Modeling and Development of Medical Information System Based on Support Vector Machine in Web Network

Chuanfu Hu[1], Caichang Ding[2*], Lu Dai[1]

[1]*School of Computer Science*
Dongguan University of Technology
Dongguan, China
E-mails: huchuanfu@dgut.edu.cn, dail@ccud.org.cn

[2]*School of Computer Science*
Yangtze University
Jingzhou, Chiina
E-mail: dail@dgut.edu.cn

[*]*Corresponding author*

Received: November 02, 2016 **Accepted: July 19, 2017**

Published: December 31, 2017

Abstract: *This paper aims at improving and utilizing the ontology information in ontology design of FOAF and vCard in real time, and the application of open relational data technology, SPARQL query information results and sending RDF/JSON data format. In addition, improve the effectiveness and efficiency of patient information extraction from the medical information website. This article includes two web search engines that are used to inform patients about medical care information. The experiment uses Drupal as the main software tool, and the Drupal RDF extension module provides some meaningful mapping. In the evaluation part, the structure of the experimental test platform is established and the system function test is carried out. The evaluation results include consumers or patients retrieving the latest doctor information and comparing search capabilities and techniques, between our system and existing systems.*

Keywords: *Modeling, Medical information system, Support vector machine, Web network.*

Introduction

Most of the patients prefer to visit online health clinics as it has become the most convenient way to get better quality of online medical consultation from experienced health professionals. Nowadays, the demand for online medical consultation service has increased these days with the advancement of Internet technology and everyday millions of healthcare consumers go online looking for online doctor information [3].

One traditional way to discover the online information, such as online doctor information, is using popular search engine (e.g. Google, Yahoo, Baidu, etc.), but patients will confront a problem when using search engine to discover online doctor information. For example, I would like to find cardiology doctor in China, they are many listing of web pages returned by a search engine in response to a keyword query which is limited in how they crawl the web and interpret content to retrieve and display the results. It made patients unable find a suitable match for them and the received information may be inaccurate and incomplete, because some of web pages returned by a search engine are out of date or broken link [6, 12].

Nowadays, more and more professional medical companies or healthcare has already provided doctor information on their web page to do online medical consultation, and a thought regarding online consultancy is produced with Video Conferencing, online doctor chat, Email address [8], but the problem is that patients always pay no attention to the site if it not top of the search engine rankings. These issues may reduce opportunities to reach consumers or patients and can also cause consumers or patients to make a service with other providers or business competitor. There is also several websites with online medical consultation provider [1, 2, 7, 11]. For example of online medical consultation provider in Fig. 1 by giving various healthcare, which are registered with online medical consultation provider and all of online doctor information are provide by healthcare that helps consumers or patients to make more easier search, find, and connect doctors around the world than using search engine, but when using search function of online medical consultation, the results may have received erroneous doctor information because there have not been any tools or systems with guaranteed of date information.

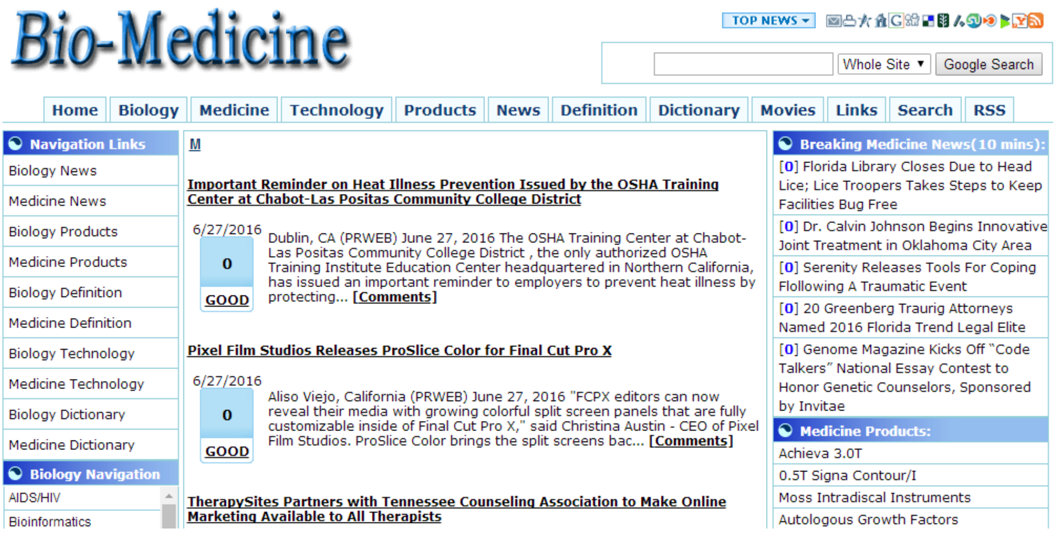


Fig. 1 An example of online medical consultation provider

One of the ways to guarantee real-time date information is getting information from healthcare's website directly. Healthcare already provides doctor information on their web page. Linked Open Data (LOD) [13] is one of hottest semantic web techniques to create typed links between data form different sources; it is a new and more efficient paradigm for sharing and connecting distributed data, permitting decentralization and interoperability. These standards comprise RDF for publishing data in a structured format with explicit semantics and the SPARQL query language and protocol for querying and accessing RDF data through an open and HTTP-based protocol [9]. As the problem of search engine and online medical consultation provider above, the approach is to improve effectiveness and efficiency for discover online doctor information by applying LOD technique.

Overview

In 2001, Tim Berners-Lee introduced their vision of the Semantic Web, as an extension of the current Web in which information has "well-defined meaning, hence it better enables computers and people to work in cooperation". The most essential part of this next-generation Web is content that is formally described via ontologies, metadata conforming to these ontologies, logic, and agents. Many definitions of the tern ontology exist. The most popular is by Gruber who defines ontology as "an explicit specification of a conceptualization".

This definition is future extended to "Ontologies are a formal, explicit specification of a shared conceptualization" [5]. Conceptualization refers to an abstract model of some part of the world which identifies the relevant concepts and relations between these concepts. "Explicit" means that the type of concepts, the relations between the concepts, and the constraints on their usage, are explicitly defined. Formal refers to the fact that the ontology should be machine readable. Finally, it shared means that the ontology should reflect the understanding of a community and should not be restricted to the comprehension of some individuals, by doing so, it captures consensual knowledge. Ontologies occur in different degrees of formality, ranging from thesauri to richly structures [3].

Tim Berners-Lee's vision of the future network consists of two parts: the first part of the network becomes a more in-depth collaboration media, and the second part of the network can understand and deal with the machine. Semantic Web learned the discipline of artificial intelligence, information theory, philosophy, logic, and calculation of the complexity of research; hope to make significant improvements to the representation and access of information on the Web, in order to solve the bottleneck of using the Web. The Semantic Web is "an extension of the Web that can read information and automatically provides powerful Web service function". From the above analysis, conclusions can be drawn: the establishment of the Semantic Web is based on the conceptualization and formal knowledge, as well as the corresponding reasoning, and it has a profound and artificial intelligence sources relation. Therefore, many analysts are needed from the perspective of the field of artificial intelligence to inspection. However, due to the two different application environments, there are some differences. Traditional artificial intelligence system that calls for compliance with some of the public, the concept of the definition are exactly the same, and they generally have their own narrow, special rules for the collection of information reasoning. Although data can be converted from one system to another one, the inference rules system are usually in the form of completely different and the rules of a system which cannot be used in other systems. From this point, the traditional artificial intelligence system is a concentrated, isolated (proprietary) system [10]. Similarly, the Semantic Web is that the conceptualization of a knowledge-based start. Semantic Web of knowledge is a series of resource modeling and description.

Resources, is a very broad concept, which may be websites, or a page of a certain part of the content. This description uses some form of symbolic expression on the web and on the resources and other relevant resources, as well as the relationship between these resources to portray. However, the traditional and artificial intelligence systems, the knowledge that the Semantic Web is the uniqueness, to comply with its own web decentralized and versatile [4]. Therefore, the Semantic Web is to create knowledge on the spread with the application of universal at the same time. This is the Semantic Web and traditional artificial intelligence system, a very important distinction.

Model and algorithm

LOD rely on three technologies: URIs, HTTP and RDF. Web resources are named with URIs and when the URI are looked up via the HTTP protocol, they provide formatted in the RDF data model. This data includes references to other datasets. A huge momentum has recently been built up in Semantic-Web research, due to the ongoing implementation of a vision of a Web of Data formulated by Time-Berners Lee. In this vision, formerly fragmented data is connected and interlinked with each other on the so-called LOD principles. Over just a few years, the so called LOD cloud, which represents a huge interconnected dataset, has been steadily growing. In early 2007, the LOD community project had been launched within the

W3C Semantic Web Education and Outreach group. It bootstraps the Web of Data by publishing datasets using the Resource Description Framework (RDF), the metadata model primarily used on the Semantic Web. RDF enables automated software to store, exchange, and use machine-readable information distributed throughout Web which, in turn, allows users to deal with the information with greater efficiency and certainty. Currently, the LOD project includes more than 200 different datasets, ranging from rather centralized ones, such as DBpedia, a structured version of WiKiPedia, to those that are very distributed, for example the FOAF-sphere.

The basic equation of the proposed algorithm is shown in the Eq. (1):

$$u_{c1}^{'} = \begin{cases} -(\omega_c t + \alpha_1 - 2\pi k)\dfrac{U_c^{'}}{\pi} + U_c^{'} & 2\pi k \le \omega_{ct} < 2\pi k + \pi \\ (\omega_c t + \alpha_1 - 2\pi k - \pi)\dfrac{U_c^{'}}{\pi} & 2\pi k + \pi \le \omega_{ct} < 2\pi k + 2\pi \end{cases}, \ k = 0,\ \pm 1,\ \pm 2, \ldots . \tag{1}$$

The equation of sinusoidal modulation wave is:

$$u_s = U_s \sin \omega_s t . \tag{2}$$

Set $F = \omega_c / \omega_s \gg 1$, modulation degree $M = U_s / U_c^{'} \le 1$. For u_{P1}, the sampling point is:

$$U_s \sin \omega_s t = -(\omega_c t + \alpha_1 - 2\pi k)\frac{U_c^{'}}{\pi} + U_c^{'} .$$

Set $\omega_s t = Y$, $\omega_c t = X$, then

$$X = 2\pi k + \pi - \alpha_1 - \pi M \sin Y .$$

As shown in the form of $X = \omega_c t$ is in the interval between $2\pi k + \alpha$ and $2\pi(k+1) + \alpha$. Between *a* and *b*, when the sine wave modulation is larger than the triangular carrier, the u_{P1} is gotten. The time function of u_{P1} is:

$$u_{P1}(X,Y) = \begin{cases} 0 & X \begin{cases} < 2\pi k + \pi - \alpha_1 - \pi M \sin Y \\ \ge 2\pi k + \pi - \alpha_1 + \pi M \sin Y \end{cases} \\ E/6 & X \begin{cases} < 2\pi k + \pi - \alpha_1 + \pi M \sin Y \\ \ge 2\pi k + \pi - \alpha_1 - \pi M \sin Y \end{cases} \end{cases} . \tag{3}$$

The double Fourier series of function $u_{P1}(X,Y)$ is given:

$$u_{P1}(X,Y)=\frac{A_{00}}{2}+\sum_{n=1}^{\infty}(A_{on}\cos nX+B_{on}\sin nY)+\sum_{n=1}^{\infty}(A_{mo}\cos mX+ \\ B_{mo}\sin mY)\sum_{m=1}^{\infty}\sum_{\pm 1}^{\pm\infty}[A_{mn}\cos(mX+nY)+B_{mn}\sin(mX+nY)].$$

In the above equation

$$A_{mn}+jB_{mn}=\frac{2}{(2\pi)^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}u_{P1}(X,Y)e^{j(mX+nY)}dXdY. \quad (4)$$

Take the equation (3) into equation (4)

$$A_{mn}+jB_{mn}=\frac{E}{6\pi^2}\int_0^{\pi}\int_{2\pi k+\pi-\alpha_1-\pi M\sin Y}^{2\pi k+\pi-\alpha_1+\pi M\sin Y}e^{j(mX+nY)}dXdY \\ =\frac{E}{j6m\pi}e^{jm(n-\alpha_1)}[\frac{1}{\pi}\int_0^{\pi}e^{jmM\pi\sin Y}e^{jnY}dY-\frac{1}{\pi}\int_0^{\pi}e^{-jmM\pi\sin Y}e^{jnY}dY].$$

By Bessel function,

$$\frac{1}{\pi}\int_0^{\pi}e^{jmM\pi\sin Y}e^{jnY}dY=J_n(mM\pi)\frac{e^{jn\pi}-1}{2}, \\ \frac{1}{\pi}\int_0^{\pi}e^{-jmM\pi\sin Y}e^{jnY}dY=J_n(mM\pi)\frac{1-e^{jn\pi}}{2}.$$

Then,

$$A_{mn}+jB_{mn}=\frac{E}{j6mn}e^{jm(\pi-\alpha_1)}[J_n(mM\pi)\frac{e^{jn\pi}-1}{2}-J_n(mM\pi)\frac{1-e^{jn\pi}}{2}] \\ =j\frac{E}{6mn}J_n(mM\pi)e^{jm(\pi-\alpha_1)}[1-e^{jn\pi}]. \quad (5)$$

Then we get:

$$\frac{\partial C}{\partial t}=D_L\frac{\partial_2 C}{\partial x^2}-u\frac{\partial C}{\partial x}, \quad (6)$$

$$C(x,t)|_{t=0}=0 \qquad 0\le x<+\infty \\ C(x,t)|_{x=0}=C_0 \qquad t>0 \\ C(x,t)|_{x\to+\infty}=0 \qquad t>0. \quad (7)$$

After Laplace change, the standard normal distribution function is solved as follows:

$$C=1-\Phi_{0.1}\left[\frac{x-ut}{\sqrt{2D_L t}}\right]. \quad (8)$$

Structures and methods

RDF can be used to describe RDF predicates and classes of doctor describe the relationship between entities. Fig. 2 shows relations between entities.

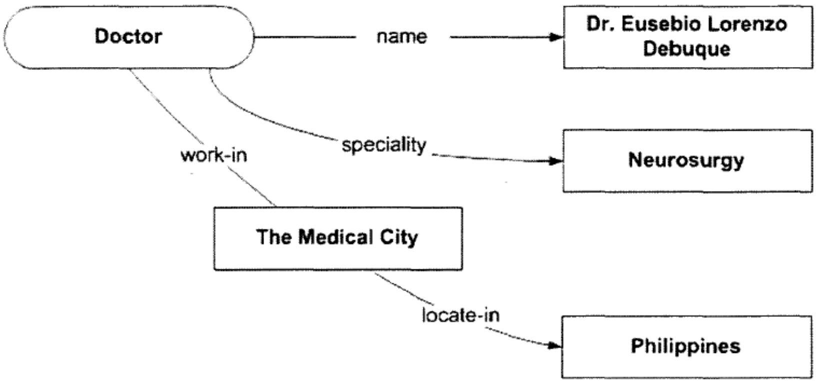


Fig. 2 Relations between entities

RDF is simply a standard way to associate properties with URLs (what RDF calls resources). A property is simply a name-value pair where the value can either be a string of text, another name or a list (of name and strings.) In RDF all names are also URLs (resources), this way they are guaranteed to be unique (the magic property of URLs). Fig. 3 shows associate properties with URLs.

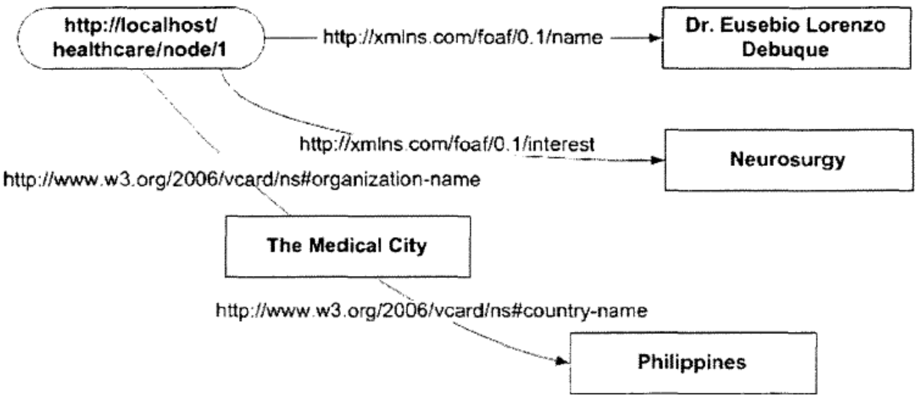


Fig. 3 RDF name things with URLs

System architecture (see Fig. 4) consists of three key areas. First, Interface represents a patient's search purpose in a structured format. The patient can choose doctor's specialty and doctor's procedure in Interface. Query Generator transforms the patient's configuration into SPARQL queries for each healthcare by using Linked Open Data technique to link data from different sites and Endpoint returns detailed information of the doctors from healthcare and shows the patient the list of doctor.

Currently a large number of websites are driven by Content Management System (CMS) including healthcare websites because it is viability in the Search Engine Optimization (SEO). One of the most popular CMS is Drupal. It is a free and open sourcing CMS written in PHP and distributed under the GNU General Public License. Drupal stems from a project by a Dutch university student. The goal of the project was to provide a mechanism for Buytaert and his friends to share news and events. Buytaert turned Drupal into an open source project in 2001 and the community readily embraced the concept and has expanded on its humble beginnings, creating what is now one of the most powerful and feature rich CMS platforms on the web. Individuals, teams, and communities leverage Drupal's features to easily publish, manage, and organize content on variety of websites, ranging from personal blogs to large corporate and government sites.

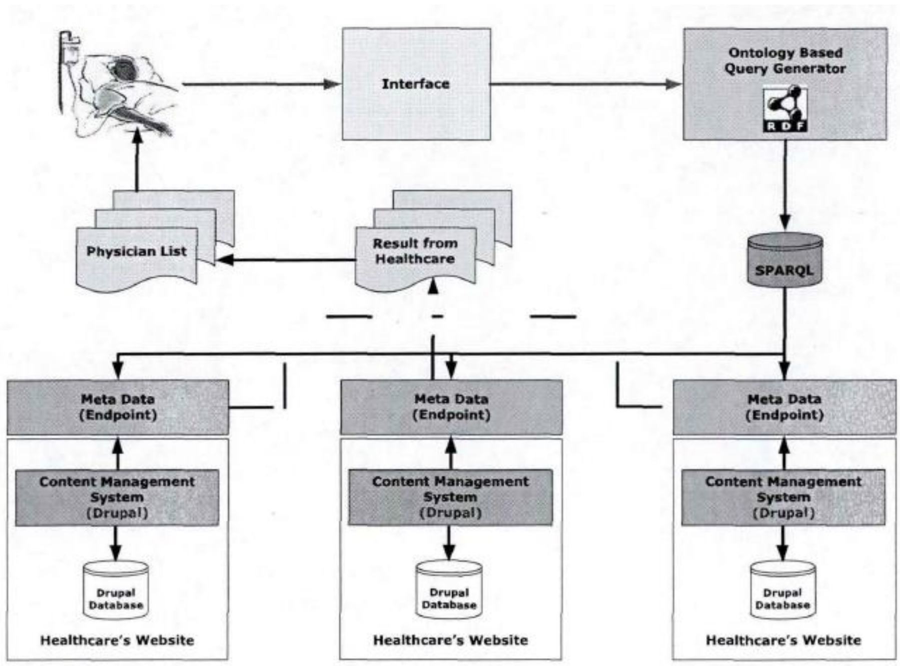


Fig. 4 System architecture

Experiment and discussion

In this section, test bed architectures is presented to evaluate a system. In this experiment consists of two computers such as doctor finder screen and doctor information screen. The specifications of two computers are listed:

(1) Doctor finder (Patient's GUI) computer specification
(2) OS: Microsoft Windows XP Professional SP3
(3) CPU: Intel(R) Core(TM) 2 Duo CPU T5670 1.80 GHz (2 CPUs)
(4) RAM: 2 GB
(5) Doctor information (Healthcare's GUI) computer specification
(6) OS: Microsoft Windows XP Professional SP3
(7) CPU: Intel(R) Pentium(R) Dual-Core Processor T2390 1.86 GHz
(8) RAM: 1 GB

The doctor information which was created by database is created as web based doctor information. With the capability of existing module provided by Drupal developer community, this project can use Drupal and its plug in modules in implementation phase. Some modules are included in Drupal installer as core modules for instance, RDF includes RDF mapping provided in each content type, while some modules are additional plug in features as follow.

RDFx is providing extra APIs and additional serialization formats such as RDF/XML, NTriples, Turtle, etc. RDF UI allows site administrator to specify the RDF mapping via a user interface SPARQL is adding to query RDF graph. Views are added to create, manage, and display lists of content. SPARQL Views is added for creating SPARQL queries on external endpoints and displaying those using Views. Entity API is added to enables work with any entity type and to provide entities. Chaos Tools (CTools) is added to create and manages page in Drupal site. The Content Construction Kit (CCK) is added to enables adding custom fields to node using a web browser on the Drupal site. RDF External Vocabulary Importer (EVOC) is used importing additional RDF vocabulary or ontology into the website. VCard is an external one that must be imported.

System features tests aim to test date doctor information from healthcare directly. Search doctor information before healthcare updating information. For example, consumers or patient would like to search doctor information in "Thailand" and specialty is "Cardiology" (see Fig. 5 the results before healthcare updating doctor information) it one record found.

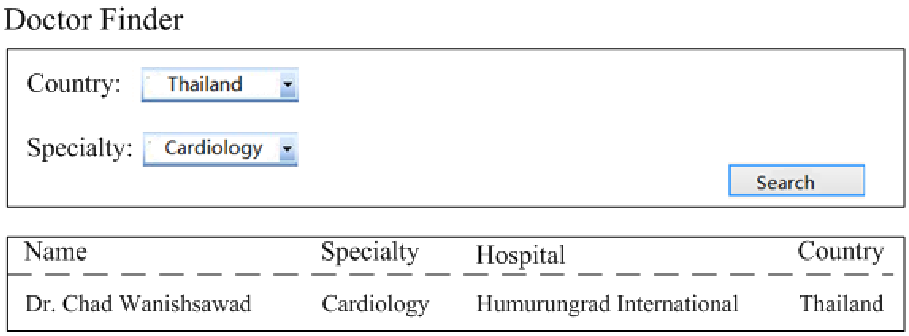


Fig. 5 Results before healthcare updating doctor information

Search doctor information after healthcare updating information. For example, healthcare inserts new doctor information record to their backend database and show output on their web page after consumers or patient search doctor information in "Thailand" and specialty is "Cardiology" again (see Fig. 6 the results after healthcare updating doctor information) it two record found.

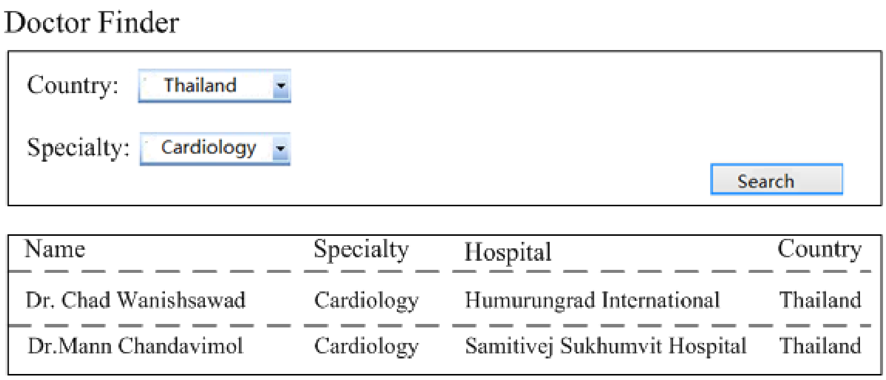


Fig. 6 Results after healthcare updating doctor information

Conclusion

This paper is intended to contribute to apply the LOD approach and semantic web technologies to improve effectiveness and efficiency for discover online doctor information. The main achievements presented in this project are:

- To study searching process of search engine (e.g. Google, Yahoo, Baidu, etc.) and analyzes problems with results of search engine.

- To study searching function process of online medical consultation system and analyzes problems with searching function.

- To develop and implement search engine for discover doctor information. This project includes two sites search engine that returns doctor information from healthcare as results and another sites is healthcare that provided doctor information on web pages.

In evaluation part, the thesis setting up experimental test bed configurations to do system features tests. The evaluation results consumers or patient retrieved up to date doctor information and comparison search features and technologies between our system and existing systems that is Medeguide, Cleveland clinic, and Medindia. All of them query doctor information from their backend database but ours system query doctor information healthcare website directly by using SPARQL to be able query more than one site in the same time.

Acknowledgements

This paper was supported by the Popular Science Series into the Campus of Guangdong Province (No. 2014A070710018), National Outstanding Engineers Training Plan "Software Engineering", No. 40 (2011), Characteristic Specialty "Software Engineering" in Guangdong Province, Guangdong Teach No. 473 (2011), Training Mode Innovation Experimental Area of Guangdong Province (Local Computer Science Classes of Applied Talents Innovation Experimental Area) Guangdong Teach High Letter No. 204 (2012).

References

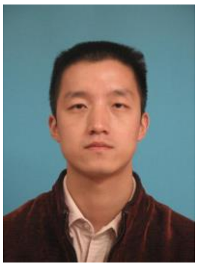
1. Arshad H., M. Nikooghadam (2014). Three-factor Anonymous Authentication and Key Agreement Scheme for Telecare Medicine Information Systems, Journal of Medical Systems, 38:136.
2. Giri D., T. Maitra, R. Amin, P. D. Srivastava (2015). An Efficient and Robust RSA-based Remote User Authentication for Telecare Medical Information Systems, Journal of Medical Systems, 39:145.
3. Guo D., Q. Wen, W. Li, H. Zhang, Z. Jin (2015). An Improved Biometrics-based Authentication Scheme for Telecare Medical Information Systems, Journal of Medical Systems, 39:20.
4. Husin N. A., N. S. Herman, B. Hussin (2012). Integrated Features by Administering the Support Vector Machine of Translational Initiations Sites in Alternative Polymorphic Context, International Journal Bioautomation, 16(1), 13-22.
5. Islam S. H., M. K. Khan (2014). Cryptanalysis and Improvement of Authentication and Key Agreement Protocols for Telecare Medicine Information Systems, Journal of Medical Systems, 38:135.
6. Kim S. H., K. Y. Chung (2015). Medical Information Service System based on Human 3D Anatomical Model, Multimedia Tools & Applications, 74(20), 8939-8950.
7. Lu Y., L. Li, H. Peng, Y. Yang (2015). An Enhanced Biometric-based Authentication Scheme for Telecare Medicine Information Systems Using Elliptic Curve Cryptosystem, Journal of Medical Systems, 39:32.
8. Mishra D., S. Mukhopadhyay, S. Kumari, M. K. Khan, A. Chaturvedi (2014). Security Enhancement of a Biometric based Authentication Scheme for Telecare Medicine Information Systems with Nonce, Journal of Medical Systems, 38:41.
9. Mishra D., S. Mukhopadhyay, A. Chaturvedi, S. Kumari, M. K. Khan (2014). Cryptanalysis and Improvement of Yan et sl.'s Biometric-based Authentication Scheme for Telecare Medicine Information Systems, Journal of Medical Systems, 38:24.

10. Siddiqui Z., A. H. Abdullah, M. K. Khan, A. S. Alghamdi (2014). Smart Environment as a Service: Three Factor Cloud based User Authentication for Telecare Medical Information System, Journal of Medical Systems, 38:9997.
11. Wang C. (2015) A Modified Machine Learning Method Used in Protein Prediction in Bioinformatics, International Journal Bioautomation, 19(1), 25-36.
12. Xu B., L. D. Xu, H. Cai, C. Xie, J. Hu, F. Bu (2014). Ubiquitous Data Accessing Method in IoT-based Information System for Emergency Medical Services, IEEE Transactions on Industrial Informatics, 10(2), 1578-1586.
13. Xu X., P. Zhu, Q. Wen, Z. Jin, H. Zhang, L. He (2014). A Secure and Efficient Authentication and Key Agreement Scheme based on Ecc for Telecare Medicine Information Systems, Journal of Medical Systems, 38:9994.

Chuanfu Hu
E-mail: huchuanfu@dgut.edu.cn

Chuanfu Hu is now a teacher at Dongguan University of Technology. His research interests include nonlinear circuit theory, chaotic signal processing, embedded systems design, social network and their applications.

Caichang Ding, Ph.D.
E-mail: dail@dgut.edu.cn



Caichang Ding received his B.Sc. degree from the School of Mechanical and Electronic Information, China University of Geosciences, in 2003, the M.Sc. degree from the School of Computer, Wuhan University, in 2006, and Ph.D. degree from the State Key Lab of Software Engineering, Wuhan University, in 2014.

Lu Dai, Ph.D.
E-mail: dail@ccud.org.cn

Lu Dai was born in 1988. She received the Ph.D. degree from Wuhan University, in 2009. She is currently working at Dongguan University of Technology. Her current research interests include cloud computing, evolutionary algorithm, social network and text mining.

© 2017 by the authors. Licensee Institute of Biophysics and Biomedical Engineering, Bulgarian Academy of Sciences. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).