Acta of Turin Polytechnic University in Tashkent, 2023, 30, 12-19
Published Online September 2023 in Acta TTPU (http://www.acta.polito.uz/)

DEVELOPMENT OF ENCRYPTION AND ARCHIVING ALGORITHMS IN ACCORDANCE WITH GOST 28147-89 IN OPEN SOURCE OPERATING SYSTEMS

N.N. Ochilov
State Testing Center under the Cabinet of Ministers Republic of Uzbekistan
Email: nizom.ochilov91@gmail.com

Abstract– Operating systems belonging to the Linux family are somewhat more secure due to the limited access to system files. However, due to the fact that unlimited access is created when logged in using the system administrator (root), it requires special encryption methods to keep some data confidential. According to the level of data confidentiality, only the GOST 28147-89 standard is a highly durable encryption algorithm.

Key words– lossless compression, cryptanalysis, concatenation of cypher text blocks, initialization vector.

I INTRODUCTION

This article is dedicated to the development of encryption algorithms. The encryption algorithm requires at least an 80386 processor, 32 MB of RAM, at least 2 GB of hard disk space, and a C compiler to compile special modules. The software requirements are set according to the requirements of the 7zip archiving software. Downloading the system archiving module is done by downloading the 7zip program. This module cannot be accessed from outside. This module has the ability to encrypt and archive system files with a password. For this reason, 7zip was chosen as a convenient archiver for encryption based on the GOST 28147-89 standard.

The article describes the principles and methods underlying the creation of an application in secure operating systems, which provides reliable data encryption. The aim of the research is to analyze and indicate the specifics of encryption methods and algorithms based on domestic standards in open-source operating systems. Cryptanalysis was used in the article, as this avoids vulnerabilities identified in previously created implementations. In the article, the authors draw attention to the fact that 7-Zip uses CBC encryption (concatenation of encrypted text blocks), but the Counter Mode is supported. The same support was provided in the encrypt implementation. Since the key expansion function initially fills the special array created by p7zip with round keys using a unique property of the domestic standard, only one round encryption function was created (performed both during encryption and decryption). This method is also used in various modes. In many cases, initialization time deviations depending on the selected mode are insignificant. The created cryptographic module was tested to meet the domestic standard, which contains several test cases. It was confirmed during the tests that the created module really implements the algorithm of the domestic standard. The article shows a way to implement a fairly convenient graphical interface for accessing the cryptographic module, which enables the user not to call the command line and remember the sequence and types of parameters passed to p7zip. This implementation also takes into account the verification of the correctness of decryption and the reading of other error codes.

The open-source operating system ensures the protection of processed and stored data through Security Policies, kernel settings, additional software. One of the requirements is to support encryption in accordance with the law. It was required to create an application that performs encryption in accordance with the domestic standard, the only encryption algorithm allowed for use when working with information containing state secrets.

However, the lack of open-source software available in the public domain that performed the assigned tasks, while having sufficient performance, with open source, necessitated the development of a cryptographic module.

The main problem in creating a cryptographic application is to provide all verification procedures — splitting the data stream into blocks, padding to the desired length, creating a pseudo-random sequence initializing vector (IV). As a result, it was decided to use ready-made software, in which these procedures were successfully implemented, [3,7].

For block ciphers, the use of IV is described by modes of operation. Randomization is also required for other primitives such as generic hash functions and derived message authentication codes. An initialization vector (IV) is introduced in CBC, CFB, and OFB encryption. Moreover, both the sender and the receiver must have the same IV at the beginning of the communication session. The IV does not have to be secret at all and can be transmitted along with the first ciphertext block. What's really important is that this value should be unpredictable in CBC and CFB modes and unique in OFB mode. Unpredictability in CBC and CFB modes can be achieved in several ways. For example, the value of a counter (say a message counter) can be converted using the same function. GPC can also be used to generate a pseudo-random sequence of the desired length.

As encryption in everyday work is more often applied to text information, especially to documents in DOC, DOCX format, the presence of standard expected sections in them creates a serious danger in the reliability of the password used for encryption. To eliminate this effect, it is recommended to pre-compress the information, thus reducing data redundancy. For this reason, the archiver was chosen as the main application, [8, 9].

The encryption algorithm in WinZIP versions earlier than 9 is not strong enough with the current computing power, but it also has a 64-bit key size. In this case, it is necessary to correct not only the encryption module, but also key extensions, check it for statistical properties, and so on. Moreover, the encryption block is one byte, [10, 11].

The most popular archiver WinRAR is proprietary, which means that it does not have open source, thereby automatically excluding this software from consideration.

7-Zip is a free, highly compressed file archiver. It supports multiple compression algorithms and multiple data formats, including proprietary 7z format with the highly efficient LZMA (Lossless Compression Algorithm) compression algorithm, [8]. The encryption algorithm used is AES with a block size of 128 bits, a key size of 256 bits.

II METHODS AND RESULTS OF EXPERIMENTAL RESEARCH

Based on the presence of an encryption mechanism in archiving programs and the feature of working in open-source operating systems, they are divided into the following classifications:

- WinRAR archiver is a version of RAR archiver developed for Windows operating systems. This program is designed to create archives in RAR and ZIP formats. In addition, this archiver provides the ability to archive and open 7Z, ACE, ARJ, BZIP2, CAB, GZ, ISO, JAR, LZH, TAR, UUE, Z formats;

- Unlike WinRAR archiver, 7zip archiver is absolutely free. 7zip is open-source software and most of its source code is licensed under the GNU LGPL. One of the main advantages of this archiving program is that it supports working with the command line. 7zip archiver can compress files 2% better than WinRAR. The compression efficiency of the 7zip archiver can be increased up to 10% using the LZW (Lempel-Ziv-Welch) algorithm. The LZW algorithm uses the idea of expanding the information alphabet. Instead of the traditional 8-bit representation of 256 characters in the ASCII table, 12 bits are used to define a table of 4096 entries.

The main purpose of the LZW algorithm is to replace a string of characters with codes using 4096 entries without analyzing the sequence of incoming characters. Each time a new character string is added, the character table is revised. A compression algorithm works when a string of characters is replaced by a code.

During encoding, the characters of the input stream are read sequentially and checked for the presence of such a line in the generated string table (Fig. 1).

A special feature of the LZW algorithm is that there is no need to save the string table to a file for decoding. The algorithm is designed in such a way that it is possible to reconstruct the table of strings using only the code flow.

As a result of using the LZMA version of the LZW algorithm in the 7zip archiver, it is possible to optimize the compression efficiency up to 10% by properly choosing the size of the dictionary, the word length and the number of streams. In open-source operating systems, it is possible to develop an encryption archiving program based on the GOST 28147-89 standard using this algorithm [4].

To organize encryption and archiving in the 7zip archiving program based on the GOST 28147-89 standard [10-11], the following modules should be developed and implemented:

- Gostvars() method does not perform any function. Used to maintain 7zip's compatibility with other archiving methods, it initializes the values of three global variables;

- GOST_SetKey_Enc method is designed to create round keys used in encryption according to GOST 28147-89 standard;

- the GOST method performs the round function of the Feistel network according to GOST 28147-89;

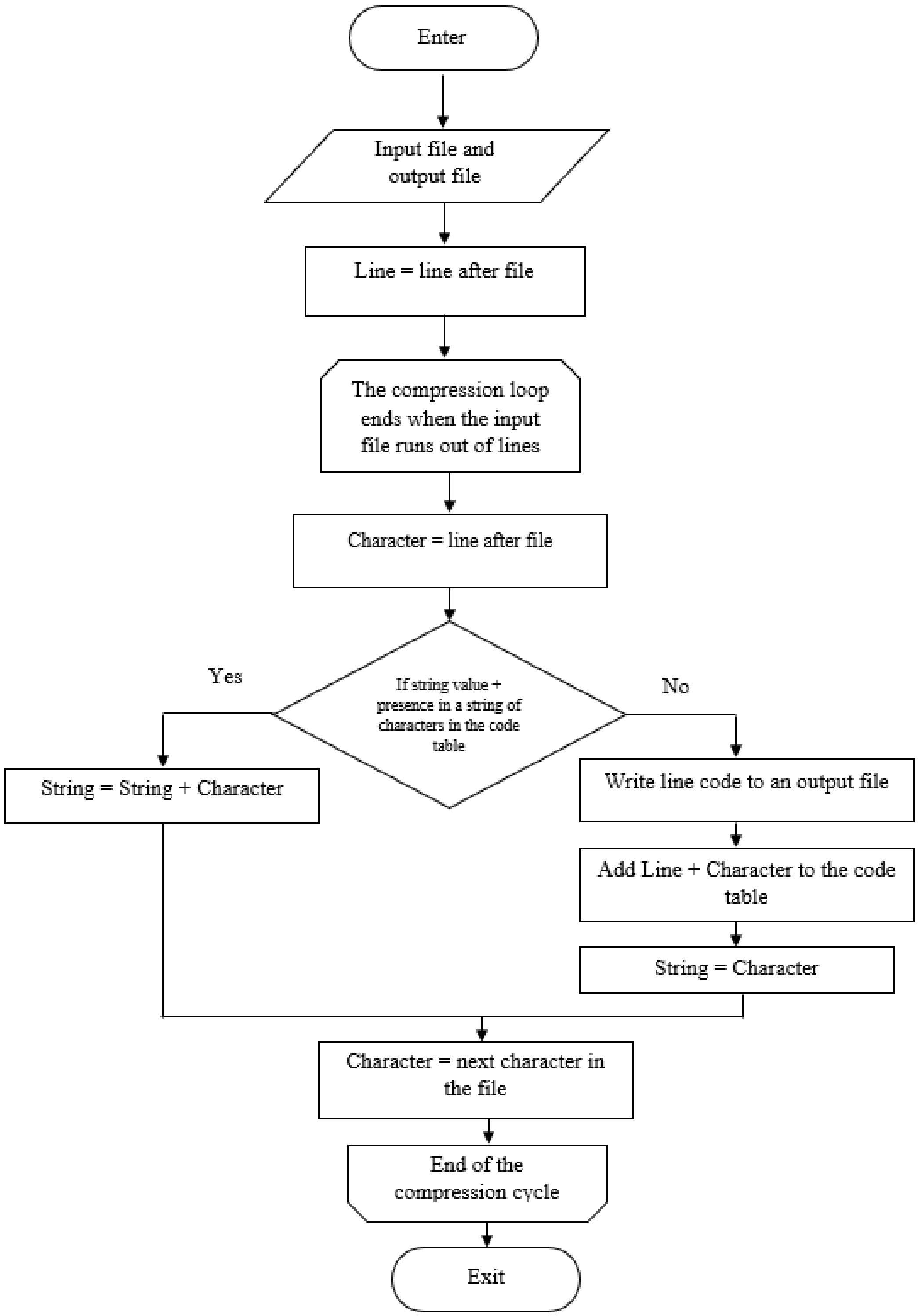

Fig. 1: Block diagram. The LZW algorithm is an information alphabet expansion algorithm.

- GOST_Code method encrypts/decrypts one message block (8 bytes) according to GOST 28147-89 standard;
- GOSTCbc_Init method is designed to write the initialization vector required for CBC mode in array form;
- GOSTCbc_Encode method is designed to encrypt a message whose length is a multiple of 16 bytes in CBC mode according to the GOST 28147-89 standard;
- GOSTCbc_Decode method is designed to decode a message whose length is a multiple of 16 bytes in CBC mode according to GOST 28147-89;
- the GOSTCtr_Code method is designed to encrypt/decrypt a message with a length of 16 bytes in STR mode according to GOST 28147-89;
- the get method is designed to convert a 4-byte unsigned byte into a single 4-byte unsigned number;
- the set method is designed to convert a 4-byte unsigned number to a 4-byte one.

The module is considered part of the 7zip archiver. Used to encrypt archive files with a password. The password is converted to an encryption key using the SHA-2 hash algorithm built into 7zip. Input data to the encryption module is prepared by 7zip according to the standard defined by the SVS encryption mode.

The characteristics of the data included in the module are as follows:

- the initialization vector iv must be a pseudo-random sequence. 64 bits are enough for GOST 28147-89, which are obtained by shifting 64 bits to the right to 128 bits. Only the 7zip archiver performs this check;
- the encryption key (byte array key) is obtained from the password entered by the user using hashing with the SHA-2 algorithm. The use of this algorithm depends only on its statistical, not cryptographic, properties. The length of the encryption key is 256 bits. Only the 7zip archiver performs this check;
- round keys (an array of unsigned 4-byte integers w) are generated from the encryption key by the module based on the key generation procedure established by the GOST 28147-89 standard. The result is a 1024-bit sequence;
- a message (byte array src) is created for encryption/decryption of files being archived. Only the 7zip archiving program is involved in creating the message and filling it with a length of 128 bits;
- the number of 128-bit blocks in the message (an unsigned 4-byte number). To use GOST 28147-89, this number in the module is multiplied by 2 to get the number of 64-bit blocks in the message [2].

When using these methods, the size of files for archiving (with encryption) should not exceed 32 GB, otherwise, it will cause an overflow of the used variables. The characteristic of the main working array is an unsigned 32-bit number for incoming and outgoing data (Fig. 2).

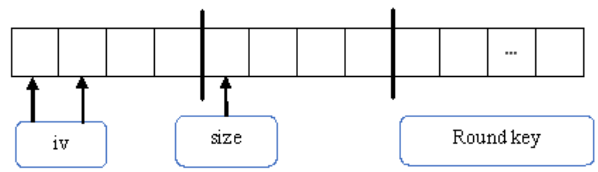

Fig. 2: Characteristics of the main working array.

Since the 7zip archiver uses the AES algorithm for encryption, there are some inconveniences in applying the GOST 28147-89 standard. Since the number of elements required for storing round keys in AES exceeds the number of elements required for GOST 28147-89 round keys, the number in the "size" field is kept the same as in AES. All unused elements of the array have no data (all data in this array is entered and used only by this module, other 7zip modules do not make any changes to it) [4].

To ensure the confidentiality of encrypted files, it is recommended to create a copy by calling the command 7za a –p <archive_name> <list_of_files_to_archive> from the command line. After this command, the archiving program will ask you to enter a password. In this case, the entered password is not saved in the history of the commands called in the system. To increase work efficiency, it is recommended to encrypt files before the end of the work session (turn off the computer) and decrypt them before working with these files. No special programs and actions are required during the operation of this module.

III ANALYSIS OF THE RESULTS

Depending on the specific archiving methods of open-source operating systems, different data formats may produce different results. For example, in many types of formats, the 7zip archiver shows the best result due to the very large size of the dictionary (32 MB). One of the important features below is that 7zip archives have 5% more archiving capacity than WinZIP archives, but are much slower to archive (Fig. 3). On average, it can be observed that the volume of archives is 30% larger than that of other types of archives (Fig. 4). A pilot test for high archiving speed

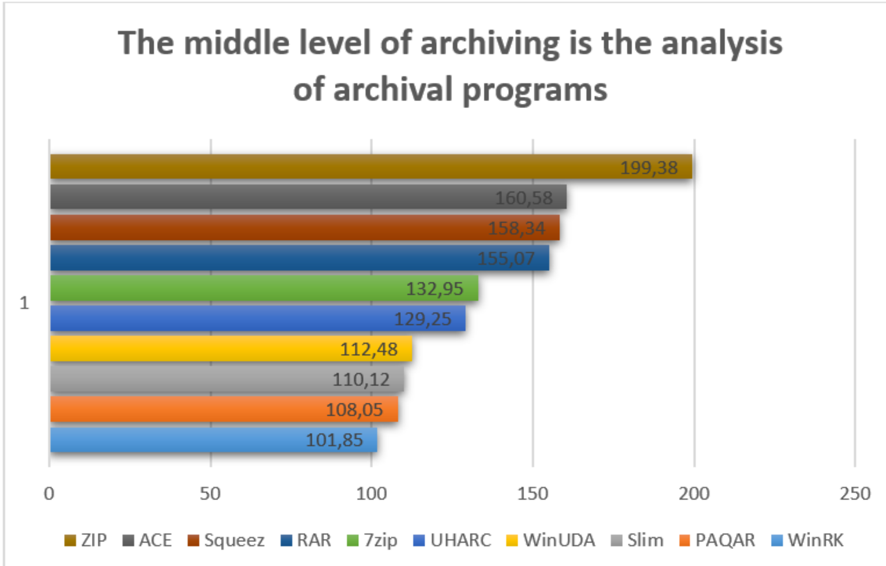

Fig. 3: The middle level of archiving is the analysis of archival programs.

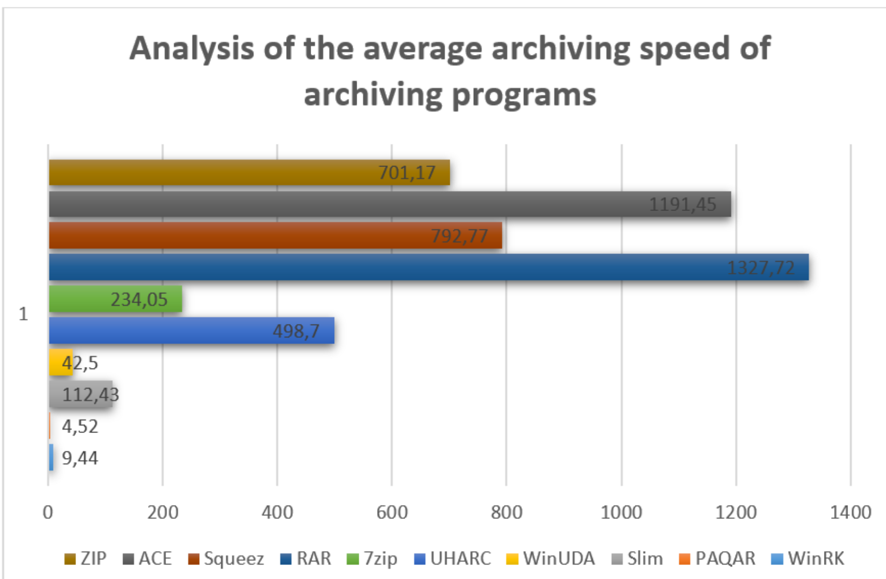

Fig. 4: Analysis of the average archiving speed of archiving programs.

was carried out. In this test, archiving software with the best archiving ratio was found to have a faster archiving speed. A 5 Mbps hard drive was chosen as the high archiving speed in the test system. This speed seems close to optimal.

Later, as speed increases, file storage may become the limiting factor of the subsystem. At the same time, for archiving 1 GB of data can be collected in 3-4 minutes. In addition, a limit on the use of RAM for archiving is set at 32 MB. According to the results of the experiment, such archiver programs can be used in systems with limited resources. The maximum archiving speed differs for different archiving programs. Since the difference between different archiving programs is relatively small at high archiving speeds, there is no need to test different datasets. The results of experimental tests can be expressed as follows (Table 1).

For many archiving programs, including 7zip, on computers with limited resources, a compression rate of 5 Mb/s was taken as the limiting speed in the test (Fig. 5).

	7zip	**ACE**	**WinZIP**	**RAR**	**Win RAR**	**Squees**	**UHARC**
Output directories	+	+	+	+	+	+	+
Creation of self-extracting (SFX) archives	+	+	+	+	+	+	-
Change the contents of the archive	+	+	+	+	+	+	+
Encryption mode	+	+	+	+	+	+	+
Restore archive	-/+	+	+	+	+	+	+
Split the archive into parts	+	+	+	+	+	+	-
Console version	+	+	+	+	+	+	+
Graphic version (GUI)	+	+	+	+	+	+	+
Asymmetry	+	+	+	+	+	+	-
RAM requirements, MB	5	14	8	8	21	8	21

TABLE 1: PROGRAM ARCHIVE "COMPARISON TABLE"

Highly compressed archive applications have some difficulty reaching the required 5 Mbps speed. The most common ZIP archiving program was chosen as a reference. Only archiving programs that provide the ability to archive directories have been tested.

The ACE, RAR and Squeeze archivers show good results but lag far behind the leading archivers in terms of maximum compression. The difference ranges from 25% to 105%, averaging 55% (Fig. 6).

7zip and UHARC can be recommended as universal archivers with a good compression ratio (30% worse than the best archivers). An important advantage of these archiving programs is that they are free and open-source. UHARC almost always archives with the best compression level, but it is a bit less functional. In the maximum mode, the UHARC archiver requires 54 MB of RAM, which allows you to use this program on computers with a limited amount of RAM. UHARC is the leader in high-speed archiving. 7zip is the best asymmetric archiver in terms of archive quality. Archives created with it can be used on almost any computer, but keep in mind that the size of the dictionary does not exceed 256 MB. In the maximum mode, the archiving speed can be increased by 2-3 times by decreasing the value of the Word size parameter. The archiving speed will decrease by more than a few percent. The 7zip archiver has a special PPMd method for archiving text. This method is similar to the RAR text compression method. 7zip uses special methods to archive media data.

WinUDA and Slim are among the programs with a high degree of archiving. Despite the low performance and relatively small functionality, the WinUDA archiving program leaves a good impression on the user. All necessary functions are available, including the ability to create self-extracting (SFX) archive files. Also, the unpacker module takes 18 KB. There is Mode-0 compression, requiring 24 MB of RAM. At the same time, the speed increases to 127 Kbps, and the compression ratio is somewhat worse. Slim has a slightly higher level of archiving (taking into account the specifics of testing processes) and speed but lacks such important features as creating offline decompressed (SFX) archives, a graphical (GUI) version and continuous archiving mode.

PAQAR, one of the best conventional archiving software, has a speed of 5 Kbps. Work is currently underway to develop special methods designed only for certain file formats.

IV CONCLUSION

In conclusion, it should be noted that several archiving programs that support data encryption have been reviewed and tested. The WinRAR archiver uses its own encryption algorithm, which works in blocks of 1 byte. The 7zip archiver program performs AES encryption and performs

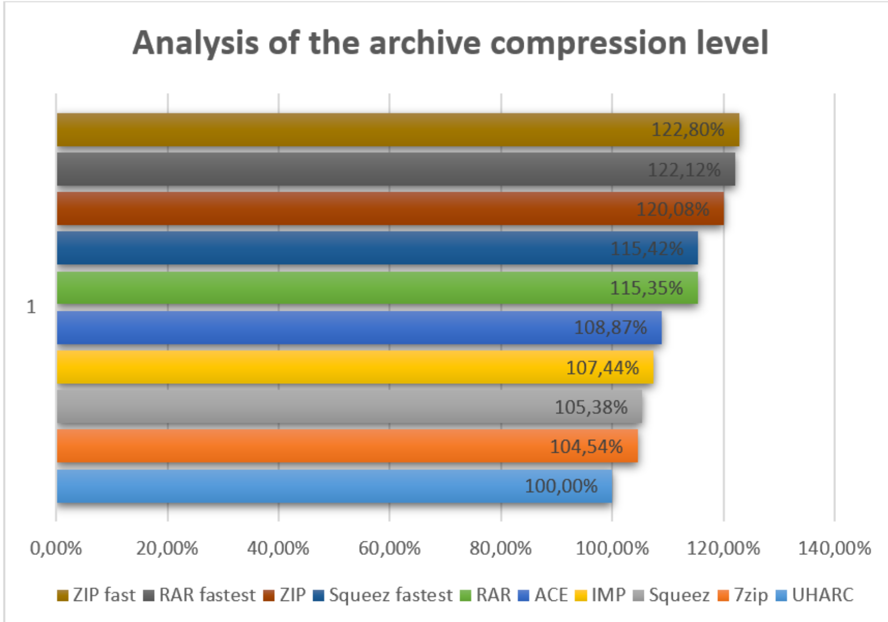

Fig. 5: Analysis of the archive compression level

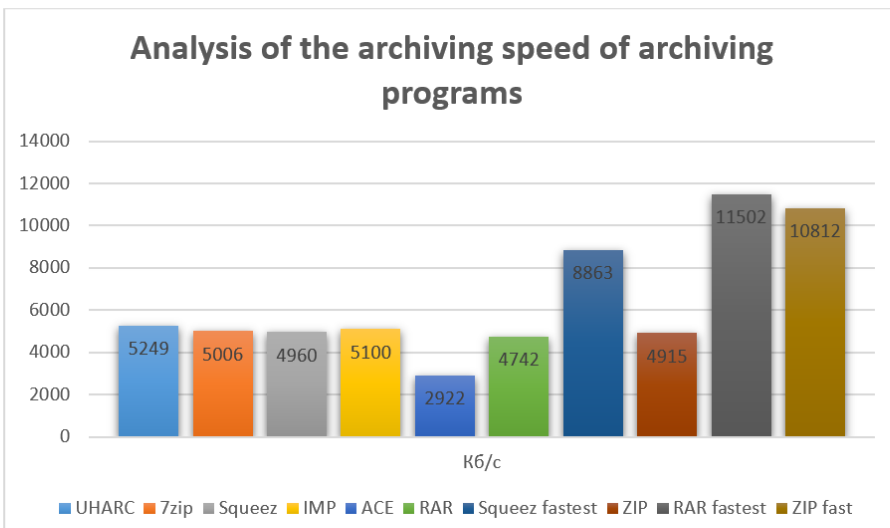

Fig. 6: Analysis of the archiving speed of archiving programs

all the necessary data preparation before encryption (creating a key from a password, adding a multiple-message to the block length by checking the decryption using an extended sequence, creating an initialization vector, etc.). Using GOST 28147-89 also slightly improves the encryption speed since the AES block size is twice the GOST 28147-89 block size. In the course of the analysis, 7zip was adopted as an archiver in which the maximum archiving speed mode can be increased up to 2-3 times by reducing the file parameters, and with the help of these tests, a specially created archiving and encryption program according to the local standard can be used on almost any computer, with that the size of the dictionary does not exceed 256 MB.

V References

[1] Qi Luo, Advancing Computing, Communication, Control and Management, Springer Science & Business Media, 2009, p.290.

[2] Algorithm of cryptographic transformation GOST 28147-89, IPK Publishing House of Standards, Moscow, (Russian), 1996.

[3] Qi Luo, Advancing Computing, Communication, Control and Management, Springer Science & Business Media, 2009., p.290.

[4] Algorithm for cryptographic transformation GOST 28147-89, IPK Standards Publishing House, Moscow, 1996.

[5] Mao, V. Modern cryptography: theory and practice: trans. from English / B. Mao. - M .: Publishing House "Williams", 2005.

[6] Stallings, V. Cryptography and protection of networks: principles and practice: trans. from English / V. Stallings. - M .: Publishing House "Williams", 2001.

[7] Ferguson, N. Practical cryptography: trans. from English / N. Ferguson, B. Schneier. - M .: Publishing House "Williams", 2005.

[8] Shannon, K. Communication theory in secret systems // Works on information theory and cybernetics: trans. from English / K. Shannon. - M .: Publishing house of foreign literature, 1963. p. 333-369.

[9] Schneier, B. Applied cryptography: protocols, algorithms, source texts in C: per. from English / B. Schneier. - M .: Triumph, 2012.

[10] GOST 28147—89 “Information processing systems. Cryptographic protection. Algorithm for cryptographic transformation ".

[11] GOST R 34.13—2015 “Information technology. Cryptographic information protection. Modes of Operation of Block Ciphers ".ISO/IEC 18033-2:2006 Information technology — Security techniques — Encryption algorithms — Part 2: Asymmetric ciphers.

[12] FIPS PUB 197. Federal Information Processing Standards Publication. Advanced Encryption Standard (AES). November 26, 2001.