International Journal of Future Generation Communication and Networking
Vol. 2, No. 1, March, 2009

Linguistic Cryptographic Threshold Schemes

Marek R. Ogiela, Urszula Ogiela
AGH University of Science and Technology
Al. Mickiewicza 30, PL-30-059 Krakow, Poland
{mogiela, ogiela}@agh.edu.pl

Abstract

In this paper a new way of secret sharing algorithms expansion based on additional, linguistic stage, introduced for the generation of the secret shadow elements are presented. Such a part shall be generated in the form of a linguistic description of the shared data, built by defined sequential grammar. The introduction of linguistic formalism in this scheme will provide additional information required for revealing the secret previously split with any (m, n)-threshold scheme.

Keywords: *Linguistic Cryptographic, Threshold Schemes, Linguistic Stage*

1. Introduction

During analysis of information processing and management tasks, at least two important questions related thereto may be defined. The first is the gathering and storage of secret information used in specific companies, institutions, or banks. In the recent years this question was intensely developed, and there are many dedicated systems for intelligent semantic querying for selected information as well as systems for archiving such data according to a variety of semantic information.

Some databases may include information of special importance or sensitivity, e.g. of strategic data. Therefore, it is worth to focus our attention to the other significant question related to intelligent information management. It is the question of the capacity to ensure secrecy and selective access to such data for the authorized persons. As such data is ever more often stored in digital form, it becomes necessary to design new solutions and algorithms that allow sharing of crucial information between appropriately authorized persons. Such a potential of managing strategic information may be acquired thanks to the use of certain mathematical techniques, originating from the fields of cryptography. In our case, the task comes down to searching for the formulas that allow intelligent splitting and sharing of information in a way that would allow its reconstruction to appropriately authorized people. The only condition here is the possibility of splitting the data and later their reconstruction by a group of appropriately authorized people.

The task of this paper is to presents such techniques. Especially the questions of information management will in this case focus on the development of linguistic extensions for the well-known threshold schemes [1] [18] [19]. This work makes an attempt to enrich such techniques with an additional stage of splitting the linguistic representation, defining the split data in the binary form. To achieve this, a simple, context free grammar is introduced to allow converting a sequence of bits into its linguistic representation. This representation will then be subject to sharing with the use of one of the known threshold schemes. To reconstruct

the entire secret, however, it will also be necessary to know a number of linguistic rules that will be assigned to one of the participants in the scheme.

2. Classic secret sharing

Algorithms for splitting and sharing secret information are a young branch of cryptography. In the most general case, their objective is to generate such parts for the data in question that could be shared by multiple authorized persons. What arises here is the problem of splitting information in a manner allowing its reconstruction by a certain n-person group interested in the reconstruction of the split information. Algorithm solutions developed to achieve this objective should at the same time make sure that none of the groups of participants in such a protocol, whose number is lesser than the required m persons, could not read the split message. The algorithms for dividing information make it possible to split it into chunks known as shadows that are later distributed among the participants of the protocol so that the shares of certain subsets of users, when combined together, are capable of reconstructing the original information. There are two groups of algorithms for dividing information, namely, secret splitting and secret sharing.

In the first technique, information is distributed among the participants of the protocol, and all the participants are required to put together their parts to have it reconstructed. A more universal method of splitting information is the latter method, i.e. secret sharing. In this case, the message is also distributed among the participants of the protocol, yet to have it reconstructed it is enough to have a certain number of constituent shares defined while building the scheme. The other type of splitting techniques are the methods for information sharing. They are information distribution methods that are somewhat more complex. The algorithms for information sharing are also known as threshold schemes. Using such a scheme allows taking any information and splitting it into n discretional parts known as shadows. In such a manner that any m (where $m \leq n$) from among them may be used to reconstruct the information. This is the so-called (m, n)-threshold scheme.

Below, this work proposes an algorithm for expanding the operation of such schemes and generation of a single additional shadow in the form of linguistic information necessary for the reconstruction of the entirely secret. The general methodologies of using the grammatical approach to the expansion of threshold systems are therefore as follows:

1. Selection of one of the classical schemes for secret sharing.
2. Transformation of the source data into the form of bit sequence.
3. Definition of grammar generating each bit position (or bit blocks) for input secret,
4. Using an syntax analyser to parse the bit sequence.
5. Acquisition of a sequence of production numbers (grammatical rules), being the result of parsing.
6. Splitting the secret represented by a sequence of production numbers, with the application of the selected threshold scheme.
7. Distribution of shadows among the participants of the protocol.

The next subsection describes a method of extending classical threshold schemes for secret splitting to include an additional linguistic stage at which binary representations of the shared secret are coded into new sequences representing the rules of a formal grammar introduced. It also presents an opportunity to generalise the binary conversion procedure into the linguistic

conversion of a larger number of bits. Such stages will introduce additional security against the unauthorised reconstruction of the information and can be executed in two independent versions of protocols for assigning created shadows to protocol participants. The first one is the version involving a trusted arbiter to mediate in the assignment and reconstruction of information. The second is the version without the arbiter (an additional trusted party), but with the assignment of the introduced grammar as a new, additional part of the secret.

3. Linguistic extension for splitting protocols

Expansion of the threshold scheme by an additional stage of converting the secret recorded in the form of a bit sequence is performed thanks to the application of context-free grammar in the following formula:

$G_{SEC}=(V_N, V_T, SP, STS)$, where:

V_N = {BIT, Z, O} – set of non-terminal symbols

V_T = {0, 1, λ } – set of terminal symbols which define each bit value.

{λ } – define an empty symbol.

STS = BIT - grammar start symbol.

A production set SP is defined in following way.

1. BIT → Z BIT
2. BIT → O BIT
3. BIT → λ
4. Z → 0
5. O → 1

The grammar presented here is context-free grammar [15], changing the bit sequences in the form of zeros and ones into a sequence of grammar production numbers that allow the generation of the original bit sequence. In practice, this means that the resulting sequence contains the numbers of rules of the grammar, that is numerical values from the range 1,.., 5.

The conversion of representation is ensured through syntax analyser (parser) that changes the bit sequence into numbers of linguistic rules of the grammar in square time. The graphic representation of using the grammar expansion in classical threshold schemes is presented in Fig. 1.

After performing such a transformation, any scheme of secret sharing can be applied to distribute the constituents among any number of n participants of the protocol. This means that at this stage, any classical (m, n)-threshold algorithm for secret sharing can be run. However, the secret being split is not a pure bit sequence, but a sequence composed of numbers of syntactic rules of the introduced grammar. Depending on its structure and type, it can contain values of two or more bits. So you can imagine a situation in which the grammar conversion will not consist in transforming single bits (as shown above) but also transforming pairs or greater numbers of bits at the same time (i.e. values of two, three, four and more bits will be considered). In that case, the structure of the grammar will be similar, but the sequence of generation rule numbers obtained will have a greater range of values (i.e. the

number of generation rules of the grammar defined for the conversion will increase). At the same time, as the number of generation rules in the grammar increases, the representations of coded bits (now understood as character sequences and not numerical values) grow shorter.

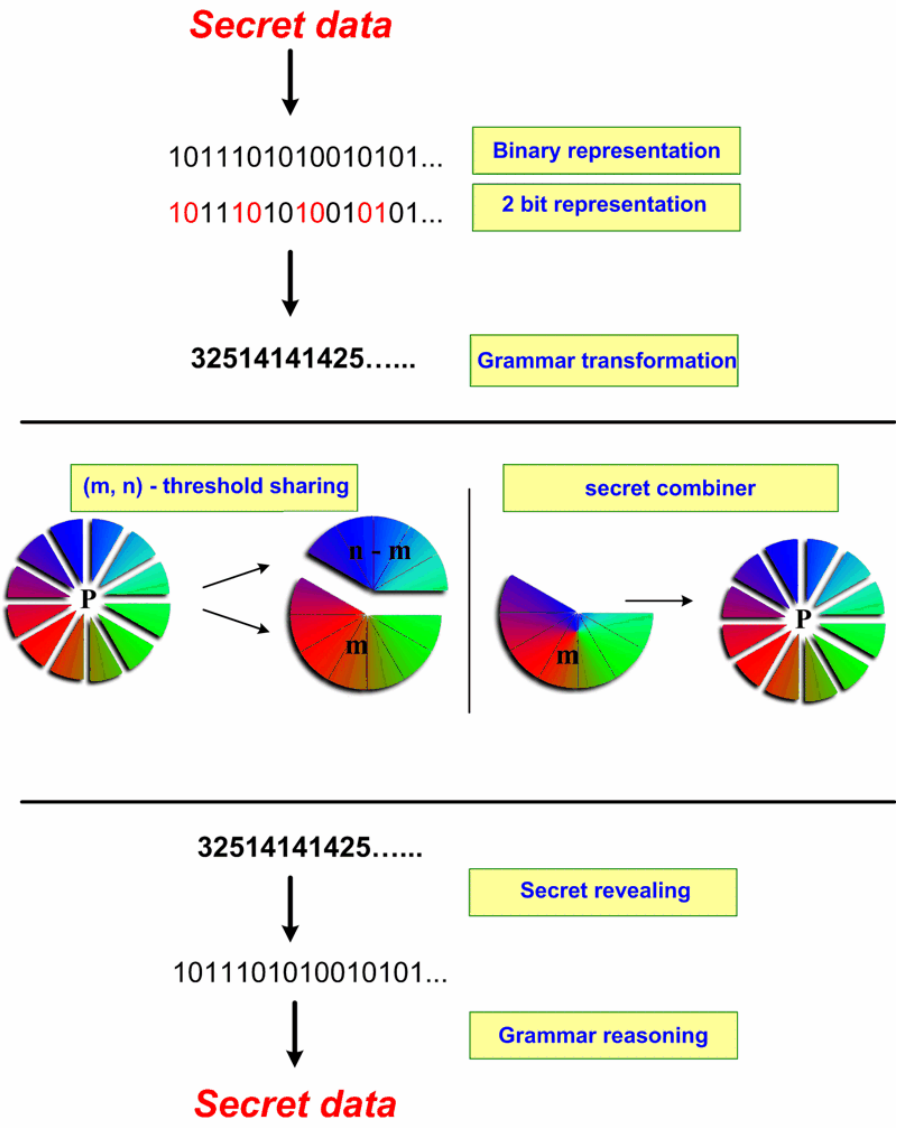

Figure 1. Linguistic threshold scheme. The expansion concerns the use of grammar at the stage of changing the bit representation into sequences of numbers of grammar rules

Executing the introduced algorithms provides an additional stage for re-coding the shared secret into a new representation using grammatical rules. The grammar itself can be kept secret or made available to the participants of the entire protocol.

If the allocation of grammatical rules is to remain secret, as mentioned earlier, what we deal with is an arbitration protocol, which – to reconstruct the secret for the authorized group

of shadow owners –requires the participation of a trusted arbiter, equipped with information about grammar rules.

Should the grammar be disclosed, the reconstruction of the secret is possible without the participation of the trusted person and only on the basis of the constituent parts of the secret kept by the authorized group of participants in the algorithm of information sharing.

4. Grammars for converting bit blocks

As pointed out in the previous section, the stage of converting the bit representation of the shared secret can also be generalised from the version coding single bits to the grammatical coding of bit blocks of various lengths. However, to avoid too many generation rules in the defined grammar, it is worth imposing a restriction on the length of coded bit blocks in the proposed scheme. It seems easy and natural to consider bit clusters no longer than 4-5 bits. Based on information theory, it is then easy to calculate that all representations of values coded with such lengths of machine words will fall within the range of 16 or 32 values, which, when combined with a few additional grammatical rules, allows us to estimate the total number of productions of this grammar as not exceeding 20 for 4-bit words and 40 for 5-bit words.

To illustrate the idea of such a broader linguistic coding, an example of a grammar that converts 2-bit clusters is presented below.

This version for 2-bit block linguistic transformation is distinguished from the original version converting single bit values only by stages which require defining the appropriate grammar and applying it during the transformation to sequences constituting the new representations.

An example of a grammar capable of converting three-bit blocks to a new representation, which constitutes the shared secret at subsequent stages, is presented below.

Such a grammar can be defined as follows:

$G_{2\,BITT}=(V_N, V_T, SP, STS)$, where:

V_N = {SECRET, A, B} – non-terminal symbols

V_T = {00, 01, 10, 11, λ } – terminal symbols which define each 2-bit value

{λ } – an empty symbol

STS = SECRET - grammar start symbol

A production set SP is defined in following way:

1. SECRET→ A
2. A → B A
3. A → λ
4. B → 00
5. B → 01
6. B → 10
7. B → 11

A grammar introduced in this way can support a quicker and briefer re-coding of the input representation of the secret to be shared. Versions for longer bit blocks can be used in the same way. However, this will require introducing a greater number of generation rules. An

obvious benefit of grouping bits into larger blocks is that during the following steps of the secret sharing protocol we get shorter representations for the data that is split and then reconstructed. This is particularly visible when executing procedures that use excessive bit representations, i.e. when single-bit or several-bit values are saved and interpreted using codes in 8 or 16-bit representations.

The level of security achieved does not depend on the length of blocks converted using the rules of the introduced grammar.

5. Conclusion

This work presents a new proposition of expanding threshold schemes with the linguistic descriptions that allow obtaining additional representations that improve the security of the information being split. Linguistic representations were achieved as a result of using context-free grammars that allow conversion of bit representation (the shared secret) to the form of a series of numbers of grammatical rules that allow the generation of the bit description. Such a conversion to the linguistic form is possible thanks to the use of a parser with polynominal complexity. The possibility of establishing new types of arbitration protocols is the result of introducing linguistic descriptions to the schemes used. The arbitration protocol operates when the rules of the introduced grammar remain secret and are stored with a trusted arbiter. In this case, however, what is necessary to reconstruct the secret is the participation of the arbiter, who will have to disclose his share (being the rules of grammar). Another solution is developing an extended scheme in the case when the grammar defined is public. In such a case, the secret split has the form of a series of grammar production numbers. Such a presentation is shared by all the participants of the protocol with the same authorisation. The authorised subset of generated shadows allows for the composition of the secret, and the knowledge of the grammatical rules allows for converting this secret into the form of a bit blocks, and later numerical or text, sequence.

The research conducted in this field by the authors is focused on the definition of methodology and effective means of using threshold techniques for information sharing for multilevel, intelligent management of strategic data in digital form. An important element of the approach presented here is the application of methods of mathematical linguistics. Recently, such methods have been widely used in semantic categorisation of various patterns in [15] [21]. Authors of this work make efforts to create a new intelligent cognitive schemes dedicated for information sharing tasks that make use of biometric authentication.

Acknowledgements

This work has been supported by the AGH University of Science and Technology under Grant No. 10.10.120.783

References

[1] C.A. Asmuth, J. Bloom, “A modular approach to key safeguarding”, IEEE Transactions on Information Theory, 29, 1983, pp. 208 – 210.
[2] G. Ateniese, C. Blundo, A. De Santis, D.R. Stinson, “Visual cryptography for general access structures”, Information and Computation, 129, 1996, pp. 86-106.
[3] G. Ateniese, C. Blundo, A. De Santis, D.R. Stinson, “Constructions and bounds for visual cryptography” Lecture Notes in Computer Science, 1099, 1996, pp. 416-28.

[4] P. Beguin, A. Cresti, “General short computational secret sharing schemes”, Lecture Notes in Computer Science, 921, 1995, pp. 194-208.
[5] A. Beimel, B. Chor, “Universally ideal secret sharing schemes”, IEEE Transactions on Information Theory, 40, 1994, pp. 786-794.
[6] G.R. Blakley, “Safeguarding Cryptographic Keys”, Proceedings of the National Computer Conference, 1979, pp. 313-317.
[7] G.R. Blakley, “One-time pads are key safeguarding schemes, not cryptosystems: fast key safeguarding schemes (threshold schemes) exist”, in "Proceedings of the 1980 Symposium on Security and Privacy", IEEE Press, 1980, pp. 108-113.
[8] B. Blakley, G.R. Blakley, A.H. Chan, J. Massey, “Threshold schemes with disenrollment”, Lecture Notes in Computer Science, 740, 1993, pp. 540-548.
[9] C. Blundo, A. de Santis, “Lower bounds for robust secret sharing schemes”, Inform. Process. Lett. 63, 1997, pp. 317-321.
[10] C. Charnes, J. Pieprzyk, “Generalised cumulative arrays and their application to secret sharing schemes”, Australian Computer Science Communications, 17, 1995, pp. 61-65.
[11] Y. Desmedt, Y. Frankel, “Threshold Cryptosystems”, Advances in Cryptology – CRYPTO’89 Proceedings, Springer-Verlag, 1990, pp. 307-315.
[12] M. van Dijk, “On the information rate of perfect secret sharing schemes”, Designs, Codes and Cryptography, 6, 1995, pp. 143-169.
[13] N. Hang, W. Zhao, “Privacy-preserving data mining Systems”, Computer, 40(4), 2007, pp. 52-58.
[14] W.-A. Jackson, K.M. Martin, C.M. O'Keefe, “Ideal secret sharing schemes with multiple secrets”, Journal of Cryptology, 9, 1996, pp. 233-250.
[15] M.R. Ogiela, R. Tadeusiewicz, Modern Computational Intelligence Methods for the Interpretation of Medical Images, Springer-Verlag, Berlin Heidelberg, 2008.
[16] M.R. Ogiela, U. Ogiela, “Linguistic Extension for Secret Sharing (m, n)-threshold Schemes”, SecTech 2008 - 2008 International Conference on Security Technology 2008, ISBN: 978-0-7695-3486-2, DOI: 10.1109/SecTech.2008.15, pp. 125-128.
[17] M.R. Ogiela, U. Ogiela, “Linguistic Approach to Cryptographic Data Sharing”, FGCN 2008 – The 2nd International Conference on Future Generation Communication and Networking 2008, DOI: 10.1109/FGCN.2008.89, Vol. 1, pp. 377–380.
[18] A. Shamir, “How to Share a Secret”, Communications of the ACM, 1979, pp. 612-613.
[19] G.J. Simmons, “An Introduction to Shared Secret and/or Shared Control Schemes and Their Application in Contemporary Cryptology”, The Science of Information Integrity, IEEE Press, 1992, pp. 441-497.
[20] S. Tang, “Simple Secret Sharing and Threshold RSA Signature Schemes”, Journal of Information and Computational Science, 1, 2004, pp. 259-262.
[21] R. Tadeusiewicz, M.R. Ogiela, Medical Image Understanding Technology, Springer Verlag, Berlin-Heidelberg, 2004.
[22] T.-C. Wu, W.-H. He, “A geometric approach for sharing secrets”, Computers and Security, 14, 1995, pp. 135-146.
[23] Y. Zheng, T. Hardjono, J. Seberry, “Reusing shares in secret sharing schemes”, The Computer Journal.,37, 1994, pp. 199-205.

Authors

Professor Marek R. Ogiela D.SC, Ph.D., works in Bio-Cybernetics laboratory at the AGH University of Science and Technology in Krakow. In 1992 graduated from the Mathematics and Physics Department at the Jagiellonian University. In 1996 for his honours doctoral thesis on syntactic methods of analysis and image recognition he was awarded the title of Doctor of Control Engineering and Robotics at the Faculty of Electrical, Automatic Control, Computer Science and Electronic Engineering of the AGH University of Science and Technology. In 2001 he was awarded the title of Doctor Habilitated in Computer Science for his research on medical image automatic analysis and understanding. In 2005 he received a professor title in technical sciences. Member of numerous world scientific associations as well as of the Forecast Committee `Poland 2000 Plus´ of the Polish Academy of Science and member of Interdisciplinary Scientific Committee of the Polish Academy of Arts and Sciences (Bio cybernetics and Biomedical Engineering Section). Author of more than 150 scientific international publications on pattern recognition and image understanding, artificial intelligence, IT systems and biocybernetics. Author of recognized monographs in the field of cryptography and IT techniques; author of an innovative approach to cognitive medical image analysis. For his achievements in these fields he was awarded many prestigious scientific honors, including Prof. Takliński´s award (twice) and the first winner of Prof. Engel`s award, nominated in the category Science to the Silver Nike award in 2003. Reviewer of world scientific periodicals.

Urszula Ogiela MBA – economist, and computer scientist. She received Master of Science degree and Master of Business Administration in Information Management from AGH University of Science and Technology in Krakow in 2002. Currently she is a Ph.D. student at Krakow University of Economics, and works at the AGH University of Science and Technology, leading her research on linguistic aspect of information data sharing, as well as grammar extensions for secret splitting threshold protocols.