Research on the implementation of a database encryption system based on R algorithm

Haiyan Xu, Jing Guo

Henan city in Zhengzhou Province Henan Polytechnic, Zhengzhou, 450046, China

Corresponding author's e-mail: xuhaiyanhp@163.com

Received 10 September 2013, www.cmnt.lv

Abstract

With the development of computer technology and database technology, more and more MIS are implemented. Database is a basic platform in MIS, it stores a plenty of information which is shared by many users. Therefore, database security technology has become the key technology in the development of MIS. According the security requirements of a MIS, this paper introduces R encryption algorithm which adapts to database, and discusses the architecture and characteristics of a database encryption system based on application layer. Also the paper gives a detailed description about the implementation methods of key technology.

Keywords: database encryption; key management: R encryption algorithm; security control.

1 Introduction

With the development and popularization of computer technology, especially broad applications in important branches of national economy, the problem of computer security has been standing out in the information society.

Information is usually stored and managed in database system, so how to guarantee and strengthen the security and secrecy of database system has been the exigent problem. The security of database system lies on two layers: one is measure of user name/password identification, view, permission control and audit from database system itself, large database systems, such as Oracle, SQL server have these functions. The other is that application systems provide. Generally, basic secure technology from database system is adaptive in generic applications. For applications in important branches and sensitive fields, the above measures are not enough. Some users, especially interior ones can also obtain user name and password illegally, use other methods to enter database exceeding their authority and get or modify information. So it's necessary to encrypt important data in database system.

Information is usually stored and managed in database system, so how to guarantee and strengthen the security and secrecy of database system has been the exigent problem. The security of database system lies on two layers: one is measure of user name/password identification, view, permission control and audit from database system itself, large database systems, such as Oracle, SQL server have these functions. The other is that application systems provide. Generally, basic secure technology from database system is adaptive in generic applications. For applications in important branches and sensitive fields, the above measures are not enough. Some users, especially interior ones can also obtain user name and password illegally, use other methods to enter database exceeding their authority and get or modify information. So it's necessary to encrypt important data in database system.

When more and more information systems are constructed in different application areas, the security of these systems becomes a very important aspect concerned especially in some critical systems. Database is center of most of these systems [1]. Database stores not only the permanent data, such as privacy data, person's credit card numbers, but also some important control data, such as some critical task states. So database becomes the hacker's main target and the main protection object in information systems [2-3]. But most information protection mechanisms are aimed to protect the perimeter of the network and to control the access to database [4]. In complicated real application environments, hackers can easily bypass the protection mechanisms of perimeter of the network and the DBMS to attack the underlying operating system (e.g. inside attacks). When hackers break into the underlying operation system, the datum stored in database would be accessed directly through the operating system's file management service [5]. In some circumstances, thefts maybe steal physical hard disks to get critical data. In front of these threats, defense of network perimeter and access control of DBMS are not enough to protect these sensitive datum stored in database [6-7]. Database encryption may occur in OS, DBMS and client. E. Goh etc. has proposed an architecture that used to encrypt P2P file systems [8]. They assume the network storage is un-trusted. All the encrypting and encrypting operations are did in client. One of challenges on database encryption is that typical indexing techniques can't be used on encrypted data. Ernesto Damiani etc. [9] proposed an architecture in which the indexing information is stored in client. That requires change the client application. In some circumstances, it is not easy. The scheme is based on trusted platform. In most systems, it is not applicable.

2 R Encryption Algorithm

Database encryption system has its own requests and characteristics compared with traditional data encryption technology. In traditional data encryption, messages are used as units; encryption and decryption are both from cover to cover. It's impossible to use the whole database as units to encrypt due to the usage of database. As concepts of file (table), record and field in database system, so files, records and fields may be encrypted units. The unit to be encrypted is smaller, the applied range is larger, but the realization is more difficult. In research actuality, no cords and fields are more used as units. Usually, the length of records and fields is short, and the time of storage is long. If the secret keys used are few, the secrecy cannot guarantee; but if the secret keys are more, the management is too complicated. So general encryption technology cannot be used in the encryption of database, and the encryption algorithm and secret key management methods which are suitable for database encryption system must be researched. Requests of database encryption technology are as follows:

(1) The cryptographic strength is first requested due to the long storage time of database information.
(2) Because the heavy use methods of database are random access, the efficiency of encryption and decryption must be high, the performance is not allowed to largely decline.
(3) The structure of database cannot be changed greatly, the size of plaintext and cipher text must be same or corresponsive to the greatest extend.
(4) For the long storage time of data and the complexity of secret keys, the secret key management must be flexible and firm.

Because of the large quantity and frequent access of data, block encryption algorithm ought to be selected in the encryption algorithm of database system, and small block encryption algorithm is better due to the length of fields. The length of DES cipher's secret key is fixed cannot be extended, and the algorithm has been broken. The length of block is at least 128 bits in AES cipher. Both of them are not suited for database encryption. Whereas RCS and RC6 algorithm are very mature block algorithms, which are offered from RSA Co., and they have not been broken, the lengths of block and secret key are variable. So they are suitable for database system encryption.

Rc5 is a fast symmetric block cipher suitable for hard or software implementations. Rc5 is word-oriented; it has a variable word size, a variable number of rounds, and a variable-length secret key. A novel feature of Rc5 is the heavy use of data-dependent rotations-the amount of rotation performed is dependent on the input data, and is not pre-determined. While no practical attack on RCS has been found, the studies provide some interesting theoretical attacks, generally based on the fact that the "rotation amounts" in Rc5 do not depend on all of the bits in a register. Rc6 was designed to thwart such attacks, and indeed to thwart all known attacks.

Rc6 is an evolutionary improvement of Rc5, designed to meet the requirements of the AES. New features of RC6 include the use of four working registers instead of two, and the inclusion of integer multiplication as an additional primitive operation. The use of multiplication greatly increases the diffusion achieved per round, allowing for greater security, fewer rounds, and increased throughput.

According to the characteristics and restriction of the database encryption technology, the author puts forward a new block cipher suitable for database encryption which is named R encryption algorithm. R algorithm expansion routine of RCS, but the encryption algorithm of RCS is modified from RCS, RC6. It inherited key is modified. R inherits the data-dependent rotations of RCS, and also inherits integer multiplication of RC6, but it has only two working registers.

The merits of RCS, RC6 are integrated in R algorithm, and small block is kept, so it is fit for database encryption. The security of R algorithm is higher than RCS, DES; exhaustive search for encryption key can be resisted due to the changeability of the length of key. The R encryption speed is faster. For R-32/16/16 on a 200MHz Pentium, a preliminary C++ implementation is compiled with the Borland C++ compiler, the encryption speed is 4.9M bytes/sec. It can fulfill the need of database encryption.

The detected impulses will be removed by R algorithm. Let $f'_{i,j}$ be the value of the noise image at pixel location (*i*,*j*). For the corrupted pixel (*i*, *j*), the filtering window of size $(2L_f+1)\times(2L_f+1)$ is used. Starting with L_f =1, this filtering window iteratively extends outward by one pixel in its four sides until the number of noise-free pixels (denoted by $P_{i,j}$) within this window is not less than 1. Let $W'_{i,j}$ denote the values of noise-free pixels in the filtering window, i.e.,

$$W'_{i,j}=\{f'_{i+s,j+t}\mid b_{i+s,j+t}=0,b_{i,j}=1,\\ (s,t)\neq(0,0),-L_f\le s,t\le L_f\} \quad (1)$$

The weighted mean value $g_{i,j}$ of the pixel values in $W'_{i,j}$ is defined as:

$$g_{i,j}=\frac{\sum_{f_{i+s,j+t}\in W'_{i,j}} w_{i+s,j+t}f'_{i+s,j+t}}{\sum_{f_{i+s,j+t}\in W'_{i,j}} w_{i+s,j+t}} \quad (2)$$

where $w_{i+s,j+t}$ means the weight of $f'_{i+s,j+t}$. Let $m'_{i,j}$ be the median value of $W'_{i,j}$. Because the median value has the least probability to be the value of the corrupted pixels [1], $m'_{i,j}$ is utilized to determine $w_{i+s,j+t}$. It is easy to understand that the smaller the absolute difference between $f'_{i+s,j+t}$ and $m'_{i,j}$, the larger the weight $w_{i+s,j+t}$ should be to strengthen the influence of $f'_{i+s,j+t}$ on $g_{i,j}$. Based on extensive simulations which indicate that $w_{i+s,j+t}$ is dependant on both above absolute difference and noise ratio, $w_{i+s,j+t}$ is chosen as:

$$w_{i+s,j+t}=R+(1-R)\sqrt{\frac{\dfrac{|f'_{i+s,j+t}-m'_{i,j}|}{f'_{\max}-f'_{\min}}}{1-\dfrac{|f'_{i+s,j+t}-m'_{i,j}|}{f'_{\max}-f'_{\min}}}}, \quad (3)$$

where $f'_{\max}$ and $f'_{\min}$ denote the maximum pixel value and the minimum one in the noise image, respectively.

The output of the DAWM filter is obtained by:

$$h_{i,j} = b_{i,j} \cdot g_{i,j} + (1 - b_{i,j}) f'_{i,j} + \Delta\omega_{i,j} . \tag{4}$$

3 Implementation of Database Encryption System based on R Algorithm

Encryption based on field is adopted in this database encryption system. Cipher text is stored in former data tables. Because the length of cipher text might be longer than the one of plaintext, the length of cipher text needs to be designed longer. In order to simplify the management of secret key and not to influence the operation efficiency, eclectic methods are used in management of secret key: one secret key for one record, so a field to store the secret keys must be added in every table.

The implemented database encryption system consists of two parts: the management module of encryption dictionary table; and the processing module of encryption and decryption. The encryption requirements are stored in the encryption dictionary table. Its maintenance work is done in management module of encryption dictionary table, which is an important tool of database manager. Processing module of encryption and decryption is the hard core of database encryption system, which processes the data encryption and decryption, and it is transparent for the users.

Management module of encryption dictionary is the tool of defining encryption data. Its functions include maintenance of encryption dictionary table, table information encryption and decryption. It completes encryption and decryption of data through processing module of encryption and decryption. This module is managed by database manager; the common user need not care about it. This module gives an interface to database manager. In the interface, the tables, fields needed to be encrypted and decrypted can be appointed. The adjustment of database structure and encryption dictionary table, and the conversion of encryption and decryption are processed. The process is as follows:

(1) A table is chosen;

(2) Click “information”, all fields’ information is got and encryption fields are showed;

(3) Click “set” to define decryption;

There are two list boxes in the form. In original state, all fields list are showed in left list box, encryption fields are showed in right list box. “>|” button can add all contents in left list box to right one, “>”button adds contents selected in left list box to right one, “<” button can delete fields selected in right list box, “<|” button can delete all fields in right list box. The four buttons can set the encryption fields.

Click “ok” to end setting. If you want to encrypt the table, click “encryption” button. First, encryption dictionary table is updated; then new encryption fields is encrypted, the fields which are the former encryption fields but now are not are decrypted.

The structure of encryption dictionary table:

```
CREATE TABLE jmzdb
(TABLENAME char (30) NOT NULL,
FIELDNAME char (30) NOT NULL)
```

The interface of data table is shown in figure 1.

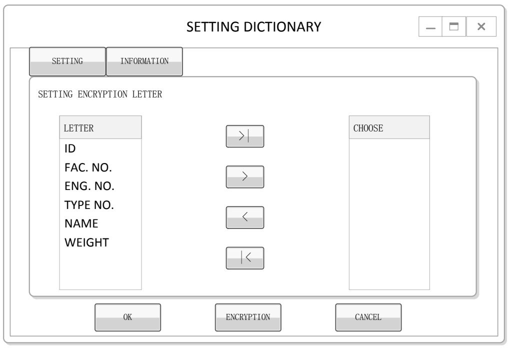


FIGURE 1 The interface of data table

Processing module of encryption and decryption deals with the encryption and decryption. It includes retrieval of encryption dictionary table, creation of secret key, analysis of SQL sentences, and realization of algorithm. All the functions are given as common functions. Because the database platform of MIS system is MS SQL SERVER7.0, development tool is DELPHI 7.0, and R algorithm is operated aimed at binary data, it's easy to realize rotation and x or operation by C. R algorithm is realized by C++, and is made as DLL files, then Delphi applications call DLL files to realize encryption and decryption.

The encryption function is encryfunc(char *s,char *key), the decryption function is decodefunc(char *s,char *key). Two DLL files are decode.dll and encry.dll o

DELPHI calling functions consists of encryption calling function nowencry (miwen: string; key: string): string and decryption calling function nowdecode (miwen: string; key: string): string, respectively call decode.dll and encry.dll to realize encryption and decryption of database fields. A field is needed to store the secret keys (Level 2 keys), the field length is 100 or longer. In case some fields are encrypted, a key is created and stored in the field, and it must be guaranteed that the keys will be exclusive.

When a record is inputted into the application system, the database encryption system checks the request. If encryption is needed, the plaintext is processed, and then cipher text is inserted into table. If encryption is not needed, plaintext is inserted into table directly. If application system need modify a record, encryption dictionary table is first retrieved to find whether there are encryption fields. If so, decrypt them and show, after that encrypt them and input into table.

When querying table, encryption dictionary table is first retrieved. If there are not encryption fields, query processes. If there are encryption fields, SQL sentences are analysed first, check the query condition (WHERE clause) and whether there are encryption fields in output fields, if so, decrypt them and execute SQL sentences, or execute SQL sentences directly.

2 levels secret key management is used in this database encryption system. Level 1 key is main key, Level 2 key is work key. There is one Level 1 key in whole system; it takes charge of Level 2 key encryption. Work key takes charge of the encryption and decryption of database infor-

mation. A work key for a record, after encryption under Level 1 key, it is stored in a key field of every table. Work key is created at random, consists of ID field and current time, the length is uncertain, 100 bytes or so. ID field is created automatically when data is inputted, each ID field is unique.

In the system, main key protects work key, and work key protects sensitive information, so the security of whole system relies on the main key. The main key is appointed as 64 bits binary data when the database encryption system is designed, it is stored in safe area after being encrypted and decrypted automatically when used.

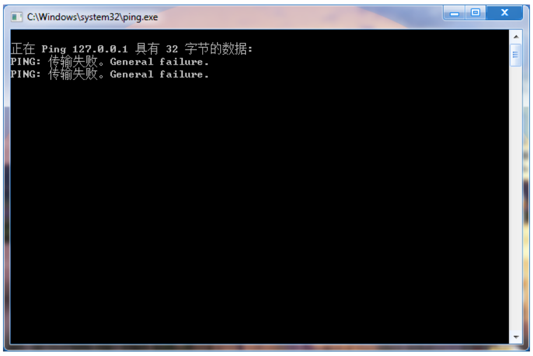
In this system, it's a tough thing to replace main key. After replacement of main key, work keys must be all replaced. There is large quantity of data in database, the data encrypted needs former work keys to decrypt, and the time will be long. The safe way is to backup all data in database before replacement of main key, then convert cipher text to plaintext. After replacement of main key, convert plaintext to cipher text. In fact, it is known from the analysis of key system that even main key is not replaced, the security of data can be guaranteed in a long time. So it's not needed to replace main key in exist period of database.

4 Experimental results

In this section, we test encryption and decryption effects on database performance. We insert 10 tuples, 50 tuples, 100 tuples, 500 tuples, 1000 tuples, 5000 tuples, 10000 tuples respectively and continually into a relation which is not encrypted and record the used time. Then clear the test table and insert same content into a relation which is encrypted and record the used time. Compare the two groups of times.

Rc5 is a fast symmetric block cipher suitable for hard or software implementations. Rc5 is word-oriented; it has a variable word size, a variable number of rounds, and a variable-length secret key. Rc6 was designed to thwart such attacks, and indeed to thwart all known attacks. Rc6 is an evolutionary improvement of Rc5, designed to meet the requirements of the AES. New features of RC6 include the use of four working registers instead of two, and the inclusion of integer multiplication as an additional primitive operation. The use of multiplication greatly increases the diffusion achieved per round, allowing for greater security, fewer rounds, and increased throughput.

The “ping” indicator is used to attack the database which is encrypted by R algorithm. The result is shown in figure 2.



```
正在 Ping 127.0.0.1 具有 32 字节的数据:
PING: 传输失败。General failure.
PING: 传输失败。General failure.
```

FIGURE 2 The result of ping indicator

The security of R algorithm is higher than RCS, DES; exhaustive search for encryption key can be resisted due to the changeability of the length of key. The R encryption speed is faster. For R-32/16/16 on a 200MHz Pentium, a preliminary C++ implementation is compiled with the Borland C++ compiler, the encryption speed is 4.9M bytes/sec. It can fulfill the need of database encryption.

The comparison of R algorithm and RC5 can be seen from figure 1. The result shows that in the same decoding time, the R algorithm achieves better performance than RC5 in encryption complication.

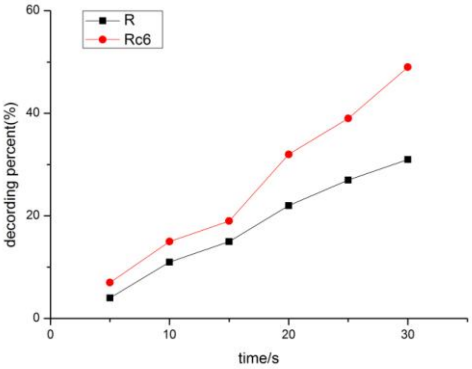


FIGURE 3 The comparison of R algorithm and RC5

The comparison of adaptive weighted mean algorithm and RC6 can be seen from figure 4. The result shows that in the same decoding time, the adaptive weighted mean algorithm achieves better performance than RC6 in encryption complication.

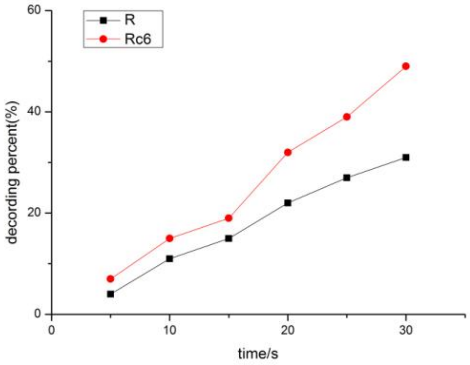


FIGURE 4 The comparison of R algorithm and RC6

5 Conclusions

The realization of database encryption system in application layer is given in this paper. A new small block encryption algorithm, whose block size and key length is variable, R algorithm is adopted. Its encryption granularity is field encryption; secret key management is 2 levels management. The database encryption system has the characterristics as high speed, high intension and simple realization, which can fulfil the need of database security.

Because of time, there are some shortages in practicability and currency of this system:

a) As the length of cipher text is multiple of 64, if the length of plaintext is not multiple of 64, cipher text will be a little longer than plaintext. A longer field is needed in design.

b) When query condition has several encryption fields, decrypting every field to find right data cost much time, the absolute speed of query and responding speed is oppressive.

The future study work will be expanded in above aspects.

References

[1] Biham E, Shamir A 1993 A Differential Cryptanalysis of the Data Encryption Standard *Springer-Verlag* 126-9

[2] Mitsuru Matsui 1994 The first experimental cryptanalysis of the data encryption standard *In Yvo G.Desmedt, editor, Proceedings CRYPTO 94, Lecture Notes in Computer Science* **839** 1-11

[3] Xu Ke, Liu Yaxiao, Liu Weidong 2001 The Design and Implementation of Security Access Proxy in Database application System *Computer Engineering and Application* **1** 105-7

[4] Zhang Jianqiang, Dai Yiqi 2002 Design and Implementation of Network Encrypted Database System Based on Proxy *Computer Engineering and Application* **18** 196-8

[5] Wang Xiaofeng, Wang Shangping, Qin Bo 2002 Research on Database Encryption and Verification *Journal of Xi'an University of Technology* **18(**3) 263-8

[6] Chen T, Ma K, Chen L 1999 Tri-state median filter for image denoising *IEEE Trans Image Process*. **8** 1834–8

[7] Chen T, Wu H R 2001 Space variant median filters for the restoration of impulse noise corrupted images *IEEE Trans Circuits Syst-II:Analog Digital Signal Process* **48** 784–9

[8] Goh E, Shacham H, Modadugu N, Boneh D, SiRiUS: Securing Remote Untrusted Storage *In Proceedings of the Internet*

[9] 2003 *Society (ISOC) Network and Distributed Systems Security (NDSS) Symposium* 133-145

Authors

Haiyan Xu, born in January 6, 1983, Henan of Puyang City

Current position, grades: The lecturer
University studies: Computer science and technology
Scientific interest: database
Publications: seven
Experience: In 2004, after graduation from the university to the Henan Polytechnic Department of information engineering as full-time teachers, in 2008 transferred to the academic affairs office engaged in teaching management work, made in 2009, Huazhong University of Science and Technology master of engineering; have full-time or part-time in computer teaching, the main direction is to: database, network, as the course "database principle and Application

Jing Guo, born in October 24, 1979, Henan of Linying County

Current position, grades: The lecturer
University studies: Computer and communication
Scientific interest: Computer network
Publications: five
Experience: In 2002, after graduation from the university to the Henan Polytechnic Department of information engineering as full-time teachers, in 2005 transferred to the Department office in zhe Department office responsible for the work, made in 2009, Huazhong University of Science and Technology master of engineering; have full-time or part-time in computer teaching, the main direction is to: database, network, as the course "database principle and Application, 2011 was responsible for the administration work in state-owned assets management department