Cognitive cryptography techniques for intelligent information management

Marek R. Ogiela⁎, Lidia Ogiela

AGH University of Science and Technology, Cryptography and Cognitive Informatics Research Group, Al. Mickiewicza 30, 30-059, Krakow, Poland

ARTICLE INFO

Keywords:
Intelligent information management
Cognitive cryptography
Semantic and personal information
Data sharing

ABSTRACT

This paper discusses the foundations of cognitive cryptography used to secure information by splitting it and distributing the split parts among selected groups of secret trustees. The process of concealing data by its splitting and distributing secret parts (shadows) with the use of cognitive techniques will be discussed. Cognitive cryptography describes the possibilities of using personal information contained in individual biometric traits. At the same time, it will be presented as an innovative solution allowing the holder of a secret shadow to be identified based on their characteristic biometrics and their semantic features. Cognitive cryptography is used to manage strategic information. By using techniques for splitting and sharing this type of data as well as utilizing individual biometric traits in the entire process of distributing all shadows of the concealed and split information makes the proposed cognitive cryptography techniques an innovative, extremely useful tool for securing data of major importance.

1. Introduction

Cryptographic techniques are used to secure information and restrict access to it. Securing the information is mainly aimed at protecting it from theft. This theft is understood as the ability of unauthorised individuals (systems) to access information and use it (Menezes, van Oorschot, & Vanstone, 2001; Ogiela, 2016; Schneier, 1996). Every piece of information is thus subject to protection whose level depends on the important message (data) contained in this information. It is, however, worth noting that in the case of information which is in public domain, this information is not subject to any special security, but if it has to be protected, cryptographic protocols can be used for this purpose. In the case of data that is confidential, secret or strategic, it is necessary to use data protection algorithms to verify the individuals having access to this data. Cryptographic algorithms are among those that are to ensure the appropriate security and protection of information (Beimel et al. 2016; Ogiela, 2015b; Ogiela & Ogiela, 2008, 2011, 2014; Tang, 2004). They are considered secure if their use ensures the complete protection of data. Data security is a necessary element in the operation of various information exchange processes. These processes can include, for instance, the exchange of information between:

• parties to or participants of a process – the exchange of information between various individuals who participate in the data/information exchange process,
• sites where process participants are located – the exchange of information between different places in which parties to the process are situated, such as company branches, representative offices, remote sites,
• process participants, but at different points in time – creating copies of information which will be reproduced after a certain period of time has expired.

In all the above situations, the complexity of the process of exchanging information between parties to the protocol should be accounted for as well (Shamir, 1979; TalebiFard & Leung, 2011). This is because there are single, simple information exchange protocols, and protocols dedicated to complex ranges of users, such as information splitting and sharing. These kinds of solutions are designed for securing data by splitting it and distributing parts of this split information (the so-called shadows) among a selected group of secrets trustees. This information is not stored by one protocol participant whose action would depend only on their own decisions, but is distributed among a specific group of participants (individuals, computers), who should act rationally as a group. This means that if a decision needs to be taken, a group of secret trustees must work in concert, and in addition:

• intent to disclose the secret in a situation in which its contents need to be disclosed,
• guard the secret if its security is threatened.

Any action aimed at operating contrary to the other participants may cause:
• disclosing secret information in a situation in which it should not have been,
• no access to the information when its disclosure has a priority, overriding nature.

In every instance, the assessment of the situation and the need to take appropriate decisions rests with the protocol participants (holders of parts of the split secret) who should ensure the complete protection of information against its unauthorised disclosure and seizure. Consequently, the choice of the appropriate data security protocol represents the main step in securing data from its unauthorised disclosure. Cryptographic protocols used for data security are divided into two classes (Beimel, Farras, & Mintz, 2016; Ogiela, 2015a, 2015b; Ogiela & Ogiela, 2008, 2011, 2014, 2016b):

• Data splitting protocols – this class includes cryptographic protocols allowing the secret information to be split between a group of participants, of whom every participant receives one part of the secret (shadow). Each separate shadow is useless on its own, because it contains no complete information. To reconstruct the split secret, all parts of the secret have to be combined. Hence, to retrieve the secret information, all protocol participants must be in full agreement.
• Data sharing protocols – this group includes cryptographic protocols which allow the secret to be split between a group of participants, and to retrieve the secret \( m (m < n) \) parts of all \( n \) must be combined. Just as in the case of data splitting protocols, each shadow does not contain any complete information, but once the required number of shadows is combined, the complete content of the secret can be retrieved. In this case, to retrieve the concealed information, it is necessary for a specific group of secret trustees – \( m \) of \( n \) participants – to agree.

Cognitive techniques aimed at securing information by splitting it have been described, among others, in publications (Beimel et al., 2016; Ogiela, 2015a, 2015b; Ogiela & Ogiela, 2008, 2011, 2014, 2016b; Schneier, 1996; Shamir, 1979; Tang, 2004)), which present the characteristic features and the method of selecting optimal solutions and cryptographic techniques dedicated to secret information sharing. The method of concealing information by splitting it and distributing individual parts of the secret (shadows) to protocol participants allows the data to be protected by restricting access to it. This is because no protocol participant owns all parts of the split secret, and therefore they cannot take an individual decision to disclose or refuse to disclose the information.

However, these types of solutions also have drawbacks, such as the ability to generate empty shadows and assign them to protocol participants without their knowledge of the contents of the shadows they receive. In the case of data sharing protocols, a kind of collusion between a selected group of secret trustees is also possible, resulting in them deciding about the fate of the data that they have been entrusted with without the knowledge and agreement of the remaining participants.

In order to eliminate this type of threat, it is necessary to introduce the ability to randomly select protocol participants who can reconstruct the secret information. Then, no participants will know whether, in the given protocol, they are appointed as a trustee deciding about the fate of the data or not.

An innovative solution eliminating the above threats is offered by cognitive cryptography, which is the main subject of this publication.

2. Cognitive cryptography

Cognitive cryptography is to improve personal identification processes using personal information (Haynes, Bawden, & Robinson, 2016; Ogiela, 2010, 2014; Ogiela & Ogiela, 2016a) characteristic for each protocol participant. This personal information is contained in individual biometrics sets describing the biometric features of each protocol participant (Ogiela, 2016a, 2016b). Because of the nature of biometrics, they constitute individual, unique information that unambiguously characterises their owners. Because of this individual and unique nature of biometric features, it is a very significant both from the scientific and practical perspective to use biometrics for the personal identification and verification. The ability to combine solutions that secure information with an unambiguous way of verifying its owner by identifying their characteristic biometrics has allowed the authors of this publication to develop an innovative solution called cognitive cryptography.

Cognitive cryptography is thus a novel approach to securing data using the individual personal features of each protocol participant.

Definition 1. Cognitive cryptography is a division of cryptography within which any information set can be secured using personal information contained in the biometric sets of information and semantic information unambiguously identifying individual features of protocol participants.

Personal information contained in individual biometrics is used during personal identification to correctly assign biometric traits to the right person, and then, during personal verification it is used to assess whether the biometric traits characterise the right person or whether

![Fig. 1. The scheme of personal verification.](image-url)
they belong to another protocol participant or even to an unknown individual.

The process of personal verification using the biometrics characteristic for a given participant is schematically presented in Fig. 1. The diagram shows the way biometric features are selected from the set of available biometrics. Then, after the characteristic biometric features have been identified, information is concealed by splitting it and distributing parts of secret data among protocol participants. At this stage, protocol participants are biometrically verified by comparing their personal features with the patterns stored in the base. If the biometric characteristic for the given protocol participant complies with the pattern corresponding to that participant, the verification process is successful. As a result, the protocol participant is assigned a part of the secret (shadow).

Otherwise, when there is no compliance between the characteristic biometric of the recipient, and the pattern stored for them, the verification process ends in failure, as a result of which the participant cannot be assigned one of the parts of the split secret (shadow). There may be various reasons for this situation:

- an incorrectly defined biometric pattern, an incorrectly measured and stored biometric,
- a biometric assigned to the wrong owner,
- an attempt to impersonate another protocol participant by using their biometric features.

Regardless of the reason for the non-compliance of the biometric features and the patterns characteristic for their owners, the process of
determining the compliance will end in failure. If so, the reasons for this situation should be determined.

If the personal verification is successful, i.e. if the biometric feature pattern complies with its owner, the protocol participant is assigned a part of the split secret. The secret trustee holds it until the reconstruction of the complete secret information. Then, it is necessary to combine the parts of the secret according to the cryptographic protocol. At this stage, the protocol participant is verified based on their biometrics. This is because it is important that when secret data is being reconstructed, the parts of the split secret should come from the correct protocol participants. This is because a protocol participant could, inappropriately, transfer their part to another person (another protocol participant or an individual from outside the group of shadow holders), who, by attempting to participate in the process of reconstructing secret information, could learn its contents. To prevent this kind of situation, every protocol participant taking part in the process of reconstructing secret data undergoes the personal verification process. This stage is aimed at protecting the secret information from disclosure to unauthorised individuals.

Both the personal identification and verification carried out using biometric information are aimed at not only verifying whether the biometric belongs to its specified owner, and, conversely, whether the specific holder of the biometric is its real owner, but they are also used to identify specific semantic information. The role of the semantic information is to determine the meaning of the data held (interpreted, analysed). Every type of biometrics, both standard and non-standard, has a specific set of semantic information (Grossberg, 2012; Koriat and Gelbard, 2014; Ogiela & Ogiela, 2015, 2016a, 2016). Biometric sets are divided as follows (Ogiela, 2016a, 2016b):

- standard biometrics – including:
  - DNA,
  - biometrics of the finger(s),
  - hand biometrics,
  - facial biometrics,
  - retinal biometrics,
  - speech biometrics,
  - gait biometrics,
  - voice biometrics,
- non-standard biometrics – including:
  - body structure biometrics,
  - coronary vessel biometrics,
  - foot biometrics,
  - nervous system biometrics.

Each of the above sets of biometrics is individual and unique. It does contain (in its structure, in the description of characteristic features, in the repeatability of specific features etc.) information that is unique and determines the significance of a given biometric feature in the entire personal verification process. The authors of this paper have proposed this type of semantic information for concealing data of a confidential, secret or strategic importance in cryptographic applications. Semantic information contained in personal biometrics becomes a type of marker for each part of the information concealed (Fig. 2).

Every holder of a part of the concealed information marks it as their individual part using a standard biometric, or if the information is of special significance, they may use information (data) representing the semantics of their non-standard biometrics for this purpose. This means that for the complete security of data, they may use e.g. a fingerprint in the process of identifying protocol participants, and in addition information describing the shape of the coronary vessels. This type of data constitutes a set of individual data held only by each of the protocol participants. In this situation, every part of the split secret is marked twice:

- at the first stage, using a standard biometric,
- at the second stage, using a feature or a set of semantic features describing non-standard biometrics.

This two-stage security of data makes its protection more certain. If an attempt is made to impersonate a specific protocol participant to steal their part of the secret, it would be necessary to have their biometrics in the form of e.g. the retina or the palm-print, and also information about non-standard biometrics, such as the mutual location of the coronary vessels characteristic for that secret holder. However, this type of data is very difficult to steal because even the secret holder is not always aware of it and does not fully know each one.

Semantic information contained in the descriptions of individual biological features can thus be used in the processes of personal identification and verification carried out by assessing the compliance of selected features of the semantic description during processes in which the compliance of personal features with their owner is established. Determining this type of compliance forms the basis for creating cryptographic solutions aimed at the strongest possible protection of secret information. Information of a secret or strategic nature in not only stored or concealed, but also managed (Buchanan and McMenemy, 2012; Ogiela & Ogiela, 2011, 2014, 2015).

3. Intelligent information management

Personal cryptography techniques ensuring the security of data that has the nature of specially protected data (strategic, secret, confidential) are to ensure the security of data regardless of the external situation. Protocols of data splitting and sharing are used in the process of protecting data from its theft or unauthorised disclosure (publication). In these types of solutions in which secret information is split between a group of its holders, its security and protection from unauthorised access is ensured.

Data sharing processes are aimed at properly securing data and preventing decisions on disclosing the secret being made by e.g. a single individual. The individual process of deciding on disclosing or preventing the disclosure of secret information depends only on the individual interpretation of the circumstances in which the data to be disclosed. It can therefore be the result of a subjective assessment of the situation. In order to rule this kind of situation out, protocols of data splitting and sharing can be used to protect data from being subject to a decision of one individual.

Using data sharing protocols makes the following possible:

- concealing the information, because no protocol participant holds the complete secret/strategic information,
- no single person decision to disclose or prevent disclosure of the secret information.

In addition, using data sharing protocols in which all participants receive parts of the split secret means that every participant feels responsible for the secret part they hold. Hence the process of distributing parts of the split information among protocol participants ensures the correct management of the secret/strategic information. Every protocol participant receives their part of the split secret. In accordance with the defined method of distributing secret parts, every protocol participant may receive one or more parts of the split secret. However, the number of parts assigned to a protocol participant depends on the cryptographic protocol used.

Consequently, the distribution of parts of the split secret (so-called shadows) may be (Ogiela, 2014, 2015a, 2015b, 2016; Ogiela & Ogiela, 2015):

- equal with every protocol participant receiving only one shadow,
- equal with every protocol participants receiving more than one shadow, but all receiving the same number of shadows,
- privileged with a specific group of protocol participants receiving a
greater number of shadows than others,
• unequal – every protocol participant receives a different number of shadows.

A novel approach to managing secret/strategic information using cognitive cryptography is the one proposed in this paper, consisting in the random selection of biometric sets for the personal identification and verification of each protocol participant. The proposed solution has been presented schematically in Fig. 3.

Fig. 3 presents the method of checking the compliance of protocol participants with their characteristic personal features identified using biometric descriptions. At every step in the operation of a given entity, sets of individual features of secret holders are randomly selected, with the full personal identification and verification carried out.

The proposed approach means that in every process aimed at verifying the protocol participant and secret holder, a set of features which will be used in the comparison process will be randomly determined. Hence the proposed solution ensures:

• the complete randomness of the choice of identification and verification features,
• the independence of the verification process and the identification process, as at each stage other biometric features may be used, so that the person trying to impersonate the protocol participant would not know which of the personal feature sets will be used,
• the possibility of any number of repetitions of assessments of compliance between the personal features and the features of protocol participants if there is doubt about the way this assessment is made.

Managing secret, strategic and confidential information requires its special protection from theft. Consequently, this process is supported by the proposed solutions for sharing data within a specific group of secrets trustees and a process of identifying all protocol participants, and when an attempt is made to reconstruct the concealed information, also a process of verifying the holders of secret information parts.

These processes can be executed independently at each level of information management: its concealment, distribution among protocol participants and its reconstruction.

Specific information management processes performed by using information concealment protocols are an example of an intelligent management process. Their special characteristic is the ability to manage/strategic/secret information using a particular type of marker consisting of the biometric and semantic features. This approach to information management points not only to the special nature of the data (secret), but mainly to the needs to protect it at every stage of the management process.

Hence intelligent data management applies to:

• managing secret, confidential and strategic data,
• managing data using the protocols for verifying all information holders,
• managing data using personal descriptions based on biometric data and the semantic information it contains,
• managing data at any level of the entity’s operation (i.e. at the organisation level, from the fog level and from the cloud level).

Intelligent management of secret data can be carried out at various levels at which such data exists (Fig. 4).

The first and basic level is the level of the entity (organisation, enterprise), where all secret data is protected from theft and unauthorised disclosure.

At this level, it is necessary to use data splitting and sharing protocols for the correct, complete protection of data. Protocols for distributing data between all holders of secret parts include the identification and verification of participants using individual or biometric features and their meaning. At this stage, the shared information is
distributed between the holders of the secret, usually the employees of the specific structure.

In hierarchical entities, the concealed and protected information can be distributed in different ways at various hierarchy levels. This means that at one level, it may be distributed among \( n \) participants, and at another among \( k \) participants, where \( n \) and \( k \) represent any different numbers of participants not greater than the number of employees at the given level of the hierarchical structure — this applies to an equal distribution of parts of the split secret, in which every protocol participant receives only one part of it. In other cases of secret distribution – equal, in which every participant receives more than one shadow and in the unequal and privileged distributions – the number of parts of the split information may be greater than the number of shadow holders. This is because, as everyone receives more than once shadow, the total number of shadows may be greater than the number of protocol participants.

The second level of management takes place at the fog level. Data is transferred to a higher level of storage called the fog. Data is held outside the entity, on external servers, where only individuals with the appropriate access certificates can access it. It is managed from the level of the entity which owns it, but without the need to also own expensive hardware for its storage.

The fog level may also be managing secret data. In this case, the data being distributed must be correctly protected. To protect the data, information sharing protocols can be used in which this information is protected by the biometric marking of every shadow at the stage at which parts of the split information are assigned to individual protocol participants, and also at this stage of transmitting and collecting data. This process ensures the unambiguous distribution of secret parts between protocol participants, who manage their parts from the level of the fog. To retrieve the concealed data, it is necessary to combine the required number of shadows, every one of which is verified with regards to its source. This process is executed by determining the compliance of biometric features with their semantic description.

From the fog level, the data can be transferred to the next level: the cloud. This level ensures access to the protected data after the full verification of every user, to properly protect the data from unauthorised access to it. From the level of the cloud, the following data management processes are conducted:

- data storage,
- data administration,
- data access, concealment and security,
- protecting data from unauthorised access to it,
- splitting data and distributing it between protocol participants,
- reconstructing the split information,
- exchanging parts of the split information between protocol participants.

From the cloud level, the data can be sent to the other management levels and the same way of distributing data among protocol participants can be applied at various levels of data management:

- at the organisational level, in connection with data concealment and its distribution between protocol participants at the organisational level,
- from the fog level, in connection with data concealment and its distribution between protocol participants at the organisational level and at the fog level,
- from the cloud level, in connection with data concealment and its distribution between protocol participants from the fog level and the cloud level.

In the process of reconstructing the concealed and shared data, the following reconstruction of secret information are possible:

- at the organisational level, taking into account the organisational structure:
  ○ in hierarchical structures, at the level of the hierarchy at which the secret data was split, or from a higher level, taking into account the secret splitting rules at this level,
  ○ in layered structures, in the specific layer or in other layers, taking into account the rules of secret splitting within the given layer,
  ○ in mixed structures, at the level of the hierarchy and in the layer in which the secret data was split, or from a higher hierarchical level and a superior layer, taking into account the secret splitting rules in these subdivisions,
- from the fog layer – for data shared at the level of the organisation and the fog,
- from the cloud layer – for data shared at the level of the fog and the cloud.

The proposed solutions of cognitive cryptographic techniques for securing sensitive data and managing this type of data make it possible to use the above techniques in various areas in which there is data that is secret or strategic in nature.

### 4. Possible applications

The areas in which the described solutions can be used are those where sensitive data exists. The most important of them definitely include defense, state security, the financial strategies of both states and enterprises or organisations as well as the directions of economic development.

Within an organisation or an enterprise, the proposed solutions can be used for:

- the entity’s security strategy,
- the financial strategy,
- the entity’s development strategy,
- this strategy for distributing confidential/secret information.

Taking into account the nature of selected entities, the proposed solutions could be used in the following sectors:

- service distribution:
  ○ financial institutions, banks, medical institutions, forwarding companies, production companies, etc.
- information distribution:
  ○ banks, hospitals, clinics, colleges and universities, schools, local government etc.
data distribution:
- business organisations, industrial companies, manufacturing companies, state administration etc.

However, regardless of the area in which data is to be secured using cognitive cryptography, one should keep in mind that the essence of this solution is the way in which the data is protected, and not its contents. This is because the proposed cognitive cryptographic methods can be used both to secure data about the financial development of an enterprise, and that characterising the financial activity strategies of the company. Their universality is due to the ability to use the protocols of concealing data by its splitting and distribution within a specific group of secret trustees, who are personally identified and verified based on characteristic biometric features and the semantic description of selected personal features.

Hence the essence of the proposed approach to data management processes is the ability to use the individual and unique nature of the biometrics of secret trustees. This is because marking parts of the secret with the individual features and determining the significance of individual personal features enables the unambiguous and unique description of each participant of the information concealment protocol. These features, because they are unique, can be used to mark any information, as a result unambiguously linking that information to the person. In addition, if it is necessary to determine the holder of a given part of the information, these features will unambiguously indicate the owner with the use of their individual traits.

Using cognitive cryptography techniques in the intelligent management of secret data makes it possible to:
- protect data from unauthorised access to it,
- share the data among a specified group of secrets trustees by assigning a specific number of shadows to every one of them,
- unambiguously distribute the shadows (parts of the split secret) among the protocol participants by determining the compliance of the biometric features of each secret trustee with the defined pattern,
- unambiguously verify each protocol participant using their characteristic biometric features and their semantic description,
- efficiently manage secret data and parts of the split secret,
- reconstruct the information when it is necessary to disclose the secret,
- protect the information in situations in which it is threatened.

5. Conclusions

Cognitive cryptographic techniques are dedicated to protecting and securing datasets of particularly high and material significance. They are dedicated to protecting data that is classified, confidential, strategic and frequently called a secret because of its significance. In cognitive cryptography, it is particularly legitimate to use personal information contained in biometric information sets, as well as semantic information which is unambiguously used to identify the individual features of all protocol users.

Personal information is contained in the datasets held by each protocol participant, because they are embedded in individual biometrics. During the personal identification process, these biometrics and their features are unambiguously used to properly assign the biometric features to the specific person. Then, in the personal verification process, they are used to assess whether the specific biometric features characterise the correct protocol participant.

Hence cognitive cryptography is used to protect information based on the biometric analysis of individual features used to conceal this information. The ability to utilise biometric features and description to identify and verify protocol participants in secret data management allows the shares to be correctly assigned to secret holders, and conversely, the holders to their features. These types of solutions used for managing information of great significance indicate the areas in which intelligent information management can be applied. The word ‘intelligent’ applies to semantic solutions which use non-standard biometric features that can be described and interpreted in processes of the semantic interpretation of personal traits.

Acknowledgment

This work has been supported by the National Science Centre, Poland, under project number DEC-2016/23/B/HS4/00616.

References


