A Guanxi Shibboleth based Security Infrastructure for e-Social Science

Wei Jie ¹  Alistair Young²  Junaid Arshad³
June Finch¹  Rob Procter¹  Andy Turner³
¹ University of Manchester, UK
² UHI Millennium Institute, UK
³ University of Leeds, UK
wei.jie@manchester.ac.uk

Abstract

An e-Social Science infrastructure generally has security requirements to protect their restricted resources or services. As a widely accepted authentication and authorization technology, Shibboleth supports the sharing of resources on inter-institutional federation. Guanxi is an open source implementation of the Shibboleth protocol and architecture. In this paper, we propose a security infrastructure for e-social science based on the Guanxi Shibboleth. This security infrastructure presents two main features. Firstly, Guanxi Shibboleth is integrated into the user-friendly Sakai collaborative and learning environment which provides an ideal place for users to access a variety of federation resources in line with the Shibboleth authentication model. Secondly, PERMIS technology is used to enhance the authorization mechanisms thus enabling a policy-driven, role-based, fine-grained access control. As a result, the security infrastructure presents the advantages of Guanxi Shibboleth, PERMIS and Sakai, and it has been applied to e-Social Science application. We believe this security infrastructure provides a promising authentication and authorization solution for e-social science applications as well as applications in other domains.

1. Introduction

The e-Infrastructure for Social Science [1] is a project funded by UK Economic and Social Research Council (ESRC) which leverages Grid technology [2] to build an e-infrastructure to provide integrated access to a variety of resources of social science research, including datasets, tools, services and easy-to-use user environments. Security underpins the e-Infrastructure for the social sciences as it is one of the foundations of any working Grid infrastructure. As an important activity of the e-Infrastructure for Social Science project, we will build a security infrastructure which enables secure access to the e-social science resources.

The e-Infrastructure allows resources to be shared between members of virtual organizations (VOs). However, if an organization is to allow its resources to be shared amongst its VO members, it needs to be able to determine who is authorized to access these resources in which ways, and who is not. Access control rests on the two related concepts of authentication (establishing the identity of an entity, e.g., a user) and authorization (determining the level of access to an authenticated entity) which are two fundamental but critical issues in constructing the security infrastructure for the e-Infrastructure.

There have been continuous efforts in the development of authentication and authorization technologies in the Grid communities.

- Typical early efforts like Grid Security Infrastructure (GSI) [3] in the Globus Toolkit [4] implement security infrastructures using Public Key Infrastructure (PKI) [5, 6] based authentication and Grid-map file based authorization. This approach has been criticized, mainly because it lacks scalability and flexibility in coping with dynamically changing users, rights and permission in the scenarios of large scale VOs.
- Thereafter a variety of advance authentication and authorization infrastructures have been developed, such as Athens [7], Shibboleth [8], CAS [9], VOMS [10] and PERMIS [11]. A core idea behind these security infrastructures is that user attributes or roles are used for expressing user’s privileges. However, these security infrastructures are predominantly used by different Grid communities.
and thus have interoperability issues. In addition, each of these solutions has its functionality focus, either on authentication or authorization, and more efforts are required to provide an integrated solution.

- GridShibPERMIS [12] and VPMan [13] are two pioneer projects conducted by UK universities aiming to provide an integrated and advanced security infrastructure. The GridShibPERMIS project integrates the Internet2 Shibboleth implementation [14] with the PERMIS infrastructure to authorize Grid jobs running with Globus Toolkit in order to provide policy driven role-based access control decision making to Grid jobs. The VPMan project integrates two advanced authorization infrastructures, i.e., VOMS and PERMIS (especially the privilege management function of VOMS and the authorization decision function of PERMIS). VPMan further seamlessly embeds into the Globus authorization framework, so that the management of grid users is made easier and fine grained access to resources can be achieved. A problem is that these infrastructures are restricted to working in a Globus environment, but in practice many applications do not work within this environment. An integrated authentication and authorization infrastructure is expected to protect non Globus-enabled resources.

We present a security infrastructure that leverages Guanxi Shibboleth technology as authentication mechanism and PERMIS technology as authorization mechanism. Guanxi Shibboleth enables the setup of federations of collaborating institutions that trust each other to authenticate their users properly. Furthermore, Guanxi Shibboleth technology is an attribute based system, and it implements secure transfer of users’ attributes from users’ home sites to remote service providers, and these attributes can be fed to appropriate authorization infrastructures, in our case, PERMIS, to make authorization decisions. This indicates that Guanxi Shibboleth provides the technical basis and linkage for the provision of fine-grained access control in a Grid environment; when integrated with PERMIS authorization technology, our Guanxi Shibboleth based security infrastructure can realize fine-grained access control for distributed resources or services, and these resources are not necessarily working in a Globus environment.

As a result, our security infrastructure offers scalable and flexible user authentication brought about by Guanxi Shibboleth. Meanwhile, making use of the user attribute assignment function of Guanxi Shibboleth, the integrated PERMIS provides policy-driven, role-based, multi-grained authorization for access to and usage of the e-Infrastructure. Besides, we employ Sakai as the principal user portal which provides a user-friendly environment to access various resources. In this paper, we focus on the architecture design of the security infrastructure. In Section 2, we introduce the background of PERMIS and Sakai technologies as well as the motivations for adopting them. Section 3 focuses on discussing the issues involved with the architecture design of the security infrastructure, including the Guanxi architecture, the integration of Guanxi and Sakai, as well the integration of Guanxi and PERMIS. In Section 4 we discuss a use case supported by the security infrastructure and the implementation issues involved. Finally Section 5 concludes the paper and gives our plan for future work.

2. PERMIS and Sakai technologies

To better understand our security infrastructure, in this section, we will give a brief overview of the two key technologies integrated into our infrastructure, i.e., PERMIS and Sakai. The advantages of these technologies will be discussed and the motivations of adopting these technologies will be addressed.

2.1. PERMIS authorization

A number of authorization technologies and infrastructures have been applied to Grid and distributed computing, typical examples include CAS, VOMS, Akenti and PERMIS. PERMIS, i.e., PrivilEge and Role Management Infrastructure Standard is an advanced authorization infrastructure. It differs from other Role Based Access Control (RBAC) authorization systems in that access control decisions are made based upon users’ attributes, not just upon their organizational roles as in conventional RBAC. This authorization model is named as attribute based access control (ABAC). PERMIS can support the ABAC model as the basis of the infrastructure is user attributes stored in Attribute Certificates (ACs) [15]: administrators issue ACs to users which reflect the users’ attributes and roles, and service providers publish policies that detail the access rights granted to each role or attribute. In other words, service providers determine a user’s privileges and access right based on the user’s attributes issued by administrators.

As illustrated in Figure 1, the PERMIS authorization decision engine consists of the Policy Decision Point (PDP) and the Credential Validation Service (CVS). Both of the components are policy driven. When a user requires access to a PERMIS
protected service, the PERMIS authorization model can be simply described as follows:

1. The Policy Decision Point (PDP) informs the Policy Enforcement Point (PEP) about the credentials required for access. The PEP then collects the user credentials from the credential provider.

2. The Credential Validation Service (CVS) of the PERMIS authorization engine will validate the user attribute assignments based on its set of policies in the form of “this authority is trusted to assign these attributes to this group of users”. As a result, only those attributes that can be validated by the credential validation rules in the policy will be recognized as valid.

3. After validation by the CVS, the user attributes are passed to the PDP which takes the decision based on its policy in the form of “users with this set of attributes are allowed this type of access to this resource, providing that the following conditions are met”. As a result, the user is either granted or denied access to the service requested.

Comparing with other authorization infrastructures, PERMIS offers a couple of advantages highlighted as follows:

- In PERMIS, user attributes are held in attribute certificates which describe a user’s rights and the target services then read the user’s AC to see if the user is allowed to perform the action being requested. This approach de-couples users’ privileges from their local identity and thus enables scalability and flexibility of access control, especially for large-scale VOs with a big number of user, service providers and access requests.

- PERMIS is a policy-driven authorization system in that a service provider makes authorization decisions using policies or rules. Service providers have the flexibility to write multi-grained and conditional access policies on the basis of user attributes, either in XML format or X.509 signed certificate.

- Finally, PERMIS can be easily integrated with attribute based authentication systems in the provision of user attributes to the PERMIS authorization engine. This indicates that service providers only need to specify the access control policy, and leave PERMIS to enforce their policy on their behalf. This reduces service provider’s authorization workload significantly.

It should be noted that PERMIS has been developed with the primary focus on policy based authorization decision functionality, and thus has concentrated much less on the privilege assignment function in which users are assigned to roles and attributes although PERMIS does provide its own tools to make attribute assignment. Indeed this shows the flexibility brought by PERMIS, i.e., any system is allowed to make attribute assignments, but PERMIS has the validation control on which attributes are trusted from which authority.

2.2. Sakai

Sakai [16] is an integrated Collaboration and Learning Environment (CLE) that can be used for a variety of purposes such as teaching, research, learning and so on. Sakai provides an open standard portal framework that hosts extensive built-in online collaboration and learning tools — discussion boards, email discussion list, announcements, chat room, grade book, resource area for any files you would like to share including web links to external sites, calendar, wiki, and much more. Sakai has been used or extended by more and more communities as their Web-based environment, e.g. the Sakai Virtual Research Environment (VRE) [17].

In our project, we intend to make the Sakai portal the principal user interface of our e-infrastructure for social science. Through the Sakai portal, users can access various tools provided by Sakai as well as distributed resources or services integrated into Sakai. Most importantly, the security infrastructure we present...
will be integrated into the Sakai system and applied to user authentication and access control of remote resources or services.

3. Architecture of the security infrastructure

The goal of our security infrastructure is to integrate Guanxi Shibboleth implementation and PERMIS to provide authentication and authorization mechanism. The infrastructure will work in the Sakai environment. In this section, we will discuss the architecture of the security infrastructure. Firstly we describe the Guanxi Shibboleth technology. Then we will address the integration of Guanxi with Sakai and PERMIS, and outline the overall architecture of the security infrastructure.

3.1. Guanxi Shibboleth

Developed by the UHI Millennium Institute (UK) in partnership with the University of Oxford and the University of Leeds (UK), Guanxi [18] is an open source implementation of the Shibboleth protocol and architecture (another popular implementation is the Shibboleth System [14] provided by Internet2 Middleware Initiative). As an authentication and authorization infrastructure to support inter-institutional sharing of resources, Guanxi Shibboleth presents a couple of features in terms of:

- User authentication
  Guanxi Shibboleth manages user authentication at Grid VO or user federation layer, and ‘glues’ various local authentication systems at users’ home institution, allowing a user to use the local authentication mechanisms to achieve federation wide authentication. In other words, the Guanxi Shibboleth infrastructure introduces a two-layer authentication scheme and is capable of supporting a wide range of heterogeneous authentication methods, and allows a user to gain authentication using local authentication systems. More importantly, a user can achieve single-sign on, that is, using the same local login information to access various cross-domain resources.

- Attribute management
  Guanxi Shibboleth manages user attributes assignment by means of its Attribute Authority. More importantly, it implements the SAML (Security Assertion Markup Language) protocol [19] for the secure passing of user information between users’ institutions and resources. This enables fine grained authorization which is the requirements of complex systems like Grid environments.

![Figure 2. Guanxi Architecture and Components](image)

The Guanxi Shibboleth architecture is composed of three main components: an Identity Provider (IdP), a Service Provider (SP) and an optional Where Are You From (WAYF) service. The architecture is summarized in Figure 2.

3.1.1. Identity Provider. The Guanxi Identity Provider (IdP) is a modular system installed at the user’s home organization. The main function of IdP is to perform user authentication as well as to manage user attributes. Guanxi IdP provides an Authenticator interface, and one can write its implementations so that IdP can work with the alternative backend authentication systems like Sakai, LDAP, and FlatFile. Guanxi IdP also provides an attributor mechanism, making IdP associated with multiple attributes providers to realize aggregation of attributes from diverse institutional data stores.

3.1.2 Service Provider. The Guanxi Service Provider (SP) uses a radically different architecture to the Internet2 SP default implementation. In the Guanxi model, the SP is split into two co-operating entities, known as an Engine and a Guard. The Engine, also known as the SAML Engine, is responsible for communicating with federation entities such as IdPs and handles all SAML communication and processing on behalf of a community of Guards that are associated with that Engine. More specifically:
### Guard

The Guard is a lightweight SP module that is designed to sit in front of a service (or resource) and trap requests to the service. Guards do not normally deal with SAML as they delegate all SAML processing to their Engine. It's the Guard's job to block access to a service until attributes arrive from the Engine, which can then be passed to the service. The Guard plays no further role in service access beyond this blocking role. What happens to access, based on the attributes, is up to the service. This means Guards are very lightweight processes that can be easily deployed in front of services.

### Engine

The Engine is the SP module that takes care of all SAML processing and XML signature and security verification. It can be situated anywhere on the internet. Each Engine can have a community of Guards with which it has metadata associations and it is this metadata that enables the secure communications among the Engine and its Guards. To prevent capture of user attributes by rogue Guards, the Engine will use a Guard's metadata to verify the session that is being requested. Once verification is complete, the Engine will represent that Guard in the federation via identity masquerading, while the Guard will return to the service, there to await the arrival of attributes, obtained by the Engine via an IdP and sent to the Guard's attribute consumer service, e.g. an authorization system like PERMIS.

### 3.1.3 WAYF Service

In a federation where many organizations share a trust relationship, the federation may manage a WAYF service. The WAYF provides a mechanism for allowing users to be forwarded to the correct home organization. This is normally presented in the form of a web page with a drop down list. Once the user selects an organization, the user is forwarded to that institution’s IdP to complete the authentication process.

As illustrated in Figure 2, the process of a user from organization 1 accessing resources (or services) hosted at organization 2 can be described as follows:

1. The user from organization 1 makes a request for a resource located on organization 2. The request will be trapped by the Guard which sits in front of the resources.
2. The Guard sets up WS-Callback with the SAML Engine situated on the federation server.
3. The Guard redirects the user to WAYF. WAYF will present a list to the user containing his/her home institutions and will ask to select his/her institution.
4. User will be redirected to his/her home institution to get authenticated by the Single-Sign On (SSO) system.
5. SSO replies to SAML Engine.
6. The SAML Engine requests attributes of the user from the Attribute Authority (AA) of the IdP.
7. The AA will look for the user attributes in the database and will send the required attributes to the SAML Engine.
8. The SAML Engine invokes WS-Callback to the Guard which retrieves user attributes.
9. The Guard makes access decision based on attributes gathered by the SAML Engine.

### 3.2. Integration of Guanxi and Sakai

Our e-infrastructure for social science project adopts Sakai as a one-stop user-friendly environment for users to access various resources or services. An issue arising from the adoption of Sakai is the integration of Sakai and Guanxi Shibboleth such that users can have the advantages of both systems.

We use Guanxi Shib Kit (GSK) [20] which is a tool developed by the Guanxi project to address the above challenge. GSK tool provides the following two ways to make Sakai seamlessly work with Guanxi system:

- **Sakai as a Service Provider**

  GSK can work in conjunction with the main Sakai portal to allow users to log in to Sakai using Shibboleth, i.e. the GSK provides Sakai with the capability to be used as a Service Provider. Normal Sakai users are not affected as the normal authentication and authorization mechanisms of Sakai are not affected by the Guanxi Shibboleth portal. The Guanxi Shibboleth portal acts as a holding area for users while they are authenticated by their Identity Provider and their attributes gathered by the Guanxi Shibboleth portal. If their attributes match what is specified in the configuration of the Guanxi Shibboleth portal they are then promoted to a full Sakai user and redirected to the main Sakai portal, where they seamlessly join other users of the system.

- **Sakai as an Identity Provider**

  The GSK also provides Sakai with the capability to be used as a Guanxi IdP, to allow Sakai to be used in Shibboleth federations to access Guanxi Shibboleth protected resources. The GSK is bundled with a Guanxi Authenticator and Attributor that are customized to work with the Sakai authentication. This allows the Guanxi Shibboleth functionality to delegate...
to the mechanisms Sakai is using to control authentication and authorization.

![Security infrastructure architecture and integration model of Guanxi, PERMIS and Sakai](image)

For our case, Sakai is integrated with Guanxi mainly as a service provider, and Sakai has its own Guanxi Guard deployed. Figure 3 illustrates the integration of Sakai and Guanxi using GSK. Let’s look at a scenario that a user from organization 1 makes a request for resources located on organization 2 via the Sakai system. Please note in the Sakai environment, remote resources can be exposed as portlet and user can access resources through WSRP (Web Services for Remote Portlets) protocol. The generic workflow of request and response across the whole security infrastructure can be described as follows:

1. The entry point of the user’s session is the GSK portal, which will interact with Guanxi on behalf of users.
2. The Sakai Guard will block user’s request to the GSK portal, followed by the establishment of user authentication session.
3. As discussed earlier, the user will get authenticated by his/her IdP SSO system, which involves a series of interactions among the Guanxi components including SAML Engine, WAYF, IdP and Guard (please refer to Figure 2 for more details). In addition, the SAML Engine will request and obtain attributes of the user from the IdP.
4. After passing authentication, the user will be redirected to the Sakai portal which is pre-registered with the GSK as the main Sakai entry point.
5. The user can now make a request for the resources represented as portlet in the Sakai. A Guard is deployed sitting in front of the resources located at organization 2.
6. User’s request will be captured by resource Guard. Then the resource Guard sets up WS-Callback with the SAML Engine, and in turn the SAML Engine sends user attributes to the resource Guard.

Above we have discussed the process of user authentication based on Guanxi Shibboleth and further logon the Sakai portal to make resource request. The next will be the authorization phase, which will be discussed in the following subsection.

Please note that the users we are talking here are from the participating institutions of a federation. Each participating institution has their own Guanxi IdP and backend authentication systems. For the normal Sakai users that may not belong to any participating institution, we can use the GSK function discussed earlier to make Sakai working as a Guanxi IdP so that Sakai users can request to access resources subject to access control policies. Especially, Sakai Authenticator and Sakai Attributor bundled with GSK need to be created and customized (see Figure 3) to work with Guanxi IdP: Sakai Authenticator delegates to Sakai authentication system (i.e., Sakai works as a backend authentication system), and Sakai Attributor uses Sakai for user information (i.e., Sakai provides information about users which can be further transformed or mapped into user attributes in the specific format of IdP and SAML Engine).

### 3.3. Integration of Guanxi and PERMIS

In Section 3.1 we have discussed the features of Guanxi Shibboleth, but Guanxi Shibboleth also has its weakness in that it has limited authorization functionality itself and thus needs to work with a powerful authorization system to protect resources. In this section, we will discuss our approach to ‘plug’ PERMIS into the Guanxi Shibboleth framework and make the two systems work seamlessly.

The core idea behind the integration of Guanxi Shibboleth and PERMIS is to leverage the attribute
management functionality of Guanxi Shibboleth, by transporting user attributes as SAML attribute assertions from the Guanxi IdP, to PERMIS authorization engine. At the technical level, a context handler needs to be implemented as an interface between the resource Guard and the PERMIS authorization engine (as illustrated in Figure 3). Following up the workflow of request and response across the whole security infrastructure as described in the previous subsection, the authorization phase based on PERMIS will be continued as follows:

(7) The context handler retrieves and extracts user attributes returned by resource Guard, and converts them into the internal format recognized by PERMIS.

(8) The context handler then passes these attributes to the PERMIS CVS which checks if the Guanxi IdP is trusted to issue the attributes it has returned in accordance with the PERMIS policy.

(9) The valid attributes are passed along with information about the user’s resource request and target to the PERMIS PDP. The PERMIS PDP makes an access control decision (either access granted or denied) as per the PERMIS policy, and this is return by the context handler to the resource Guard.

(10) And finally, the resource Guard will access the resource if the request is granted, or simply take no action if the request is denied.

As per the PERMIS technology, the context handler can accept attribute certificate in the form of X.509 to be returned by the Guanxi IdP. The PERMIS engine reads the attribute certificates and makes the authorization decision.

4. Use case and implementation issues

Geo-linking is about linking data using geographical attributes such as a postcode or address rather than using a geometry, such as the boundary of a postcode or a coordinate of an address. Much data is geographic in that it contains geographical attributes, but does not have the spatial geometry necessary to generate geographical maps. However, often spatial geometry does exist in auxiliary data and by linking this, it becomes possible to do spatial analysis and produce geographical maps. A Geo-Linking Service (GLS) was developed as a part of the Secure Access to Geospatial Services (SEE-GEO) project [21]. The GLS allows users to query the service with appropriate parameters about datasets, and generates output of a particular query. The SEE-GEO GLS implementation uses client-server architecture with the client being a JSR-168 portlet.

Figure 4. The GLS Framework

Two major security concerns of the GLS service are that users may come from different institutions and the data being used may be confidential/restricted. The former issue means that it is appropriate to implement a federated authentication mechanism, while the later one requires appropriate authorization mechanisms to secure/access the data servers (see Figure 4). Furthermore, the access granted to data might be dependent on the date and time constraints, e.g., the sharing could be restricted to specific time limits such as the working hours of the organization. This means that fine-grained authorization that supports conditional access control is wanted. In addition, the GLS service is integrated into Sakai portal for a user-friendly access purpose.

The security infrastructure we present in the previous section provides a solution and it has been applied to the GLS service to fulfill its security requirements. We have set up a testbed which consists of a cluster of virtual machines. On this testbed the GLS was deployed and the security infrastructure was implemented as follows:

- We set up Guanxi IdP version 1.4 on two virtual machines representing two institutions, and these institutions forms a Guanxi Shibboleth federation. The users will be authenticated by their backend LDAP authentication systems.
- We also set up Guanxi SP version 1.3 including a federation-wide SAML Engine on one virtual machine, a Sakai Guard working with the latest Sakai version 2.5 on one machine, as well as a resource Guard to protect the GLS service deployed on one virtual machine. Besides, the WAYF service version 1.2 was deployed on the same virtual machine as the SAML engine.
• The GSK tool has been deployed and configured to work with Sakai and Guanxi. The integration of Sakai and Guanxi has been achieved.
• The open PERMIS package [22] version 4.0, an open source implementation of PERMIS, was deployed with the GLS service site. We implemented a context handler as interface between the resource Guard and PERMIS. The integration of Guanxi and PERMIS was accomplished.

An initial set of core user attributes has been defined and recommended for the IdPs of each institution such that the PERMIS systems can subsequently use for authorization decisions to access the GLS service. We also use the policy editing tools provided by the open PERMIS package to write policies and rules to access the GLS service. The operations of the security infrastructure show that it works well with the GLS service and achieved its functionality.

5. Conclusion and future work

Security is a fundamental requirement in the e-Infrastructures for the social sciences project. The emergence of Shibboleth and advanced authorization technologies like PERMIS provides a promising solution to address the authentication and authorization issues in building a security infrastructure. In this paper, we present an integrated security infrastructure that provides the advantages of both Guanxi Shibboleth cross-organization identity federation and attribute management with PERMIS policy driven role-based fine grained authorization. This security infrastructure has been implemented and applied to some use case. We believe that this security infrastructure will provide a general framework for other applications that have similar authentication and authorization requirements.

On the whole, the security infrastructure presented in this paper is yet a concept-proof authentication and authorization framework, and there are still a number of issues to be further addressed. Basically our future work will be carried out in the following aspects:
• Improvement of the security infrastructure – we will further address the issues such as user attribute management in terms of attribute definition and format, authorization policy management to support complicated policy composition, federation membership management and attribute assignment (possibly introducing VOMS technology to work with Guanxi Shibboleth and PERMIS), etc.
• Application of the security infrastructure – we plan to apply the security infrastructure to more e-social sciences applications as well as other domain applications that have security requirements.
• Performance study of the security infrastructure – we plan to set up a cross-institution federation and deploy our security infrastructure on a large scale testbed. Then we will conduct experiments to investigate the scalability and performance of the security infrastructure. This will help further optimization and enhancement of the security infrastructure.

References