasant et al. (Eds.): ICO 2018, AISC 866, pp. 505–514, 2019.
https://doi.org/10.1007/978-3-030-00979-3_53

Cryptography is the science of information and communication security. Cryptographic system transforms a plaintext into a cipher text using a key generated by a cryptographic algorithm. RC4 is a stream cipher that is used to protect internet traffic as part of the Secure Socket Layer (SSL). RC4 stream cipher is used to protect data confidentiality for SMS transmitted over mobile networks.

The purpose of this paper is to provide data confidentiality during SMS message transmission period in order to prevent the SMS message from being illegally intercepted by illegal sources and to ensure the origin of the message from the legitimate sender. The structure of this paper is as follows. The Sect. 2 includes basic concepts of SMS technology, SMS mobile network communication system, introduction to cryptography and RC4 cipher. In Sect. 3, we discuss design and implementation of mobile applications that are used to protect data confidentiality of SMS message transmitted on mobile networks. The Sect. 4 describes how statistical tests suite is used to measure data confidentiality. Finally, in Sect. 5 we conclude our discussion by describing data confidentiality level of pseudorandom number sequence generated by RC4 cipher and by suggesting RC4 cipher should be used for data confidentiality of SMS message transmitted on mobile network communication system.

2 Background Theory

2.1 Basic Concepts of SMS Technology

SMS messages are created by mobile phones or other devices (e.g. personal computers). These devices can send and receive SMS messages by communicating with the GSM network. All of these devices have at least one MSISDN number. They are called Short Messaging Entities (SMEs). The SMEs are the starting points (the sender) and the end points (the receiver) for SMS messages. They always communicate with a Short Message Service Center (SMSC) and never communicate directly with each other [3]. An SME can be a mobile telephone. An SME can also be a computer equipped with a messaging software, such as Ozeki NG - SMS Gateway, which can communicate directly with the SMSC of the service provider. Depending on the role of the mobile phone in the communication, there are two kinds of SMS messages: Mobile-originated (MO) messages and Mobile-terminated (MT) messages. MO messages are sent by the mobile phone to the SMSC. MT messages are received by the mobile phone. The two messages are encoded differently during transmission. The functions of Short Message Service Center (SMSC) are shown in Fig. 1 [5].

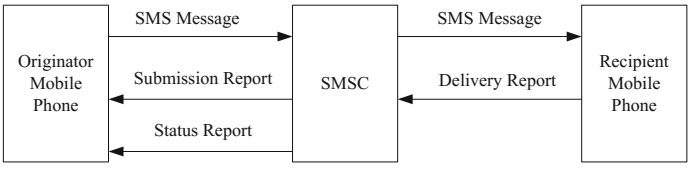

Fig. 1. The functions of Short Message Service Center (SMSC)

2.2 SMS Mobile Network Communication System

SMS messages are transmitted over the Common Channel Signaling System 7 (SS7). SS7 is a global standard that defines the procedures and protocols for exchanging information among network elements of wire line and wireless telephone carriers [2]. These network elements use the SS7 standard to exchange control information for call setup, routing, mobility management, etc. Figure 2 shows the mobile network architecture for SMS communication. Conceptually, the general SMS mobile network architecture consists of two segments that are central to the SMS model of operation: the Mobile Originating (MO) part, which includes the mobile handset of the sender, a base station that provides the radio infrastructure for wireless communications, and the originating Mobile Switching Centre (MSC) that routes and switches all traffic into and out of the cellular system on behalf of the sender. The other segment, the Mobile Terminating (MT) part, includes a base station and the terminating MSC for the receiver, as well as a centralized store-and-forward server known as SMS Centre (SMSC). The SMSC is responsible for accepting and storing messages, retrieving account status, and forwarding messages to the intended recipients [5].

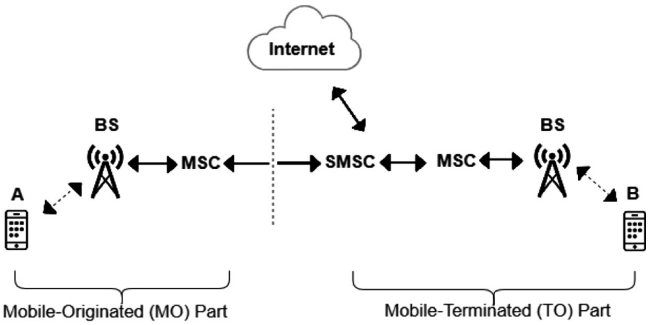

Fig. 2. Mobile network architecture for SMS communication

2.3 Cryptography

Cryptography is the science of using mathematics to encrypt and decrypt data. Cryptography enables you to store sensitive information or transmit it across insecure networks like the Internet so that it cannot be read by anyone except the intended recipient. Encryption is the process of converting ordinary information (called plain text) into unintelligible gibberish (called cipher text). Decryption is the reverse, in other words, moving from the unintelligible cipher text back to plain text.

Cryptographic algorithms can be divided into:

- Symmetric key algorithms.
- Asymmetric key algorithms.

Symmetric key algorithms have the property that same secret keys are used for encryption and decryption. It is also called as private key algorithms. Asymmetric key

algorithms use two different keys: public key for encryption and private key for decryption.

There are two types of symmetric-key algorithm: *block cipher* and *stream cipher*. (1) Stream Cipher - In a stream cipher, encryption and decryption operate on the basis of one symbol (a bit or byte) at a time, (2) Block Cipher - In a block cipher, encryption and decryption operate on the basis of a block of symbols of particular size [1]. The general concept of a stream cipher is shown in Fig. 3.

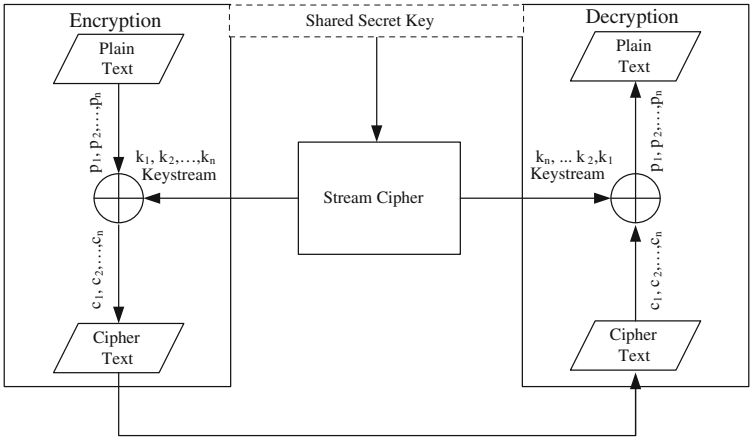

Fig. 3. Stream cipher

2.4 RC4 Cipher

RC4 cipher is one of the most used software-based stream ciphers in the world. Rivest Cipher 4 was designed by Ron Rivest in 1987 and is known as RC4 cipher. The general logic design structure of RC4 cipher is shown in Fig. (4). It is a standard of IEEE 802.11 within WEP, Wireless Encryption Protocol, and generates a keystream. This stream cipher consists of two parts.

- key-scheduling algorithm (KSA).
- Pseudo-random generation algorithm (PRGA).

The key-scheduling algorithm (KSA) is used to initialize the permutation in the array box "S". The length of key is number of bytes in the key and is in the range 1 to 256. First, the array "S" is initialized to the identity permutation. The array box "S" is then processed for 256 iterations in a similar way to the main PRGN, but also mixes in bytes of the key at the same time. The key-scheduling algorithm (KSA) [1] is listed below.

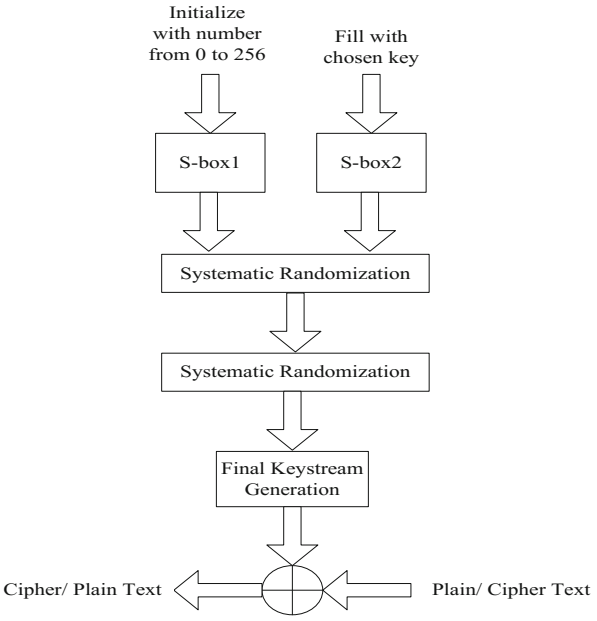

Fig. 4. General logic design structure of RC4 cipher

Key-Scheduling Algorithm (KSA)

```
begin
for i from 0 to 255
    S[i] := i
endfor
j := 0
for i from 0 to 255
    j := (j + S[i] + key[i mod keylength]) mod 256
    swap values of S[i] and S[j]
endfor
end
```

For as many iterations as are required, the pseudo-random generation algorithm (PRGA) modifies the state and generates a byte of the keystream. In every iteration, the PRGA increments i, looks up the *i*th element of the array box "S", S[i], and adds that to j, swaps the values of S[i] and S[j], and then uses the sum S[i] + S[j] (modulo 256) as an index to obtain a third element of array box "S", (the keystream value K below) which is bitwise exclusive ORed (XORed) with the next byte of the plain text to generate the next byte of cipher text. Each element of the array box "S" is exchanged with another element at least once in each of 256 iterations. The Pseudo-Random Generation Algorithm (PRGA) [1] is listed below.

Pseudo-Random Generation Algorithm (PRGA)

```
begin

i := 0

j := 0

while GeneratingOutput:

    i := (i + 1) mod 256

    j := (j + S[i]) mod 256

    swap values of S[i] and S[j]

    K := S[(S[i] + S[j]) mod 256]

    output K

endwhile

end
```

3 Design and Implementation

First, we implement RC4 stream cipher by using *Key-Scheduling Algorithm* (KSA) and *Pseudo-Random Generation Algorithm* (PRGA) in Java programming language. Then we implement two android mobile applications: *SendSMS* and *ReceieveSMS*. *SendSMS* mobile application is used for the sender to transform from SMS plain text to cipher text, and to send it confidentially to the receiver through SMS mobile network communication system while *ReceieveSMS* mobile application is used for the receiver to receive cipher text that is passed through SMS mobile network communication system and to transform from cipher text to SMS plain text.

The design for implementation of two android mobile applications, *SendSMS* and *ReceieveSMS,* is shown in Fig. 5. For *SendSMS* mobile application, at first password is used in RC4 cipher to generate keystream and it is XORed with SMS plain text to

output cipher text. The *Sending* process sends the cipher text to the phone number accepted by this application. Correspondingly, in *ReceieveSMS* mobile application the *Receving* process receives the cipher text from the phone number accepted by this application. Then the cipher text is XORed with the keystream generated by RC4 cipher that uses the same password to output original SMS plain text.

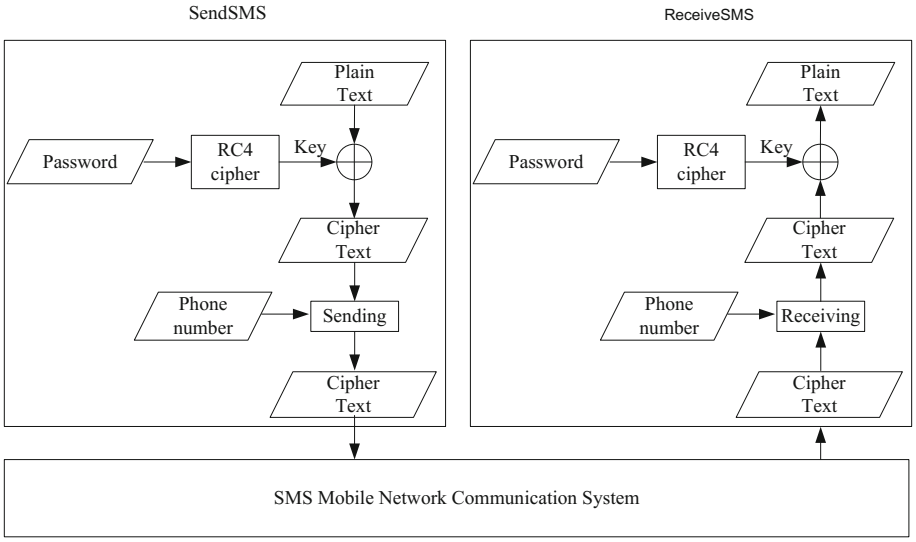

Fig. 5. Design for implementation

The general data flow diagram of these two mobile applications is shown in Fig. 6. *SendSMS* mobile application accepts SMS plain text, password and phone number of the receiver as inputs and outputs cipher text. The cipher text is passed through mobile network communication system. *ReceieveSMS* mobile application accepts cipher text that passed through mobile network communication system, password and phone number of the sender as inputs and outputs SMS plain text.

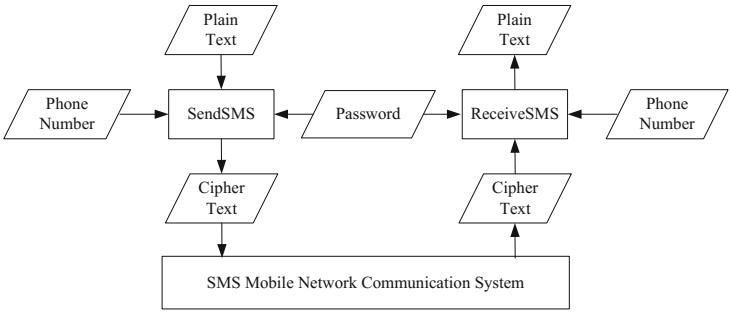

Fig. 6. General data flow diagram

The system user interfaces for *SendSMS* and *ReceieveSMS* mobile applications is shown in Fig. 7. *SendSMS* mobile application is used at the side of the sender and *ReceieveSMS* mobile application is used at the side of the receiver. The sender must input phone number of the receiver, password and SMS message to *SendSMS* mobile application and press *Send Message* button. The receiver must input phone number of the sender and the same password used by the sender to *ReceiveSMS* mobile application and press *Receive Message* button. Then SMS message of the sender is appeared in the display window screen of *ReceiveSMS* mobile application.

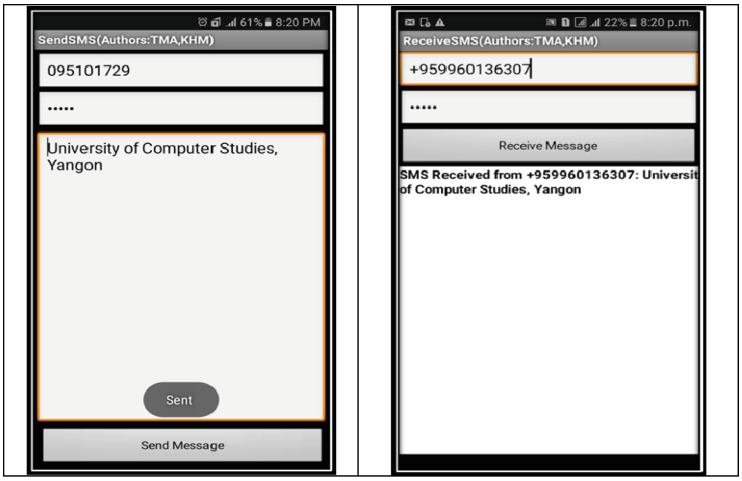

Fig. 7. System user interfaces

4 Confidentiality Measurement

RC 4 stream cipher generates keystream that is pseudorandom number sequence. The pseudorandom number sequence can be used for data confidentiality mechanism during data transmission. The quality of this data confidentiality mechanism depends on the randomness of pseudorandom number sequence generated by RC4 stream cipher. The randomness of pseudorandom number sequence can be measured by using following statistical tests recommend by NIST, National Institute of Standards and Technology. These tests focus on a variety of different types of non-randomness that could exist in a sequence. Some tests are decomposable into a variety of subtests. The 15 tests are:

1. The Frequency (Monobit) Test,
2. Frequency Test within a Block,
3. The Runs Test,
4. Tests for the Longest-Run-of-Ones in a Block,
5. The Binary Matrix Rank Test,
6. The Discrete Fourier Transform Test,
7. The Non-overlapping Template Matching Test,

8. The Overlapping Template Matching Test,
9. Maurer's "Universal Statistical" Test,
10. The Linear Complexity Test,
11. The Serial Test,
12. The Approximate Entropy Test,
13. The Cumulative Sums Test,
14. The Random Excursions Test,
15. The Random Excursions Variant Test.

A statistical test is formulated to test a specific *null hypothesis* (H0). The null hypothesis under test is that the pseudorandom number sequence being tested is *random*. Associated with this null hypothesis is the alternative hypothesis (Ha), that is, the pseudorandom number sequence is *not random*. For each applied test, a decision or conclusion is derived that accepts or rejects the null hypothesis, i.e., whether the pseudorandom number generator is (or is not) producing random values, based on the pseudorandom number sequence that was produced [4].

For each test, a relevant randomness statistic must be chosen and used to determine the acceptance or rejection of the null hypothesis. Under an assumption of randomness, such a statistic has a distribution of possible values. A theoretical reference distribution of this statistic under the null hypothesis is determined by mathematical methods. During a test, a test statistic value is computed on the pseudorandom number sequence being tested. This test statistic value is compared to the critical value. If the test statistic value exceeds the critical value, the null hypothesis for randomness is rejected. Otherwise, the null hypothesis (the randomness hypothesis) is not rejected (i.e., the hypothesis is accepted) [4].

Each test is based on a calculated test statistic value, which is a function of pseudorandom number sequence. If the test statistic value is S and the critical value is t, then the Type I error probability is $P(S > t \,\|\, Ho \text{ is true}) = P(\text{reject } Ho \mid H0 \text{ is true})$, and the Type II error probability is $P(S \leq t \,\|\, H0 \text{ is false}) = P(\text{accept } H0 \mid H0 \text{ is false})$. The test statistic is used to calculate a *P-value* that summarizes the strength of the evidence against the null hypothesis. For these tests, each *P-value* is the probability that a perfect random number generator would have produced a pseudorandom number sequence less random than the pseudorandom number sequence that was tested, given the kind of non-randomness assessed by the test. If a *P-value* for a test is determined to be equal to 1, then the pseudorandom number sequence appears to have perfect randomness. A *P-value* of zero indicates that the pseudorandom number sequence appears to be completely non-random. A significance level (α) can be chosen for the tests. If *P-value* α, then the null hypothesis is accepted; i.e., the pseudorandom number sequence appears to be random. If *P-value* $< \alpha$, then the null hypothesis is rejected; i.e., the pseudorandom number sequence appears to be non-random. The parameter α denotes the probability of the Type I error. Typically, α is chosen in the range [0.001, 0.01] [4].

An α of 0.001 indicates that one would expect one sequence in 1000 pseudorandom number sequences to be rejected by the test if the sequence was random. For a *P-value* ≥ 0.001, a pseudorandom number sequence would be considered to be random with a confidence of 99.9%. For a *P-value* < 0.001, a pseudorandom number sequence would be considered to be non-random with a confidence of 99.9% [4].

An α of 0.01 indicates that one would expect one sequence in 100 pseudorandom number sequences to be rejected. A *P-value* $\geq$ 0.01 would mean that the pseudorandom number sequence would be considered to be random with a confidence of 99%. A *P-value* < 0.01 would mean that the conclusion was that the pseudorandom number sequence is non-random with a confidence of 99% [4].

5 Conclusion

The pseudorandom number sequence generated by RC4 stream cipher is measured by the statistical test suite developed by NIST. According to P-value of each test, the pseudorandom number sequence may be considered to be random with a confidence of 99%. Moreover, RC4 stream cipher possesses better performance among stream ciphers. Therefore, we suggest that the pseudorandom number sequence generated by RC4 stream cipher should be used for data confidentiality of SMS message transmitted on mobile network communication system.

References

1. Forouzan, B.A.: Cryptography and Network Security. International Edition, McGrawHill, ISBN: 978-007-126361-0 (2008)
2. Agoyi, M., Seral, D.: SMS security: an asymmetric encryption approach. In: IEEE 6th International Conference on Wireless and Mobile Communications (2010)
3. Medani1, A.G., Zakaria, O., Zaidan, A.A., Zaidan, B.B.: Review of mobile short message service security issues and techniques towards the solution. Sci. Res. Essays Acad. J. **6**(6), ISSN 1992-2248 (2011)
4. NIST: A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications. National Institute of Standards and Technology Special Publication 800–822 (2010)
5. Katankar, V.K., Thakare, V.M.: Short message service using SMS gateway. Int. J. Comput. Sci. Eng. **02**(04) (2010)