Base64 Character Encoding and Decoding Modeling

Article · December 2016

CITATIONS
12

READS
1,536

3 authors:

Isnar Sumartono
Universitas Pembangunan Panca Budi
4 PUBLICATIONS **34** CITATIONS
SEE PROFILE

Andysah Putera Utama Siahaan
Universitas Pembangunan Panca Budi
326 PUBLICATIONS **1,394** CITATIONS
SEE PROFILE

Arpan Arpan
Universitas Pembangunan Panca Budi
1 PUBLICATION **12** CITATIONS
SEE PROFILE

International Journal of Recent Trends in Engineering & Research
ISSN(ONLINE) : 2455-1457
IMPACT FACTOR : 3.344

Base64 Character Encoding and Decoding Modeling

Isnar Sumartono[1], Andysah Putera Utama Siahaan[2], Arpan[3]
Faculty of Computer Science,Universitas Pembangunan Panca Budi
Jl. Jend. Gatot Subroto Km. 4,5 Sei Sikambing, 20122, Medan, Sumatera Utara, Indonesia

Abstract: Security is crucial to maintaining the confidentiality of the information. Secure information is the information should not be known to the unreliable person, especially information concerning the state and the government. This information is often transmitted using a public network. If the data is not secured in advance, would be easily intercepted and the contents of the information known by the people who stole it. The method used to secure data is to use a cryptographic system by changing plaintext into ciphertext. Base64 algorithm is one of the encryption processes that is ideal for use in data transmission. Ciphertext obtained is the arrangement of the characters that have been tabulated. These tables have been designed to facilitate the delivery of data during transmission. By applying this algorithm, errors would be avoided, and security would also be ensured.
Keywords: Base64, Security, Cryptography, Encoding

I. INTRODUCTION

Security and confidentiality is one important aspect of an information system [9][10]. The information sent is expected to be well received only by those who have the right. Information will be useless if at the time of transmission intercepted or hijacked by an unauthorized person [7]. The public network is one that is prone to be intercepted or hijacked [1][2]. From time to time the data transmission technology has developed so rapidly. Security is necessary for an organization or company as to maintain the integrity of the data and information on the company.

Base64 is not an encryption method, but it is the standard encoding [8]. Base64 history begins with electronic mail. At that time, the email is sent by SMTP (simple mail transfer protocol) to a mail server, then sent to the mailbox at the mail server destination. A Protocol is an ordinance of the computer to communicate with each other via a network [3][4][5]. Sometimes Base64 is used to encrypt the plaintext, but it does not have the key. Everyone can decrypt the message by knowing the table pattern. This study describes the process of encryption and decryption using Base64. Although it cannot be the standalone algorithm, it might be combined with another method to make the security level increases.

II. THEORIES

A. Description

Base64 is an algorithm that uses a concept of modern encryption algorithms [6]. It is a block cipher algorithm that operates on a bit, but the Base64 mode is easier in its implementation than others. Base64 is a general term for some similar encoding scheme that encodes binary data and translates it into a representation of the base 64. The term comes from the Base64 MIME encoding specific content. The base64 encoding scheme is typically used when there is a need to encode binary data that needs to be stored and transferred through media designed to deal with textual data. This is to ensure that the data remains intact without modification during shipping. Base64 is used commonly in multiple applications including email through MIME and storage of complex data in XML. Base64 needs to be learned because the transformation base64 widely used on the Internet as a medium to transmit data format data. Due to the result of the transformation base64 be plain text, then this value will be much more easily shipped, compared to the form of binary data format.

B. Usage

Base64 is used in many scopes. For example, some of the use of transformation in the field such as PEM base64, MIME, UTF-7, and OpenPGP [6].

Privacy-Enhanced Mail (PEM)

PEM protocol is the first protocol that uses base64 transformation technique, which is based on RFC 989 in 1987, which consists of 7-bit characters used by the mail server (SMTP) to transfer data. The current version of PEM used was based on RFC 1421, which uses the characters (A..Z, a..z) and numbers (0..9) and two symbols, namely the "+" and "/" /.

Multipurpose Internet Mail Extension (MIME)

Multipurpose Internet Mail Extensions, using the "base64" as one of the two encoding schemes binary-to-text (the other is "quoted-printable Base64 encoding"). MIME is based on RFC 1421 version of PEM. Using a 64-character alphabet are the same and encoding mechanism as PEM, and use the "=" symbol for output padding in the same manner, as described in RFC 1521. MIME did not specify a fixed length for Base64-encoded channels but did not specify a maximum length of 76 characters. It also stipulates that any alphabetical characters must be ignored by the decoder compatible, although most implementations use a pair CR / LF to limit rows encoded. Thus, the actual length of MIME-compliant Base64-encoded binary data is usually around 137% of the length of the original data, albeit for a very short message overhead can be much higher due to the overhead of the header. The final size Base64-encoded binary data equal to 1.37 times the size of the original data + 814 bytes (for the header). In other words, you can estimate the size of the data decoded by this formula:

$$byte = (string\_length\,(encoded\_string) - 814) / 1{,}37$$

UTF-7

Encoding UTF-7 is based on RFC 2152, which is generally referred to as "Modified Base64" / UTF-7 uses a character MIME base64, but do not wear padding "=" / Character "=" is used as an escape character for encoding "quoted printable" / UTF-7 is also used as a MIME header.

OpenPGP

OpenPGP (PGP - Pretty Good Privacy) in RFC 2440, which uses a 64-radix encoding, or sometimes referred to as "! SCII! Mor" / encoding technique based on the technique of MIME encoding, but coupled with a 24-bit CRC checksum. The checksum value is calculated from the input data, before the encoding process.

Base64 use the characters A - Z, a - z and 0-9 for 62 the first value, while the second last value is used for symbol (+ and /). Several other encoding methods such as uuencode and bin hex using 64 different characters to represent six binary digits but is not mentioned as methods of Base64 encoding.

III. METHODOLOGY

The transformation of Base64 is one of the algorithms for encoding and decoding data into ASCI format, which is based on the number 64. The characters generated from Base64 consist of "A..Z", "a..z" and "0..9", and the last two characters are "/" and "+".

The Base64 encoding technique is simple. There are severla steps to be done to complete the Base64 algorithm is:

- Look for the ASCII code of each text.
- Search binary number 8 bits of the ASCII code exist.

- Combine the last 8 bits to 24 bits.
- Then, broke a 24 bit earlier to 6 bits. It will produce four fractions.
- Each fragment is converted into a decimal value.
- Lastly, make value - the decimal value to an index to choose a character constituent of base64 and the maximum is 63 or 64 to the index.

Value	Char	Value	Char	Value	Char	Value	Char
0	A	16	Q	32	g	48	w
1	B	17	R	33	h	49	x
2	C	18	S	34	i	50	y
3	D	19	T	35	j	51	z
4	E	20	U	36	k	52	0
5	F	21	V	37	l	53	1
6	G	22	W	38	m	54	2
7	H	23	X	39	n	55	3
8	I	24	Y	40	o	56	4
9	J	25	Z	41	p	57	5
10	K	26	a	42	q	58	6
11	L	27	b	43	r	59	7
12	M	28	c	44	s	60	8
13	N	29	d	45	t	61	9
14	O	30	e	46	u	62	+
15	P	31	f	47	v	63	/

Fig. 1 Base64 characters table

Figure 1 shows the characters table. The ASCII code is mapped into the new 64-characters table.

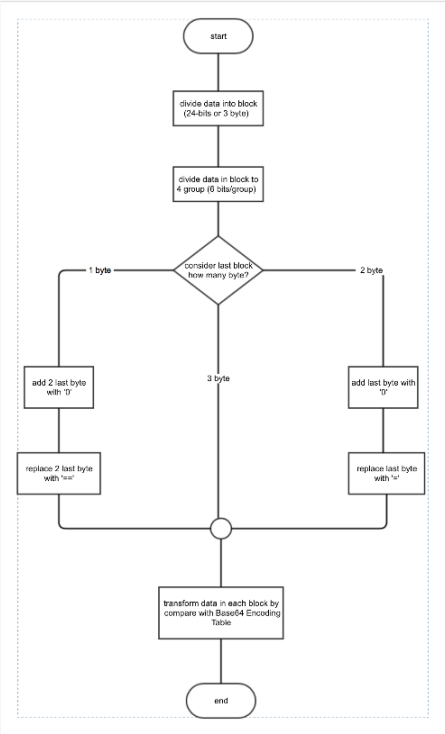

Fig. 2 Encoding flowchart

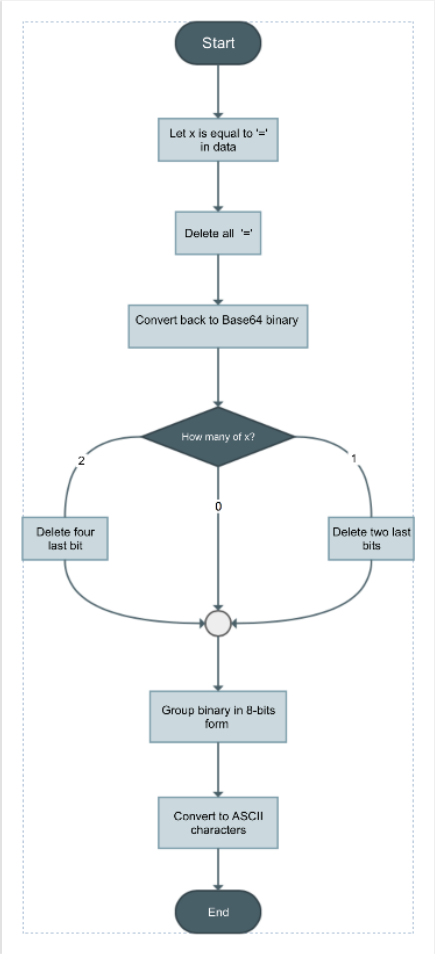

Fig. 3 Decoding flowchart

Figure 2 shows the ecoding flowchart and Figure 3 shows the decoding flowchart.

IV. EVALUATION

This section will analyze the Base64 algorithm. The plaintext will be encoded into ciphertext and then decoded into plaintext back. Assume the plaintext is "Putera". It has six characters. The illustration in Table 1 shows the explanation of Base64 example.

TABLE 1 BASE64 EXAMPLE

Index	1	2	3	4	5	6
Char	P	u	t	e	r	a
Decimal	80	117	116	101	114	97

The calculation:

Index 1 : P
ASCII : 80
Binary : 01010000
Index 2 : u
ASCII : 117
Binary : 01110101

Index 3 : t
ASCII : 116
Binary : 01110100

Index 4 : e
ASCII : 101
Binary : 01100101

Index 5 : r
ASCII : 114
Binary : 01110010

Index 6 : a
ASCII : 97
Binary : 01100001

The concatenation of the binaries is
010100000111010101110100011001010111001001100001

TABLE 2 BINARY 6-BIT

Index	**Binary 6-bit**	**Char**
1	010100	20
2	000111	7
3	010101	21
4	110100	52
5	011001	25
6	010111	23
7	001001	9
8	100001	33

Table 2 show the final binary bit of the previous example. The amout of the plaintext is six characters. The total bit is 6 x 8 bits = 48 bits. The 48-bits is divided into eight parts of 6-bits characters. The will be eight characters in Base64 format. The ciphertext is **UHV0ZXJh** based on Base64 character table in Figure 1.

V. CONCLUSION

The research was successfully performed. The calculation above concludes that the Base64 algorithm is good at information security system for encryption and decryption. Base64 encryption function using the method goes well from testing conducted subsystem. Base64 is easy to apply. It transforms the 8-bits into 6-bits character. It is not used for encryption actually, but sometimes it looks like encryption. The Base64 algorithm transform the character to limited character to make the transmission well-transmitted. No password or key in Base64 algorithm. The algorithm must be improved to make the security level increase.

REFERENCES

[1] H. Anton dan C. Rorres, Elementary Linear Algebra, 2011: John Wiley & Sons.

[2] R. Bhanot dan R. Hans, "A Review and Comparative Analysis of Various Encryption Algorithms," International Journal of Security and Its Applications, vol. 9, no. 4, pp. 289-306, 2015.

[3] S. K. Das, G. Sharma dan P. K. Kevat, "Integrity and Authentication using Elliptic Curve cryptography," Imperial Journal of Interdiscliplinary Research, vol. 2, no. 5, 2016.

[4] D. Shah, "Digital Security Using Cryptographic Message Digest Algorithm," International Journal of Advance Research in Computer Science and Management Studies, vol. 3, no. 10, pp. 215-219, 2015.

[5] K. D. Lewis dan J. E. Lewis, "Web Single Sign-On Authentication using SAML," International Journal of Computer Science Issues, vol. 2, no. 1, pp. 41-48, 22 3 2009.

[6] P. Guwalani, M. Kala, R. Chandrashekar, J. Shinde dan D. Mane, "Image File Security using Base-64 Algorithm," Pooja Guwalani et al, Int.J.Computer Technology & Applications, vol. 5, no. 6, pp. 1892-1895, 2014.

[7] K. Solanki, V. Vankani, P. Pukle dan S. Iyer, "Multimedia Encryption Using Visual Cryptography," International Journal of Recent Trends in Engineering & Research, vol. 2, no. 9, pp. 261-264, 2016.

[8] L. Cantara, METS: The Metadata Encoding and Transmission Standard, vol. 4, Cleveland: The Haworth Press, 2005.

[9] S. M. Poonkuzhali dan M. Therasa, "Data Hiding Using Visual Cryptography for Secure Transmission," International Journal of Advanced Research in Computer and Communication Engineering, vol. 4, no. 4, pp. 440-441, 2015.

[10] A. P. U. Siahaan, "A Three-Layer Visual Hash Function Using Adler-32," International Journal of Computer Science and Software Engineering, vol. 5, no. 7, pp. 142-147, 2016.

