

# The Application of SWRL Based Ontology Inference for Privacy Protection

Qiang Ge

Department of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China

E-mail: renzheng0521@163.com

Guohua Shen, Zhiqiu Huang and Changbo Ke

Department of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China

E-mail: ghshen@nuaa.edu.cn, zqhuang@nuaa.edu.cn, kcb1984@163.com

**Abstract**—The phenomena of illegal disclosure of user privacy has become more serious considerably in the last years. How to protect user privacy and prevent user privacy from illegal disclosure has become of great interest to researchers. In this paper, we propose an approach of user privacy protection based on ontology inference. We create privacy ontology through detailed analysis of privacy domain. Then, we reason the privacy ontology with Semantic Web Rule Language (SWRL) rule of user privacy preference. By creating ontology and SWRL rules, then reasoning logically, we can find adequate service whose privacy policy coincides with user privacy preference and protect user privacy effectively. The result of experiment shows that the approach proposed in this paper is feasible and effective.

**Index Terms**—Ontology Inference, Semantic Web Rule Language (SWRL), Rules, Reasoning

## I. INTRODUCTION

As the development of Internet technology and the popularity of network application, a large amount of users enjoy the pleasure and convenience of Internet. However the problem caused by user privacy disclosure has also become more serious than ever before. Recently the issue of user privacy protection has drawn attention of government departments, academics as well as the legislative.

Privacy protection faces massive problems because of two main factors: first, “the inherently open, nondeterministic the nature of the Web”; second the “complex leakage-prone information flow of many Web based transactions that involve the transfer of sensitive, personal information.”[1].

There are some standard for privacy preservation on the Web, such as P3P [2] (Platform for Privacy Preferences), conjunction with APPEL [3] (A P3P Preference Exchange Language), and EPAL (Enterprise Privacy Authorization Language). P3P defines a platform by which service providers could describe their privacy policies. APPEL provides the mechanism through which service users could match their own privacy preferences to the provider’s privacy policy and decide whether to use

a given service. Although P3P automates the decision process for matching service provider’s privacy policy to user’s privacy preferences, it can’t let users control their privacy information once they’ve entered it. Another shortcoming is P3P’s disability to interlink concepts from different domains together and match user preferences accordingly. EPAL makes it possible for software developers to set up the execution of security and privacy policy on the enterprise software applications. But it has a restricted vocabulary and do not offer semantic support.

In this paper we present our privacy protection methodology based on ontology inference [4]. Firstly, we analyze the model of privacy domain and confirm constituent elements of privacy domain. Then we construct an ontology which contains privacy semantic adequately. Depending on user privacy preference, we make out rules based on SWRL [5, 6, 7]. We choose Jess as inference engine, and reason the privacy ontology with the SWRL rules. We can find an adequate service provider whose privacy policy is corresponding to user privacy preference through the reasoning result. According to this approach, we can prevent user privacy from illegal disclosure by service provider and protect user privacy effectively. The rest of this paper is organized as follows. Section 2 introduces relevant background and related work. Section 3 provides the model of privacy domain and creates a privacy ontology. Section 4 describes how to transform user privacy preferences to SWRL rules. Section 5 is experimental analysis. Finally, in section 6 we give out the conclusion and state the future work.

## II. BACKGROUND AND RELATED WORK

Gruber [8] point out that an “ontology is explicit specification of conceptualization”. Guarino [9] uses formal logic represent explicitly what something is, and the relationship among the components composed of the thing. In this paper, we accept that ontology is “formal explicit description of concepts in domain discourse (classes or concepts), properties of each concept describing various features and attributes of the concept (roles or properties) and restrictions.” [10]. Both domain

experts and ontology engineers are needed in the process of building domain ontology. Domain experts provide their knowledge about the domain, and ontology engineers build the ontology using the knowledge provided by the domain expert.

Our work in this paper aims at user privacy protection based on privacy domain ontology inference. The privacy ontology is modeled considering the basic components of privacy domain, and relationships among main concepts of privacy domain. Reasoning this privacy ontology with SWRL rule based on user privacy preference, an appropriate service whose privacy policy coincides with user privacy preference could be found. This approach follows the means proposed by Garcia [11] who develop a privacy ontology for services and a privacy framework for service-oriented architecture. The ontology provides a base vocabulary for a given service domain and operating environment. Through the ontology, the framework support service selection based on the consumer's privacy requirements. In our work, the privacy ontology is created through analysis of model of privacy domain and hierarchical structure of the components of privacy domain, and service selection is achieved by ontology inference.

Knutson [12] presents some principles that organizations should follow to create privacy awareness. He points out that a privacy core team with technical and legal experts must define a privacy terminology to achieve a common understanding of the meaning of rules. Another recommendation is to create guidelines to help developers to become independent from privacy experts with respect to basic tasks. Similar concerns for software design are endorsed within other works on privacy awareness [13, 14]. In our work, these requirements are carried out with the definition of the privacy ontology.

Spyns [15] presents an approach to apply automatic evaluation procedure for triples as material for a privacy ontology. Denker [16] presents an ontology for semantic Web service. This ontology focuses on security aspects. Cheung [17] proposes an approach of privacy protection based on ontology, however he only consider specific area, that is e-science.

### III. PRIVACY DOMAIN AND PRIVACY ONTOLOGY

#### A. Privacy Domain

Privacy domain has four components as follows: Data, Policy, Entity and Operation. Data in privacy domain represents privacy information to be protected strictly. Policy represents privacy protecting policy made out to prevent Data from illegal discourse. Entity consists of service provider and service user. Service user provides Data relevant to privacy of himself which need to be protected. Service provider applies to operating on Data

to provide service for service user. Policy is formulated by Entity. Policy decides whether service provider could operate on Data or not. Operation represents action that service provider could take on Data. The model of privacy domain is shown in Figure 1.

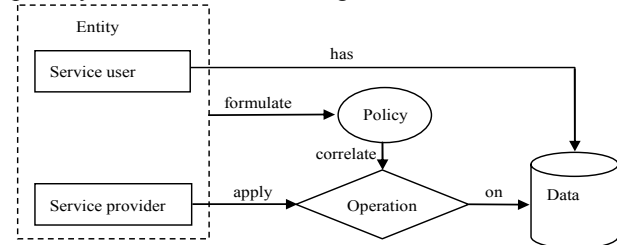


Figure 1. The model of privacy domain.

#### B. Privacy Ontology

In section III(A), we analyze the model of privacy domain and confirm the main components of privacy domain. In this section, we adopt Protégé [18] to create ontology for privacy domain. Protégé is a free, open source ontology editor and knowledge-base framework developed by Stanford University. Privacy domain ontology has four parallel basic classes as described in section III(A): Data, Policy, Entity and Operation. These four basic classes constitute main module of privacy ontology. The Figure 2 shows the structure for basic classes of privacy ontology.

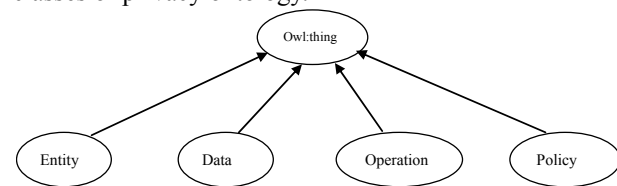


Figure 2. The structure graph for basic class of privacy ontology.

The class of Data includes eight subclasses: Identification, Basic, Contact, Commerce, Finance, Health, Faith and Other\_Data. The class of Identification is used to indicate identity which could identify service user, such as passport and identification card. The class of Basic is used to describe basic information of service user. The class of Contact represents contacting means and physical address. The class of Commerce indicates the information of business for service user. Finance in Privacy ontology is used to represent the finance information for service user. Health describes health condition of service user. Faith in privacy ontology indicates religion belief of service user and which political association service user is affiliated to. Other\_Data is used to describe other information relevant to privacy of service user. The structure for subclasses of Data is illustrated in Figure 3.

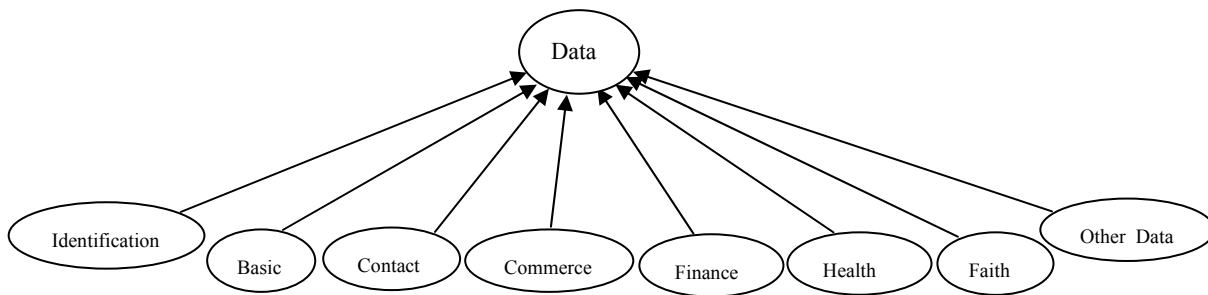


Figure 3. The structure graph for class of Data.

The basic class of Entity contains two subclasses: User, Site. The class of User represents service user whose privacy need to be protected appropriately. The class of Site indicates service provider applying user privacy data to provide service for service users.

The basic class of Policy includes six subclasses: Purpose, Condition, Obligation, Consent, Retention as well as Judge. The class of Purpose describes the purpose for which service provider apply privacy information of service users. The class of Condition indicates the condition which must be satisfied before operating on the privacy information of service users. The class of Obligation indicates what the service providers should perform after finishing the service. The class of Consent indicates whether service user agree service provider operating on the privacy information or not. Only when service provider get agreement of service user, can service provider operate the privacy information. The class of Retention describes how long service provider could keep the privacy information of service users. The class of Judge describes whether service provider is able to supply service for service user depending on comprehensive elements. The structure for subclasses of Policy is illustrated in Figure 4.

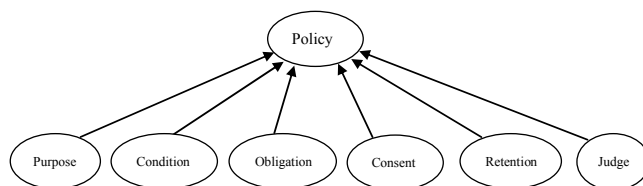


Figure 4. The structure graph for class of Policy.

The basic class of Operation includes four subclasses: Collect, Access, Publish and Other\_Operation. These four subclasses describe common operations which service provider could take on the privacy information.

After finishing defining classes of privacy ontology, we define datatype properties and object properties of these classes through the “Properties” label of Protégé.

The privacy ontology includes datatype properties as follows: site\_address, site\_url, site\_contact, identification\_card, passport, telephone, user\_address, user\_birthday, user\_email, user\_id, user\_name, user\_sex, zipCode, payment\_pattern, bill, account, access card, access\_card\_record, psychology, physical, party, religion, length, et al.

The privacy ontology includes object properties as follows: collect\_name, collect\_telephone, collect\_address,

judge\_is, purpose\_is, obligation\_is, retention\_is, condition\_is, has\_identification, has\_basic, has\_contact, has\_commerce, has\_finance, has\_health, has\_faith, publish\_identification, publish\_basic, publish\_contact, other\_operation\_identification, other\_operation\_basic, other\_operation\_contact, et al.

TABLE 1.

THE DETAILED INFORMATION FOR PARTIAL OBJECT PROPERTIES OF PRIVACY ONTOLOGY.

Object property	Domain	Range	Explanation
collect_name	Site	Consent	Whether to allow to collect name.
collect_telephone	Site	Consent	Whether to allow to collect telephone.
purpose_is	Site	Purpose	What is the purpose of operating the privacy information.
obligation_is	Site	Obligation	What is the obligation service provider must carry out.
judge_is	Site	Judge	Service provider judge whether it can provide service user with service.

So far, we have finished constructing the privacy ontology mainly. However, in order to facilitate our study and make the ontology corresponding to the real state of privacy domain, we should progressively refine the ontology in later work.

#### IV. CREATION OF SWRL RULES

##### A. Atoms of SWRL

SWRL is abbreviation of Semantic Web Rule Language, which combines OWL DL and OWL Lite. SWRL extends the set of OWL axioms to include Horn-like rules. It thus enables Horn-like rules to be combined with an OWL knowledge base [6].

A high-level abstract syntax extends the OWL abstract syntax. An extension of OWL model-theoretic semantics provide a formal meaning for OWL ontology, which includes rules written in this abstract syntax [6].

The SWRL rules consist of an antecedent(body) and consequent(head). There are several forms for atoms in SWRL rules, such as  $D(x)$ ,  $P(x, y)$ .  $D$  is an OWL description in  $D(x)$ , while  $P$  is an OWL property and  $x, y$  are one of three situations as below: variables, OWL individuals or OWL data values.

According to standard of SWRL language and the

privacy ontology created at section 3, the partial atoms of SWRL rules based on privacy ontology is listed in Table2:

TABLE 2.

THE DETAILED INFORMATION FOR PARTIAL ATOMS OF SWRL BASED ON PRIVACY ONTOLOGY.

Atom	Description
Site(?x)	X is an instance of Site.
collect_name( ?x,yes)	Agree X to collect name. ("yes" is an instance of Consent )
collect_telephone( ?x, no)	Deny X collecting telephone. ("no" is an instance of Consent)
purpose_is(?x, "providing service")	Purpose is "providing service".("providing service" is an instance of Purpose)
obligation_is(?x, "deleting privacy information immediately after service is finished")	Obligation is "deleting privacy information immediately after service is finished".("deleting privacy information immediately after service is finished" is an instance of Obligation)
judge_is(?x, "agree to provide service")	Judgment is "agree to provide service" depending on comprehensive elements.("agree to provide service" is an instance of Judge)

### B. SWRL Rules

We suppose two scenarios which are buying flowers online and selecting suitable express mail service separately. The user privacy preferences are as follows.

#### (1)scenario 1(buying flowers online)

Privacy preference of someone who attempt to buy flowers online is composed of five elements as follows:

- Flower shop could collect his name.
- Flower shop could collect family address.
- Flower shop could collect telephone number.
- Flower shop must delete privacy information relevant to him immediately after finishing service.
- Flowers booked online should arrive in 24 hours.

Depending on the object properties of privacy ontology described in section III(B) and description of SWRL in section IV(A), we can translate the five elements of user privacy preference into atoms of SWRL described briefly in Table 3:

TABLE 3.

THE DETAILED INFORMATION OF ATOMS OF SWRL FOR USER PRIVACY PREFERENCE IN SCENARIO 1.

Privacy preference	Atom of SWRL
Flower shop could collect his name.	collect_name(?x, yes)
Flower shop could collect family address.	collect_address(?x, yes)
Flower shop could collect telephone number.	collect_telephone(?x, yes)
Flower shop must delete privacy information relevant to him immediately after service is finished.	obligation_is(?x, "deleting privacy information immediately after service is finished")
Flowers booked online should arrive in 24 hours.	condition_is(?x, "goods should arrive in 24 hours")

User privacy preference in the scenario of buying flowers online could be translated into SWRL rule as shown in Figure 5:

```
Rule1: Site(?x) ∧ collect_name(?x, yes) ∧
collect_telephone(?x, yes) ∧ collect_address(?x, yes) ∧
obligation_is(?x, "deleting privacy information immediately after
service is finished") ∧ condition_is(?x, "goods should arrive in
24 hours") → judge_is(?x, "agree to provide service")
```

Figure 5. Rule1 based on user privacy preference in scenario 1.

#### (2)scenario 2(selecting express mail service)

Privacy preference of someone who choose an appropriate express mail service among several express mail services is composed of five elements as follows:

- Express company could collect his name.
- Express company could collect family address.
- Express company could collect telephone number.
- Only for the purpose of providing service, could privacy information relevant to him be operated by express company.
- Express company must delete privacy information of him immediately after service is finished.

As analyzed in scenario 1, the privacy preference listed above could also be translate into atoms of SWRL as shown in the Table 4:

TABLE 4.

THE DETAILED INFORMATION OF ATOMS OF SWRL FOR USER PRIVACY PREFERENCE IN SCENARIO 2.

Privacy preference	Atom of SWRL
Express company could collect his name.	collect_name(?x, yes)
Express company could collect family address	collect_address(?x, yes)
Express company could collect telephone number	collect_telephone(?x,yes)
Only for the purpose of providing service, could privacy information be operated by express company.	purpose_is(?x, "providing service")
Express company must immediately delete privacy information of him after service is finished.	obligation_is(?x, "deleting privacy information immediately after service is finished")

User privacy preference in the scenario of selecting express mail service could be translated into SWRL rule as shown in Figure 6:

```
Rule2: Site(?x) ∧ collect_name(?x, yes) ∧
collect_telephone(?x, yes) ∧ collect_address(?x, yes) ∧
obligation_is(?x, "deleting privacy information immediately after
service is finished") ∧ purpose_is(?x, "providing service")
→ judge_is(?x, "agree to provide service")
```

Figure 6. Rule2 based on user privacy preference in scenario 2.

## V. EXPERIMENT

Choosing adequate service whose privacy policy is corresponding to privacy preferences of user can prevent user privacy from illegal disclosure by service provider. In this paper, by reasoning the privacy ontology based on SWRL rules, we can find the appropriate service and protect user privacy efficiently.

We choose Jess [19, 20, 21] as reasoning engine. Jess is small, light, and one of the fastest rule engines available. The reasoning process through Jess engine based on SWRL rule is as follows: Firstly, we create privacy ontology as knowledge base needed in reasoning process. We also register corporations and their privacy policy on privacy ontology. Secondly, according to privacy ontology created in first step, we make out SWRL rule based on privacy preferences of users. Finally, we import privacy ontology and SWRL rule into Jess engine, then reason privacy ontology with SWRL rule.

Through the reasoning result, we can find appropriate service whose privacy policy is corresponding to privacy preferences of user.

In scenario 1(buying flowers) at section IV(B), we suppose that there are three flower shops online called “Flower Artistry”, “Rose Boutique” and “Floral Paradise” respectively. The privacy policy of flower shop called “Flower Artistry” is as follows:

1. “Flower Artistry” need to collect user name.
2. “Flower Artistry” need to collect user telephone number.
3. “Flower Artistry” need to collect user address.
4. “Flower Artistry” deletes user privacy information immediately after finishing service.
5. “Flower Artistry” assures user that flowers booked online arrive in 48 hours.

“Rose Boutique” has privacy policy also composed of five elements. The first four elements of privacy policy are the same as privacy policy of “Flower Artistry”. While the fifth element is: “Rose Boutique” assures user that flowers booked online arrive in 24 hours. The first three elements of “Floral Paradise” are the same as privacy of “Flower Artistry”. The fourth element of “Floral Paradise” is: user privacy information is reserved for next service. The fifth element of “Floral Paradise” is: Flowers sent by “Floral Paradise” arrive in 24 hours.

We register “Flower Artistry”, “Rose Boutique”, “Floral Paradise” and their privacy policies on privacy ontology created at section III(B). We load Rule1 created at section IV(B) and privacy ontology on Jess engine. The process and result of reasoning is as shown in Figure 7 and Figure 8. The result of reasoning shows that the privacy policy of “Rose Boutique” coincides with user privacy preference and user agrees “Rose Boutique” to provide service.



Figure 7. Importing the privacy ontology and rule1 into reasoning engine.



Figure 8. The reasoning result of scenario 1.

In scenario 2(selecting express mail service) at section IV(B), we suppose there are three express companies called FastEx, SFT and DSH respectively. The privacy policy of FastEx is as follows:

1. FastEx need to collect user name.
2. FastEx need to collect user telephone number.
3. FastEx need to collect user address.
4. FastEx operates user privacy information only for the purpose of providing service.

5. FastEx deletes user privacy information immediately after finishing service.

The privacy policy of SFT also is composed of five elements. The first four elements are the same as the privacy policy of FastEx. The fifth element of SFT is: SFT reserves user privacy for next service. The DSH's privacy policy consists of five elements too. The first four elements are the same as FastEx's privacy policy. The fifth element of DSH's privacy policy is: DSH reserves user information as corporate performance.

We register FastEx, SFT, DSH and their privacy policies on privacy ontology, and then load privacy ontology and Rule 2 created at section IV(B) on Jess engine. The process and result of reasoning is as shown in Figure 9 and Figure10. The result of reasoning indicates that the privacy policy of FastEx coincides with user privacy preference and user selects FastEx for express service.

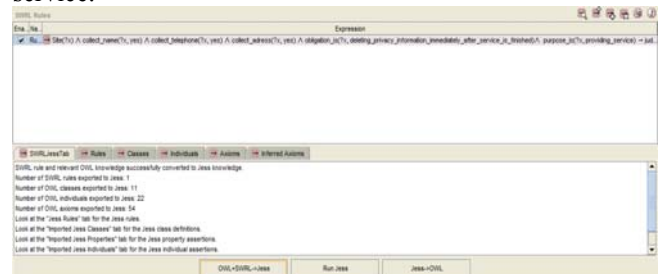


Figure 9. Importing the privacy ontology and rule2 into reasoning engine.

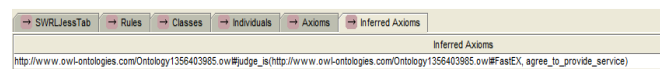


Figure 10 The reasoning result of scenario 2.

According to the results of experiments, we could find appropriate service whose privacy policy corresponds to privacy preferences of users though reasoning privacy ontology with SWRL rules of user privacy preferences. We could prevent privacy disclosure by service whose privacy policy does not satisfy privacy preferences of users. The results of experiments show that the approach of privacy protection based on ontology inference could protect privacy of users from illegal disclosure effectively.

## VI. CONCLUSION

In this paper, we analyze privacy domain and present the model of privacy domain. According to explicit analysis for model of privacy domain, we create a privacy ontology using protégé. Depending on the privacy ontology, we create SWRL rule of user privacy preference in service. Based on SWRL rule of user privacy preference, we reason the privacy policy with Jess engine. Through the result of reasoning, we can find appropriate service whose privacy policy coincides with user privacy preference. The result of experiment shows that this approach proposed in this paper could find adequate service whose privacy policy satisfies user privacy preference and prevent user privacy from illegal disclosure and protect user privacy effectively.

The future work is application of the approach proposed in this paper in the situation of service combination and in the situation of access control.

#### ACKNOWLEDGMENT

This work was supported by the National Natural Science Foundation of China under Grant Nos. 61272083, 61100034, 61170043, the Funds for Outstanding Doctoral Dissertation in NUAA under Grant No. BCXJ12-14, and the Fundamental Research Funds for the Central Universities of China under Grant Nos. 56XAA12022, NZ2013306.

#### REFERENCES

- [1] Bouguettaya ARA, Eltoweissy MY, "Privacy on the Web: Facts, challenges, and solutions", *Security & Privacy*, Vol. 1, pp. 40-49, 2003.
- [2] Reagle J, Cranor L F, "The platform for privacy preferences", *Communications of the ACM*, Vol. 42, pp. 48-55, 1999.
- [3] Cranor L, Langheinrich M, Marchiori M, A P3P preference exchange language 1.0 (APPEL1.0), W3C working draft, 2002.
- [4] Wenying GUO, "SWRL based semantic relevant discovery and its application on ontology mapping and integration", DSc, Zhejiang University, Zhejiang, China, 2006.
- [5] O'connor M, Knublauch H, Tu S, Grosz B, Dean M, Grosso W, Musen M, "Supporting rule system interoperability on the semantic web with SWRL", *The 4th International Semantic Web Conference*, pp. 974-986, 2005.
- [6] Bener A B, Ozadali V, Ilhan E S, "Semantic matchmaker with precondition and effect matching using SWRL", *Expert Systems with Applications*, Vol. 36, pp. 9371-9377, 2009.
- [7] Chen J F, Wang Y H, Liao J C, et al. Content Adaptation For Context-Aware Service. *Journal of Software*, 2012, 7(1): 176-185.
- [8] Gruber TR, "A translation approaches to portable ontology specifications", *Knowledge Acquisition*, Vol. 5, pp. 199-220, 1993.
- [9] Guarino N, *Formal Ontology and Information Systems*, IOS Press, 1998.
- [10] Noy NF, McGuinness DL, *Ontology development 101: A guide to creating your first ontology*, Stanford University, 2001.
- [11] Garcia D, Toledo MBF, Capretz M, Allison D, "Towards a base ontology for privacy protection in service-oriented architecture", *2009 IEEE International Conference on Service-Oriented Computing and Applications (SOCA)*, pp. 1-8, 2009.
- [12] Knutson TR, "Building Privacy into Software Products and Services", *Security & Privacy*, Vol. 5, pp. 72-74, 2007.
- [13] Troncoso C, Danezis G, Kosta E, Balasch J, Preneel B, "PriPAYD: Privacy-Friendly Pay-As-You-Drive Insurance", *IEEE Transactions on Dependable and Secure Computing*, Vol. 8, pp. 742-755, 2011.
- [14] Ye X, Zhu Z, Peng Y, Xie F, "Privacy Aware Engineering: A Case Study", *Journal of Software*, Vol. 4, pp. 218-225, 2009.
- [15] Spyns P, Hogben G, "Validating an automated evaluation procedure for ontology triples in the privacy domain", *The 18th Annual Conference on Legal Knowledge and Information Systems*, pp. 127-136, 2005.
- [16] Denker G, Kagal L, Finin T, Paolucci M, Sycara K, "Security for daml web services: Annotation and matchmaking", *Second International Semantic Web Conference*, pp. 335-350, 2003.
- [17] Cheung W, Gil Y, "Towards privacy aware data analysis workflows for e-science", *AAAI Workshop on Semantic e-Science*, pp. 22-26, 2007.
- [18] Knublauch H, Fergerson RW, Noy NF, Musen MA, "The protégé OWL Plugin: An Open Development Environment for Semantic Web applications", *Third International Semantic Web Conference*, pp. 229-243, 2004.
- [19] Friedman E, *Jess In Action: Rule-Based Systems in Java*, Manning Publications Co., 2003.
- [20] Guo W Y. Reasoning with semantic web technologies in ubiquitous computing environment. *Journal of Software*, 2008, 3(8): 27-33.
- [21] Hung J C. Forward of Special Issue on "Mobile System, Agent Technology, and Web Services". *Journal of Software*, 2008, 3(8): 1-2.



**Qiang Ge** was born in 1981, pursues his masters degree in Computer Science and Technology in Nanjing University of Aeronautics and Astronautics. His research interest include domain engineering, semantic Web and description logic.



**Guohua Shen** was born in 1976. He received his Ph.D. degree in Computer Software from Nanjing University of Aeronautics and Astronautics. Now he is an associate professor at Nanjing University of Aeronautics and Astronautics. His research interests include domain engineering, software metrics, semantic Web and description logic.



and software metrics.

**Zhiqiu Huang** was born in 1965. He receives his Ph.D. degree in Computer Software from Nanjing University of Aeronautics and Astronautics in 1999. Now he is a professor and Ph.D. supervisor at Nanjing University of Aeronautics and Astronautics, and the senior member of CCF. His research interests include software engineering



**Changbo Ke**, born in 1984, come from Ankang City Shaanxi Province, is a Ph.D candidate of Nanjing University of Aeronautics and Astronautics. Major research interests include security and privacy of information system and ontology-based software engineering.