Authentication control algorithm for long-term keeping of digital data

A V Solovyev
Institute for Systems Analysis Federal Research Center “Computer Science and Control” of Russian Academy of Sciences, 44/2 Vavilova str., Moscow, 119333, Russia
E-mail: soloviev@isa.ru

Abstract. The article is devoted to solving the problem of keeping the authenticity of digital data during long-term keeping in the digital economy. The study gives a brief description of the problem stated in the topic of the article. The existing methods of confirming the authenticity of digital data, their advantages and disadvantages are described. Definitions of authenticity, long-term keeping and other concepts and definitions are given. The article describes an algorithmic method for solving the problem of keeping the authenticity of digital data during their long-term storage. An algorithm for inventory of electronic signatures is described, which is the key to solving the problem of authenticity of digital data using electronic signatures. An assertion is made that the proposed algorithm, which has been tested by practice, will solve the problem posed in the article. The prospects of practical application of the algorithm are considered.

1. Introduction
The problem of the long-term keeping of digital data (DD) in the digital economy is extremely relevant, as shown in many scientific papers (for example, [1, 2, 3]). The leading world economies are busy solving the problem of DD security (Universal Electronic Records Management (ERM) Requirements. U.S. National Archives and Records Administration (2017) https://www.archives.gov/records-mgmt/policy/universalemrequirements, National Archives Announces a New Model for the Preservation and Accessibility of Presidential Records. U.S. National Archives and Records Administration (2017) https://www.archives.gov/press/press-releases/2017/nr17-54). For the Russian economy, the solution to this problem is also relevant, especially since the DD is “a key factor of production” (Program “Digital Economy of the Russian Federation” approved by the order of the Government of the Russian Federation at July 28, 2017 № 1632-r). Of course, the problem of the long-term keeping of DD is a complex problem, and it is wrong to reduce it only to the problem of maintaining authenticity. The problems of long-term keeping also include problems: independence, interpretability, authenticity, reliability, loss of semantics, resistance to external influences, security. More details about the general statement of the DD keeping problem are described in [4, 5]. However, one of the most important problems in the long-term keeping of DD is the problem of maintaining their authenticity, i.e. confidence after years and decades that DD is unchanged and possibly guaranteed confirmation of authorship.
This article is devoted to solving the problem of maintaining the authenticity of DD during long-term storage. An algorithm for keeping authenticity is proposed, with the help of which the posed problem can be solved within the framework of software for hardware and software DD storage systems.

We give the basic concepts and definitions used in the text of the article. Definitions marked with the symbol “*” were introduced as part of this study.

- **Electronic archive (EA)*** – a structured repository of immutable digital data: electronic originals of documents (electronic images of paper documents), created on the basis of laws and rules for maintaining archives in a specific territory (in a specific country);
- **Long-term keeping (storage)*** – keeping over a period of time during which digital data “is maintained as an accessible and authentic evidence (evidence)” (GOST R 54989-2012 /ISO TR 18492:2005 2013 Ensuring long-term keeping of electronic documents);
- **Authentic digital data** – digital data “the accuracy, reliability and integrity of which is maintained over time” (GOST R 54989-2012 /ISO TR 18492:2005);

The definitions “Digital Signature (DS)”, “Enhanced Qualified DS (EQDS)”, “Certification Authority (CA)” correspond to Federal Law of the Russian Federation of April 6, 2011 N 63-FZ «About electronic signature».

2. Literature and problem review

From a review of the technologies for building EA in different countries, one can see that there are three approaches to solving the problem of maintaining authenticity.

The first approach involves organizing a “super-protected” DD archive in which multiple backups of EA software and hardware, organizational measures to limit access to data should completely solve the problem of maintaining authenticity (examples of the USA, Australia, Great Britain see [6, 7, 8, 9, 10]). The second approach is related to the fact that ensuring authenticity rests with an DS (German example [11]). The third is associated with the use of blockchain technologies [2, 12, 13].

The disadvantages of the first approach are related to the fact that the "superprotected" EA:

- may “lose” part of the data during data migration;
- there is also the opportunity to change data during migration;
- operation logs may also be lost (changed) during migration;
- availability of data hashing functions in the system, storage of checksums, etc. in the Russian Federation, for example, will not be a guaranteed confirmation of data integrity since Software for government agencies in the Russian Federation, containing crypto-conversion of information, must be certified;
- high costs of technical support due to the need to constantly ensure a high degree of protection and redundancy of technical equipment and communication channels.

The disadvantages of the second approach (the use of DS):

- data loss during migration, however, its probability will be lower, because not only DD, but also operations with them can be certified by DS;
- the need to integrate EA with third-party certified DD encryption tools;
- certificates and public keys of DS have a limited validity period.

Disadvantages of the third approach (use of blockchain technologies):

- high energy intensity of crypto conversions;
- lack of certified solutions for use in the Russian Federation;
- problems with the long-term storage of block chains, similar to the problems of long-term keeping of DD: authenticity, interpretability, independence from the cryptographic protection algorithms used in a specific implementation of the blockchain.

It can be argued that in the context of the emerging digital economy of the Russian Federation, the second approach looks much more practical. Really:

- has certified cryptographic tools, technologies of work with them over the years worked;
• a third party (in relation to the developers and users of EA) appears uninterested in the violation of authenticity (cryptographic provider), which allows more to guarantee a solution to the problem;
• DD certified by electronic devices are relatively independent of EA software and can migrate to any software and hardware environment, which increases the integrity of data authenticity, even if the EA software environment fails;
• no need to separately organize long-term storage, authenticity, interpretability of block chains;
• relative independence from cryptographic protection algorithms is achieved.

However, to rely only on the DS is not necessary. There are several reasons for this.

Firstly, certificates and public keys of an DS have a limited validity period, so after a year or 5 years when accessing DD from an expired DS, you can receive a message about the incorrectness of the DS, which will cast doubt on the authenticity of the data.

An DS is convenient to use in electronic document management systems, since in them the time for working with data is limited by the life cycle of documents, however, in systems that provide long-term storage, problems with expired certificates and signature keys are guaranteed to occur.

The recently adopted Federal Law of the Russian Federation of December 21, 2013 N379-FZ 2013 «On amendments to certain legislative acts of the Russian Federation» actually equates an electronic document certified by a strengthened qualified notary public with a paper signed and stamped document. However, it does not stipulate how to store the electronic document later, how long it will be valid, how notarial registries maintained in electronic form will be stored for a long time, etc.

It can be argued that only the use of EQDS is not enough, because to the end there is no certainty about the point in time at which the DS was installed. Thus, it is necessary that the DS contains a confirmed time stamp. Only such an electronic notary will confirm the point in time when the electronic signature was installed and prove its validity at the specified point in time.

Secondly, without a certificate of DS verification of the signature is impossible. The certificate is stored in the CA, but there is no guarantee that the CA (and the certificate base) will not cease its activities before the expiration of the DD storage period.

Thus, it can be argued that the key certificate chain of DS must be contained within the DS or transferred to the EA together with the DS. Only in this case, in the presence of all certificates of electronic signature, is there at least some guarantee that after decades the authenticity of digital certificates certified by electronic signature can be confirmed. It should also be taken into account that when checking the ES, you may need a certificate revocation list (CRL) that is current at the time of signing. It should also be stored in EA, as according to the law N 63-FZ «About electronic signature», the CA must keep certificates for at least their expiration date, which is no more than 5 years. However, this period may be significantly lower than the period of storage of DD.

Thirdly, over the period of long-term keeping of DD, cryptographic protection standards may change, and the cryptographic means of electronic signature verification, which was established several decades ago, may disappear or cease to function.

3. Inventory algorithm for digital signatures

To solve the problem of ensuring the keeping of authenticity, an algorithm for inventory of DS was proposed. The main idea of the algorithm is the organization of periodic new certification of a new DS while maintaining the details of the author of the DS from the old DS. The limiting value of the frequency of certification with a new DS is equal to the validity period of the certificate of DS. Since cryptographic protection tools change more slowly, this approach guarantees the independence of DD from specific cryptographic protection tools.

The proposed algorithm is presented at figure 1.

The DS inventory algorithm can be described as follows.

Step 1. The “old” DS installed on DD is checked using the available certified cryptographic protection tools.
**Figure 1.** Digital signature inventory algorithm.

**Step 2.** In case of confirmation of its authenticity, the time stamp (TmSt) of the DS is verified, which confirms the moment of installation of the DS.

**Step 3.** In the case of a positive result of checking the TmSt, the fact of a positive DS verification is recorded in the electronic audit log and certified by a special archive DS indicating authorship of the EA checking officer.

**Step 4.** Next, the author information is extracted from the certificate of the "old" DS.

**Step 5.** A “new” DS is generated, in which data about the author of the “old” DS is recorded, a new DS is re-certified the DD, including the “old” DS. The chain of “old” DSs is maintained in case of the need for a detailed analysis of all stages of the inventory.

**Step 6.** The fact of certification of a “new” DS is recorded in the electronic audit log of the electronic inventory system, log information is certified by the special archive DS indicating the authorship of the EA checking officer.

**Step 7.** If at least one check at the steps of the Step 1 - Step 4 algorithm did not give a positive result, then the fact of authenticity violation is recorded in the DS audit log, DD is marked as having
not passed the DS inventory procedure. In this case, a detailed investigation of the causes of the violation of authenticity is necessary.

It should be noted that the DS inventory procedure does not exclude the substitution or destruction of DD by administrative personnel operating the EA, but guarantees the impossibility of carrying out this operation by input operators.

Additional protection against malicious acts by administrative personnel is the mandatory automatic maintenance of an audit log. Replacing DD and the audit log is a more difficult task, especially when dividing access rights to these objects.

In the process of performing the inventory algorithm, the correctness of the DS of DD is confirmed. The new DS, as more cryptographic resistant, will eliminate or at least significantly reduce the risk of the appearance of fake DD in the future, certified by the old "correct" DS.

The power of computers is constantly increasing, so theoretically over time it is possible to fake DD (a collision of the first kind), when DD are selected for electronic devices in an acceptable time. Fighting this collision requires constant complication of cryptographic algorithms and increasing bit depth of DS keys. By certifying the DD of the new DS, it is possible to guarantee their authenticity during long-term keeping.

Of course, questions still remain, for example, the interaction of EA and CA. Especially often they encounter it when digital information is stored in EA, signed by DS, which are issued by different CAs, for example, in various regions of the Russian Federation. In this case, situations arise when the EA cannot verify the DS of the received DD due to the lack of CA root certificates. As one of the intermediate solutions, it can be proposed to organize the storage in EA of all certificates of DS, CRL and many other additional information on the basis of which an investigation can be conducted and the authenticity of DD can be established.

4. Implementation of algorithm

The ideas laid down in the creation of the DS inventory algorithm were put to practical use in the framework of creating a large geographically distributed information system – the electronic archive for the Pension Fund of the Russian Federation. This information system has been operating for over 16 years in 80 regions of the Russian Federation. The system stores more than 50% of the personalized records of the Pension Fund of the Russian Federation.

The results of practical application in the framework of a large geographically distributed information system give reason to argue about the successful solution of the problem of ensuring the integrity of DD authenticity over a sufficiently long storage period. During the operation, cryptographic protection algorithms, technical and software of the information system have changed. However, the DD has been completely retained.

Practical testing of the algorithm was also performed in a number of implementations of long-term keeping systems of primary electronic documents.

5. Conclusion

Despite the fact that recently there have appeared regulatory documents regulating many issues of long-term keeping (for example, GOST R 54989-2012 /ISO TR 18492:2005), nevertheless, all technological issues of ensuring long-term keeping, in particular, the keeping of authenticity, are assigned to information system developers.

As a solution to the problem of keeping the authenticity of DD, this article proposes an algorithm for DS inventory, which can be implemented in a software and hardware environment for long-term keeping.

The proposed algorithm has been tested by practice in a number of software projects of electronic archives intended for long-term keeping of digital data of electronic documents.

In further studies, it is planned to implement software implementations of the presented algorithm for various types and formats of DD, as well as to develop an inventory algorithm for digital storage media for long-term keeping systems.
In the future, the proposed algorithm may become part of an integrated technology for the long-term keeping of digital data.

References

[1] Suvorovtseva N G 2017 Storage of electronic documents: foreign experience Bulletin of culture and art 4(52) 17-23
[2] Swan M 2015 Blockchain: Blueprint for a New Economy (Sebastopol, CA, US: O'Reilly Media Inc.)
[3] Danilenko A Y, Pashkina E V, Pashkin M A and Solovyev A V 2018 The use of blockchain technology in information systems. Part 2. Confirmation of authorship and integrity High Availability Systems (vol 14) ed V I Budzko (Moscow: Radiotechnika) chapter 1 pp 9–11
[4] Bakanova N B and Solovyev A V 2019 Problems of long-term keeping of big data Information Technology and Computing Systems 2 44-53
[5] Solovyev A V 2020 Long-term storage technology of digital documents Lecture Notes in Electrical Engineering 641 901-11
[6] Ryskov O I 2005 On the main activities of foreign archival bodies in the field of research and regulatory work with electronic documentation Secretarial business 3 76
[7] Afanasyeva L P 2005 Automated Archive Technologies Federal Agency for Education. State Educational Institution of Higher Professional Education Russian State University for the Humanities
[8] Ryskov O I 2004 The main activities of the national archives of the United States and the United Kingdom of Great Britain and Northern Ireland in the field of management of electronic documents of government agencies Russian archives 3
[9] Miller J 2012 NARA to suspend development of ERA starting in 2012 http://www.federalnewsradio.com/?sid=2204570&nid=35, last accessed 21.04.2020
[10] Lipowicz A 2011 NARA officials defend searchability of electronic archive http://fcw.com/articles/2011/11/01/nara-officials-defending-searchability-of-electronic-archive.aspx, last accessed 21.04.2020
[11] Preservation of Evidence of Cryptographically Signed Documents 2015 BSI Technical Guideline TR-03125 (Version 1.2) (Bonn, Germany: Federal Office for Information Security) p 183
[12] Anderson L, Holz R, Ponomarev A, Rimba P and Weber I 2016 New kids on the block: an analysis of modern blockchains https://arxiv.org/pdf/1606.06530.pdf, last accessed 21.04.2020
[13] Danilenko A Y and Akimova G P 2018 Features of using blockchain technology Proc. 27th Scientific and Technical Conf: Methods and Technical Tools of Information Security (St. Petersburg, Russia: Publishing House of the Polytechnic University) pp 73–5