A Lightweight Encoding Mechanism for Encrypted User Notification on Mobile Device in Power Grid System

T. Zhang
State Grid Smart Grid Research Institute
China

Z.X. Sun, Y.C. Jin
Key Lab of Broadband Wireless Communication and Sensor Network Technology
Ministry Education
Nanjing University of Post and Telecommunications
China

Abstract-**As a large amount of users are now using the short message service (SMS) for communication, the demand for security is growing much stronger than ever before. User notification encrypted with traditional encrypt mechanism can be transported via SMS only if the encrypted message is encoded into text form. There are still many defects on existing data encoding mechanism in terms of efficiency. This paper provides a lightweight encoding mechanism for encrypted short message based on UCS-2 Chinese character set which has a lower data extension rate than commonly used encoding mechanism. Analyze for functionality and efficiency proves the new mechanism improves the performance for user notification transmission via SMS on mobile device in power grid system.**

Keywords-user notification; short message service (SMS); text encryption; data encoding; UCS-2 Chinese character set

I. INTRODUCTION

With the development of technology, mobile phone has become one of the most important tools for communication in people's daily lives. Short message service (SMS), with the advantages of fast, convenient, and low cost, is favoured by people. With the increasing applications of SMS, people's awareness of information security becomes stronger and people demand security methods to protect the safety of SMS. SMS encryption technology emerges as the times require solving the problem and becomes a hot spot for researchers.

This paper mainly discusses in the mechanisms to encode the data, which encrypted with encryption algorithm, to the text structure that is supported by SMS in existed mobile phone communication environment. By analyzing existed data coding mechanisms, we will studies its characteristic and coding performance and then put forward a lightweight encoding mechanism for encrypted short message based on UCS-2 Chinese character set. By comparing with other data coding mechanisms, the mentioned novel mechanism has good effectiveness and efficiency for the transmit of encrypted short message which contain user notification.

II. BACKGROUND TECHNOLOGY

A. Encoding Mechanisms for SMS

SMS provides service according to the Global System for Mobile Communications (GSM) standards which established by the European Telecommunications Standards Institute (ETSI). In the standard group of GSM, GSM 03.38[1], GSM 03.40[2] and GSM 07.05[3] define the detail of encoding mechanisms for SMS. Each message has a maximum length of 140 bits which supports 3 kinds of coding mode: 7-bit coding for a single ANSII letter, 8-bit coding for a single ANSII character and 16-bit coding for a single UCS-2 character. On the condition of using 7-bits coding mode, a short message can contain 160 Latin letters while in the country using UCS-2 encoded character, the maximum number of characters for a short message is 70.

B. Cipher Text Encoding Mechanisms

Cipher text encoding mechanisms are applied which mapping and rearranging the bit stream to meet the limited of the transmission channel. Base16, Base32, Base64 and ASCII85 [4, 5, 6, 7] encoding mechanism were designed to covert the bit streams of cipher text to text form data for the transmission of SMS. The main purpose of implementing cipher text encoding mechanisms is to mapping the date to text format and confirming the data can be successfully transmitted by the channel. However, the existing encoding mechanisms often lead to the encoded data becomes a long string which reduce the information capacity of a single short message. Applying text data compression maybe a way to solve this problem, but extra computation and resources requirement makes it becomes another burden for SMS. Furthermore, text compression algorithms on mobile device [8, 9] tend to focus on reducing resource and time consumption in the compression process rather than deal with tiny amount of data. So the goal that makes the encrypted short message to be transmitted smoothly is still unachieved.

III. ENCODING SCHEME BASED ON UCS-2 CHINESE CHARACTER SET

A. Single Character Encoding Scheme

The core idea of the designed encoding scheme is to mapping the encrypted data to a UCS-2 character. As mentioned earlier, the coding region of Chinese character in UCS-2 encoding mechanism is within 0x4E00-0x9FBF, to fulfill the condition that the lower byte can take any values within 0x00-0xFF, the higher byte can take 81 different values (0x4E-0x9E). Considering the 2 bytes cipher text may

distributed in 0x0000-0xFFFF, we design the encoding scheme for a single character as follows:

Step1: select a 14-bit data from the cipher text marked as $C = (c_1, c_2, c_3, \ldots, c_{13})$.

Step2: form the last 8 bits of $C\ (c_7 - c_{13})$ to a byte marked as B_1.

Step3: insert two 0-bit before the first 6 bits of $C\ (c_0 - c_6)$ and form a byte marked as B_0.

Step4: plus B_0 with 0x4E to guarantee its value is within 0x4E-0x9E.

Step5: take the new B_0 as the higher byte and B1 as the lower byte to form a 2-byte data and convert the 2-byte data into a UCS-2 Chinese character.

After all the steps, 14-bit original data is mapping to a UCS-2 Chinese character, the whole process is shown in fig. 1.

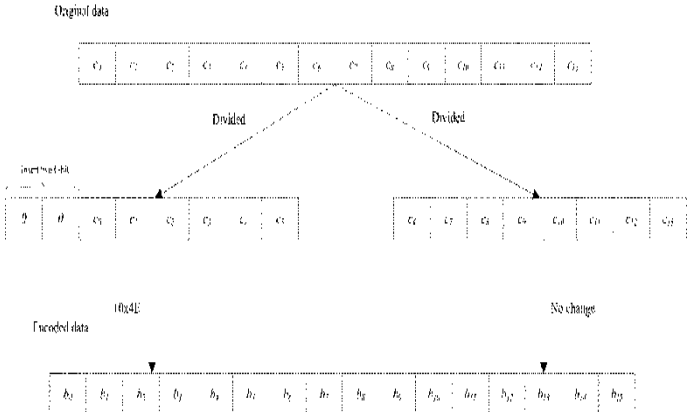

FIGURE I. BITS MAPPING SCHEME FOR SINGLE CHARACTER.

B. Encoding Process

On the basis of the single character encoding scheme we proposed above, we design the encoding process for encrypted data as follows:

Step 1: divide the original data into blocks with 14-bit data and then form an encoding wheel which contains 8 continuous blocks.

Step 2: for the block less than 14 bits (usually the last block of the last encoding wheel), insert 0-bits at the end of the block to make it contains 14 bits data.

Step 3: for all the blocks in an encoding wheel, insert two 0-bit before the block to expand the block to 16 bits.

Step 4: plus each block with 0x4E to make the value of the block within 0x4E00 to 0x8DFF (details of step 3 and 4 is shown in fig. 2).

Step 5: connect the entire encoding wheel sequentially and convert all the blocks to a UCS-2 Chinese character.

After all the steps, here comes a string as the result of the encoding process.

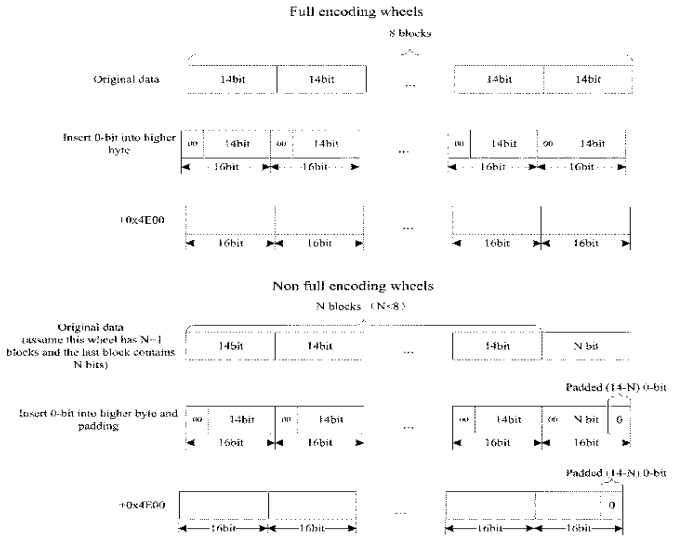

FIGURE II. ENCODING PROCESS FOR ENCODED WHEELS.

C. Decoding Process

Designed decoding process contains several steps as follow:

Step 1: convert the received string to bit format according to UCS-2 encoding rules.

Step 2: divide the bit stream into blocks with 16-bit data and then form decoding wheels with 8 continuous blocks.

Step 3: for the decoding wheel that has 8 blocks, subtract 0x4E00 from each block and remove the first two bits to reduce the length of the block to14.

Step 4: if the last decoding wheel has less than 8 blocks (assume that wheel has n blocks, $n < 8$), on the condition the length of original data is multiple of 32, we can calculate the length of original data in this decoding wheel is $16 \times (n-1)$ bits. Under this premise, we process the blocks in the last decoding wheel in the same way as we mentioned in step 2 and 3. According to the calculated length of the original data, the padded 0-bit can be removed and the cipher text would be restored. The essential steps of decoding process is shown in fig. 3.

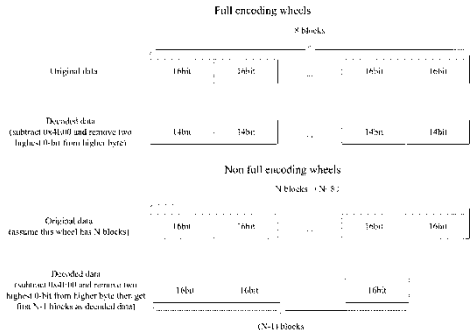

FIGURE III. DECODING PROCESS FOR ENCODED WHEELS.

After the decoding process, the cipher text can be decrypted by the corresponding encryption algorithm and this process is beyond the encoding domain, so we don't talk too detail about that.

IV. ENCODING EFFICIENCY ANALYSIS

In order to analyze the efficiency of the proposed encoding mechanism, we will check the ratio of volume between the original data and processed original data. This ratio is marked as E_p and defined in eqn(1):

$$E_p = \frac{DV_{ori}}{DV_p} \times 100\% \qquad (1)$$

Where DV_{ori} represents the volume of original data, DV_p represents the volume of processed data and encoding efficiency is better where E_p is closer to 100%.

Considering that a single short message can only contain 140 bytes data, the maximum cipher text capacity becomes another factor affecting the efficiency. So we compare the mentioned encoding mechanism with Base16 and Base64 encoding mechanism and details is shown in table 1.

We assume the length of cipher text that generated by encryption algorithm is multiple of 128 (such as AES-128 encryption algorithm). From Table 1 we can figure out that the maximum cipher text capacity of proposed encoding mechanism is 896 bits which is better than Base 16 (512 bits) and Base 64 (768 bits). according to the data above, the proposed mechanism is the best (more than 80%).

V. CONCLUSIONS

Comparing with improving the security of SMS from the basic structure, applying encryption algorithm to encrypt data of message is a more convenient and easier way. However, the data redundancy caused by encrypting process reduce the information capacity of SMS. On the other hand, additional steps must be implemented to guarantee the successful transfer of encrypted data. This article proposal an encoding mechanism base on UCS-2 encoding mechanism to map the cipher text to the UCS-2 Chinese character. The major advantage of this mechanism is that it has a low data expansion ratio with simple encoding and decoding process. Hence, it can be used to enhance the security of user notification in power grid system via SMS cooperate with existed encryption algorithm and improves the performance of communication between client and server.

REFERENCES

[1] GSM 03.38 (Version 5.3.0). Digital cellular telecommunications system (Phase 2+); Alphabets and language-specific information, ESTI, http://www.etsi.org, 1996.

[2] GSM 03.40 (Version 5.3.0). Digital cellular telecommunications system (Phase 2+); Technical realization of Short Message Service (SMS) Point-to-Point (PP), ESTI, http://www.etsi.org, 1996.

[3] GSM 07.05 (Version 5.5.0). Digital cellular telecommunications system (Phase 2+); Use of Data Terminal Equipment - Data Circuit terminating Equipment (DTE - DCE) interface for Short Message Service (SMS) and Cell Broadcast Service (CBS), ESTI, http://www.etsi.org, 1998.

[4] JOSEFSSON, S., RFC3548: The Base16, Base32, and Base64 Data Encodings. IETF, http://www.ietf.org, 2003.

[5] Wu, P. C., Using plain base32 ASCII-compatible encoding in the local part of E-mail addresses. 2002 Symposium on Applications and the Internet (SAINT 2002), IEEE: Nara, Japan, pp. 214-219, 2002.

[6] LINN, J., RFC1421: Privacy Enhancement for Internet Electronic Mail: Part I: Message Encryption and Authentication Procedures, IETF, http://www.ietf.org, 1993.

[7] ELZ, R., RFC1924:A Compact Representation of IPv6 Addresses, http://www.ietf.org, 1996.

[8] ISLAM, M. R. & AHSAN RAJON, S. A., An enhanced scheme for lossless compression of short text for resource constrained devices, 2011 14th International Conference on Computer and Information Technology (ICCIT), IEEE: Dhaka, Bangladesh, pp. 292-297, 2011.

[9] REIN, S., GUHMANN, C. & FITZEK, F. H. P., Low-complexity compression of short messages, 2006 IEEE Data Compression Conference(DCC 2006),IEEE: Snowbird, USA, pp. 123-132, 2006.

TABLE I. COMPARISON BETWEEN THREE ENCODING MECHANISMS.

Length of cipher text[bit]	Length of encoded data [bit]			Encoding efficiency [%]		
	proposed UCS-2 encoding	Base16	Base64	proposed UCS-2 encoding	Base16	Base64
128	160	256	192	80	50	66.67
256	304	512	352	84.21	50	72.73
384	448	768	520	85.71	50	73.85
512	592	1024	712	86.49	50	71.91
640	736	-	872	86.96	-	73.39
768	880	-	1040	87.27	-	73.85
896	1024	-	-	87,50	-	-