BIOMETRIC AUTHENTICATION DEVICES AND SEMANTIC WEB SERVICES

An Approach for Multi Modal Fusion Framework

L. Puente Rodríguez
Universidad Carlos III de Madrid, Avda de la Universidad, 30, Leganés, Madrid, Spain

M. J. Poza
Universidad Francisco de Vitoria. Ctra. Pozuelo-Majadahonda Km. 1.800. Pozuelo de Alarcón, Madrid, Spain

J. M. Gómez, B. Ruiz
Universidad Carlos III de Madrid, Avda de la Universidad, 30, Leganés, Madrid, Spain

Keywords: Biometrics, Authentication Devices, Emerging Technologies, Data Integration, Semantic Web Services, Ontologies.

Abstract: Identity verification is now a days a crucial task for security applications. In the near future organizations dedicated to store individual biometric information will emerge in order to determine individual identity. Biometric authentication is currently information intensive. The volume and diversity of new data sources challenge current database technologies. Biometric identity heterogeneity arises when different data sources interoperate. New promising application fields such as the Semantic Web and Semantic Web Services can leverage the potential of biometric identity, even though heterogeneity continues rising. Semantic Web Services provide a platform to integrate the lattice of biometric identity data widely distributed both across the Internet and within individual organizations. In this paper, we present a framework for solving biometric identity heterogeneity based on Semantic Web Services. We use a multimodal fusion recognition scenario as a test-bed for evaluation.

1 INTRODUCTION

Identity recognition is performed now a days by the use of traditional techniques such as PINs, passwords, digital signatures, etc. Biometrics promise to offer a new alternative, portable, easy to use, free of memory, loss or theft problems. A global solution will be based on the creation of specialized organizations offering authentication services. This Biometric Accreditation Entities (BAE) will base their services on previously acquired biometric data.

Biometrics authentication usually refers to the identification of an individual based on his or her distinguishing traits. In principle, a biometric identity is based on the premise that a measurable physical or behavioural trait is a more reliable indicator of identity than the traditional systems such as pairs composed by password and username, Personal identification numbers (PIN) and the akin. Particularly, since biometric identity technologies deal with security and privacy issues, the challenge for the research community is to attain integrated solutions that address the entire problems from sensors and data acquisition to biometric data analysis and system design.

Presently, the lack of performance of biometric systems is being alleviated by the use of multiple biometric indicators for identifying an individual in order to increase its accuracy when using a technique called Multimodal Fusion (Kittler, J. Hatef, R. Matas, J. G., 1998) (Jain R. and J. Quian, 2001). As a result of this, biometric information has grown exponentially and algorithms for feature extraction, matching score or decision levels handle a tremendous amount of data. Furthermore, the

Puente Rodríguez L., J. Poza M., M. Gómez J. and Ruiz B. (2008).
BIOMETRIC AUTHENTICATION DEVICES AND SEMANTIC WEB SERVICES - An Approach for Multi Modal Fusion Framework.
In *Proceedings of the First International Conference on Biomedical Electronics and Devices*, pages 95-100
DOI: 10.5220/0001046300950100
Copyright © SciTePress

recent years have provided with a lattice of duplicated efforts in building test databases such as face recognition databases (e.g. FERET, PIE or BANCA) (Bailly-Baillière E., S. Bengio et al., 2003) and a lack of uniform standards and granted open access to these databases, as discussed in (Ming, A. Ma, H., 2007).

Hence, arguably the most critical need in biometric identity recognition is to overcome semantic heterogeneity i.e. to identify elements in the different databases that represent the same or related biometric identities and to resolve the differences in database structures or schemas, among the related elements. Such data integration is technically difficult for several reasons. First, the technologies on which different databases are based may differ and do not interoperate smoothly. Standards for cross-database communication allow the databases (and their users) to exchange information. Secondly, the precise naming conventions for many scientific concepts in fast developing fields such as biometrics are often inconsistent, and so mappings are required between different vocabularies.

Hence, we present in this paper a framework for solving multimodal fusion oriented biometric identity data heterogeneity problems, keeping the structure of databases created with the aim of being used for identity accreditation and distributed over the Web. Our approach is based on the breakthrough of adding semantics to Web Services which perform a role of entry points for such databases. Fundamentally, this implies that our framework enables different biometric identity data to be discovered, located and accessed since they provide formal means of leveraging different vocabularies and terminologies and foster mediation.

The remainder of this paper is organized as follows. In Section 2, a brief state-of-the-art on the technologies employed in our research is given. Section 3 defines some terms we use along this paper. Section 4 identifies the heterogeneity of data involved in the biometric identification process. Section 5 describes the framework for solving problems using Semantic Web Services. Finally, conclusions and related work are discussed in Section 6.

2 STATE-OF-THE-ART

Semantic Web Services and Ontologies are the cornerstone technologies applied in our research. On the one hand, data interoperability between different information sources is achieved by means of ontologies and their mapping. On the other hand, Web Services semantically annotated are the software entities responsible for providing a normalized interface to disparate functionality and data sources. In this section, a brief description of each of these technologies is put forward.

2.1 Ontologies

Although a number of different ontology definitions can be found currently in literature, in this work we use Borst's one (Borst, W.N., 1997): "an ontology is a formal specification of a shared conceptualization", where 'formal' refers to the need of machine-understandable ontologies. This definition emphasizes the need of agreement in carrying out a conceptualization. On the other hand, 'shared' refers to the type of knowledge contained in the ontologies, that is, consensual, non-private knowledge. In this work, this definition of ontology has been adopted.

Ontologies have become the de-facto standard knowledge representation technology after the emergence of the Semantic Web along with Semantic Web Services and the Semantic Grid. For all these new research branches, ontologies are the cornerstone technology. Knowledge in ontologies is mainly formalized using five kinds of components: classes, relations, functions, axioms and instances (Gruber, T. R., 1993). There are several formal languages used to construct ontologies, that is, ontology languages, including KIF, OCML and F-Logic. Along with the Semantic Web, new markup ontology languages have come out such as SHOE, DAML+OIL, and the current *de facto* standard, OWL (Web Ontology Working Group, 2004).

2.2 Semantic Web Services

Semantic Web Services are a new technology resulting from the combination of other two technologies, namely, the Semantic Web and Web Services. On the one hand, the Semantic Web (SW) aims at adding semantics to the data published on the Web (i.e., establish the meaning of the data), so that machines are able to process these data in a similar way a human can do (Berners-Lee, T., Hendler, J., Lassila, O., 2001). Ontologies are the backbone technology of the SW as they provide structured vocabularies that describe the relationships between different terms, allowing computers (and humans) to interpret their meaning flexibly yet unambiguously.

On the other hand, Web Services (WS) technology extends the Web from a distributed source of information to a distributed source of functionality. It is based on a set of standard protocols, namely, UDDI (Universal Description, Discovery and Integration), SOAP (Simple Object Access Protocol), and WSDL (Web Services Description Language). Therefore, WS provide the means to develop globally accessible loosely-coupled applications. However, as the Web grows in size and diversity, there is, composition and invocation can be carried out by autonomous software entitiesan increased need to automate traits of WS such as discovery, selection, composition and execution. The problem is that current technology around UDDI, WSDL, and SOAP provide limited support for all that (Fensel, D. & Bussler, C., 2002). As a consequence, the principles behind SW technology were applied to WS leading to what we know as Semantic Web Services (SWS) technology. It consists of annotating WS with semantic content so that service discovery.

The W3C is currently examining various approaches with the purpose of reaching a standard for the SWS technology: OWL-S (OWL Web Ontology Language for Services) (OWL-S W3C Submission, 2004), WSMO (Web Service Modeling Ontology) (WSMO W3C Submission, 2005), SWSF (Semantic Web Services Framework) (SWSF W3C Submission, 2005), and WSDL-S (Web Service Semantics) (WSDL-S W3C Submission, 2005). The two most widespread approaches are OWL-S and WSMO. OWL-S is an ontology for services that makes it possible for agents to discover, compose, invoke, and monitor services with a high degree of automation. Similarly, WSMO provides a conceptual framework for semantically describing all relevant traits of WS in order to facilitate the automation of discovering, combining and invoking electronic services over the Web.

3 SOME DEFINITIONS

A trait is defined as any physical, motor or psychomotor human characteristic capable of being used in biometric identification.

A user is any person for the system to recognize, and whose traits are stored somehow in the database.

A donor is every person (user or not) whose trait is captured, voluntary or involuntary, by a sensor of the system.

A sample is defined in (Mansfield, J. Wayman, J.L., 2002) as a biometric measure presented by the donor which eventually results in an image or signal.

4 IDENTITY HETEROGENEITY PROBLEMS

A typical biometric system presents a well defined structure (Mansfield, J. Wayman, J.L., 2002) that includes two phases: enrolment and testing. Enrolment faces the creation of a type of model representing the user in a univocal way, while testing tries to determine if the donor matches or not the model.

This two phases share the steps related to sample acquisition, pre–processing signal and feature extraction. Enrolment completes its chain with a model creation, while testing do it with a matching step.

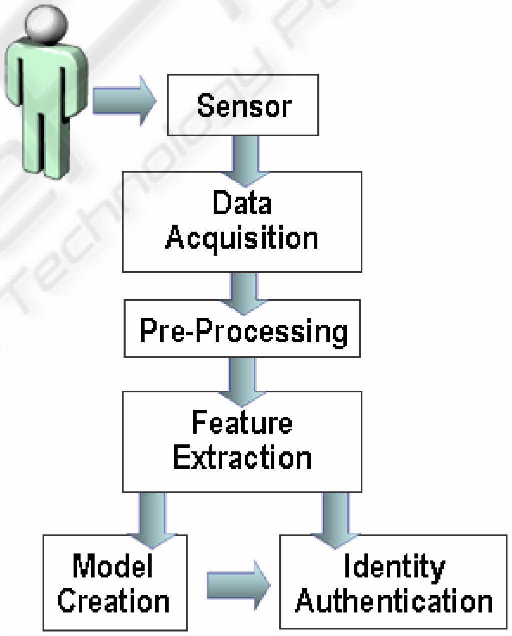

Figure 1: The process.

Acquisition implies that one or more sensors acquire one or more samples of certain donor biometric traits presented to the biometric systems (e.g. fingerprint, face, iris image). Different kind of sensors capturing different biometric donor traits generate different kind of samples.

After capture, samples are pre–processed by cleaning and normalizing in order to adapt the signal to further data extraction, which means signal filtering, enhancement, energy detection, image centring...

The feature extraction module obtains certain information values supposedly related to the donor in a univocal way. These values are collected in sets called “feature vectors”. Different feature extraction algorithms generate different vectors types.

During enrolment a set of homogeneous vectors are used to create a new model of the related trait of the user. While testing the identity claim user model is compared with another set of vectors obtained from the current donor.

For mono–modal biometric systems only one trait is scanned, and of course only one model is generated for each user.

In our multi–modal biometric fusion approach, multiple user models are generated: one for each modality. Every kind of sensor, scanning different traits, generates different signals. Every one of them flows through different acquisition–preprocessing–extraction chains. The testing phase matches the resulting vectors of every chain with the correspondent model, obtaining as a result a confidence level which informs of the probability that the sample belongs to the user identity claim. Finally, these results are collected by a decision module which will decide to validate or reject the donor.

Different capture, pre-processors or feature extractor algorithms generate different kind of models. Then, other kind capture–process–extractor chain generate feature vectors that should not match properly the model. That’s why semantic information should be added to raw data in order to identify such a variety.

5 THE FRAMEWORK

All over the world we can find heterogeneous databases as a set of biometric recorded data. In this section, we present an integrated approach that address the entire problem of enabling entities (organizations) to confirm individual authentication, based on biometric models (traits) stored in different databases all over the Web. The main requirements of our framework are as follows:

- Provide a platform that allows data matching of acquired biometric samples against individual biometric models (traits) stored in certain databases

- Provide catalogues of data that allows to determine a given model location and make these catalogues available via the Web.

- Use of the Web Service technology to make this access a reality and provide the plumbing communication technology over the wire.

- Use of Semantics to find the most accurate model source for the biometric testing that is taking place.

- And finally use available Web Services in order to ask for autentication for acquired samples.

In our framework, we have addressed this process, taking into account the growing complexity of having a multimodal biometrics test. For that, we notice that our work is mostly oriented to multiple biometric fusion strategies, where multiple biometric measures are utilized (Kittler, J. Hatef, R. Matas, J. G., 1998).

A graphical representation of our framework is depicted in the figure 2.

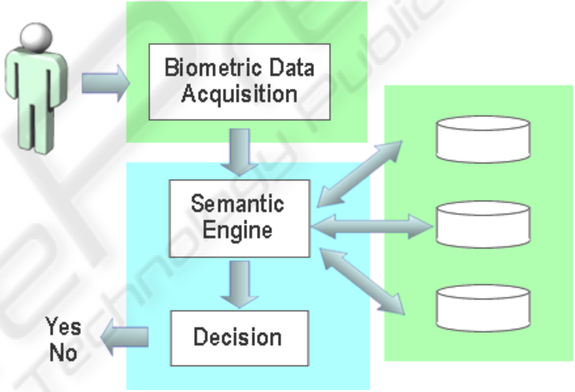

Figure 2: The framework.

The first part of the aforementioned steps i.e. the capture of the biometric samples and certain steps of the signal-process is addressed by the Biometric Data Acquisition System. The goal of this system is to encapsulate the biometric measure with a “data cover” that includes the type of the measure (voice, image, fingerprint, etc.) and a number of significant attributes that can describe the measure and the identity claim.

The second one, the Semantic Service Engine, takes the measure and its “data cover” and accesses the different data storages looking for the best fitting for the cover along the Web. A Semantic Service is an execution environment for the Semantic Web Services initiatives described in section 2. Along with some of the W3C submissions, different tools have been implemented that bring together the major Semantic Web Services functionalities in an integrated framework. One example is WSMX (Web Services Execution Environment), an execution environment to perform dynamic discovery, selection, mediation, invocation and inter-operation of Semantic Web Services. Another example is IRS (the Internet Reasoning Service), a Semantic Web Services framework which allows applications to semantically describe and execute Web Services. The IRS system supports the provision of semantic reasoning services within the context of the Semantic Web.

In both cases, a Semantic Service Engine would take the measure and its "data cover" to hook it up with the Web Service that fits the most and behaves as an entry point of the aforementioned databases, as it is show in Figure 1.

Finally, the Decision Module component produces a typical index called matching score. A match or no-match decision can be made according to whether this score exceeds a decision threshold or not. This implies an attempt to validate or not the claim of identity, which outcomes a final decision about the biometric identity.

6 CONCLUSIONS AND RELATED WORK

Since Biometric technologies are intended to tackle security and privacy issues, the integration of access control mechanisms and information security are also areas of growing interest.

The creation of Biometric Accreditation Entities will be an alternative in the near future to the current digital certification organisms.

In this environment the heterogeneity of sample capture and data process should not become a barrier for the use of this identification technology.

As the use of Semantic Web Services grows, the problem for searching, interacting and integrating relevant services is becoming increasingly a hurdle for the leverage of existing Semantic Web technologies which have reached a certain level of maturity.

In this paper, we have proposed a conceptual approach for a effective solution in this heterogeneous environment. It is based on the application of the Semantic Web Services properties. It requires to add semantic to the data stored in the biometric accreditation entities databases, and provide the adequate services to the SWS servers.

It has already been proposed the idea of using agents. They can take advantage of the machine-processable metadata provided by the Semantic Web Services. In (Hendler, J., 2001), the author points out how the ontology languages of the Semantic Web can lead to more powerful agent-based approaches for using services offered on the Web. A more practical approach is shown in (Gandon, F. and Sadeh, N., 2004), where the authors describe an application where intelligent agents, aided by context information provided by Semantic Web Services, assist their users with different sets of tasks.

Finally, our future work will focus on creating a complete adapted ontology and to define a standard for required services on SWS, identifying real–world scenarios and validating the efficiency of our approach and to determine its feasibility. This work is related to existing efforts about ontology merging and alignment. A future version of our framework will be orientated towards that direction.

ACKNOWLEDGEMENTS

This work is founded by the Ministry of Science and Technology of Spain under the PIBES project of the Spanish Committee of Education & Science (TEC2006-12365-C02-01).

REFERENCES

Bailly-Baillière E., S. Bengio et al., 2003. "The BANCA Database and Evaluation Protocol," in Springer LNCS-2688, 4th Int. Conf. Audio- and Video-Based Biometric Person Authentication, AVBPA'03. 2003, Springer-Verlag.

Berners-Lee, T., Hendler, J., Lassila, O., 2001. The Semantic Web. Scientific American, May 2001, pp. 34-43.

Borst, W.N., 1997. Construction of Engineering Ontologies for Knowledge Sharing and Reuse. PhD Thesis. University of Twente. Enschede, The Netherlands.

Fensel, D. & Bussler, C., 2002. The Web Service Modeling Framework WSMF. Electronic Commerce Research and Applications, 1(2).

Gruber, T. R., 1993. A translation approach to portable ontology specifications. Knowledge Acquisition Vol. 5:199-220.

Web Ontology Working Group, 2004. OWL Web Ontology Language Guide.

OWL-S W3C Submission, 2004. OWL Web Ontology Language for Services. Available at: http://www.w3.org/Submission/2004/07/

WSMO W3C Submission, 2005. Web Service Modeling Ontology. Available at: http://www.w3.org/Submission/2005/06/

SWSF W3C Submission, 2005. Semantic Web Service Framework. Available at: http://www.w3.org/ Submission/2005/07/

WSDL-S W3C Submission, 2005. Web Service Semantics. Available at: http://www.w3.org/ Submission/2005/10/

Hendler, J., 2001. Agents and the Semantic Web. IEEE Intelligent Systems, 16(2): 30-37, March/April 2001.

Gómez, J. M., Rico-Almodóvar, M., García-Sánchez, F., Martínez-Bejar, R. & Bussler, C., 2004. GODO: Goal-driven Orchestration for Semantic Web Services. WSMO Implementation Workshop, September 2004.

Jain, K. Bolle, R. et al. Biometrics, 1999: Personal Identification in Networked Society. Kulwer Academic. 1999.

Jain R. and J. Quian, 2001: Information Fusion in Biometrics. Proc. 3rd International Conference on Audio and Video Based Person Authentication (AVBPA) pp. 354-391,Sweden, 2001.

Gibbins, N., Harris, S., Shadbolt, N., 2003. Agent-based Semantic Web Services. In Proc. of the 12th Int. World Wide Web Conference, May 2003.

Gandon, F. and Sadeh, N., 2004. Semantic Web Technologies to Reconcile Privacy and Context Awareness. Web Semantics Journal, 1(3), 2004.

Ming, A. Ma, H., 2007. An Algorithm Tested for the Biometrics Grid. Proceedings of the Second International Conference in Grid and Pervasive Computing (GPC07). Paris, France. 2007.

Mansfield, J. Wayman, J.L., 2002. Best Practices in Testing and Reporting Performance of Biometric Devices. National Physics Lab for Mathematics and Scientific Computing. 2002.

Kittler, J. Hatef, R. Matas, J. G., 1998. On Combining Classifiers. IEEE Transactions on PAMI, vol. 12 (1998). Pp. 226-339.

SciTePress
Science and Technology Publications