

Analysis of Personalized Information Mining and Recommendation Technology in E-Commerce based on Semantic Web Ontology

Hongsheng Xu^{1,2 a *}, Ke Li^{1,2} and Ganglong Fan^{1,2}

¹Luoyang Normal University, Luoyang, 471934, China

²Henan key Laboratory for Big Data Processing & Analytics of Electronic Commerce, Luoyang, 471934, China

^a85660190@qq.com

Keywords: Semantic Web; Ontology; E-commerce; Personalized recommendation system; Web mining

Abstract. In this paper, we first introduce the hierarchical structure of semantic Web, the goal of semantic Web, and describe the definition of ontology and the properties of the elements such as concepts, relations, functions, axioms and instances. In this paper, the domain ontology is integrated into the Web mining and personalized recommendation process, which can make full use of the semantic knowledge in preprocessing, pattern discovery and online recommendation. The paper presents analysis of personalized information mining and recommendation technology in e-commerce based on Semantic Web Ontology.

Introduction

With the popularity of the Internet, e-commerce has been an unprecedented development, to businesses and customers with more and more information. However, with the increase of information, information overload / product becomes more and more serious, to the customers and merchants brought a lot of inconvenience, mainly in: the customer often submerged in a mass of information, can not find the goods they need quickly; and businesses will lose their customers to sales difficult [1]. Therefore, in the e-commerce environment, need a recommendation function of e-commerce system to help customers to buy the needs of customers for the goods, the system can dynamically capture the customer's hobby, predict the purchase preference.

The field of knowledge space, learning resource is the main source of knowledge and learning objects, because learning resources with regional diversity, complexity and characteristics of cultural diversity, making a large number of learning resources can not be shared, different education systems are difficult to communicate with each other. The key to solve this problem lies in the standardization of educational information. The standardization of the educational information can be made between different platforms with interactive learning resources, the same learning resources can be reconstituted into different learning processes, improve the reusability, and through a unified standard that resources are consistent, can reduce the cost of other learning abroad.

The Web information integration method based on ontology is far more complicated than that of using structured method. The main reason for this complexity is that the Web information integration based on Ontology involves many research fields, such as ontology construction, ontology integration, and ontology query and so on. As a new field of artificial intelligence, ontology engineering has not formed a mature and complete system in many fields. This immaturity and incompleteness have brought many difficulties to ontology based applications.

Ontology is a philosophical concept, which is a systematic explanation of the nature of things. But in recent decades, the word has been applied to the computer industry, and plays an increasingly important role in artificial intelligence, computer language and database theory. With the development of artificial intelligence technology, ontology has been given a new definition, which is developing with the development of information technology.

At present, almost all of the major e-commerce systems, such as Amazon, eBay, etc., are used in a variety of personalized recommendation technology. Personalized recommendation technology has been developed rapidly in both theory and practice. However, with the expansion of the application scale, personalized recommendation technology is also facing a series of challenges, such as: the number of customers and the number of goods showed exponential growth. The paper presents analysis of personalized information mining and recommendation technology in e-commerce based on Semantic Web Ontology.

Analysis of Personalized Data Mining in E-commerce Based on Ontology

Ontology evolution, ontology, ontology matching verification, ontology engineering, ontology modeling and ontology mapping and other fields, although in the number of search results ranking by comparison, but this does not mean that the research focus is not the potential of these research direction, full of vitality and development prospects of the potential research direction, but rather a result of the research carried out on they are tools, ontology learning, ontology, ontology and ontology language translation of the basic research can be carried out based on the results.

The single ontology method uses a global ontology corresponding to each distribution, heterogeneous information sources, as a universal semantic model for all information sources [2]. When the field of view of the basic information source is the same when using this method is more suitable, but if from the perspective of various information sources in different areas, to create a unified global ontology is very difficult, and when the source of information is changed, the global ontology must also be modified.

Specification for design of IMS learning provides a description method that learners in the learning environment, in a certain sequence of learning activities to achieve learning goals, it can simulate the multi roles in teaching and learning process, the application also supports the individual learning process and t online teaching mode [3]. The teaching mode is complete, convenient abstraction and formal description in a flexible language design and learning, including email, discussion or announcement of specific services, especially in the learning activities in the description of the activities and activities of the organization structure of the phase separation and phase separation, and the role and resources, improve reusability. The system that follows the learning design specification is able to understand these descriptions to ensure a high degree of sharing and reuse of learning resources, as is shown by equation(1) .

$$w_{i+1}^1(t+1) = (1 - wd_i^1(t))x_i^1(t) - rs_i\alpha N^1(t) \quad (1)$$

A relationship usually contains two parts: the domain and the range, where w is define the scope of the relationship between the two. In ontology, the definition domain of a relation is usually a concept C, and the range can be a concept or a specific range. The relationship between attributes: the attribute of the relation described further restrictions on the relationship, such as a relationship between the age of "Has Age", if its range is integer, then further through its properties specified range is 1 to 9 between the integer.

Machine readable views - the semantic Web is a vision of the idea that Web data can be defined and linked in some way [4]. In addition to displaying data, the machine is able to automatically process and integrate data and reuse the data among multiple applications. The purpose of semantic Web is to make the existing network better machine readable, so that the intelligent agent can acquire and operate the relevant information.

At the same time, Internet has developed into the world's largest library of information resources, in order to achieve from massive data, find their required data and useful information, put forward the concept of Web mining. Web mining is a challenging task, it searches for the structure, in order to determine the importance of Web content, discover the regularity and dynamic of Web content, and explore the access pattern of Web.

In many query Web information integration oriented multi domain integrated drive information source, and user queries often involve multiple sources of information, so users often want only according to a middle layer Ontology (mediator ontology) will be able to query the information they are interested in, and do not care about the answer from the query what information sources and how after treatment. In order to is adapt to this demand.

Unified user model for cross system personalization rich ontology store user personalization analysis system and its application in the information, and preferences, and it is to provide personal access. However, the stored user model is a proprietary format, as is shown by equation (2), where x is the exchange of information and re-uses caused inconvenience, and information duplication. In addition, the user model is often trapped in the system, causing the user to control the user model. In this paper, we propose a general user context model based on ontology, which can be used as a basis for cross system personalization [5].

$$\begin{cases} x_t = \phi_0 + \phi_1 x_{t-1} + L + \phi_p x_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - L - \theta_q \varepsilon_{t-q} \\ \phi_p \neq 0, \theta_q \neq 0 \end{cases} \quad (2)$$

The emergence of learning resource standards represents a significant advance in the e-learning framework, but they have some limitations [6]. The lack of semantic metadata, due to the lack of a unified description of the lexical database, metadata is usually used to describe ordinary text, so the learning resource information is difficult to clearly express the semantic description and relation model of concrete. Some metadata data can not be clearly described and defined in the design phase.

At present, there is no uniform standard for the specific methods and methods of building ontology. Therefore, the construction of ontology is a direction for further research. However, in the process of building specific domain ontology, one thing is to be recognized, that is, the construction of ontology is completed by the joint efforts of domain experts and ontology creators. Ontology construction method, any kind of ontology construction method by a general process and each step of the operation rules, only the design details of the operation and the principle of scientific and reasonable process and accurate and controllable, can better reveal the domain knowledge.

$$v_k(t) = e^{-\lambda t} \frac{(\lambda t)^k}{k!}, k = 1, 2, 3, \dots \quad (3)$$

The semantic web is the next generation of web, as is shown by equation(3), where $v(t)$ makes the application on the Web can automatically collect Web documents from different data sources, information integration and information processing, and to accomplish complex tasks assigned by people and other collaborative applications. A better annotation view - the idea of semantic Web is to add and link the machine - to - machine labels to the Web we know. Enhanced search perspective - it is possible to access Web resources through content rather than keywords in semantic Web.

Personalized recommendation technology based on Web usage mining has become a hot issue in current e-commerce technology [7]. On the one hand, Web usage mining can improve the recommendation quality in a certain extent; however, this method still has many deficiencies, such as: poor quality recommendations, after all is because the semantic recommendation system can well understand the goods. In the current Internet environment, the main content of the page is presented to the user; the machine can not understand the meaning of the information. The computer can skillfully analyze the page layout, effectively distinguish the title and link.

Personalized Information Mining and Recommendation Technology in E-Commerce based on Semantic Web Ontology

There are two problems in the personalized recommendation of the traditional adaptive multimedia. The first one is the inherent dynamics of the news, the other is the modeling of the user interest and how the user interest model is linked with the news. Using ontology is as a semantic bridge between

RSS push news and user interest. The combination of adaptive hypermedia technology and ontology is reasoning to construct dynamic personality on news service.

Predicate logic is a form of language, is currently a most accurate expression of human thinking, it is close to the human natural language, so the first-order predicate logic representation of a method is the earliest use of knowledge. It is simple, natural, accurate, flexible, easy to implement and so on. Predicate logic can be used to represent the factual knowledge of the status, property, concept, and so on.

Ontology based Web information extraction, the need to locate the ontology of Web information sources. As a general semantic model, domain ontology may correspond to multiple sources of Web information, inevitably there is a concept of ontology have different expressions in different sources of information, and automatic identification method for using only keyword matching to complete source of information in the ontology concept.

For the development of open source Java semantic Web application environment, interface ontology data it supports application called Jena to provide support; mainstream ontology description language, such as RDF, RDF (S) and OWL [8]; to support a variety of methods such as XML method, ontology encoding N3 and three element formula; storage model, support a variety of ontology such as memory model and database model; support RDFs based reasoning [9]. In addition, Jena also supports the use of SPARQL language in the semantic query data ontology.

Content based recommendation algorithm is based on information retrieval and information filtering. Recommendation algorithm based on content analysis first, users have score items, establish the user access preference file, and then select match user preferences file items from all projects, and higher scores of projects to recommend to customers.

Experiments and Analysis

The key of information extraction based on ontology is to establish the mapping of the concept in the data source to the ontology. In some projects, the mapping relation is established by hand, and a method of extracting reliable Web data based on ontology is discussed. For different Web sites, it is necessary to map the Web data to the ontology concept by manually writing different mapping files, which can only be used to extract data from Web sites with relatively fixed structure.

The sequence model overcomes the problem that the information is independent from each other, which is based on the rules and the classification model. But the sequence model has a long training time. In addition, the existing sequential model is a first-order linear model, which can only describe the linear features, but can not describe more complex nonlinear characteristics [10]. This makes the linear model can not get a good result when there is a nonlinear dependence between the annotations.

The problem of information sharing and exchanging in semantic Web needs to be solved by ontology. Ontology (ontology) layer is the semantic layer of semantic Web. The Logic layer: the logic (Logic) layer further enhances the capabilities of the ontology language and allows for the writing of descriptive knowledge for specific applications, mainly through the definition of stronger logical rules. Logical reasoning is mainly used to establish the consistency and correctness of data sets, but also to derive implicit consistent conclusions from a set of known data sets, as is shown by equation(4).

$$\mu(x) = e^{\int P(x)dx} = e^{\int -\frac{n}{x+1}dx} = (x+1)^{-n} \quad (4)$$

The results of Web structure mining can provide information about the content, quality and structure of Web pages. Citing documents helps to infer the authority of Web pages. The so-called authority page is a web page that is highly referenced or referenced in a topic. Another concept associated with it is the central page - a page that points to many authoritative pages. The authority page and the center page are mutually reinforcing, and the central page plays an important role in establishing the authoritative page of the theme.

The electronic commerce recommendation system salesman recommend the products to customers, to help customers find the products, thus the successful completion of the purchase process, it can effectively retain customers, improve the electronic commerce system of the level of sales; businesses can through the recommendation system to maintain contact with customers, to establish good relationship with customers, maintain customer loyalty.

For the preprocessing of Web log case, due to the stateless HTTP protocol, and by configuring the e-commerce site Web server logs using the original data integrity is relatively poor, and irregular, it is difficult to meet the need of Web usage mining. In addition to the data integration, data cleaning, data transformation and data reduction of four aspects of the operation, but also the need for data filtering, user identification, session identification, transaction identification and path of it.

Summary

The paper presents analysis of personalized information mining and recommendation technology in e-commerce based on Semantic Web Ontology. This paper analyzes the electronic commerce recommendation technology based on domain ontology, domain ontology is integrated into the Web mining and personalized recommendation process, which can effectively use semantic knowledge in preprocessing, pattern discovery and online recommendation stage.

Acknowledgements

This paper is supported by Henan key Laboratory for Big Data Processing & Analytics of Electronic Commerce, and also supported by the science and technology research major project of Henan province Education Department (13B520155, 17B520026).

References

- [1] Berners-Lee T, James Hendler , Ora Lassila. The Semantic Web. Scientific American,2001,5.
- [2] Daoud, M. and L.Tamine, et al. Learning ImPlicit User Interests Using Ontology and Search History for Personalization. Web Information Systems Engineering-WISE 2007 Workshops:325-336.
- [3] Maedche A, Staab S.Ontology Learning for the Semantic Web. Special Issue on the Semantic Web,IEEE Intelligent System,2001,16(2):72-79.
- [4] Hongxin Wan, Yun Peng, Clustering and Evaluation on Electronic Commerce Customers Based on Fuzzy Set, IJACT, Vol. 5, No. 3, pp. 199 ~ 206, 2013.
- [5] LEE CS ,KAO YF,KUO YH. Auomated ontology construction for unstructured text documents. Data & Knowledge Engineering.2007;60:547-566.
- [6] Stojanovic L,Stojanovic N,Volz R.Migrating Data-intensive Web Sites into the Semantic Web[C].In:Proc.of the 17th ACM symposium on applied computing. ACM press,2002:1100-1107.
- [7] Schmidt, K.U.and L. Stojal, ovie, etal. On Enriching Ajax with Semanties: The Web Personalization Use Case. The Semantie Web :Research and Applications:4th European Semantic Web Conference, ESWC 2007:686-700.
- [8] Shamsfard M,Barforoush A A.Learning Ontologies from Natural Language Texts.Int’l Journal of Human-Computer Studies,2004,archive Volume 60:17-63.
- [9] Missikoff M, Navigli R,Velardi P.Integrated Approach for Web Ontology Learning and Engineering.IEEE Computer,2002.35(11):60-63.
- [10] ZHANG Rui-ling, XU Hong-sheng .Building and mapping ontology of e-business based on fuzzy rough concept lattices, Journal of Convergence Information Technology Vol. 6, No. 9 2011:81-88.