

The design of an electronic pedigree system for food safety

Weili Han · Yun Gu · Wei Wang · Yin Zhang ·
Yuliang Yin · Junyu Wang · Li-Rong Zheng

© Springer Science+Business Media, LLC 2012

Abstract The problem of food safety is a critical issue in recent years. To address the issue, the technologies of the Internet of Things are used to offer the possibilities to easily track the processes in the production, storage, transportation, sale, and even using phases of foods. This paper, therefore, introduces the design of an electronic pedigree system for food safety, which uses electronic pedigrees to enhance the safety of food supply. The system implements an extension of the pedigree standard of EPCglobal, and offers a more trustworthily tracking service to monitor and supervise the production and supply of food. We discuss the key issues of the design, and implement a prototype to evaluate the feasibility of the design. Finally, we analyze the trustworthiness assurance and security of our electronic pedigree system.

Keywords Electronic pedigree · Food safety · Internet of things · Trustworthiness

1 Introduction

Food safety is a critical problem in China because of the food scandals in recent years. E.g., the “2008 Chinese milk scandal” (Xin and Stone 2007) is a serious food incident where melamine is added to infant milk powder. The added melamine led to more than 300,000 victims, especially

babies in a-few-month old. In 2011, the meat from Shuanghui Group, the China’s largest meat supplier, was detected to contain clenbuterol hydrochloride, a chemical forbidden in food. To address the food safety problem, the end consumers and governments require a faster and easier way to monitor the traces left by foods in the food supply chain (Muckstadt et al. 2001; Kumar et al. 2011; Li 2012; Zdravković et al. 2011).

The development of the Internet of Things (Zheng et al. 2011; Atzori et al. 2010) enables more and more physical objects to easily leave their traces in the cyber space. Similar to the concept of pedigrees in the physical world, the electronic pedigrees (EPCglobal 2007; Tan and Li 2006; Harrison and Inaba 2008), which consist of the traces of these objects, are also proposed. They can be verified in trusted ways, then be used to anti-counterfeit. Pedigrees in the physical world record the traces of physical objects, including human beings, usually in paper format. The pedigrees in paper format are usually signed and can be verified by authorities. For instance, as a typical segment of a pedigree, a certificate of a bachelor’s degree from Fudan University (as authority) can be signed to a person. Then an employer can verify the certificate on-line or off-line according to the stamp, signature, and even series number in the certificate.

The current trustworthy technologies (EPCglobal 2007; Tan and Li 2006) based on electronic pedigrees focus on the processes of supply of drugs rather than foods. In addition, these technologies mainly focus on the basic steps of supply chains, and pay less attention to the relationship between a certified object and other objects. For example, the current electronic pedigrees do not include the environment monitor data, such as temperature and humidity data. As a result, many key requirements cannot be met in the food supply chain management.

This paper, therefore, introduces the design of an electronic pedigree system for food safety, which basically

W. Han · Y. Gu · W. Wang · Y. Zhang · Y. Yin
Software School, Fudan University,
Shanghai, China

W. Han
e-mail: wlhan@fudan.edu.cn

J. Wang · L.-R. Zheng (✉)
School of Information Science and Engineering, Fudan University,
Shanghai, China
e-mail: lrzheng@fudan.edu.cn

follows the concept of IIIE (short for Industrial Information Integration Engineering) (Xu 2011), and comprises methods to collect, analyze, distribute, and utilize the information in order to offer services for users who are involved in the food industry. The contributions in this paper are as follows:

- 1) Our design extends the standard (EPCglobal 2007) of EPCglobal (now GS1) (GS1 2011a), and offers more featured functions to control the processes in production and supply of foods. These functions include: the *Initial Environment Pedigree* and *Environment Pedigree* which record the monitor data of the production environment; the *Transportation Pedigree* which records the sensing data in food transportation. This is particularly important when the fresh foods are transported. The *Processing Pedigree* which records the transaction data when raw foods are processed. The *Inspection Pedigree* which records the inspection data of foods. To the best of our knowledge, these featured functions are firstly applied in the food supply chain management.
- 2) Our design introduces a master–slave architecture for electronic pedigree management to meet the storage and search requirements of the massive pedigrees. In the architecture, the slave sides are tightly coupled with business systems by using web services technologies. They create, verify electronic pedigrees according to business demands; and the master side is an independent service. It coordinates the slave sides, such as registering the slave sides, allocating serialization codes for the slave sides, processing the massive electronic pedigrees, and offering the search and discovery services of electronic pedigrees for end consumers.
- 3) We analyze the trustworthiness assurance in our electronic pedigree system. The analysis shows that our design can assure more trustworthiness than previous works.

The rest of the paper is organized as follows: section 2 introduces the background of electronic pedigree and related work; section 3 introduces the extension of the EPCglobal standard (EPCglobal 2007) to meet the demands of food safety; section 4 introduces the master–slave architecture of the electronic pedigree system; section 5 discusses the key issues in the electronic pedigree system; section 6 analyzes the trustworthiness assurance and security of our electronic pedigree system; finally, section 7 summarizes our work and introduces our future work.

2 Related work

The technologies, including body sensors, of the Internet of Things (Atzori et al. 2010) offer a more intuitive, simpler view of the movement of objects, including patients

(Domingo 2012) and Taxi (Pan et al. 2012). These technologies offer a chance to track and trace goods movement in productions and supply chain (Xu 2011; Kumar et al. 2011), including food supply (Gu and Jing 2011; Yin et al. 2011). The tracking and trace services for food supply can also enhance the safety of food, because the end users can easily see the whole process from production to sale. However, what they see could be forged. That is, the malicious suppliers can forge the goods and information at any phase in the food supply chain. The concept of electronic pedigree (EPCglobal 2007) would help us resolve this problem.

The concept of electronic pedigree was initially proposed to anti-counterfeit in the drug network (EPCglobal 2007; Tan and Li 2006; Thompson 2004; Kwok et al. 2008; Lehtonen et al. 2007). Because of the threat of drug counterfeit, the EPCglobal Healthcare and Life Sciences Pedigree Task Force, together with other contributions of schema from Cyclone Commerce, Raining Data and Verisign, proposed a standard (EPCglobal 2007). The motivation of the works partly comes from the laws whereby Florida and California require the companies to track and trace their products as the products move through the supply chain. These researches usually record the basic data of suppliers and receivers, and use the digital signature to ensure the integrity of the data.

Afterward, electronic pedigree was used in manufacturing (Kwok et al. 2008), where Kwok et al. leveraged the technologies of electronic pedigree and RFID (short for Radio Frequency Identification) to deliver global and accurate supply chain visibility to the processes of storage and materials transportation.

The RFID technology is usually used as an infrastructure in the electronic pedigree systems (Zheng et al. 2011; Tan and Li 2006; Kwok et al. 2008; Meng et al. 2010; Kumar et al. 2011). A tag which is compatible with RFID standards stores an identifier code, such as an EPC code, and is attached to an object. Once the movement of tag is detected, an EPCIS (GS1 2011b) (short for EPC Information Service) system will process the event, and would create the electronic pedigree segment according to pre-defined policies. Furthermore, Tan et al. (Tan and Li 2006) proposed schemes to store the information of electronic pedigrees to RFID tags.

However, our paper does not focus on storing the information of electronic pedigrees into RFID tags though these RFID-enable methods can be applied in our electronic pedigree system. Furthermore, our work focuses on the processes of food production and supply, which are missed in the above works. First, our work focuses on the production process and record the environment data in electronic pedigrees; second, our work considers the complex processes, e.g., transportation, inspection; third, our work discusses how to coordinate the electronic pedigree systems and

serialization modules in the electronic pedigree systems and how to store and manage the massive electronic pedigrees.

3 The extension of the electronic pedigree standard

3.1 Lifecycle of an electronic pedigree

As is shown in Fig. 1, the lifecycle of an electronic pedigree for food safety consists of multiple mappings of six key steps:

Generate: There are two ways to generate an electronic pedigree in our system, which are the initialization and nested generation. The former way means that an electronic pedigree is generated without dependence on other electronic pedigrees. But other electronic pedigrees can be linked to the initialization generated pedigree by using their pedigree ID. The latter means that an electronic pedigree is generated on the basis of another electronic pedigree, which means this kind of electronic pedigrees, contain some other electronic pedigrees, including initialization and nested generated electronic pedigrees.

For an electronic pedigree, in which way it should be generated depends on the pedigree’s type. In general, *Initial Environment Pedigree*, *Initial Pedigree*, *Birth Pedigree*, *Repacking Pedigree* and *Processing Pedigree*, which are described in section 3.2 and Appendix A, are generated in the initialization way, whereas the rest types of pedigrees in our electronic pedigree system are usually nested generated. Note that, the latter type may not be nested generated if exceptions happen. For instance, if we lose the build-up information of an environment, we still need record the data of the environment, which leads to an environment pedigree generated in the initialization way.

Import: When food products are transported from one company to another, their electronic pedigrees should be transmitted simultaneously. There are two ways to

import electronic pedigrees, which are one-by-one importation and enveloped importation. Here, the enveloped importation means that several electronic pedigrees can be enveloped as a whole file and be imported together.

Export: When products are shipped or transported from one company to another, their electronic pedigrees should be transmitted as well. So the system must be capable of exporting electronic pedigrees to storage devices. There are two ways to export electronic pedigrees, which are one-by-one exportation and enveloped exportation, as well.

Upload: the EPSServs (short for Electronic Pedigree Sub Server, a slave) need upload the electronic pedigrees to the CEPSEv (short for Center Electronic Pedigree Server, a master) in order to support the CEPSEv to deal with the local query, which is proposed in section 4.2. Acquiescently, after the EPSServ imports electronic pedigrees, it can upload them to the CEPSEv. The electronic pedigrees saved in the EPSServs are uploaded in order to maintain the accordance with those saved in the CEPSEv so that the results of the local query can keep fresh.

Verify: verification should be applied to the electronic pedigrees to ensure the trustworthiness of them. Generally, electronic pedigrees are verified when they are imported to the electronic pedigree system or when a user launches a request to query information about a product.

We verify an electronic pedigree by the digital signatures attached to it. We firstly fetch the relevant public key in local database. Then, we use the public key to verify the validation of the signature. If the digital signature on the electronic pedigree is validated to be true, then we determine the information in the electronic pedigree can be trusted. Otherwise, we believe that the information could be counterfeited and untrustworthy. Note that, we will verify the signatures one by one according to electronic pedigree’s nested layers.

Query: A user can launch a query to see a certain electronic pedigree through a web interface provided by the CEPSEv. The system supports three kinds of search keys, which are ① *product code type + product code + item serial number*, ② *product code type + product code + lot number* and ③ *electronic pedigree’s identifier*. When the queried electronic pedigrees are found, the CEPSEv will verify the electronic pedigrees at first, and then combine the verification results and the electronic pedigrees to the user.

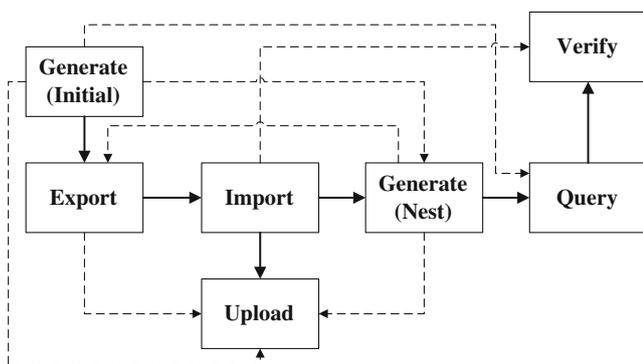


Fig. 1 Lifecycle of an electronic pedigree

Figure 1 illustrates the concrete mappings in the lifecycle of an electronic pedigree. The solid lines in Fig. 1 show the key path of the lifecycle while the dotted lines show the

operations are optional. At the beginning of the lifecycle, an electronic pedigree is initialized. This step occurs when the product is produced in the factory. Next, the electronic pedigree is exported from the company and imported to another company. During the import process, uploading the electronic pedigree to the CEPSEv is required. After the importation, new electronic pedigrees may be nested generated in the new system. For example, when the factory transports the product to an inspection and quarantine organization, a new inspection pedigree will be generated by nesting the old electronic pedigree. Then, when a product arrives at a customer, the customer can query the electronic pedigree of this product. The CEPSEv will receive the query from the user and search the electronic pedigree in its local database. If there is such an electronic pedigree, the CEPSEv will verify it and return the electronic pedigree and its validation to the user.

3.2 Types of electronic pedigree in the extension

Our design extends the types in the standard (EPCglobal 2007). Seven new types of electronic pedigree are proposed, which are *Initial Environment Pedigree*, *Birth Pedigree*, *Processing Pedigree*, *Environment Pedigree*, *Transportation Pedigree*, *Unsigned Transportation Pedigree*, and *Inspection Pedigree*. As the final result, there are thirteen kinds of pedigrees in our electronic pedigree system. They can be divided into two groups according to the ways they generated, which are shown in Fig. 2.

The initialization group, where electronic pedigrees are usually generated without dependence of other electronic pedigrees, has six types of electronic pedigrees (*Initial Pedigree*, *Repacking Pedigree*, and *Alt Pedigree* are introduced in Appendix A):

Initial Environment Pedigree: This kind of electronic pedigree is used to describe the initial data of production environment. When a production environment is found, its initial environment pedigree needs to be generated. For example, when a pigpen is established, its location, capability, foundation time and other initial environment information should be recorded to generate an initial environment pedigree.

Birth Pedigree: This kind of electronic pedigree is used to describe the birth information of the product. When several products are produced, there may be only one birth pedigree generated since their birth information is same. For example, when two batches of pork chops are produced by one pig, only one birth pedigree of the pig is available. For fruits and vegetables, a birth pedigree includes seed condition, its previous generation and other birth information. For livestock, birth pedigree includes products' parent information and other birth information.

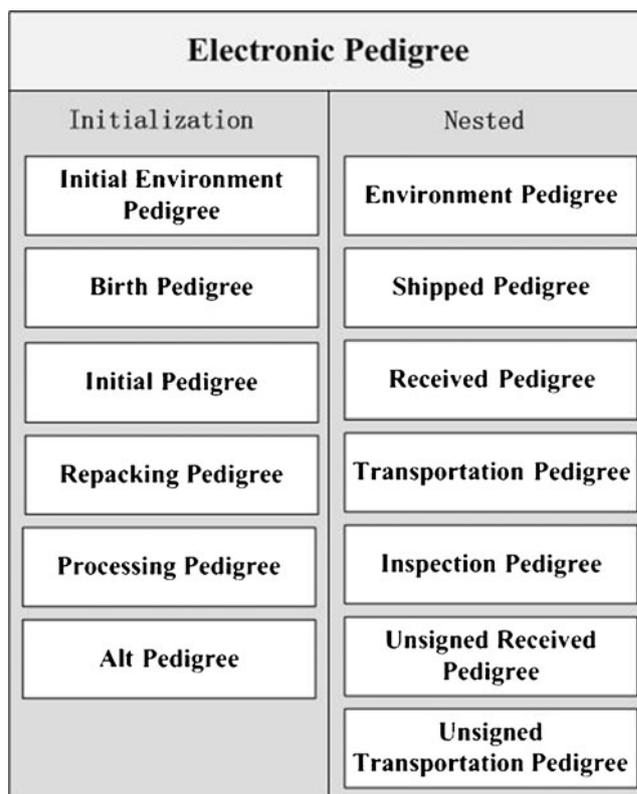


Fig. 2 Types of electronic pedigree

Processing Pedigree: This kind of electronic pedigree is used to record the processing information of products. Additives, ingredients, time and other process information should be recorded. When a processing pedigree is generated, the previous electronic pedigrees need to be related to it by their identifiers. So when a user asks for the product information, we can trace the previous information by the previous electronic pedigree identifiers.

In the nested group where electronic pedigrees are generated on the basis of other electronic pedigrees, there are seven types of electronic pedigrees (*Shipped Pedigree*, *Received Pedigree*, and *Unsigned Received Pedigree* are introduced in Appendix A):

Environment Pedigree: This kind of electronic pedigree is used to describe the changing environment. At set intervals, an environment pedigree needs to be generated to record the changing environment. The environment pedigree includes temperature, humidity, air quality and other environment information. Another environment pedigree or an initial environment pedigree may be embedded in the environment pedigree. An environment pedigree is often related to other electronic pedigrees to show the environment of that time.

Transportation Pedigree: This kind of electronic pedigree is used to record the condition in the transportation. Since the food may go bad during transportation, the transportation conditions need to be recorded. A transportation pedigree includes temperature, accelerated speed, and humidity. Transportation data are usually recorded at intervals during the transportation. When the transportation finishes, the transportation pedigree will be generated.

Unsigned Transportation Pedigree: This kind of electronic pedigree is similar to transportation pedigree, except that it has not a digital signature. This electronic pedigree is used when transportation companies are not able to generate electronic pedigrees in special conditions. For example, the transportation transaction is taken over by the company who receives goods. As a result, the unsigned transportation pedigree will be included into a received pedigree, and signed.

Inspection Pedigree: This kind of electronic pedigree is used to record the inspection information of the product. The inspection information includes time and inspection data.

Among these new pedigrees, the Unsigned Received Pedigree and Unsigned Transportation Pedigree do not contain a digital signature while other types of electronic pedigrees must have digital signatures. Note that, all types of electronic pedigrees have a pedigree unique identifier and relative information in it. Nested electronic pedigrees usually contain other electronic pedigrees while initialized electronic pedigrees do not.

4 Architectures of the electronic pedigree system

4.1 Master–slave architecture

The electronic pedigree system uses the master–slave architecture, which is shown in the Fig. 3. The slaves, referred as to EPSServs, are the subsystems which integrated with companies' business systems. They are mainly responsible for creating, verifying, importing, exporting and uploading electronic pedigrees. It is worthwhile to note that an EPSServ can upload its electronic pedigrees to the CEPSServ only after the EPSServ has registered in the CEPSServ. The EPSServ integrates with an EPCIS server and has a serial number management component, which is in charge of generating unique serial numbers used for electronic pedigrees.

The master in our architecture is the CEPSServ, which is mainly responsible for treating customers' query and search the massive electronic pedigrees. After the slaves (EPSServs) have registered in the website offered by the CEPSServ, the CEPSServ can coordinate the slaves, such as allocating the

serialization codes in the slave sides and processing the massive electronic pedigrees. The CEPSServ has a massive data storage to save the electronic pedigrees, and can achieve an efficient search.

4.2 Architecture of the master server

As is shown in the Fig. 4, the CEPSServ, which is the master server, has a three-layer architecture. The bottom layer is the file server and database which store the massive electronic pedigrees and relevant structural information of EPSServs. The middle layer is the function modules dealing with the search services, electronic pedigree management and EPSServs coordination. The top layer is the web interface for customers and the system interface for EPSServs.

The bottom layer is responsible for data storage. It has two types of data, which are massive electronic pedigrees and structural information of the EPSServs registered in the CEPSServ. The massive electronic pedigrees are stored in the file server while the EPSServs' information is stored in a database.

Massive data (Electronic pedigrees) Billions of or even trillions of electronic pedigrees need to be stored in the CEPSServ if the electronic pedigree system is applied to the whole food market. So the massive storage and processing technologies must be adopted in our design. The massive electronic pedigrees are stored in distributed servers. Besides, the storing technology is able to distribute the storage load of the massive data in balance among multiple servers.

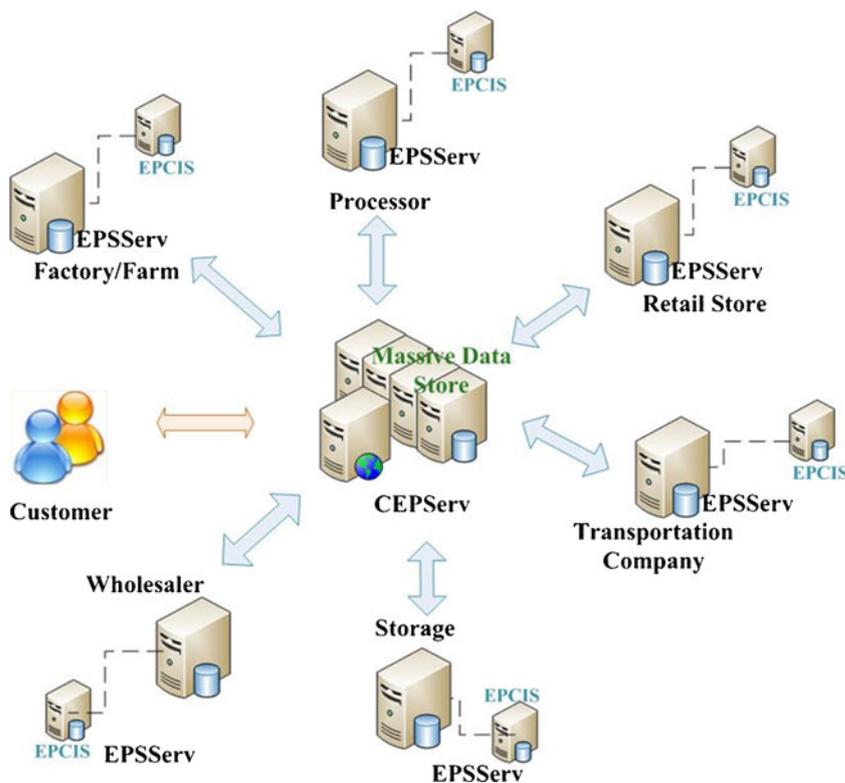
EPSServ info After an EPSServ registers in the CEPSServ, its basic information such as company name, tax number, corporate representative, address and other company information should be recorded. At the same time, the CEPSServ will send an authentication code to the EPSServ via a secure channel. The EPSServ needs to use this authentication to update its company information and upload electronic pedigrees.

The middle layer is a key layer of the architecture. It is responsible for managing the data in the bottom layer and offering services for the top layer. The components in this layer can be separated into three parts according to their functions.

The first part is in charge of the management of massive data. There are two components in this part.

Indexing Since there are massive electronic pedigrees stored in the distributed file system, a fine indexing mechanism is needed to find the exact electronic pedigree efficiently. When the CEPSServ stores the electronic pedigrees, it records the physical address of the electronic pedigrees'

Fig. 3 Architecture of the electronic pedigree system



store position, and set up an indexing structure for searching. When we want to find an electronic pedigree, we can use its identifier and the indexing component to obtain the electronic pedigree.

Receive electronic pedigrees The EPSServs will upload electronic pedigrees to the CEPserv in order to help the CEPserv run its local search, so the CEPserv needs a component to receive electronic pedigrees. Before the CEPserv receives electronic pedigrees from an EPSServ, the EPSServ needs to offer its authentication code to show its validity. Then the EPSServ can upload its electronic pedigrees.

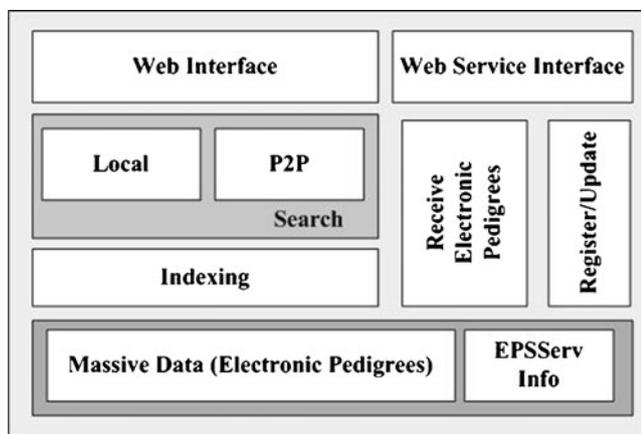


Fig. 4 Architecture of the CEPserv

The second part is in charge of the search. There are two components:

Local Search: Since the EPSServs have already uploaded some electronic pedigrees to the CEPserv, the local search is feasible. End users input the search keywords to get the corresponding electronic pedigrees' identities. Then the CEPserv uses the indexing mechanism to get electronic pedigrees with these identities efficiently.

P2P Search: This search method help end users executes retrievals in the EPSServs. The CEPserv analyses the search keywords and gets the corresponding electronic pedigrees' identities. Then, it sends these identities to all the registered EPSServs, the EPSServs will have a search in their own file servers and return the matched electronic pedigrees to the CEPserv. The CEPserv receives the electronic pedigrees, verifies, and then show them to the end users. The P2P search can offer more efficiency than the local search when a user wants the timely and integrated search results.

The third part is in charge of the information coordination (uploading electronic pedigrees is excluded) between CEPserv and the EPSServs. There is only one component in this part.

Register/update When an EPSServ wants to register in the CEPserv, it submits a register application. Then, after it fills in its company information, the CEPserv will send an authentication code to it via a secure channel.

The top layer has two interfaces: a web interface which processes the query of customers, and a system interface which coordinates with the EPSServs.

Web interface The interface will call the search component to run the search and return the hit electronic pedigrees and their validation condition to the customers. By using this interface, end users can input the keywords to have a search in the web interface. There are three kinds of keywords, which are described in 3.1. The keyword *product code type + product code + item serial number* is used to get an item's related electronic pedigrees. The keyword *product code type + product code + lot number* is used to get the information of the products with the given lot number. The third type of keyword *electronic pedigree's identifier* is used to get the electronic pedigree with the given identifier. The customers can choose the search methods by inputting different keywords. Besides, they can choose local search P2P search.

Web service interface The web service interface is related to the Receive Electronic Pedigrees and Register/Update components. It offers the interfaces for the EPSServs to use the two components.

4.3 Architecture of the slave server

An EPSServ, which is the slave server, consists of three modules as are illustrated in the Fig. 5. The three modules are the Storage, Information Module and Electronic Pedigree Module.

The Storage has two components:

Info Database: This database stores the three kinds of basic information in the system. Those are staff

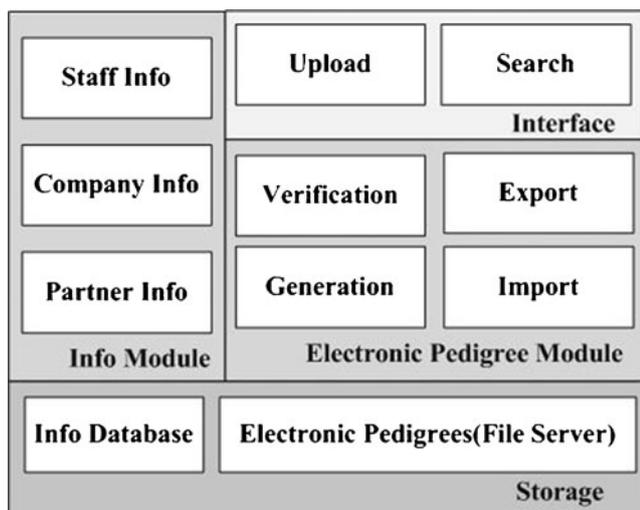


Fig. 5 Architecture of the EPSServ

information, company information and partner information. All these information need to be used in the operations of the Information Module and Electronic Pedigree Module.

Electronic Pedigree File Server: Our design uses file systems to store the electronic pedigrees. There are indexing tables in the database to record the mapping from the pedigree identifier to the physical address of the electronic pedigree's location.

Furthermore, the Information Module is composed of three components:

Staff Info: Staff Information Management includes the adding, removing and modifying the staff information of the system. The staff information includes the staff's identifier, name, password, address, telephone number and his or her role. It is worthwhile to note that the role refers to the set of rights of the staff to operate on the electronic pedigrees in this system.

Company Info: Company Information Management includes the initialization and modification of the company information. Only the administrator has right to operate on the company information. When the EPSServ is first used, the administrator must fill in the company information at first. Otherwise, other functions of the EPSServ are forbidden. Note that the old company information will be reserved instead of being deleted if it is modified.

Partner Info: Partner Information Management includes adding, removing and modifying the partner information. The partner information refers to the basic information of the companies that have trade with the present company. The Basic Information usually includes the company name, register number, corporate representative, address, telephone number and the public key certificate, which is used to verify the digital signature of the partner.

The Electronic Pedigree Module is the key part of the EPSServ. It is responsible for the functions related to the electronic pedigrees. As are described in section 3.1 in detail, the functions include generation, verification, import, export, upload and search. Among them, upload and search functions are two interfaces offered to the CEPSServ.

4.4 Prototype

The electronic pedigree system's web interfaces are developed in the SSH (Struts + Spring + Hibernate) framework with Eclipse 3.5. Besides, MySQL is used as the database to store the index of electronic pedigrees, and other system data. Hadoop is used to store massive electronic pedigrees. Figure 6 is the portal page in the CEPSServ, in which end

食品安全电子履历系统

Fig. 6 Electronic pedigrees query page of CEPSServ

users can input the keywords to query the corresponding electronic pedigrees.

5 Key issues in the design

5.1 Serialization coordination

The electronic pedigree system has a serial number management component, which is the basic component to identify an object and to generate unique serial number so that the CEPSServ can use this number to have retrieval.

In the system, four types of numbers need to be managed, which are pedigree identifier, item serial number, document serial number and envelope serial number. In each type, two identical serial numbers are forbidden.

Table 1 shows the serialization format for the four types of serial numbers.

The *pedigree identifier* is composed of the EPSServ’s name, the electronic pedigree’s type and a twelve-digit number. The EPSServ’s name uniquely identifies the company globally. Both the EPSServ’s name and the electronic pedigree’s type are English words or their abbreviations. The three parts of the pedigree identifier are separated by underlines. For example, if an EPSServ’s name is ‘fudan’ and it generates a shipped pedigree, the serial number management component will allocate Fudan_SP_000000000001 to the new shipped pedigree if no shipped pedigree exists in the database before. The twelve-digit number increases one by one every time a pedigree of the same type is generated in the same system.

Table 1 Serialization format

Serial number type	Format	Example
Pedigree identifier	EP_Type_12Number	Fudan_SP_000000000001
Item serial number	EPNum + TypeNum + 12Number	000101000000000003
Document serial number	EPNum + TypeNum + 12Number	000103000000000024
Envelope serial number	EPNum + TypeNum + 12Number	102109000000000221

The *item serial number* is composed of the EPSServ’s globally unique number, the corresponding number of an electronic pedigree’s type and a twelve-digit number. All the three parts of the item serial number are numbers and there is no separate among them. For example, we suppose an EPSServ’s number is 0001 and the pedigree is an initial environment pedigree. The item serial number can be 000101000000000003 if there is already an initial environment pedigree with a serial number 000101000000000002 in the database.

The formats of *document serial number* and *envelope serial number* are the same as that of item serial number. They are composed of the EPSServ’s globally unique number, the corresponding number of an electronic pedigree’s type and a twelve-digit number.

Note that, there may be duplicate numbers among the four types of serial numbers, which means a serial number such as 000101000000000001 may be a pedigree serial number, a document serial number and an envelope serial number at the same time.

Table 2 describes the abbreviation and numbers for all types of electronic pedigrees used in the electronic pedigree system.

5.2 Consistency of electronic pedigrees

The electronic pedigree system will face some inconsistent problems because of the exceptions, such as the mis-operation and delay. In order to keep the consistency of the system, measures are taken to overcome these problems.

When more than one verified electronic pedigrees are found for one event, we choose to trust the latest one. This rule is always helpful to solve the problem caused by staff’s mis-operation. For example, a staff member generates a shipped pedigree in which the receiver is company A. But it happened that the receiver changes before the product are shipped out. So the staff member generates another shipped pedigree in which the receiver is company B. In this condition, there are two conflict pedigrees when customer wants to trace the shipped process of the product. When such a problem occurs, we choose to trust the pedigree with the latest time. In this case, the second pedigree in which the receiver is B will be chosen.

Besides the mis-operation, delay is another problem which can be solved by the EPCIS servers. The delay is a

Table 2 Abbreviation and numbers for all types of electronic pedigrees

Electronic pedigree type	Abbreviation	Corresponding number
Initial environment pedigree	IE	01
Environment pedigree	EN	02
Birth pedigree	BR	03
Initial pedigree	IN	04
Inspection pedigree	IS	05
Repacking pedigree	RP	06
Transportation pedigree	TS	07
Shipped pedigree	SP	08
Received pedigree	RC	09
Processing pedigree	PC	10
Unsigned received pedigree	UR	11
Unsigned transportation pedigree	UT	12
Alt pedigree	AT	13

common condition in the electronic pedigree system because electronic pedigrees usually reach a company faster than products. We have to guarantee that the follow-up operations on the electronic pedigrees are consistent with the facts. For example, when an electronic pedigree is sent to a company but the corresponding product has not reached, the staff cannot generate a received pedigree based on the electronic pedigree. In order to face this problem, we need integrate EPSServs with the EPCIS servers. When the electronic pedigrees reach but the products do not, the electronic pedigree system would not let the staff see the arrival of the electronic pedigrees. Only after the products have arrived and the EPCIS servers has recorded the information, the staff can see the electronic pedigrees of the products and the EPCIS servers will offer the information to help the staff generate the received pedigrees.

5.3 Integration with EPCIS servers

The EPSServs are integrated with the EPCIS servers, which enable disparate applications to leverage Electronic Product Code (EPC) data via EPC-related data sharing, according to the concept of IIIIE (Xu 2011). In our design, major information, such as temperature, time and location, is gained from the EPCIS servers to create electronic pedigrees.

The EPCIS servers can offer standardized information, because a lot of information needs to be recorded when the system generates an electronic pedigree. If the company forces a staff member to fulfill all the information, mistakes are likely to be made because of carelessness. Besides, every staff member has different preferences, so the same thing may be recorded in different forms. However, the EPCIS systems automatically input the information and have a

standard for the forms of the information. Then all the electronic pedigrees can be generated in a unified format and the electronic pedigrees can be generated conveniently.

In addition, the EPCIS servers help to save the time for the generation of electronic pedigrees. If the EPSServs let a staff fulfill an electronic pedigree, it will take several hours for them to finish it even if they are experienced. Thus, entering data by a human is expensive. However, it may cost only several seconds when the EPSServs are integrated with the EPCIS servers.

We use the event driven method for the EPSServ with the EPCIS. We design a plugin of the EPCIS to monitor the modification of relevant data, and create an integration event which includes relevant information. Then the EPCIS will send the event to the EPSServ, and the EPSServ create an electronic pedigree.

5.4 Store and process massive data in the master server

The electronic pedigree system uses the distributed file system to store the massive data in the master server. When an electronic pedigree needs to be stored, the storing location depends on its pedigree identifier, which is defined in section 5.1. For example, if there are sixteen file servers: $f_0, f_1, f_2, \dots, f_{15}$. The pedigree identifier of the electronic pedigree needed to be stored is Fudan_SP_000000000072. We use the twelve-digit number to mod the number of file servers, which means $72 \bmod 16$. Then the result is 8. So the electronic pedigree should be stored in f_8 . After the electronic pedigree is stored in a file server, the mapping between the pedigree identifier and the physical address of its location should be recorded.

When a search query approaches, the system uses the pedigree identifier as the keyword to find the physical address in the indexing table. Then, it uses the address and pedigree identifier to get the electronic pedigree.

If some of the file servers have to be removed, we move the electronic pedigrees on them to the rest file servers. We just need store these electronic pedigrees as new ones. We use their pedigree identifiers to mod the new number of the file servers, and then decide the new location of the electronic pedigrees.

6 Discussion

6.1 Trustworthiness assurance

This section analyzes the trustworthiness assurance of our electronic pedigree method based on comparison with other types of tracking methods. We argue that our electronic method has more advantages than other types of *tracking* methods, especially for cross-organization

applications. These advantages are the main contributions of our paper.

Many companies use tracking systems based on RFID to track object (Zheng et al. 2011). However, the trustworthiness of the data stored in the information system is not ensured. That is, the data can be modified or deleted due to multiple reasons, including counterfeit. This is a critical issue for the drug supply, and is with food safety.

As is shown in Fig. 7, the electronic pedigree offers trustworthiness assurance for the movement of an object in the Internet of Things. That is, the integrity of the data transmitted into cyber space will be ensured by using digital signatures.

According to the comparison with the standard of EPCglobal (EPCglobal 2007), the design of our electronic pedigrees offers more completeness to cover the most phases in the food production, transportation, and sale. As is shown in Fig. 7, the light blue circles are used to describe the trustworthiness assurance by the standard of the EPCglobal, which cannot ensure the integrity of the data in the phases of production, processing, inspection and transportation. These phases are very important for food supply. Thus, we design more electronic pedigree types to ensure the food safety.

The green circles, therefore, are used to describe the trustworthiness assurance given by our electronic pedigree system. Obviously, our design covers the most phases of the food supplies.

6.2 Vulnerabilities analysis

Because our electronic pedigree system depends on the strength of PKI (short for public key infrastructure), all vulnerabilities of PKI (Boldyreva et al. 2007; Shon and Choi 2007) especially the cryptographic issues are also exist in the system. Our design depends on the security of digital signatures. Once the digital signature schemata are broken, an attack can fabricate a same signature with different content. Then the security of our design will be broken. In addition, the security of algorithm and key size are two other issues in our system. The algorithm, e.g. RSA, and

key size, e.g. 1024 will determine the security level of our design. Last but not least, the key management, especially private key management, is a key issue in our system. The trustworthiness boundary is determined by the active range of the private key, once a signed electronic pedigree is generated. That is, when an external system, or an external software module, or an unauthorized person can access the private key, the digital signature may be forged. As a result, we must ensure that the private key is accessed in as small a range as possible.

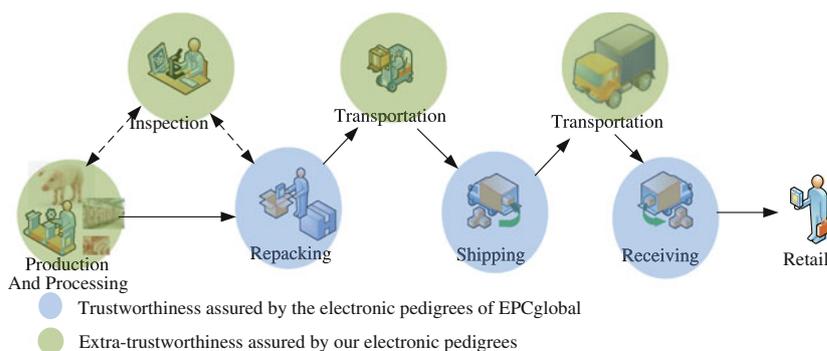
In addition, a physical link between a real object and a sensor or a tag is another vulnerability. It is a special issue in the applications of the Internet of Things. Usually, the tracking data in electronic pedigree will be gathered through an EPCIS system, where a RFID tag will be tightly attached to an object, and a RFID reader can read the tag to create the movement event for the EPCIS system. But if the tag is transferred from one object to another object without any trace (Mitrokotsa et al. 2010), the recorded data in the electronic pedigree will be inconsistent with the truth. Thus, the techniques to tightly attach a tag to an object are very important for our electronic pedigree system.

Next, the vulnerabilities of RFID and sensors can affect the security of our design. Due to the computation limitation of RFID and sensors, the used ciphers and authenticate protocols are usually light-weight, even ultra-light weight. The security strength is usually traded off due to the limitation. Thus, the data from the RFID tags or sensors might be modified without authorization.

Finally, although our system provides more trustworthiness assurance as is shown in section 6.1, the assurance requires more cost. We must design and deploy the relevant EPSServs to the nodes where the trustworthiness will be ensured. The operators should be well trained. In addition, more trustworthiness assurance will lead to long electronic pedigrees, where includes many phases of the processes of production and supply of foods. To address this issue, a friendly UI to display electronic pedigrees could be a good solution.

As a countermeasure to the above vulnerabilities, we can propose a risk adaptive mechanism (MITRE 2004) or other

Fig. 7 Trustworthiness assured by the electronic pedigrees



security evaluation approach (Yan 2008) to ensure the food safety. The risk adaptive mechanism will gather massive data in the Internet of Things, use a quantified risk model to assess the risk of an object or a sensitive operation, activate one or more risk mitigation action(s) to mitigate the risk, and make a decision based on the measured risk.

6.3 Conclusion and future work

This paper introduces an extension design of electronic pedigree system for food safety. The system, which adopts the master–slave architecture, uses digital signatures to ensure the trustworthiness of electronic pedigree. And the system designs new electronic pedigree types to meet the requirements of the food safety. Comparing with the traditional tracking systems without the support of electronic pedigrees, our electronic pedigree system offers the trustworthiness assurance for the generation of data in the Internet of Things. Furthermore, comparing with the standard of EPCglobal, our electronic pedigree system offers more trustworthiness assurance. As a result, the proposed system is more suitable for safe food supply.

We are trying to deploy the system in our project, to evaluate the system availability. The project is supported by the ministry of science and technology of the People's Republic of China, and focuses on topics of the agricultural Internet of Things and food safety.

Our future work includes three aspects: First, from the aspect of the design methodology, we will leverage Multidisciplinary Design Optimization (MDO) (Li and Liu 2012), which is a suitable framework for designers geographically dispersed to work together, to meet the requirements of the development of the electronic pedigree system. The framework can guarantee the development with fewer conflicts. Second, from the aspect of privacy protection, the current design only considers a little privacy issue. The master of the electronic pedigree system can monitor all sensitive data of all involved users, including manufacturers, wholesalers, retailers, and end-users. A privacy-preserved framework based on the blind signature technology (Sun et al. 2005) can help our system to protect the privacy data. And privacy aware security policies (Han and Lei 2012) can guide the involved users to protect their owned privacy data. Third, from the aspect of system design, we will extend the design of CEPSErv to support more analysis on massive data. Because the current design can only support the storing and retrieval according to the series numbers or electronic pedigree identifier, but cannot support the data mining and other more complex massive data processing. And we plan to use Cloud Computing Technologies (Li et al. 2012) as our future supporting technology. Furthermore, we will more leverage the concept of IIIIE to provide a next-level highly sophisticated trustworthily tracking system, where ERP

(Entire Resource Planning) (Xu 2011) is one of the most important features.

Acknowledgment This paper is supported by the 863 project (Grant NO: 2011AA100701), the Project-sponsored by SRF for ROCS, SEM, and the project of Natural Science Foundation of Shanghai (Grant NO: 12ZR1402600). The corresponding author is Lirong Zheng. The documents can be viewed at <http://crypto.fudan.edu.cn/epedigree/homepage.html>. Thank Ms. Min Li for her English polish.

Appendix A: Pedigree types in the previous pedigree standard of EPCglobal

Initial pedigree This kind of electronic pedigree is used to describe the initial condition of the products. When a product is produced, its initial pedigree should be generated. Initial pedigree includes a product's serial number and its initial information. Environment pedigrees and birth pedigrees can be related to initial pedigrees, which record the initial environment and birth condition of the product.

Repacking pedigree This kind of electronic pedigree is always used to record the repacking information. When a product is split into several smaller products or several products are merged into one product, a repacking pedigree is required. Repacking pedigree includes the repacking information such as repacking time and location. Besides, the previous pedigree's identifier should be related to the new repacking pedigrees so that we can trace the complete product information through it.

Alt pedigree This kind of electronic pedigree is used to link accessories. Other electronic pedigrees can only contain word information. But sometimes pictures, audio or video are necessary. Besides, in some cases, certificates with much data need to be related into an electronic pedigree while the amount of information in an electronic pedigree is limited. So we use the alt pedigrees to link the accessories.

Shipped pedigree This kind of electronic pedigree is used to record the shipping information. When the present company ships the products to another company, shipped pedigree should be generated. The shipped pedigree includes shipping time, location, receiver and other shipping information.

Received pedigree This kind of electronic pedigree is used to record the receiving information. When the company receives products, their received pedigrees should be generated. The received pedigree includes receiving time, location, recipient and other receiving information.

Unsigned received pedigree This kind of electronic pedigree is similar to the received pedigree, except that it has not a digital signature. In some cases, the product is refused by

other companies for some reasons after shipment. It means the company will generate a received electronic pedigree based on a shipped pedigree which is generated by the same company. So the digital signature is redundant since the product is always under the control of the present company. In this case, we may generate an unsigned received pedigree instead of a received pedigree.

References

- Atzori, L., Iera, A., & Morabito, G. (2010). The internet of things: a survey. *Computer Networks*, 54, 2787–2805.
- Boldyreva, A., Fischlin, M., Palacio, A., & Warinschi, B. (2007). A closer look at PKI: security and efficiency. *PKC 2007 (LNCS 4450)*, 458–475.
- Domingo, M. (2012). An overview of the internet of things for people with disabilities. *Journal of Network and Computer Applications*, 35, 584–596.
- EPCglobal. (2007). Pedigree ratified standard, 2007, http://www.gs1.org/gsm/kc/epcglobal/pedigree/pedigree_1_0-standard-20070105.pdf. Accessed April 2012.
- GS1. (2011a). The global language of business, 2011, <http://www.gs1.org/>. Access April 2012.
- GS1. (2011b). EPCIS - EPC Information Services Standard, 2011, <http://www.gs1.org/gsm/kc/epcglobal/epcis>. Access April 2012.
- Gu, Y., & Jing, T. (2011). The IOT research in supply chain management of fresh agricultural products. In: *Proceedings of the 2nd international conference on Artificial Intelligence, Management Science and Electronic Commerce (AIMSEC) (7382–7385)*.
- Han, W., & Lei, C. (2012). A survey on policy languages in network and security management. *Computer Networks*, 56(1), 477–489.
- Harrison, M., & Inaba, T. (2008). Improving the safety and security of the pharmaceutical supply chain. *Networked RFID Systems and Lightweight Cryptography, III*, 223–246.
- Kumar, S., Kadow, B., & Lamkin, M. (2011). Challenges with the introduction of radio-frequency identification systems into a manufacturer's supply chain – a pilot study. *Enterprise Information Systems*, 5(2), 235–253.
- Kwok, S., Tsang, A., Ting, J., Cheung, W., & Cheung, B. (2008). An intelligent RFID-based Electronic Anti-Counterfeit System (InRECS) for the manufacturing industry. In: *Proceedings of the 17th world congress the international federation of automatic control*. Seoul, Korea.
- Lehtonen, M., Michahelles, F., & Fleisch, E. (2007). Trust and security in RFID-based product authentication systems. *IEEE Systems Journal*, 1(2), 129–144.
- Li, L. (2012). Effects of enterprise technology on supply chain collaboration: analysis of China-linked supply chain. *Enterprise Information Systems*, 6(1), 55–77.
- Li, L., & Liu, J. (2012). An efficient and flexible web services-based multidisciplinary design optimisation framework for complex engineering systems. *Enterprise Information Systems*, 6(3), 345–371.
- Li, S., Xu, L., Wang, X., & Wang, J. (2012). Integration of hybrid wireless networks in cloud services oriented enterprise information systems. *Enterprise Information Systems*, 6(2), 165–187.
- Meng, S., Chiu, D., Kafeza, E., Wenyin, L., & Li, Q. (2010). Automated management of assets based on RFID triggered alarm messages. *Information Systems Frontiers*, 12, 563–578.
- MITRE. (2004). Horizontal Integration: broader access models for realizing information dominance. *JASON Report*, JSR-04-132.
- Mitrokotsa, A., Rieback, M., & Tanenbaum, A. (2010). Classifying RFID attacks and defenses. *Information Systems Frontiers*, 12(5), 491–505.
- Muckstadt, J., Murray, D., Rappold, J., & Collins, D. (2001). Guidelines for collaborative supply chain system design and operation. *Information Systems Frontiers*, 4, 427–453.
- Pan, G., Qi, G., Wu, Z., Zhang, D., & Li, S. (2012). Land-use classification using taxi GPS traces. *IEEE Transactions on Intelligent Transportation Systems*, doi:10.1109/TITS.2012.2209201.
- Shon, T., & Choi, W. (2007). An analysis of mobile WiMAX security: vulnerabilities and Solutions. *Network-based information systems (LNCS 4658)*, 88–97.
- Sun, H., Hsieh, B., & Tseng, S. (2005). On the security of some proxy blind signature schemes. *Journal of Systems and Software*, 74(3), 297–302.
- Tan, C., & Li, Q. (2006). A robust and secure RFID-based pedigree system. In: *Proceedings of the 8th international conference on information and communications security (LNCS 4307)* (21–29).
- Thompson, C. (2004). Radio frequency tags for identifying legitimate drug products discussed by tech industry. *American Journal of Health-System Pharmacy*, 1430.
- Xin, H., & Stone, R. (2007). Chinese probe unmasks high-tech adulteration with melamine. *Science*, 322(5906), 1310–1311.
- Xu, L. (2011). Enterprise systems: state-of-the-art and future trends. *IEEE Transactions on Industrial Informatics*, 7(4), 630–640.
- Yan, Q. (2008). A security evaluation approach for information systems in telecommunication enterprises. *Enterprise Information Systems*, 2(3), 309–324.
- Yin, J., Zhang, X., Lu, Q., Xin, C., Liu, C., & Chen, Z. (2011). IOT based provenance platform for vegetables supplied to Hong Kong. *Recent Advances in CSIE 2011 (LNEE 127)*, 591–596.
- Zdravković, M., Panetto, H., Trajanović, M., & Aubry, A. (2011). An approach for formalising the supply chain operations. *Enterprise Information Systems*, 5(4), 401–421.
- Zheng, L., Zhang, H., Han, W., Zhou, X., He, J., Zhang, Z., Gu, Y., & Wang, J. (2011). Technologies, applications, and governance in the Internet of Things. In: *Internet of things - Global technological and societal trends. From smart environments and spaces to green ICT*. River Publishers.

Dr. Weili Han is an associate professor at Fudan University. His research interests are mainly in the fields of Policy Based Management, IoT Security, Information Security, and Distributed Systems. He is now the members of the ACM, SIGSAC, IEEE and CCF. He received his PhD of Computer Science and Technology at Zhejiang University in 2003. Then, he joined the faculty of Software School at Fudan University. From 2008 to 2009, he visited Purdue University as a visiting professor funded by China Scholarship Council and Purdue.

Yun Gu is a graduate student in Fudan University. Her research interests are mainly focus on the internet of things, risk analysis and trust evaluation. She received her bachelor of software engineering at Fudan University and bachelor of computer science at Dublin National University of Ireland in 2011.

Wei Wang is a graduate student at Fudan University. Her research interests are mainly in the fields of electronic pedigree system, IoT security.

Yin Zhang is a student in Fudan University, major in Software Engineering. From April 2011 to April 2012, he attended Fudan's Undergraduate Research Opportunities Program. Now he is studying in Cryptography and information Security lab.

Yuliang Yin is an incoming graduate student of CMU. He received his B.Eng. of Software Engineering at Fudan University in 2012. He has been an undergraduate research assistant for two years in the Laboratory of Cryptography and Information Security with emphasis on Access Control and Role Mining.

Dr. Junyu Wang is an associate professor at Fudan University. His research interests include the Internet of Things, RFID, Information Security. He received his PhD at University of Science and Technology

Beijing in 2002. From 2008 to 2009, he visited MIT as a visiting Associate professor funded by China Scholarship Council.

Dr. Li-Rong Zheng received his Ph. D. degree in electronic system design from the Royal Institute of Technology (KTH), Stockholm, Sweden in 2001. Since then, he was with KTH as a research fellow and project leader in Laboratory of Electronics and Computer Systems. He became an associate professor in electronics system design in 2003 and a full professor in media electronics at KTH in 2006. He is the founder and director of iPack VINN Excellence Center, and Senior Specialist of Ericsson Networks in Stockholm, Sweden. He is a guest professor of the state key laboratory of ASICs and Systems at Fudan University in China since 2008, and a distinguished professor of Fudan University since 2010. His research experience and interest includes electronic circuits and systems for ambient intelligence and media applications, wireless sensing and identification, and signal integrity issues in electronic systems. He has authored and co-authored over 300 international reviewed publications, covering areas from electronic devices and thin film technologies, VLSI circuit and system design, to electronics systems and wireless sensors.