Encoding the Record Database of Computer Based Test Exam Based on Spritz Algorithm

Taronisokhi Zebua
AMIK STIEKOM North Sumatera, Indonesia
taronizeb@gmail.com

Abstract

Computer utilization in the execution of the computer-based test is currently no strange. Almost all government agencies and companies at the time of conducting the test acceptance of new employees have been using computer-based test system online or often referred to as Computer Based Test (CBT). One of the important aspects to be considered and must be maintained in the execution of computer-based exams is the problem of question security and exam answers to abuse actions. One technique that can be done to solve the problems above is the use of cryptographic techniques. This research describes the use of spritz algorithm which is one of the cryptographic algorithms to encode the text database record of the computer-based test. The results of the encoding process can make it harder for the attackers to know the original text of the exam, so as to minimize the abuse of the exam.

Keywords: cryptography, CBT, database, exam, spritz.

1. Introduction

Implementation of computer-based exams online nowadays has been done by various agencies both government and private companies. This is done because the implementation process is more effective when compared with the Paper-Based Test (PBT) test or better known as a conventional test. Testing Computer-based or often known as Computer Based Test (CBT) is no longer excluded exams to come to the location of the exam or exam executive agencies, slowly can appear online anywhere or place that has been determined by the executive committee. One aspect that must be considered in making the test. Based on research conducted by Rejito and Setiana, who said things that must be considered in the implementation of CBT is the confidentiality of the exam database because it should not be published either from the side of the test participants (client) or database manager (administrator) [1].

Utilization of cryptographic algorithm techniques is one alternative solution to solve the above problems. Another study conducted by Setyaningsih say that the application of cryptographic techniques is one way that can be done to secure the data by encrypting it [2]. Text exam records that have been stored in the database can be encoded by the procedure or the cryptographic algorithm. It can have a significant impact on the business to minimize the abuses of the exam from a party other than the legitimate parties. Counterparts, aiming to avoid the misuse of information by other parties who are not eligible [3].

Spritz algorithm is a variant of the RC4 algorithm produces sponge-base construction in generating a key in the encryption and decryption process. This algorithm works based on the concept of a stream cipher that is encryption one by one. One of the advantages of this algorithm is the process of generating the keys used in the process of encryption and decryption. The next generated key always depends on the flow of the previous key [4]. The high complexity of the performance of the spritz algorithm led to the complexity of the cryptanalysts to find the key and solve this algorithm.

This research describes how the security text records a computer-based exam conducted online. The safeguards do is minimize the abuses of the exam to encode the original text of exam questions are stored as database records based algorithm spritz. Records that have been
encrypted is what will be accessed by the examinee (client) when accessing the exam. Exam encrypted will be decrypted automatically by the application test so that the original text of proficiency level exam questions can be understood by the examinee.

2. Research Methodology
The methodology that used in conducting this research is:
   a. Literature Review
      Search and study relevant literature or references to topics covered either through books or electronic journals.
   b. Analysis
      Analyzing the security problems in the implementation of the computer-based test, especially security exam on either the server or client computers. This is done to determine the solutions provided to solve problems that have been identified.
   c. Implementation
      Using spritz algorithm to encrypt the online test records computer-based database.

3. Literature Review
3.1. Computer Based Test
The Computer-based test known as Computer Based Test (CBT) has been done since 1960 [5]. Until now, government agencies and companies are using CBT as a model of implementation of the various examination techniques or hiring new employees, because in addition to effectively and efficiently can reduce operating costs required in the implementation of the test. Computer-based exams involve client-server systems. Computer servers to act as a provider of the exam as well as a central controller for the client implementation of the test. Implementation of computer-based test can give participants more accurate test results because everything is done by the system. In addition to this, the level of fraud participants in working on the test can be minimized.

3.2. Cryptography
Cryptography is a term of one of the commonly used data security techniques. This technique works by encoding data to be secured so it is not easy to fall into the hands of others who are not the real recipient [6]. Along with its development, the term cryptography is defined as a science that studies mathematical techniques relevant to data security aspects including confidentiality, integrity, authentication, and non-repudiation[9]. Cryptographic techniques have several algorithms such as GOST, TDES, RC4, Spritz, Triangle Chain Cipher and others. The application of cryptographic algorithms must achieve the principle of confusion (confusion/confusion) and diffusion (diffusion/melting)[7][10]. The basic functions of cryptographic algorithms are encryption, decryption, and keys. The elements of the cryptographic system are the original file/data (plaintext), the encrypted file (ciphertext), the encryption process, the process of converting ciphertext to plaintext (decryption) and key [6].

Spritz algorithm is an update of the RC4 algorithm performed by Ron Rivest and Jacob Schultz in 2014. Spritz as a variant of encryption RC4 cows including messages or data one by one using relatively short time-dependent transformation encryption [4][8]. The addition of a relatively prime element to the N value of the pseudo-random generation algorithm is the difference with the RC4 algorithm. In addition to stream ciphers, the spritz algorithm can also be used as a hash function and the Message Authentication Code (MAC) by using the sponge function in securing data. The main procedure of the spritz algorithm as a stream cipher consists of three processes: Key Scheduling Algorithm (KSA), Pseudo-Random Generation Algorithm (PRGA) and encryption or decryption process.
   a. Key Scheduling Algorithm (KSA)
      The key scheduling process is a process that is done to make the S-Box table (array S) and table permutation in the array S. The length of the array that is required is 256 which starts from an index of 0 to 255. The purpose of KSA is the process of permutation array values as much as 256 times which is initialized with variables i and j with integer types.
Pseudo-code of KSA is:
for $i = 0$ to $N - 1$
    $S[i] = i$
next i
$i, j = 0$
for $i = 0$ to $N - 1$
    $j = (j + S[i] + K[i \mod \text{Key.length}]) \mod N$
    swap $(S[i], S[j])$
    $j = j$
next i
where $N$ is the size of the array to be mutated, i.e. $0 – 255$.

b. Pseudo-Random Generation Algorithm (PRGA)
The pseudo-random generator algorithm process is performed to derive a new key number of plain elements. The value of $w$ is a new variable added to the spritz algorithm that corresponds to the RC4 algorithm. The value of the variable $i, j, k$ and $z$ starting at 0 and will change according to results at each iteration. This process involves an array of $S$ values that have been permuted in the KSA process.
Pseudo-code PRGA is:
for $i = 0$ to plain.length
    $i = (i + w) \mod N$
    $j = (k + S[j + S[i]]) \mod N$
    $k = (i + k + S[j]) \mod N$
    swap $S[i], S[j]$
    $z = (S[j + S[i + S[z + k]]]) \mod N$
output $z$
next i
where $w$ is a relatively prime value of integer with $N$ and the value of $i, j, k, z$ starts from 0.

c. Encryption and Decryption
The encryption and decryption process is done by XOR-binary each output $z$ with each plain element in a stream.
Formulation of the encryption process:
$$C_i = P_i \oplus z_i$$ (1)
Formulation of the decryption process:
$$P_i = C_i \oplus z_i$$ (2)
Description of the formula above:
$P_i$ = plain element
$C_i$ = cipher element
$z_i$ = key element (the result of PRGA process)

3.3. Database
A database can simply be defined as a system that serves to store and process data into useful information. One of the data that should be maintained and maintained by the owner of the information system is the database. Information on a system can be updated by using the database management process [7]. A database is filled with one or more tables, and each table is filled with some record. These records which shall be processed and processed into information for the users of the system. MySQL is one of the applications that can be used to create and manage databases. Through commands (query) owned by MySQL, the management of the database to generate information do.

4. Results and Discussion
Based on the description of the background above, the problem being analyzed is the issue of security text database record computer-based exam. One important aspect to be considered in the implementation of the computer-based test is the security of the exam. If the analogy database security exam conducted without securing the database, then it is very easy to be
attacked by the other party, because if an attacker manages to get the exam can access the database, it is clear the record about the exam can be easily manipulated or leaked.

This study describes how the security database records are secured by encryption of the text record exam questions are then stored into the database exam application.

A Based on the figure 1 above, it is known that the process is carried out starting with the process of encryption (encryption) exam conducted by the maker of the exam committee. The exam that was encrypted stored in the database exam application (server). That is, the text of records stored in the database password from the exam is about the original text. This database to be accessed by the client (examinees). Exam application that is accessed by the participants automatically perform the decryption process (returns cipher into a plain), so the exam can be understood by the client.

### 4.1. The process of Computer Based Test Database Encryption

The process of encrypting the computer-based test database is done by the test team and then stored into the exam application database (server). The encryption process is based on the spritz algorithm to generate exam ciphers. The schema of encryption process can be illustrated in Figure 2 below.

The following database records will serve as an example of the encryption process in this research. Database created using MySQL application.

**Table 1. Table of Record Database CBT Exam**

<table>
<thead>
<tr>
<th>Id</th>
<th>Text of Exam</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Kepanjangan KSA adalah....</td>
</tr>
<tr>
<td>2</td>
<td>Defenisi dari kriptografi yang benar adalah.....</td>
</tr>
</tbody>
</table>

Encryption Key : CRYPTEx
a. KSA Process

Based on the KSA process algorithm, it appears that there are two main processes that are done, namely, generate an array S and do permutations of the contents of the S array that has been formed. The contents of the Permuted S Array will be used in the PRGA process to generate a random key element. The value of i, j, in this case, start from 0 to 255, while the value of N is 256.

The initial step is the formation of an array of initial key:

\[ K[0] = C \text{ (dec 67)} \]
\[ K[1] = R \text{ (dec 82)} \]
\[ K[2] = Y \text{ (dec 89)} \]
\[ K[3] = P \text{ (dec 80)} \]
\[ K[4] = T \text{ (dec 84)} \]
\[ K[5] = E \text{ (dec 69)} \]
\[ K[6] = x \text{ (dec 120)} \]

The next is the manufacture of the array S, by following the KSA's pseudo-code array S, so that the resulting table array with integer values ranging from 0-255.

| Index | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|-------|---|---|---|---|---|---|---|---|---|---|----|----|----|----|----|
|       | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
| Index | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 |
|       | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 |
| Index | 32 | 33 | 34 | 35 | 36 | 37 | 38 | 39 | 40 | 41 | 42 | 43 | 44 | 45 | 46 |
|       | 32 | 33 | 34 | 35 | 36 | 37 | 38 | 39 | 40 | 41 | 42 | 43 | 44 | 45 | 46 |

The text colored red is the index of the array S.

Based on table 2 above, it appears that the value of N is 256, since the number of arrays generated is 256 integer values.

The next step is to permutate the initial array S values (table 2 values) based on pseudo-code KSA.

If \( i = 0; j = 0 \), then:

\[ j = (0 + S[0] + K[0 \text{ mod 7}]) \text{ mod 256} \]
\[ i = 0; j = 0; S[0] \text{ is array S value on index-0; } K[0 \text{ mod 7}] \text{ is value of array key on index 0 modulus 7 (number of initial key characters)} \]
\[ = (0 + 0 + K[67 \text{ mod 7}]) \text{ mod 256} \]
\[ = (0 + 0 + 4) \text{ mod 256} \]
\[ j = 4 \]

swap ( S[0], S[4] )

\( j = 4 \)

exchange (swap) the array S there is at index 0 to the value of the array S at index 4 and conversely. The result of this iteration produces a value \( j \) as 4.

<table>
<thead>
<tr>
<th>Index</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>4</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>Index</td>
<td>240</td>
<td>241</td>
<td>242</td>
<td>243</td>
<td>244</td>
<td>245</td>
<td>246</td>
<td>247</td>
<td>248</td>
<td>249</td>
<td>250</td>
<td>251</td>
<td>252</td>
<td>253</td>
<td>254</td>
<td>255</td>
</tr>
<tr>
<td></td>
<td>240</td>
<td>241</td>
<td>242</td>
<td>243</td>
<td>244</td>
<td>245</td>
<td>246</td>
<td>247</td>
<td>248</td>
<td>249</td>
<td>250</td>
<td>251</td>
<td>252</td>
<td>253</td>
<td>254</td>
<td>255</td>
</tr>
</tbody>
</table>

If \( i = 1; j = 4 \) (\( j \) value taken from the value of the previous \( j \)), then:
The index with a value 1 results. The process is the values of array $S$ that has resulted from the previous permutation.

### Example

Let's consider the following example:

#### Table 4. Value of Array $S$ Tabel in Iteration 1 ($i = 1$) ($j = 10$

<table>
<thead>
<tr>
<th>Index</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>4</td>
<td>10</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>1</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
</tr>
</tbody>
</table>

This process will be done up to the value of $i = N - 1$ or equal to 255 (255th iteration). During the iteration process, there are times when the contents of an array experiencing a process swap (exchange) more than once. The array $S$ values that are permuted in the next process are the values of array $S$ that has resulted from the previous permutation. The result of the key scheduling process (array $S$) as a whole is shown in Table 5.

#### Table 5. The Result of Key Scheduling Algorithm (KSA) Process

<table>
<thead>
<tr>
<th>Index</th>
<th>240</th>
<th>241</th>
<th>242</th>
<th>243</th>
<th>244</th>
<th>245</th>
<th>246</th>
<th>247</th>
<th>248</th>
<th>249</th>
<th>250</th>
<th>251</th>
<th>252</th>
<th>253</th>
<th>254</th>
<th>255</th>
</tr>
</thead>
<tbody>
<tr>
<td>240</td>
<td>241</td>
<td>242</td>
<td>243</td>
<td>244</td>
<td>245</td>
<td>246</td>
<td>247</td>
<td>248</td>
<td>249</td>
<td>250</td>
<td>251</td>
<td>252</td>
<td>253</td>
<td>254</td>
<td>255</td>
<td></td>
</tr>
</tbody>
</table>

#### PRGA Process

The pseudo-random process will generate a new key at random which is equal to the number of plain elements.

The value of $i$, $j$, $k$, and $z = 0$ and the value $w$ is selected one of the relatively prime values with 256, for example, $w = 29$. Suppose, PRGA process to encrypt exam number 1.

The text of the test (plain): Kepanjangan KSA adalah ....

The number of text characters about the test is 26 characters, meaning that the key will be raised as much as 26. The key value obtained will be used to perform the encryption process of each text character on the exam in a stream.

The iteration process when the value of $i$, $j$ and $k = 0$, then:
i = (0 + 29) mod 256
i = 29
j = (0 + S[0 + S[29]]) mod 256
let's see index 29 value on permutation of array S values (table 5), then we get :
= (0 + S[0 + 161]) mod 256
= (0 + S[161]) mod 256
See index 161 value on permutation of array S (table 5), then we get :
= (0 + 120) mod 256
j = 120
k = (0 + 0 + S[120]) mod 256
let's see array 120 (j) value on permutation of array S (table 5), then we get :
= (0 + 0 + 160) mod 256
k = 160
swap S[29], S[160]
exchange the array S value in index 29 with the value of the array S index 160, then the
contents of the array S that has been permutated based on KSA process will be
recovered based on the swap value.
z = (S[10 + S[29 + S[0 + 160]]) mod 256
= (S[10 + S[29 + 160] mod 256]) mod 256
= (S[10 + S[189]]) mod 256
let's see index 10 and index 189 on the result of array S permutation table (table 5),
then we get :
= S[10 + 175] mod 256
= S[185], see array S value in index 185, then we get:
z_1=185 (The key used to encrypt the first character of the exam question is the binary of the
decimal 185).

The process of iteration when the value of i = 29, j = 120, k = 160, z = 185 (value i, j, k
and z taken from the previous process value) then:
i = (29 + 29) mod 256
i = 58
j = (160 + S[120 + S[58]]) mod 256
= (160 + S[120 + 61]) mod 256
= (160 + S[181]) mod 256
= (160 + 232) mod 256
j = 136
k = (58 + 160 + S[136]) mod 256
= (((58 + 160) mod 256) + 160) mod 256
= (218 + 160) mod 256
k =122
swap S[58], S[136]
z = (S[136 + S[58 + S[185 + 122]]) mod 256
= (S[136 + (S[58 + S[51] mod 256])]) mod 256
= (S[136 + (S[58 + 228] mod 256)]) mod 256
= (S[136 + S[30]]) mod 256
= (S[136 + 63]) mod 256
= S[199]

z_2 =109 (the key used to encrypt the second character of the exam question is the binary
of the decimal 109).
This process is done until the 26th iteration (corresponding to the number of exam text).
c. Encryption Process
The process of encrypting text characters the exam is done based on equation (1) so that the cipher is obtained as follows:
Exam Question (plain): Kepanjangan KSA adalah...
4.2. Decryption Process of CBT Exam Database

The process of decrypting the text of a database record of an encrypted test is done in the same way as in the encryption process. Beginning with the KSA process, then the PRGA process and the last is the process of decryption. Decryption is performed based on the formulation in equation (2), which perform the XOR operation between the binary elements with a binary cipher each generated key to the process resulting PRGA to the original text of the record exam. The decryption process is done automatically by the application exam already available on the client computer when accessing a matter examinees. The decryption process scheme is shown in Figure 3.

![Figure 3. Schema of Decryption Process](image-url)
If examinees legitimate access the exam, then the exam application will automatically generate the decryption key to the process of KSA and PRGA based on the initial key used in the encryption process. KSA and PRGA process in the decryption process carried out in the same manner as in the encryption process, because of this algorithm including the symmetric key algorithms (the same key). Keys are generated from the process PRGA (equal the number of records about the ciphertext) is used as a key in the decryption process. Decryption process performed by equation (2), which perform XOR process between the cipher key element to element test database records.

First Key: CRYPTEx
The key generated from the PRGA process is 185 109 ..... 
Biner of key is:
K₁ = 185 (in biner is 10111001)
K₂ = 109 (in biner is 01101101)

If we assume, the decrypted cipher is ő then the decryption process is:
C₁ = ő (decimal is 242 or in biner is 11110010)
C₂ = ę (decimal is 8 or in biner is 00001000)

The next step is to XOR binary ciphers with binary keys generated from KSA and PRGA processes based on equation (2), so:

\[ C₁ \oplus K₁ = P₁ \]
\[ C₂ \oplus K₂ = P₂ \]

The same process will be done to decrypt other record characters, so get the record database exam that same as the original. The overall result of the decryption process is shown in table 7.

<table>
<thead>
<tr>
<th>Table 7. Record of Database Exam After Decrypted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Id</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
</tbody>
</table>

Computing performance measurement results from the key generation process, the decryption process encryption and obtained the following results:

a. Key Generation Performance

Based on the tests performed, the key characters generated by the spritz algorithm are very random, since the keys generated for the encryption and decryption process are no longer the same as the initial key characters but will generate new key characters equal to the number text character test record. But the process of generation of key characters that many would take a long time. This is one of the weaknesses of this algorithm. This occurs because of the number of key characters that are raised to be equal to the number of text characters exam.

<table>
<thead>
<tr>
<th>Tabel 8. Key Generation Performance Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Early Key</td>
</tr>
<tr>
<td>-----------------</td>
</tr>
<tr>
<td>CRYPTEx</td>
</tr>
<tr>
<td>AMIK_STIEKOM</td>
</tr>
<tr>
<td>StiekomSU</td>
</tr>
<tr>
<td>STIEKOM</td>
</tr>
<tr>
<td>Sumatera</td>
</tr>
</tbody>
</table>
According to the table 8 above, it appears that more and more about the encrypted text characters, so the more time it takes to generate the key. Based on the generated key, it appears that the repetition of the same character with a very little initial key character that occurs at intervals that are not adjacent letters.

The performance level of the key generation graph shown in figure 4 below.

![Measurement of Key Generation Process Time](image)

**Figure 4.** Key Generation Performance Testing

Based on figure 4 above, it appears that the more the number of text characters exam questions are encrypted, the higher the time needed in the key generation process, but the character is getting randomly generated key.

b. Computing Performance of Encryption and Decryption Process

Performance computing process of encryption and decryption based on spritz algorithm in this research, conducted by measuring the time of encryption process and decryption of five database record about an exam. The measurement results are shown in table 9 below.

<table>
<thead>
<tr>
<th>No</th>
<th>Number of early key characters</th>
<th>Number of exam text character</th>
<th>Key Processing Time (second)</th>
<th>Time of Encryption/Decryption Process (Second)</th>
<th>Total Processing Time (Second)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>7 char</td>
<td>15 char</td>
<td>0.253</td>
<td>1.405</td>
<td>1.658</td>
</tr>
<tr>
<td>2</td>
<td>12 char</td>
<td>80 char</td>
<td>1.087</td>
<td>4.455</td>
<td>5.542</td>
</tr>
<tr>
<td>3</td>
<td>9 char</td>
<td>150 char</td>
<td>1.460</td>
<td>4.673</td>
<td>6.133</td>
</tr>
<tr>
<td>4</td>
<td>7 char</td>
<td>180 char</td>
<td>5.435</td>
<td>8.246</td>
<td>13.681</td>
</tr>
<tr>
<td>5</td>
<td>8 char</td>
<td>200 char</td>
<td>6.157</td>
<td>12.159</td>
<td>18.316</td>
</tr>
</tbody>
</table>

Based on table 9 above, it appears that the processing time required to perform both encryption and decryption is the same because the process is the same. Time measurement process performed in this research does not include the time required to access the exam by the participant (client) on online question bank server. Based on the measurement results obtained, it was concluded that the more the number of characters exam encrypted or decrypted, the more time the process takes. This has become one of the characteristics of the algorithms that work on the principle of stream cipher (encryption or decryption on an individual basis) include spritz this algorithm.
5. Conclusion

Based on the description of the results and discussion of this research, it was concluded that the text encoding database record exam computer-based algorithm based spritz can minimize the abuses of the exam by parties who are not responsible for cipher generated by this algorithm is able to obscure the meaning of the exam original, so the principle of confusion and diffusion can be realized. Performance algorithms spritz in a random key generation process are quite reliable but requires a long processing time both encryption and decryption. Simple operation in the process of encryption and decryption in spritz algorithm becomes one of the weaknesses of this algorithm of attack types such as know-plain attack or cipher-only attack.

References