IC-Service: A Service-Oriented Approach to the Development of Recommendation Systems

Aliaksandr Birukou, Enrico Blanzieri, Vincenzo D'Andrea,
Paolo Giorgini, Natallia Kokash, Alessio Modena
Department of Information and Communication Technology, University of Trento - Italy
{birukou, blanzier, dandrea, pgiorgio, kokash, modena}@dit.unitn.it

ABSTRACT

Recommendation systems have proven to be useful in various application domains. However, current solutions are usually ad-hoc systems which are tightly-coupled with the application domain. We present the *IC-Service*, a recommendation service that can be included in any system in a loosely coupled way. The implementation follows the principles of service oriented computing and provides a solution to various problems arising in recommendation systems, e.g. to the problem of meta-recommendation systems development. Moreover, when properly configured, the *IC-Service* can be used by different applications (clients), and several independent instances of the *IC-Service* can collaborate to produce better recommendations. Service architecture and communication protocols are presented. The paper describes also ongoing work and applications based on the *IC-Service*.

Keywords

recommendation systems, service-oriented design, application, web information systems

1. INTRODUCTION

Recommendation systems have recently become a powerful means to help users in knowledge sharing. They have multiple applications from traditional information to more recent e-commerce systems, where they try to predict users' preferences and prune large information spaces in searching for items of interest. Recommendation domains include, but are not limited to, movies (MovieLens[1]), music (JUKE-BOX [14]), books (Amazon, [6]), web links (Implicit, [2]), and hotels (TripAdvisor[2]). eBay provides recommendations on items similar to those the user bid but have not bought. However, current solutions are usually tightly-coupled with

---

[1]MovieLens - movie recommendations. `http://movielens.umn.edu/`

[2]TripAdvisor. Reviews of hotels, resorts and vacations. `http://www.tripadvisor.com/`

the applications and only in rare cases they are *domain-independent*. This limits their inclusion in different applications and thus their reuse. Meta-recommendation systems [12] have been introduced to overcome these limitations. These systems produce recommendations processing data from multiple information sources. So far there are no tools that facilitate the development of such systems.

Interoperability is a fundamental quality for information systems aiming at enabling large-scale data exchange. Behind technical issues, interoperability includes also cultural and organizational aspects. Being interoperable means being able to manage the culture of an organization and consequently maximizing the opportunities of information sharing. This is the underpinning principle of the *Implicit Culture* theory [3], a new paradigm in the area of recommendation system development. The Implicit Culture approach aims at providing users with suggestions based on behavioral patterns extracted from users' actions. In other words, recommendations are inferred from observations of actions that a group of agents perform on a set of objects under certain conditions.

We introduce the *IC-Service*, a multi-purpose web-based recommendation service. This service provides simple and configurable access to the *System for Implicit Culture Support (SICS)* [3], which is implemented adopting the Service Oriented Architecture (SOA). Therefore, it can be easily added to existing software in any domain described in terms of *agents*, *objects*, *actions* and *attributes*. This allows for the development of complex applications that include recommendation tools using the idea of software *composability* instead of building them from scratch. The use of SOA provides technical *interoperability*, and guarantees seamless communication of various agents around the world. An instance of the *IC-Service* can be configured to be accessed by different applications. Several *IC-Services* can collaborate to facilitate knowledge sharing (e.g. between communities). Flexible configuration mechanism allows for the highest level of *reusability* providing a rapid and cheap way to embed the recommendation service in any system. To the best of our knowledge, the *IC-Service* is the unique service oriented solution to the problem of providing multi-purpose and domain-independent recommendations, which occurs, for example, in meta-recommendation systems. The *IC-Service* is a general solution that can be easily deployed both in applications under development and in existing systems, in particular, in complex software applications where user personalization and item/action recommendation have been excluded from requirements specification because of
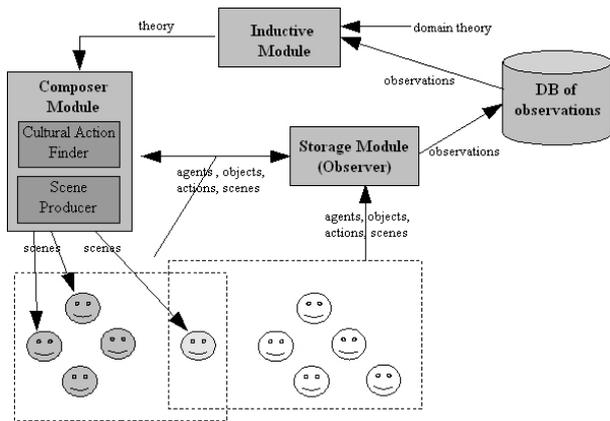
Figure 1: The general architecture of the System for Implicit Culture Support

additional costs. The *IC-Service* can be also applied to the problem of cross-selling recommendations [11], where the data about the customer come from different retailers and must be integrated.

The rest of the paper is organized as follows. Section 2 describes briefly the Implicit Culture theory. Section 3 illustrates the architecture of the SICS, whereas some application perspectives are discussed in Section 4 and 5. Section 6 gives some concluding remarks and outlines future work.

2. IMPLICIT CULTURE

When a person has to act in an unknown social environment his/her behavior is far from the optimal. We can think of many situations where, due to the lack of knowledge, it becomes very difficult for the person to take the right decision. This might not be the case for people that have previously faced similar situations. Indeed, they may have acquired the necessary knowledge to act effectively in the environment. This knowledge is usually implicit and it represents a sort of "community culture".

Implicit Culture is based on the assumption that it is possible to elicit the community culture by observing the interactions of people with the environment and to encourage the newcomer(s) to behave similarly to more experienced people. Implicit Culture assumes that *agents* perform *actions* on *objects* in the *environment* (see [3] for more details). The actions are considered in the context of *situations*, so agents perform *situated actions*. The "culture" contains information about actions and their relation to situations, namely which actions are usually taken by the observed group and in which situations. This information is then used to provide newcomers with information about the others' behavior in similar situations. When newcomers start to behave similarly to the community culture, it means that we have a knowledge transfer. "Implicit Culture is a relation between a set and a group of agents such that the elements of the set behave according to the culture of the group" [3]. This transfer of knowledge is performed by a SICS [3].

The general architecture of the SICS is depicted in Figure 1. The SICS consists of three components: the *Observer*, which uses a database of observations to store information about actions performed by users in different situations; the *Inductive Module*, which analyzes the stored observations

and applies data mining techniques to find a *theory* about the community culture; the *Composer Module*, which exploits the observations and the theory in order to suggest actions in a given situation. The Composer Module operates with the architectural abstraction of the situation — the *scene*, which contains a set of objects and a set of actions that can be performed on these objects. The Composer Module includes the following two sub-modules: the *Cultural Action Finder*(CAF), which filters those actions that satisfy the theory; and the *Scenes Producer*, which searches for scenes where the actions (found by the CAF) are likely to be performed.

Figure 1 shows that the *theory* (representing the "community culture") is extracted from the *observations* on the set of agents and is then used to produce recommendations for another set of agents. These two sets may be disjoint, may overlap or coincide. The *theory* is rule-based and contains two parts. A part of the theory used by the Composer Module can be specified a priori (*domain theory*), while the other part is learnt by the Inductive Module and can evolve over time. For instance, for the general problem of providing people with recommendations the domain theory may say that the system must recommend items which are likely to be accepted by the user, while the theory learnt by the Inductive Module may contain information about which items are accepted.

As an example of Implicit Culture, let us consider a person who would like to use a book-selling service, but does not know which service is used by other people in the same location. Obviously, people who used to sell books know the name of the service and it may be the case that they have tried several services before choosing the best one. If the system is able to use previous history to suggest that the person access the service used by others and he/she actually does it, then it is possible to say that he/she behaves in accordance with the community culture and that the Implicit Culture relation is established.

3. THE IC-SERVICE ARCHITECTURE

This section contains the details of the SICS implementation used in the *IC-Service* and justifies why particular tools and architectures have been adopted.

3.1 Implementation

The *IC-Service* is the remote part of the SICS which provides the recommendation service. The SICS architecture consists of three main layers (Figure 2):

- The *SICS Core* provides the implementation of the Implicit Culture approach. This layer is responsible for storing observations, managing theory and facilitating actions by suggesting scenes. In particular, this layer implements the *Composer Module* and the *Inductive Module* functionality.

- The *SICS Remote Module* defines protocols for information exchange with the client and converts the objects of the SICS Core in the format compatible with these protocols.

- The *SICS Remote Client* provides a simpler interface for the remote clients. It presents a wrapper that hides protocols used for information exchange.
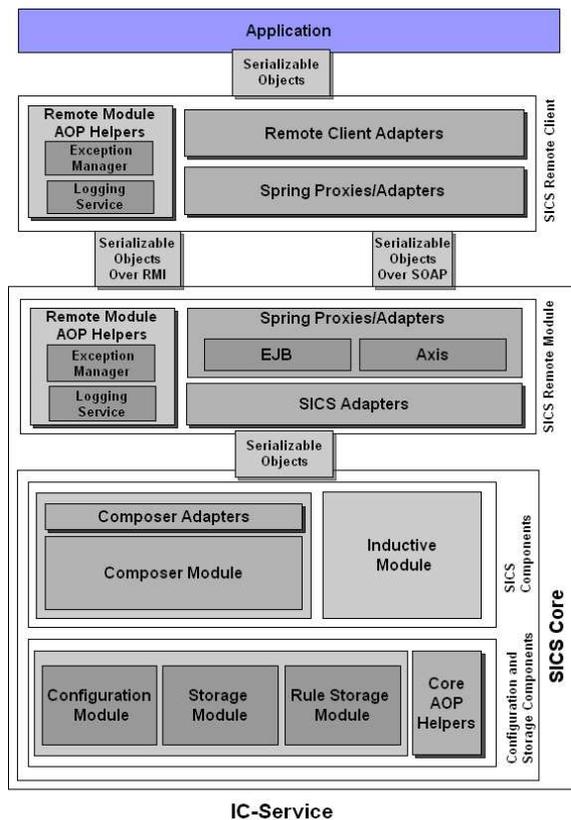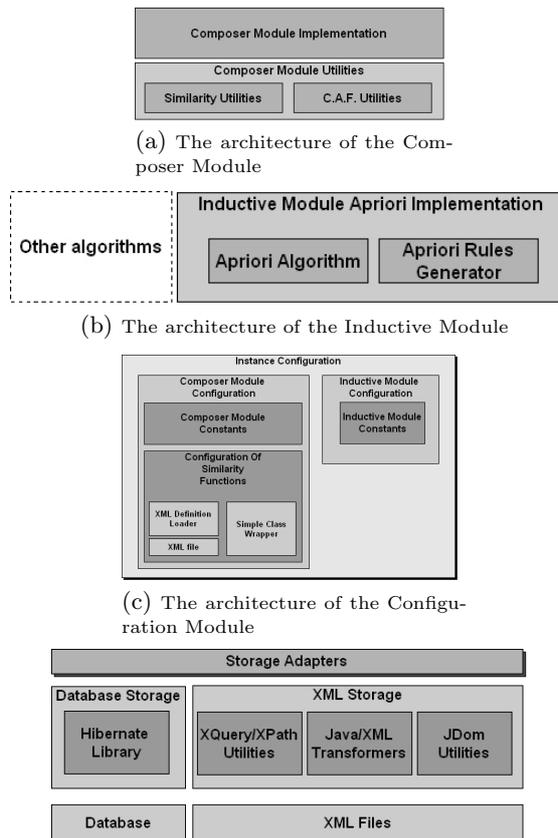
Figure 2: The detailed SICS architecture

(a) The architecture of the Composer Module

(b) The architecture of the Inductive Module

(c) The architecture of the Configuration Module

(d) The architecture of the Storage Module

Figure 3: The detailed architecture of SICS modules

The SICS Remote Client in Figure 2 is composed of *Remote Client Adapters*, *Spring*[3] *Proxies/Adapters*, and *Aspect-Oriented Programming tools (AOP Helpers)*. *Remote Client Adapters* are responsible for the asynchronous invocation of the SICS Remote Module. *Spring Proxies/Adapters* provide the connection to the SICS Remote Module via SOAP (Simple Object Access Protocol) or RMI(Remote Method Invocation). *AOP Helpers* provide logging, validation and exception management. SOAP[4] is a lightweight XML-based protocol for exchanging information in a distributed environment. It consists of an envelope that defines a template for describing the message contents and the way to process it, a set of encoding rules for expressing instances of application-defined datatypes, and a convention for representing remote procedure calls and responses.

The SICS Remote Module in Figure 2 includes the *Spring Proxies/ Adapters* for the remote invocation of the Composer Module using SOAP or RMI. *Apache Axis* is used in addition to the Spring framework that allows the Composer Module to be available as a SOAP web service. *EJB (Enterprise JavaBeans)* part of the Remote Module extends integration classes of the Spring framework to allow for the use of the Composer Module as an EJB component in J2EE environment. *SICS Adapters* provide the connection between the SICS Remote Module and the SICS Core. Finally, *AOP Helpers* deal with logging, validation and exception management.

---

[3]http://www.springframework.org
[4]http://www.w3.org/TR/soap/

Let us describe the components of the SICS Core (2) in detail. **Composer Adapters** are auxiliary modules, in particular, responsible for the asynchronous execution of the Composer services and cache management. The architecture of the **Composer Module** is shown in Figure 3(a). Besides the main functionality of providing recommendations, it contains *Similarity Utilities*, which implement the algorithms for calculating the similarity between objects, actions, etc., and *CAF Utilities* used by the Cultural Action Finder.

To discover theory about users' behavior, the **Inductive Module** incorporates the implementation of the *Apriori Algorithm* for the association rule mining [1] and its extension for generating rules in the *Apriori Rule Generator* (Figure 3(b)). With the dashed line it is shown that the functionality of the module can be extended with the implementation of other learning techniques.

All parameters of a SICS instance are setup in the **Configuration Module** (Figure 3(c)). Each instance of the SICS can have different configurations of the Composer Module (*Composer Module Constants*), the mechanism of processing the theory in the Inductive Module (*Inductive Module Constants*), and parameters of the algorithm for calculating similarity between elements such as objects, actions, etc. (*Configuration of Similarity Functions*). The following two modules are responsible for the configuration of a SICS instance: the *XML Definition Loader*, which loads the configuration of the similarity algorithm from the corresponding

XML file; and the *Simple Class Wrapper*, which loads the configuration of the similarity algorithm from the hierarchy of classes used by the Spring framework.

The architecture of the **Storage Module** is depicted in Figure 3(d). The Storage Module is responsible for storing information about the application domain, i.e., adding or deletion of agents, managing groups, and saving observations. The SICS can be configured to use either of two modules to store data: the *Database Storage Module* is responsible for the management of database storage whereas the *XML Storage Module* stores the information in XML files. *Storage Adapters* provide asynchronous execution of methods of the Storage Module and cache management. A powerful high performance query service for database storage is provided by the *Hibernate*[5] library. The Storage Module also includes a set of tools to work with an XML representation of the SICS information: *XQuery/XPath Utilities* are used to read data from an XML repository, *Java/XML Transformers* convert SICS objects into XML format and *JDom Utilities* deal with editing of XML files. The **Rule Storage Module**, which is responsible for the management of the theory (adding or removal of rules), is organized in a similar way. As opposed to the Storage Module, it supports only XML storage facilities. **Core AOP Helpers** provide logging, validation and exception management.

3.2 Usage Scenarios

The *IC-Service* can be used within an application in three different ways (Figure 4): (i) SICS can be included in the application as a library (Figure 4(a)). In this case the *SICS Core* deals directly with the objects, actions, etc. of the applications. This way should be chosen when the application is not necessarily distributed and can be tightly-coupled with the library. (ii) To enable remote access (Figure 4(b)), the SICS core can be invoked via the *SICS Remote Module* as a SOAP web service or EJB component (using SOAP/RMI). This scenario should be followed when the service is a part of a distributed system, but for some reasons (e.g. limited resources of the client, such as in portable devices) there is no need or opportunity for using the *SICS Remote Client*. However, in this case the application must take care of communicating with the service. (iii) The easiest way to add recommendation service in an application is to access the *IC-Service* via the *SICS Remote Client* (Figure 4(c)) that hides the technical details of the communication mechanism from the application designer. This way should be adopted when we deal with complex applications and the *IC-Service* must be introduced in a fully decoupled way.

The described scenarios illustrate the possibility of including the *IC-Service* in various applications ranging from small-size applications to complex distributed systems. The *IC-Service* is developed with JAX-RPC (Java API for XML-based Remote Procedure Calls[6]), a programming model that enables invocation of web services across heterogeneous platforms. The SICS modules are built using the Spring framework, which allows assembling of loosely-coupled components in a complex system via XML configuration files. All modules apart from the Storage Module and the Rule Storage Module communicate through Java function calls and serializable objects. By avoiding Java collections, the easier interoperability with SOAP is enabled.

---

[5]http://www.hibernate.org/

[6]http://java.sun.com/webservices/jaxrpc/

(a) SICS as a library

(c) SICS as a service via SICS Remote Client
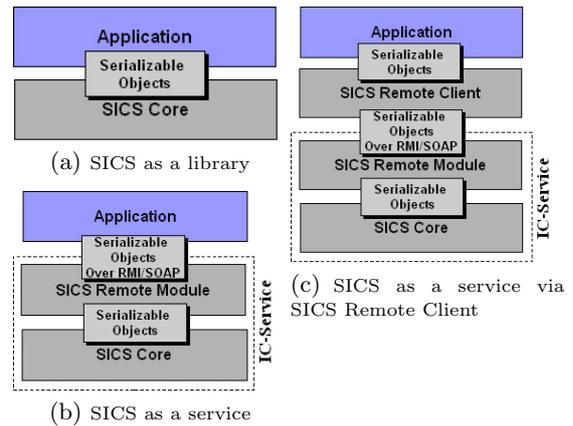
(b) SICS as a service

Figure 4: SICS invocation scenarios

SOA has been chosen among the possible architectures because it supports principles of universal access and platform independence and allows recommendation service to be transparently located inside or outside the enterprise. Support of EJB technology simplifies the use of the *IC-Service* in applications developed with Java technology. The *Storage Module* supports two possible storage facilities: XML files and the database storage. XML files provide a simple, easily deployable, and portable solution for applications where the observation history is not big and must not be accessed frequently. The database variant should be chosen with more complex applications involving heavy data processing.

The *IC-Service* can be added in an application in a fully *decoupled* way, and accessed from anywhere at any time. This guarantees *ubiquity*, allowing the system to produce sound recommendations using data collected from different sources. For instance, ubiquity is very useful in the problem of providing cross-selling recommendations. Several communicating *IC-Services* can be seen as building blocks in the development of an efficient and robust decentralized recommendation system. At the same time, the *IC-Service* is a general-purpose and domain-independent application that provides means for storing, analyzing and reasoning about the observed behavior. It presents a higher *granularity* than specialized recommendation modules. Once deployed, the *IC-Service* can be used by several applications. Changes and extensions can be smoothly embedded in the working system by modifying XML-based domain description or the theory. This leads to minimizing efforts on development and reducing overheads on support of heterogeneous systems.

4. IC-SERVICE-BASED APPLICATIONS

In this section we briefly describe ongoing projects which use the *IC-Service*: QUIEW, IC-SWSD, and the system that supports the work of biologists.

The QUIEW project[7], lead by ITC-IRST, aims at providing methods for an appropriate ordering of the Web content according to a list of categories, which represent topics of interests. The relevance feedback is used to provide more effective organization of the information. In the context of this project, the *IC-Service* provides recommendation on

---

[7]QUIEW. Quality-based Indexing of Web Information. http://quiew.itc.it/

web documents classification. In particular, the knowledge engineer is offered a set of potentially relevant categories for a given document. These recommendations are based on the history of previous interactions of users with the system. The QUIEW application accesses the *IC-Service* as a web service via the SICS Remote Client and the SICS Remote Module. The SOAP protocol is used for the information exchange between the SICS Remote Client and the SICS Remote Module. The Storage Module uses XML files to store the data.

The second application is a framework for supporting web service discovery, IC-SWSD (Implicit Culture Support for Web Service Discovery). In this scenario, users are interested in finding services that can provide a predefined functionality and guarantee a certain level of quality. The discovery process consists of two steps: (1) *matching*, i.e., meeting the functionality required by a user with specifications of existing services, and (2) *selection*, i.e., choosing a service with the best quality among those able to satisfy a user's goal. Service *clients* are software applications that rely on external services to fulfill some intervening tasks. The results of the operations produced by external services are either verified automatically or analyzed by a human. This evaluation can be used to augment the knowledge about the existing services and their features that may not be explicitly stated in service documentation and interface description (even with ontology-based semantic extensions). We suppose that virtual *communities* of the users with similar interests exist, and their members will benefit from recommendations produced by the *IC-Services* based on the observations of the actions of other community members.

In our preliminary tests, we extended the Apache Axis framework with the ability of monitoring service invocations using the *IC-Service*. For each invocation we save the identifier of the client with attributes (e.g., user name and location), the identifier of the web service with attributes (e.g., business category), name of the invoked operation with input and output parameters, time of the invocation and service response time. This allows for taking into account parameters such as average response time, throughput, success rate, etc. The client can add other important information such as report about cases of contract violation and store domain-specific parameters like book prices for book-selling services. Quality of Service (QoS) ontologies can be involved to enable formal knowledge-based reasoning [8]. The collected data are used to map the needs of new clients with the services that might satisfy those needs.

In case of several *IC-Services*, they collaborate in order to recommend more suitable web services in the context of a certain community. Thus, search results are complemented with the recommendations produced by the system. The basic data transfer process is shown in Figure 5. Community members develop applications that use external web services. The communication between these applications and web services is monitored using the *IC-Service*. When a new user submits a query, the *IC-Service* matches the user's request with the specifications of the registered services, analyzes the monitored data, communicate with other *IC-Services* and selects a set of potentially useful services with the best quality regarding the user's personal preferences. In addition, the presented framework can be used for extraction of interaction patterns and distributed semantic caching for web services [13].
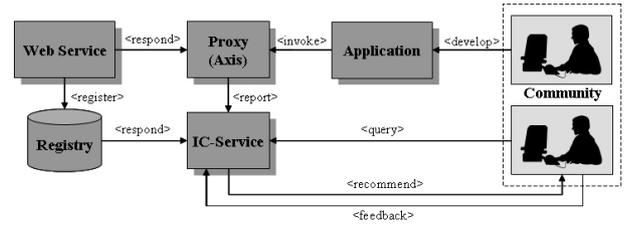

Figure 5: The application of the *IC-Service* to web service discovery

We are currently working on the development of a system based on the *IC-Service* to support the work of biologists in their laboratories, adopting the approach presented in [10]. It is aimed at assisting the scientists during Polymerase Chain Reaction (PCR) experiments [9]. The PCR, being an important part of many pieces of the state-of-the-art research, is still capricious and problematic procedure and each laboratory has its own PCR specifics. The system being created will help inexperienced biologists or laboratory newcomers to increase success/failure rate of the experiments by providing them with the suggestions induced from observed actions of their more experienced colleagues.

In all above-mentioned applications, the use of the *IC-Service* has shown the following advantages: (i) the service can be accessed from any workplace, enabling distributed collection of observations; (ii) several clients can use the same service that adapts for their needs; (iii) it is a unique solution for highly distributed communities; (iv) the knowledge transfer within/between communities is facilitated. Moreover, in the context of web service discovery, it is one of the first implementations of a system that recommends web services. The need for such kind of systems has been announced in [8]. The application of collaborative filtering to web service selection has been previously proposed by Manikrao and Prabhakar [7].

We performed preliminary performance tests of the *IC-Service*, which have shown that a reasonable response time can be achieved even in case of using XML storage with more than 10,000 observations and having about 100 clients continuously querying the system.

The experience of these projects in using the IC-Service is briefly summarized in Section 5.

5. LESSONS LEARNT

Web service technology simplifies the development of recommendation systems and allows for the integration of the recommendation service to existing systems. However, there are several open questions regarding the design of services to be used as *long-lived* loosely-coupled components of distributed systems. What makes the *IC-Service* different from standard information services such as book-selling service is that it (i) is oriented on the use in various application domains, (ii) processes client data according to the rules defined for a particular application domain, (iii) supports storage of potentially huge amount of clients' data, (iv) analyzes the collected information in order to adapt the provided functionality to the needs of a particular client. The principles underlying the design of such services are not well-established yet. Curbera et al. [4] describe *customization* of SOA components as one of the key characteristics. They ar-

gue that "a SOA programming model should enable building services and modules that programmers can customize without source code modification". Indeed, it is unlikely that a service can be reused by different applications without reconfiguration. For its nature, the *IC-Service* has a direct dependence on the context of application and must be customizable. Therefore, configurability and extensibility without code modification were the main focus of the design process. To reach the necessary properties such as adequate level of granularity, flexible configuration mechanism, powerful storage and data management facilities, etc., we used state-of-the-art tools and solutions, namely, the combination of the original Implicit Culture theory with design patterns ("Adapter", "Proxy", "Facade", "Abstract Factory", "Factory Method", etc.) [5], Aspect-Oriented Programming and auxiliary frameworks such as Spring and its principle of "designing to interfaces".

Multilevel organization of features and support of both XML and database storages are involved to satisfy the *portability* and *scalability* requirements. XML storage format imposes restrictions on the number of observations that can be stored. These restrictions can be overcome using database storage or deploying several instances of the *IC-Service*. To increase the performance, operations responsible for storing observations run in separate threads or JMSs (Java Message Services) under J2EE environment. Independent and configurable cache[8] is used at each functional level.

6. CONCLUSION AND FUTURE WORK

In this paper a service oriented architecture for the development of recommendation systems has been proposed. We presented the *IC-Service*, a general-purpose web service that uses the ideas of the Implicit Culture theory to produce recommendations. Our application has a tangible motivation for SOA that provides a way to increase the level of organization and management of systems embedding recommendation services and supporting autonomous members of communities.

We believe that the synergetic compatibility between SOA and SICS can guarantee success of the SICS in many application domains. Along with the uniform mechanisms to store and retrieve observations, analyze actions, extract behavioral patterns and produce recommendations, our approach becomes a cheap and transparent solution in the area of recommendation systems.

As future work, we would like to develop a wizard to configure and deploy *IC-Services*. In addition, we are going to build networks of several *IC-Services* and evaluate mechanisms for combining their recommendations.

7. ACKNOWLEDGEMENTS

---

[8]http://ehcache.sourceforge.net/

8. REFERENCES

[1] R. Agrawal and R. Srikant. Fast algorithms for mining association rules in large databases. In *VLDB: Proc. of the 20th Int. Conference on Very Large Data Bases*, pages 487–499. Morgan Kaufmann, 1994.

[2] A. Birukov, E. Blanzieri, and P. Giorgini. Implicit: An agent-based recommendation system for web search. In *AAMAS: Proc. of the 4th Int. Joint Conference on Autonomous Agents and Multiagent Systems*, pages 618–624. ACM Press, 2005.

[3] E. Blanzieri, P. Giorgini, P. Massa, and S. Recla. Implicit culture for multi-agent interaction support. In *CoopIS: Proc. of the 9th Int. Conference on Cooperative Information Systems*, volume 2172 of *LNCS*, pages 27–39. Springer, 2001.

[4] F. Curbera, D. F. Ferguson, M. Nally, and M. L. Stockton. Toward a programming model for service-oriented computing. In *ICSOC: Proc. of the 3d Int. Conference on Service-Oriented Computing*, volume 2172 of *LNCS*, pages 33–47. Springer, 2005.

[5] E. Gamma, R. Helm, R. Johnson, and J. Vlissides. *Design patterns: elements of reusable object-oriented software.* Addison-Wesley, Boston, MA, USA, 1995.

[6] G. Linden, B. Smith, and J. York. Amazon.com recommendations: Item-to-item collaborative filtering. *IEEE Internet Computing*, 7(1):76–80, 2003.

[7] U. S. Manikrao and T.V.Prabhakar. Dynamic selection ofweb services with recommendation system. In *NWESP '05: Proceedings of the International Conference on Next Generation Web Services Practices*, page 117, Washington, DC, USA, 2005. IEEE Computer Society.

[8] E. M. Maximilien and M. P. Singh. A framework and ontology for dynamic web services selection. *IEEE Internet Computing*, 8(5):84–93, 2004.

[9] K. B. Mullis, F. A. Faloona, S. Scharf, R. K. Saiki, G. Horn, and H. A. Erlich. Specific enzymatic amplification of dna in vitro: the polymerase chain reaction. In *Cold Spring Harbor Symposia on Quantitative Biology*, volume 51, pages 263–273, 1986.

[10] M. Sarini, E. Blanzieri, P. Giorgini, and C. Moser. From actions to suggestions: supporting the work of biologists through laboratory notebooks. In *COOP: Proc. of 6th Int. Conference on the Design of Cooperative Systems*, pages 131–146. IOSPress, 2004.

[11] J. B. Schafer, J. A. Konstan, and J. Riedl. E-commerce recommendation applications. *Data Mining and Knowledge Discovery*, 5(1-2):115–153, 2001.

[12] J. B. Schafer, J. A. Konstan, and J. Riedl. Meta-recommendation systems: user-controlled integration of diverse recommendations. In *Proc. of the 11th Int. Conference on Information and Knowledge Management*, pages 43–51. ACM Press, 2002.

[13] S. Seltzsam, R. Holzhauser, and A. Kemper. Semantic cashing for web services. In *ICSOC: Proc. of the 3d Int. Conference on Service-Oriented Computing*, volume 3826 of *LNCS*, pages 324–340. Springer, 2005.

[14] H. Tremblay-Beaumont and E. Aïmeur. Jukeblog : A recommender system in the music weblogs. In *Proc. of the IADIS Int. Conference on e-Commerce*, pages 274–280, December 2005.