Privacy Preserving Composition of Government Web Services

Athman Bouguettaya Brahim Medjahed Abdelmounaam Rezgui
Mourad Ouzzani Zhaoyang Wen

Department of Computer Science, Virginia Tech
7054 Haycock Road, Falls Church, VA 22043
{athman,brahim,rezgui,mourad,zwen}@vt.edu

Project URL: http://www.nvc.cs.vt.edu/~dgov

1 Introduction

The core of our research in digital government is the development of techniques to efficiently access government services and databases. As a case study, we use social and welfare services within *Indiana*'s *Family and Social Services Administration* (FSSA). The FSSA is composed of dozens of autonomous and geographically distant departments. Each department provides several rehabilitation programs to help disadvantaged citizens. It also contains a myriad of databases that store government and citizens information. We propose a framework that enables uniform access to a large number of government services and databases. A Web-based prototype called *WebDG* (*Web Digital Government*) has been implemented. The main features of *WebDG* are the following:

1. **Querying e-government databases**: The large number of FSSA databases makes it difficult to query the available information space if an efficient infrastructure is not available. To deal with this problem, we organize FSSA databases as *distributed ontologies*. An *ontology* defines a taxonomy based on the semantic proximity of information interest [Bouguettaya et al., 2001]. Each ontology contains databases that share the same domain of interest (e.g., pregnancy). The use of distributed ontologies accelerates the discovery of FSSA databases.

2. **Organizing e-government services**: To satisfy citizens' needs, FSSA case officers must manually execute several applications (e.g., *Indiana Client Eligibility System*). Since these applications exist in large numbers and are distributed over different FSSA departments, locating the ones that best fit citizens' needs is usually a tedious, frustrating, and cumbersome task. To tackle this problem, we wrap FSSA applications by *Web services*.A *Web service* is a functionality that is programmatically accessible via the Web [Medjahed et al., 2002]. The use of Web services caters for the dynamic *discovery* and *composition* of welfare programs.

3. **Composing e-government services**: FSSA case officers deal with different situations that depend on the particular needs of each citizen (health, children, etc). For each situation, they must determine those services that best meet citizen's needs. Then, they need to figure out how to access and invoke each service. Finally, they must combine the results returned by the different services. To deal with this problem, we propose a *declarative* framework for defining *composite services*. A *composite service* aggregates multiple e-government services to provide a *value added* service [Benatallah et al., 2000]. The use of composite services provides one-stop social services that outsource from a variety of services located in geographically distant bureaus.

4. **Preserving citizens' privacy**: Privacy is one of the major issues that may impede the wide public acceptance of digital government. To preserve citizens' privacy, we propose a model based on the citizen's privacy preference [Rezgui et al., 2002]. Users (citizens, case officers, etc.) are granted *privacy*

credentials that define their access scope (what information they can access, when, etc). A request for a citizen's private information goes through a *filter agent* that determines the *valid* access scope for that request.

In this demo, we showcase the main functionalities of the *WebDG* prototype. We mainly focus on service organization, service composition, and privacy preservation. Details about querying e-government databases can be found in [Bouguettaya et al., 2001]. We present a scenario to illustrate how e-government services are composed and citizens' privacy preserved. Detailed descriptions about the proposed approach are presented in [Medjahed et al., 2002, Rezgui et al., 2002].

2 Implementation

WebDG prototype uses emerging Web service standards. These include (i) *WSDL* (*Web Service Description Language*) [WSDL, 2002] for describing operational features of e-government services (messages, operations, host and port numbers, etc), (ii) *UDDI* (*Universal Description, Discovery and Integration*) [UDDI, 2002], a programmatic interface for publishing and discovering services, and (iii) *SOAP* (*Simple Object Access Protocol*) [SOAP, 2002], a messaging framework for exchanging XML formatted data among services.

Without loss of generality, the current implementation includes ten (10) e-government services described using WSDL and *HP*'s *e-speak* [E-SPEAK, 2002]. These database-backed services are published in UDDI and *e-speak* registries. Government and citizens information is stored in *Informix* and *Oracle* databases. Table 1 summarizes the different services and databases implemented in *WebDG*.

E-government Service	Function	Description	Registry	Database
Women, Infant, Children (WIC)	Provides high-quality nutritional care and food to needy citizens	WSDL	UDDI	Oracle
Medicaid	Provides health care to low-income citizens	WSDL	UDDI	Oracle
Teen Outreach Pregnancy (TOP)	Provides childbirth and postpartum educational support to pregnant teens	WSDL	UDDI	Oracle
Temporary Assistance for Needy Families (TANF)	Provides cash assistance and supportive services to low-income families	WSDL	UDDI	Informix
Food Stamps (FS)	Supplementing low income households with food stamps	WSDL	UDDI	Informix
Blind Registry (BR)	Enables the registry of blind people	WSDL	UDDI	Informix
Family Participation Day (FPD)	Helps families of visually impaired citizens develop a realistic outlook towards blindness	E-speak	E-speak	Oracle
Communication Skills (CS)	Teaches communication techniques needed by a visually impaired person	E-speak	E-speak	Informix
Job Placement (JP)	Helps citizens find employment consistent with their disabilities	E-speak	E-speak	Informix
Independent Living (IL)	Maximizes the independence and integration of disabled citizens in community leadership	E-speak	E-speak	Oracle

Table 1: FSSA Services and Databases

3 System Demonstration

In the following, we present a scenario to illustrate the main features of *WebDG*. Consider a pregnant teen visiting a case officer at a local FSSA agency to collect social benefits to which she is entitled. The teen

needs to apply for a government funded health insurance program. She also needs to meet with a nutritionist to get help in maintaining an appropriate diet during her pregnancy. As the teen is not able to take care of the future newborn, she is interested in finding a foster family.

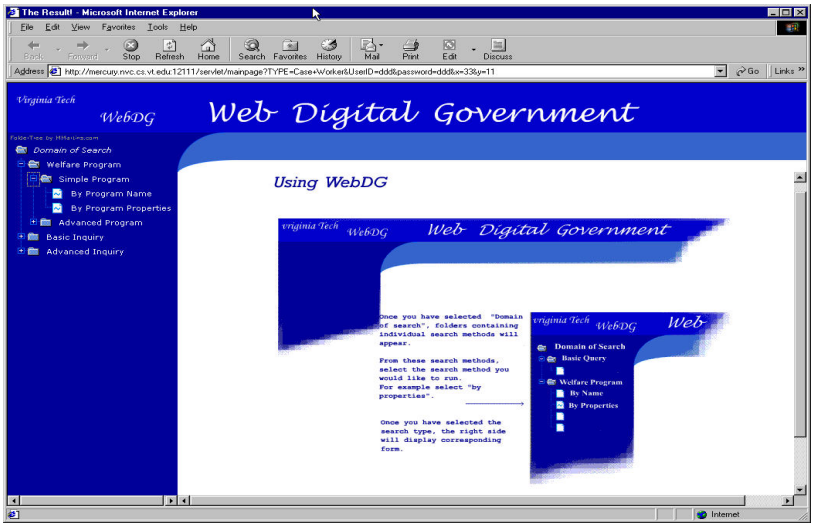

Figure 1: WebDG Graphical User Interface

The fulfillment of the teen's needs requires accessing different services scattered in and outside the local agency. For that purpose, the case officer accesses *WebDG* interface (Figure 1) to build a *Pregnancy Benefits* (*PB*) composite service based on the teen's needs. The major steps for defining *PB* service are as follows:

- **Step 1: Composite service specification** – The case officer enters a specification of the composite service using CSSL language [Medjahed et al., 2002]. The specification includes high level descriptions of the operations needed by the *PB* service. An example of operation is `insuranceEligibility` to check whether the teen is eligible for health coverage. The *type* and *category* of the operation are "eligibility" and "insurance" respectively. The case officer also defines the *input* and *output* parameters of the operations such as `monthlyIncome` and `eligibilityStatus`.

- **Step 2: Matchmaking with e-government services** – *PB* specification is handled by the *WebDG manager*. The *WebDG manager* implements a *matchmaking algorithm* [Medjahed et al., 2002]. The general premise of the algorithm is to "plug" each operation of *PB* with one or more e-government service operations. Since the category of `insuranceEligibility` is "insurance", the algorithm looks for e-government services , say *Medicaid*, whose category is "insurance". Then the algorithms looks for operations within *Medicaid* that are *compatible* with `insuranceEligibility`. For example, the operation `medicaidEligibility` of *Medicaid* is found compatible with `insuranceEligibility`.

- **Step 3: Generating composition plans** – The *WebDG manager* returns *composition plans* that conform to *PB* specification. By *composition plan* we refer to the list of component services and how they interact with each other. Figure 2 depicts an example of a composition plan. The operations offered by *PB* are outsourced from *Medicaid*, *WIC*, and *TOP* services. For example, the case officer would invoke `PB::registerClient` and `PB::insuranceEligibility` to provide health insurance coverage to the teen.

- **Step 4: Using the composite service** – Assume that the teen is now looking for a foster family for her future newborn. She would invoke `PB::getFamily` operation. This operation is "plugged into"

`searchFamily` provided by *TOP* service. It returns the family name, address, and size. Assume that the teen would also like to know the household income. She would ask the case officer to retrieve that information from *TOP* database.

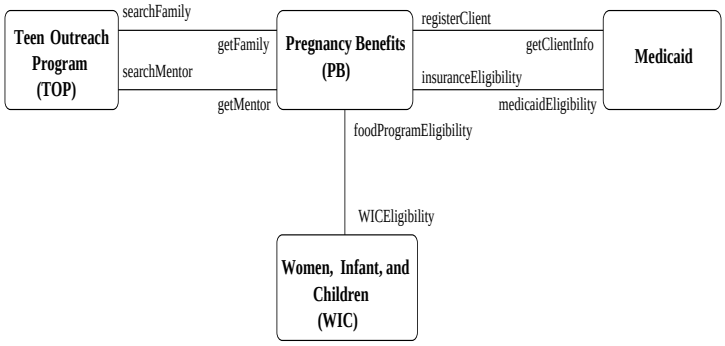

Figure 2: Composition Plan for the Pregnancy Benefits Service

- **Step 5: Preserving privacy** – The query is sent to *TOP* database along with the officer's *privacy credentials* [Rezgui et al., 2002]. A local *filtering agent* first determines whether the case officer is authorized to submit such query. Depending on whether the family is willing to release its income information, two cases are considered. If that information is kept private by the family, the filtering agent returns a warning that the data is not available. Otherwise, the query is processed and a result is returned.

Acknowledgment. This research is supported by the National Science Foundation grant 9983249-EIA and the Commonwealth Information Security Center (CISC). We would like to acknowledge the contribution of Hao Long and Hongmei Hao who participated in the implementation of the system.

References

[Benatallah et al., 2000] Benatallah, B., Medjahed, B., Bouguettaya, A., Elmagarmid, A., and Beard, J. (2000). Composing and Maintaining Web-based Virtual Enterprises. In *1st TES VLDB Workshop*, Cairo, Egypt.

[Bouguettaya et al., 2001] Bouguettaya, A., Elmagarmid, A., Medjahed, B., and Ouzzani, M. (2001). Ontology-Based Support for Digital Government. In *VLDB Conference*, Roma, Italy.

[E-SPEAK, 2002] E-SPEAK (2002). E-speak Developer Site, http://www.e-speak.net.

[Medjahed et al., 2002] Medjahed, B., Ouzzani, M., and Bouguettaya, A. (2002). Using Web Services in E-Government Applications. *Submitted to dg.o 2002 Conference (long paper).*

[Rezgui et al., 2002] Rezgui, A., Wen, Z., and Bouguettaya, A. (2002). Enforcing Privacy in Interoperable E-Government Applications. *Submitted to dg.o 2002 Conference (long paper).*

[SOAP, 2002] SOAP (2002). *Simple Object Access Protocol, http://www.w3.org/TR/soap.*

[UDDI, 2002] UDDI (2002). *Universal Description, Discovery, and Integration, http://www.uddi.org.*

[WSDL, 2002] WSDL (2002). Web Services Description Language, http://www.w3.org/TR/wsdl.