International Journal of Production Research, 2013
Vol. 51, No. 22, 6702–6719, http://dx.doi.org/10.1080/00207543.2013.832839

Combining social network and collaborative filtering for personalised manufacturing service recommendation

W.Y. Zhang[a]*, S. Zhang[a], Y.G. Chen[a] and X.W. Pan[b]

[a]*School of Information, Zhejiang University of Finance and Economics, Hangzhou, China;* [b]*School of Economic and Management, Zhejiang Sci-Tech University, Hangzhou, China*

(*Received 19 February 2013; accepted 27 July 2013*)

Owing to the rapid proliferation of Web service technologies in cross-enterprise manufacturing collaborations, information overload is becoming a major barrier that hinders the effective discovery of the shared manufacturing services provided by collaborative partners for supply chain deployment. Thus, we aimed to identify a different approach for discovering manufacturing services by making personalised service recommendations that are suited to the specific needs of active service users based on usage data from previous retrievals made by past service users. The proposed approach combines social network and collaborative filtering techniques in a unified framework to predict the missing Quality of Service (QoS) values of manufacturing services for an active service user, thereby improving the effectiveness of personalised QoS-aware service recommendations. The social network explores the usage of preference and tagging relationships among service users and manufacturing services in making personalised recommendation, which alleviates the data sparsity and the cold start problems that hinder the traditional collaborative filtering techniques. A case study and experimental evaluation demonstrate that the proposed approach can achieve the practicality and accuracy to personalised manufacturing service recommendations in a real application.

Keywords: collaborative filtering; manufacturing service; service discovery; service recommendation; social network

1. Introduction

The exponential growth of the Internet and the increasing globalisation of manufacturing enterprises have changed the nature of cooperative information processing mechanisms for supply chain deployment. For example, Service-Oriented Architectures (SOAs) (Duke, Davies, and Michardson 2005) and their corresponding Web services provide dynamic methods that allow manufacturing enterprises to communicate with their partners, including suppliers and customers (Cai, Zhang, and Zhang 2011). Each enterprise becomes a service user, which plays the role of a service provider to provide the manufacturing services to other enterprises, and the role of service consumer to consume the manufacturing services provided by other enterprises. The rapid proliferation of Web service technologies in cross-enterprise manufacturing collaboration has led to a growing number of manufacturing services on the open Internet. This has resulted in a serious information overload problem, which is becoming one of the major barriers that hinders the effective discovery of the shared manufacturing services provided by collaborative partners for supply chain deployment.

In particular, the problem search space is increased exponentially when non-functional Quality of Service (QoS) properties of manufacturing services are considered during service selection, because most of QoS values (e.g. availability, performance and reliability) are subjective and rated differently by different service users. If we consider the performance of a lathe finish service provided by service user U_a as an example, it can be rated '8' by service user U_b, but '6' by service user U_c. The different ratings of the performance QoS may be because user U_b has a lower standard when rating the manufacturing QoS compared with user U_c. The user-dependent characteristics of the QoS properties of manufacturing services make it impractical, if not impossible, for a service user to rate all of the QoS values of the candidate manufacturing services in advance. Thus, the missing QoS values of the manufacturing services for an active service user need be predicted to facilitate personalised QoS-aware service discovery.

Recently, various kinds of recommender systems have gained much attention in overcoming the information overload problem by providing the different users with personalised recommendations based on user preferences and experiences. Among these recommender systems, collaborative filtering (Adomavicius and Tuzhilin 2005) is highly desirable

*Corresponding author. Email: wyzhang@e.ntu.edu.sg

to recommend items to users by collecting and comparing the usage data from other similar users or items. It has been successfully applied in some famous commercial Web-based systems including Amazon.com, Ebay.com and Moviefinder.com to make the personalised product recommendation based on the usage data of previous decisions made for similar objectives. It has also been applied in the context of e-service (Zheng et al. 2011), e-government (Shambour and Lu 2011) and e-business (Shambour and Lu 2012) to make the personalised service recommendation by predicting the missing user-dependent QoS values.

To our best knowledge, no previous studies have explored collaborative filtering techniques for personalised manufacturing service recommendations to support service-oriented manufacturing supply chain deployment effectively. There exist three main problems: (1) There is a severe information overload problem because of the lack of a method for managing social networks with vast numbers of manufacturing services and service users, which usually have significant effects on the user-dependent service recommendation results produced by collaborative filtering; (2) There is a severe data sparsity problem because there are usually too few QoS values of manufacturing services for different service users to identify similar users or services during collaborative filtering because of the expensive or time-consuming real-world execution of manufacturing services; and (3) There is a severe cold start problem because lots of new joining service users or new published services have no historical usage data that can be used to identify similar users or services during collaborative filtering due to the dynamicity of service-oriented manufacturing supply chain.

To address these drawbacks when exploring collaborative filtering to make personalised manufacturing service recommendation, we combine social network information with collaborative filtering to enhance the effectiveness of personalised QoS-aware service recommendations. First, the social network is built with the service users and manufacturing services as the nodes, while their preferences and tagging relationships form the links, which facilitate the calculation of the preference and tagging similarities between service users and between manufacturing services, respectively. Second, the collaborative filtering is employed to calculate the rating similarities of the QoS data between service users and between manufacturing services, respectively. Third, the preference, tagging and rating similarities between service users or between manufacturing services are combined and evaluated to find the nearest neighbours with the highest similarities for an active service user or target manufacturing service. Fourth, the missing QoS values of manufacturing services for active service users are predicted using collaborative filtering and the suggested neighbour groups. Finally, with both available and predicted QoS values, the QoS-aware service recommendation specific to the needs of active service user can be made to return the most interested manufacturing service to the active service user.

The social network information is combined with collaborative filtering to identify the nearest neighbours with the highest similarities for an active service user or a target manufacturing service, so even if most of the QoS values of manufacturing services are not available for different service users, the data sparsity and cold start problems can be alleviated.

A case study is given to demonstrate how the proposed approach is practical for personalised manufacturing service recommendation in the real application. The experimental results show that our proposed approach provides more accurate QoS predictions than the traditional approach with different neighbourhood sizes during personalised manufacturing service recommendation for supply chain deployment.

2. Related work

With SOA rapidly proliferated all over the world, the diverse service discovery and recommendation technologies have been explored widely in the past decade to support service-oriented enterprise application integration in distributed environments including manufacturing area. Many studies have focused on the semantics-based matchmaking of service capabilities between service requests and advertised services, using different service description languages, such as Manufacturing Service Description Language (MSDL) in Ameri and Dutta (2008), Web Service Description Language (WSDL) in Bouzakis et al. (2009), Web Ontology Language for Services (OWL-S) in Georgios and Nick (2010) and Semantic Annotations for WSDL (SAWSDL) in Omrana, Belouadha, and Roudies (2012). Our previous works have also employed the semantics-based service matchmaking to discover Semantic Grid service in Zhang and Yin (2009), Deep Web service in Zhang et al. (2012) and Semantic Web service in Zhang et al. (forthcoming) for cross-enterprise manufacturing collaboration. However, the semantics-based matchmaking may find some Web services with similar functions without considering the non-functional QoS information, giving users the difficulty in recommending most optimal services from these similar services that achieve the same functional task.

QoS-aware service discovery and recommendation has been explored recently, allowing optimal services with good QoS values to be selected from vast functionally similar services. Wang, Lee, and Ho (2007) embedded both objective and subjective QoS properties in Universal Description, Discovery and Integration (UDDI), which considers not only the objective factors described by service providers but also the subjective information with trustability evaluations from past users. Li et al. (2008) incorporated a QoS model in a Grid Service search engine, to filter functionally matched services with appropriate QoS values to maximise user satisfaction in service discovery. Wang, Chao, and Lo (2010) proposed a QoS-aware service selection model based on fuzzy linear programming technologies, in order to identify the dissimilarity on service alternatives, and assist service consumers in selecting most suitable services with consideration of their expectations and preferences of QoS. Dou et al. (2012) presented a QoS-aware service evaluation method for a shared service's co-selection, by solving a multi-criteria decision-making problem among multiple QoS properties. Our previous work (Zhang et al. 2013) has presented a QoS-aware Bayesian approach for recommending a few optimal manufacturing services. It adds a QoS model in manufacturing service ontology to support QoS-aware semantic matchmaking for service recommendation in the dynamic manufacturing environment. However, the above QoS-aware service discovery and recommendation works assume that all QoS values are user-independent, i.e. same and known to all service users, while it is not true in reality, because most of QoS values (e.g. response time, lead time, availability, performance and reliability) are subjective, rated differently by or missing to the diverse service users.

Following its successful application in making the personalised product recommendation in commercial Web-based systems, the collaborative filtering method has recently been adopted in the community of service science, management and engineering to discover and recommend the personalised Web services by predicting the missing user-dependent QoS values. Zheng et al. (2011) presented a collaborative filtering approach for predicting QoS values of Web services and making Web service recommendation by taking advantages of past usage experiences of service users, by systematically combining the user-based filtering approach and item-based filtering approach. Shambour and Lu (2011) proposed an intelligent trust-enhanced collaborative filtering recommendation approach to recommend the personalised e-government services, by integrating both the trust values derived from the user-rating data and the proportions of the common and uncommon ratings derived from the computation of the similarity between each pair of users. Shambour and Lu (2012) extended their previous work by incorporating additional information from the users' social trust network and the items' semantic domain knowledge to alleviate the data sparsity and cold start problems in collaborative filtering-based e-business service recommendation. Zhong et al. (2012) proposed a unified collaborative filtering model as a learning classification problem for Web service recommendation to help users select the suitable Web service, by combining the latent and external features of users and services through probabilistic semantic analysis. Chan, Gaaloul, and Tata (2012) applied collaborative filtering technique on Web service operations and service users to build a recommender system for Web service discovery, by taking into account users' interactions, which reflect users' behaviours and interests via the interaction matrixes of users' IDs and their used Web service operations. Wu et al. (2013) presented a neighbourhood-based collaborative filtering approach to predict unknown QoS values for QoS-based Web service selection, by adjusting the cosine-based similarity calculation to remove the impact of different QoS scale.

The work described in this paper is inspired by the above collaborative filtering-based service recommendation approaches, but has extended them by exploring the usage of preference and tagging relationships among service users and manufacturing services in making personalised manufacturing service recommendation. The important enhancement facilitates the implementation of the methodology in a computer program to generate a more comprehensive social network to alleviate the data sparsity and cold start problems the traditional collaborative filtering techniques have suffered from.

3. Overview of the proposed approach

The goal of this research is to incorporate social network information into collaborative filtering to enhance the effectiveness of personalised QoS-aware manufacturing service recommendation. Figure 1 illustrates the overall system architecture, which is divided into three stages: the offline manufacturing service modelling phase, the offline similarity computation phase and online service recommendation phase. To reduce the time complexity of the online computation, the offline manufacturing service modelling phase and similarity computation phase model, store and preprocess the offline data. The online service recommendation phase makes the personalised manufacturing service recommendation based on the usage data of previous retrievals from past service users.

3.1 *The offline manufacturing service modelling phase*

The recent popularity of Semantic Web technology (Berners-Lee, Hendler, and Lassila 2001) has made it possible the utilisation of ontological knowledge to represent the shared manufacturing services formally to facilitate both semantics-based service matchmaking and non-functional service recommendation, by annotating manufacturing services with generalised ontologies. Our previous works (Cai, Zhang, and Zhang 2011; Zhang et al. 2013) have developed a rich

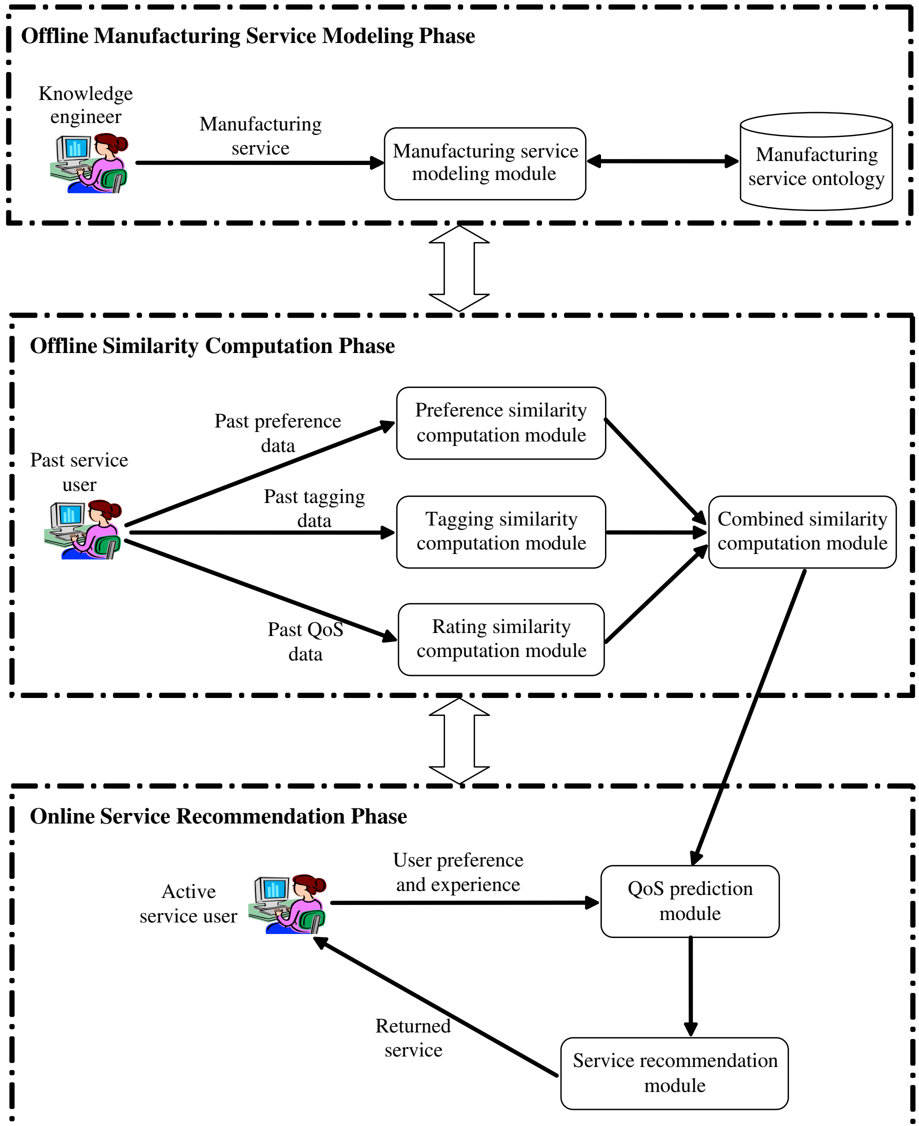

Figure 1. The overall system architecture.

body of OWL-based (McGuinness and Harmelen 2004) manufacturing service ontologies, which are represented by semantic properties and QoS properties. The semantic properties model the manufacturing capabilities including *Manufacturing type*, *Manufacturing operation*, *Manufacturing object*, *Manufacturing feature* and *Manufacturing equipment*. The QoS properties include *Response time*, *Lead time*, *Service cost*, *Time cost*, *Availability*, *Integrity*, *Performance*, *Reliability*, *Interoperability* and *Security* (Mathew, Shields, and Verma 2004).

The semantic properties are used to facilitate semantics-based manufacturing service matchmaking through semantic similarity calculation (Cai, Zhang, and Zhang 2011; Zhang et al. 2013), which may find some manufacturing services with semantically similar capabilities. Then comparison of the QoS properties can be used to recommend most optimal services from these similar services that achieve the same functional task.

3.2 *The offline similarity computation phase*

The offline similarity computation phase consists of four interactive functional modules: preference similarity computation module, tagging similarity computation module, rating similarity computation module and combined similarity computation module. The below are the logically interactive steps of four functional modules:

(1) *Preference similarity computation module*. Based on the preference data of previous retrievals from past service users, the social preference network is built, with service users and manufacturing services as nodes, and their preference relationships as links, to calculate the preference similarities between service users and between manufacturing services, respectively.
(2) *Tagging similarity computation module*. Based on the tagging data of previous retrievals from past service users, the social tagging network is built, with service users and manufacturing services as nodes, and their tagging relationships as links, to calculate the tagging similarities between service users and between manufacturing services, respectively.
(3) *Rating similarity computation module*. Based on the QoS data of previous retrievals from past service users, the collaborative filtering is employed to calculate the rating similarities of QoS data between service users and between manufacturing services, respectively.
(4) *Combined similarity computation module*. The above preference, tagging and rating similarities between service users or between manufacturing services are combined to produce the combined similarities so as to find the nearest neighbours with the highest similarities for each service user or manufacturing service.

3.3 *The online service recommendation phase*

For online personalised QoS-aware manufacturing service recommendation, the semantics-based manufacturing service matchmaking is firstly adopted to find some manufacturing services with semantically similar capabilities through semantic similarity calculation, as can be referred to our previous work (Cai, Zhang, and Zhang 2011; Zhang et al. 2013), and will not be repeated here for conciseness. Then the two interactive functional modules, i.e. QoS prediction module and service recommendation module, will be used to compare the QoS properties and recommend most optimal services from these similar services that achieve the same functional task. The below are the logically interactive steps of these functional modules:

(1) *QoS prediction module*. Based on different user preferences and experiences, the missing QoS values of manufacturing services for the active service user are predicted, employing collaborative filtering and the suggested neighbour groups, and utilising combined similarities between both service users and manufacturing services.
(2) *Service recommendation module*. Based on the predicted QoS values of diverse manufacturing services for the active service user, the personalised manufacturing service recommendation can be generated to the active service user through single QoS comparison or multi-objective decision-making among multiple QoSs.

4. Offline similarity computation and nearest neighbours selection

To utilise the usage data from other similar service users and similar services as nearest neighbours to support the personalised manufacturing service recommendation, both user similarity sets and service similarity sets, i.e. nearest neighbours, based on the preference, tagging and QoS data of previous retrievals from past service users can be discovered and combined in offline time to expedite the online recommendation process.

4.1 *Preference similarity computation*

The more the number of times that a service user has used a manufacturing service, the more preference the service user has on the manufacturing service. Based on the preference data of previous retrievals from past service users, there exist some preference links between service users and manufacturing services that are represented as social nodes. Given a data-set consisting of M service users and N manufacturing services, the preference relationships between service users and manufacturing services are denoted by an $M \times N$ user-service preference matrix. Every entry in this matrix, $e_{m,n}$ represents the number of times that the service user U_m has used the manufacturing services S_n. If the service user U_m has not used the manufacturing services S_n before, then $e_{m,n} = 0$. A popular similarity calculation method called Pearson Correlation Coefficient (PCC) (Rodgers and Nicewander 1988) is used to compute the preference similarity based on the linear correlation between service users and manufacturing services.

4.1.1 *Preference similarity computation between service users*

PCC is employed to compute the preference similarity between two service users U_{m1} and U_{m2} using the following equation:

$$\Pr e\_Sim(U_{m1}, U_{m2}) = \frac{\sum_{i=1}^{I1} (e_{m1,i} - \bar{e}_{m1}) \times (e_{m2,i} - \bar{e}_{m2})}{\sqrt{\sum_{i=1}^{I1} (e_{m1,i} - \bar{e}_{m1})^2} \times \sqrt{\sum_{i=1}^{I1} (e_{m2,i} - \bar{e}_{m2})^2}} \tag{1}$$

where $I1$ denotes the number of manufacturing services that can be used by both service users; $e_{m1,i}$ and $e_{m2,i}$ denote the numbers of times that the service users U_{m1} and U_{m2} have used the manufacturing services S_i, respectively; and $\bar{e}_{m1}$ and $\bar{e}_{m2}$ denote the average numbers of times that the service users U_{m1} and U_{m2} have used the manufacturing services, respectively. From this definition, the preference similarity between two service users $\Pr e\_Sim(U_{m1}, U_{m2})$ is normalised on interval [−1, 1], with a larger value indicating they are more similar. If $I1 = 0$, $\Pr e\_Sim(U_{m1}, U_{m2}) = null$.

4.1.2 *Preference similarity computation between manufacturing services*

Similarly, PCC is employed to compute the preference similarity between two manufacturing services S_{n1} and S_{n2} using the following equation:

$$\Pr e\_Sim(S_{n1}, S_{n2}) = \frac{\sum_{j=1}^{J1} (e_{j,n1} - \bar{e}_{n1}) \times (e_{j,n2} - \bar{e}_{n2})}{\sqrt{\sum_{j=1}^{J1} (e_{j,n1} - \bar{e}_{n1})^2} \times \sqrt{\sum_{j=1}^{J1} (e_{j,n2} - \bar{e}_{n2})^2}} \tag{2}$$

where $J1$ denotes the number of service users that can use both manufacturing services; $e_{j,n1}$ and $e_{j,n2}$ denote the numbers of times that the service users U_j has used the manufacturing services S_{n1} and S_{n2}, respectively; and $\bar{e}_{n1}$ and $\bar{e}_{n2}$ denote the average numbers of times that the service users have used the manufacturing services S_{n1} and S_{n2}, respectively. From this definition, the preference similarity between two manufacturing services $\Pr e\_Sim(S_{n1}, S_{n2})$ is normalised on interval [−1, 1], with a larger value indicating they are more similar. If $J1 = 0$, $\Pr e\_Sim(S_{n1}, S_{n2}) = null$.

4.2 *Tagging similarity computation*

A service user may mark a manufacturing service with a list of tags to express his/her interest to it. Based on the tagging data of previous retrievals from past service users, there exist some tagging links between service users and manufacturing services that are represented as social nodes. Given a data-set consisting of M service users and N manufacturing services, the tagging relationships between service users and manufacturing services are denoted by an $M \times N$ user-service tagging matrix. Every entry in this matrix, $t_{m,n}$ represents a list of tags, with which the service user U_m has marked the manufacturing services S_n. If the service user U_m has not marked the manufacturing services S_n before, then $t_{m,n} = null$.

4.2.1 *Tagging similarity computation between service users*

The tagging similarity between two service users U_{m1} and U_{m2} is computed using the following equation:

$$Tag\_Sim(U_{m1}, U_{m2}) = \frac{\sum_{i=1}^{I2} Sem\_Sim(t_{m1,i}, t_{m2,i})}{I2} \tag{3}$$

where $I2$ denotes the number of manufacturing services that have been co-tagged by both service users; $t_{m1,i}$ and $t_{m2,i}$ denote two lists of tags, with which the service users U_{m1} and U_{m2} have marked the manufacturing services S_i, respectively; and $Sem\_Sim(t_{m1,i}, t_{m2,i})$ denotes the semantic similarity between two lists of tags $t_{m1,i}$ and $t_{m2,i}$.

To calculate $Sem\_Sim(t_{m1,i}, t_{m2,i})$, let's assume $t_{m1,i}$ and $t_{m2,i}$ comprise of u and v tags, respectively, and $u \leqslant v$. They are represented as $t_{m1,i} = (Tag_{m1,i,1}, Tag_{m1,i,2}, \ldots, Tag_{m1,i,u})$ and $t_{m2,i} = (Tag_{m2,i,1}, Tag_{m2,i,2}, \ldots, Tag_{m2,i,v})$, respectively. $Sem\_Sim(t_{m1,i}, t_{m2,i})$ can be transformed by finding a subset $t'_{m2,i}$ of $t_{m2,i}$ such that $t_{m1,i}$ and $t'_{m2,i}$ comprise of equal u tags and the semantic similarity between them, i.e., $Sem\_Sim(t_{m1,i}, t'_{m2,i})$ can be maximised. It is expressed in the following equation:

$$Sem\_Sim(t_{m1,i}, t_{m2,i}) = \frac{u}{v} \times \text{Maximise}(Sem\_Sim(t_{m1,i}, t'_{m2,i})) \tag{4}$$

Let $t_{m1,i}$ and $t'_{m2,i}$ be represented as $t_{m1,i} = (Tag_{m1,i,1}, Tag_{m1,i,2}, \ldots, Tag_{m1,i,u})$ and $t'_{m2,i} = (Tag'_{m2,i,1}, Tag'_{m2,i,2}, \ldots, Tag'_{m2,i,u})$, respectively. $Sem\_Sim(t_{m1,i}, t'_{m2,i})$ is defined as the average of corresponding semantic similarities of involved tags, and expressed in following equation:

$$Sem\_Sim(t_{m1,i}, t'_{m2,i}) = \frac{1}{u} \sum_{k=1}^{u} Sem\_Sim(Tag_{m1,i,k}, Tag'_{m2,i,k}) \tag{5}$$

Mature ontology-based semantic similarity calculation methods of general concepts exist in the literature (Jiang and Conrath 1997; Li, Bandar, and McLean 2003). Wordnet can be used as the fundamental ontology to construct the lexical taxonomy among concepts to disambiguate the concepts and capture the semantic relationships among them (Xu et al. 2012). Our previous work has also developed a manufacturing ontology for semantic similarity calculation of manufacturing concepts in manufacturing service discovery (Cai, Zhang, and Zhang 2011). The above ontology-based semantic similarity calculation methods of general concepts and manufacturing concepts can be used to calculate the semantic similarities of tags.

$Sem\_Sim(Tag_{m1,i,k}, Tag'_{m2,i,k})$ can be normalised on interval [0, 1] through ontology-based semantic similarity calculation. Then, $Sem\_Sim(t_{m1,i}, t'_{m2,i})$ will be normalised on interval [0,1] too, so do $Sem\_Sim(t_{m1,i}, t_{m2,i})$ and $Tag\_Sim(U_{m1}, U_{m2})$. In other words, the tagging similarity between two service users $Tag\_Sim(U_{m1}, U_{m2})$ is normalised on interval [0, 1], with a larger value indicating they are more similar. If $I2 = 0$, $Tag\_Sim(U_{m1}, U_{m2}) = null$.

4.2.2 *Tagging similarity computation between manufacturing services*

Similarly, the tagging similarity between two manufacturing services S_{n1} and S_{n2} is computed using the following equation:

$$Tag\_Sim(S_{n1}, S_{n2}) = \frac{\sum_{j=1}^{J2} Sem\_Sim(t_{j,n1}, t_{j,n2})}{J2} \tag{6}$$

where $J2$ denotes the number of service users that have co-tagged both manufacturing services; $t_{j,n1}$ and $t_{j,n2}$ denote two lists of tags, with which the service users U_j has marked the manufacturing services S_{n1} and S_{n2}, respectively; and $Sem\_Sim(t_{j,n1}, t_{j,n2})$ denotes the semantic similarity between two lists of tags $t_{j,n1}$ and $t_{j,n2}$.

To calculate $Sem\_Sim(t_{j,n1}, t_{j,n2})$, let's assume $t_{j,n1}$ and $t_{j,n2}$ comprise of u and v tags, respectively and $u \leqslant v$. They are represented as $t_{j,n1} = (Tag_{j,n1,1}, Tag_{j,n1,2}, \ldots, Tag_{j,n1,u})$ and $t_{j,n2} = (Tag_{j,n2,1}, Tag_{j,n2,2}, \ldots, Tag_{j,n2,v})$, respectively. $Sem\_Sim(t_{j,n1}, t_{j,n2})$ can be transformed by finding a subset $t'_{j,n2}$ of $t_{j,n2}$ such that $t_{j,n1}$ and $t'_{j,n2}$ comprise of equal u tags and the semantic similarity between them, i.e. $Sem\_Sim(t_{j,n1}, t'_{j,n2})$ can be maximised. It is expressed in the following equation:

$$Sem\_Sim(t_{j,n1}, t_{j,n2}) = \frac{u}{v} \times Maximise(Sem\_Sim(t_{j,n1}, t'_{j,n2})) \tag{7}$$

Let $t_{j,n1}$ and $t'_{j,n2}$ be represented as $t_{j,n1} = (Tag_{j,n1,1}, Tag_{j,n1,2}, \ldots, Tag_{j,n1,u})$ and $t'_{j,n2} = (Tag'_{j,n2,1}, Tag'_{j,n2,2}, \ldots, Tag'_{j,n2,u})$, respectively. $Sem\_Sim(t_{j,n1}, t'_{j,n2})$ is defined as the average of corresponding semantic similarities of involved tags, and expressed in following equation:

$$Sem\_Sim(t_{j,n1}, t'_{j,n2}) = \frac{u}{v} \times \sum_{k=1}^{u} Sem\_Sim(Tag_{j,n1,k}, Tag'_{j,n2,k}) \tag{8}$$

The mature ontology-based semantic similarity calculation methods of general concepts and manufacturing concepts can be used to calculate the semantic similarities of tags.

$Sem\_Sim(Tag_{j,n1,k}, Tag'_{j,n2,k})$ can be normalised on interval [0, 1] through ontology-based semantic similarity calculation. Then, $Sem\_Sim(t_{j,n1}, t'_{j,n2})$ will be normalised on interval [0, 1] too, so do $Sem\_Sim(t_{j,n1}, t_{j,n2})$ and $Tag\_Sim(S_{n1}, S_{n2})$. In other words, the tagging similarity between two manufacturing services $Tag\_Sim(S_{n1}, S_{n2})$ is normalised on interval [0, 1], with a larger value indicating they are more similar. If $J2 = 0$, $Tag\_Sim(U_{m1}, U_{m2}) = null$.

4.3 *Rating similarity computation*

A service user may rate the QoS values of a manufacturing service through numerical numbers (e.g. 1, 3, 6, etc.) or fuzzy numbers (e.g. good, fair, worse, etc.) to express his/her usage feedback to it. The fuzzy numbers can be converted to the numerical numbers to facilitate computer processing. Based on the rating data of previous retrievals from past service users, there exist some rating links between service users and manufacturing services that are represented as social nodes. Given a data-set consisting of M service users and N manufacturing services, the rating relationships between service users and manufacturing services are denoted by an $M \times N$ user-service rating matrix. Every entry in this matrix, $r_{m,n}$ represents a vector of QoS values (e.g. response time, lead time, availability, performance and reliability), with which the service user U_m has rated the manufacturing services S_n. If the service user U_m has not rated the manufacturing services S_n before, then $r_{m,n} = null$. A popular similarity calculation method called PCC (Rodgers and Nicewander 1988) is used and extended to compute the rating similarity based on the linear correlation between service users and manufacturing services.

4.3.1 *Rating similarity computation between service users*

PCC can be employed to compute the rating similarity of QoS data between two service users U_{m1} and U_{m2} using the following equation:

$$Rat\_Sim(U_{m1}, U_{m2}) = \frac{\sum_{i=1}^{I3} (r_{m1,i} - \bar{r}_{m1}) \times (r_{m2,i} - \bar{r}_{m2})}{\sqrt{\sum_{i=1}^{I3} (r_{m1,i} - \bar{r}_{m1})^2} \times \sqrt{\sum_{i=1}^{I3} (r_{m2,i} - \bar{r}_{m2})^2}} \tag{9}$$

where $I3$ denotes the number of manufacturing services that have been co-rated by both service users; $r_{m1,i}$ and $r_{m2,i}$ denote the vector of QoS values that the service users U_{m1} and U_{m2} have rated the manufacturing services S_i, respectively; and $\bar{r}_{m1}$ and $\bar{r}_{m2}$ denote the vector of average QoS values that the service users U_{m1} and U_{m2} have rated the manufacturing services, respectively. From this definition, the rating similarity between two service users $Rat\_Sim(U_{m1}, U_{m2})$ is normalised on interval $[-1, 1]$, with a larger value indicating they are more similar. If $I3 = 0$, $Rat\_Sim(U_{m1}, U_{m2}) = null$.

4.3.2 *Rating similarity computation between manufacturing services*

Similarly, PCC is employed to compute the rating similarity of QoS data between two manufacturing services S_{n1} and S_{n2} using the following equation:

$$Rat\_Sim(S_{n1}, S_{n2}) = \frac{\sum_{j=1}^{J3} (r_{j,n1} - \bar{r}_{n1}) \times (r_{j,n2} - \bar{r}_{n2})}{\sqrt{\sum_{j=1}^{J3} (r_{j,n1} - \bar{r}_{n1})^2} \times \sqrt{\sum_{j=1}^{J3} (r_{j,n2} - \bar{e}_{n2})^2}} \tag{10}$$

where $J3$ denotes the number of service users that have co-rated both manufacturing services; $r_{j,n1}$ and $r_{j,n2}$ denote the vector of QoS values that the service users U_j has rated the manufacturing services S_{n1} and S_{n2}, respectively; and $\bar{r}_{n1}$ and $\bar{r}_{n2}$ denote the vector of average QoS values that the service users have rated the manufacturing services S_{n1} and S_{n2}, respectively. From this definition, the rating similarity between two manufacturing services $Rat\_Sim(S_{n1}, S_{n2})$ is normalised on interval $[-1, 1]$, with a larger value indicating they are more similar. If $J3 = 0$, $Rat\_Sim(U_{m1}, U_{m2}) = null$.

4.4 *Combined similarity computation and nearest neighbours selection*

The preference, tagging and rating similarities between service users or between manufacturing services are combined to produce the comprehensive combined similarities, so as to find the nearest neighbours with the highest similarities for each service user or manufacturing service.

4.4.1 *Combined similarity computation between service users*

The combined similarity between two service users U_{m1} and U_{m2} is computed using the following equation:

$$Com\_Sim(U_{m1}, U_{m2}) = \alpha_1 \times \Pr e\_Sim(U_{m1}, U_{m2}) + \alpha_2 \times Tag\_Sim(U_{m1}, U_{m2}) + \alpha_3 \times Rat\_Sim(U_{m1}, U_{m2}) \tag{11}$$

where α_1, α_2 and α_3 denote the weights of partial similarities to indicate their respective importance towards combined similarities, with $\alpha_1+\alpha_2+\alpha_3=1$, if none of partial similarities has *null* value. If one partial similarity has *null* value, for example, if $Rat\_Sim(U_{m1}, U_{m2})=null$, then $\alpha_3=0$ and $\alpha_1+\alpha_2=1$. If two partial similarities have *null* values, for example, if both $Rat\_Sim(U_{m1}, U_{m2})=null$ and $Tag\_Sim(U_{m1}, U_{m2})=null$, then $\alpha_2=\alpha_3=0$ and $\alpha_1=1$. If all three partial similarities have *null* values, then $Com\_Sim(U_{m1}, U_{m2})=null$.

For a service user U_m, its similarity set contains all service users whose similarities with U_m are greater than a preset threshold between 0 and 1. Thus, the user-based nearest neighbours are obtained by calculating the similarities of all service users with U_m, respectively, and filtering out those that are smaller than or equal to the threshold.

Because the preference, tagging and rating similarities between service users are combined to produce the combined similarities, even if one or two partial similarities are not available, the combined similarities can still be obtained. Therefore, the user-based data sparsity and cold start problems can be alleviated.

4.4.2 *Combined similarity computation between manufacturing services*

The combined similarity between two manufacturing services S_{n1} and S_{n2} is computed using the following equation:

$$Com\_Sim(S_{n1}, S_{n2}) = \beta_1 \times Pre\_Sim(S_{n1}, S_{n2}) + \beta_2 \times Tag\_Sim(S_{n1}, S_{n2}) + \beta_3 \times Rat\_Sim(S_{n1}, S_{n2}) \tag{12}$$

where β_1, β_2 and β_3 denote the weights of partial similarities to indicate their respective importance towards combined similarities, with $\beta_1+\beta_2+\beta_3=1$, if none of partial similarities has *null* value. If one partial similarity has *null* value, for example, if $Rat\_Sim(S_{n1}, S_{n2})=null$, then $\beta_3=0$ and $\beta_1+\beta_2=1$. If two partial similarities have *null* values, for example, if both $Rat\_Sim(S_{n1}, S_{n2})=null$ and $Tag\_Sim(S_{n1}, S_{n2})=null$, then $\beta_2=\beta_3=0$ and $\beta_1=1$. If all three partial similarities have *null* values, then $Com\_Sim(S_{n1}, S_{n2})=null$.

For a manufacturing service S_n, its similarity set contains all service users whose similarities with S_n are greater than a preset threshold between 0 and 1. Thus, the service-based nearest neighbours are obtained by calculating the similarities of all manufacturing services with S_n, respectively, and filtering out those that are smaller than or equal to the threshold.

Because the preference, tagging and rating similarities between manufacturing services are combined to produce the combined similarities, even if one or two partial similarities are not available, the combined similarities can still be obtained. Therefore, the service-based data sparsity and cold start problems can be alleviated.

5. Online manufacturing service recommendation

After the nearest neighbours with the highest similarities for each service user or manufacturing service are found through similarity computation, the missing QoS values of manufacturing services for the active service user can be predicted, followed by final service recommendation through single QoS comparison or multi-objective decision-making among multiple QoSs.

5.1 *QoS prediction*

Given a data-set consisting of M service users and N manufacturing services, the rating relationships between service users and manufacturing services are denoted by an $M \times N$ user-service rating matrix. Every entry in this matrix, $r_{m,n}$ represents a vector of QoS values, with which the service user U_m has rated the manufacturing services S_n. If the service user U_m has not rated the manufacturing services S_n before, then $r_{m,n}=null$. Therefore, using the weighted sum of deviations from the neighbour-based collaborative filtering algorithm (Herlocker, Konstan, and Riedl 2002), the proposed approach employs both similar users through user similarity sets and similar services through service similarity sets to predict the missing QoS values for the active service user.

The user-based filtering method employs the similar service users to predict the missing QoS values for the active service user using the following equation:

$$P_u(r_{m,n}) = \bar{r}_m + \frac{\sum_{k=1}^{K} Com\_Sim(U_m, U_k) \times (r_{k,n} - \bar{r}_k)}{\sum_{k=1}^{K} Com\_Sim(U_m, U_k)} \tag{13}$$

where $P_u(r_{m,n})$ denotes a vector of predicted QoS values of the entry $r_{m,n}$ in the $M \times N$ user-service rating matrix using the user-based filtering method; K denotes the number of the nearest neighbours with the highest similarities for the active service user U_m; $\bar{r}_m$ and $\bar{r}_k$ denote the vector of average QoS values that the service user U_m and its similar service user U_k have rated different manufacturing services, respectively; $r_{k,n}$ denotes the vector of QoS values that the service user U_k has rated the manufacturing services S_n; and $Com\_Sim(U_m, U_k)$ denotes the combined similarity between the service user U_m and its similar service user U_k.

The service-based filtering method employs the similar manufacturing services to predict the missing QoS values for the active service user using the following equation:

$$P_s(r_{m,n}) = \bar{r}_n + \frac{\sum_{l=1}^{L} Com\_Sim(S_n, S_l) \times (r_{m,l} - \bar{r}_l)}{\sum_{l=1}^{L} Com\_Sim(S_n, S_l)} \tag{14}$$

where $P_s(r_{m,n})$ denotes a vector of predicted QoS values of the entry $r_{m,n}$ in the $M \times N$ user-service rating matrix using the service-based filtering method; L denotes the number of the nearest neighbours with the highest similarities for the target manufacturing service S_n; $\bar{r}_n$ and $\bar{r}_l$ denote the vector of average QoS values that different service users have rated the manufacturing service S_n and its similar manufacturing service S_l, respectively; $r_{m,l}$ denotes the vector of QoS values that the service user U_m has rated the manufacturing services S_l; and $Com\_Sim(S_n, S_l)$ denotes the combined similarity between the manufacturing service S_n and its similar manufacturing service S_l.

When a missing QoS value does not have similar service users that have rated the target manufacturing service, we use the service-based filtering method to predict the missing QoS value. When a missing QoS value does not have similar manufacturing services that have been rated by the active service user, we use the user-based filtering method to predict the missing QoS value.

When a missing QoS does have both similar service users that have rated the target manufacturing service, and similar manufacturing services that have been rated by the active service user, using different filtering method will result in different predicted QoS value. To systematically exploit both similar users and similar services to predict the missing QoS values, and reflect their respective importance towards combined predicted values, we can use combined filtering method to predict the missing QoS value according to following equation:

$$P_{com}(r_{m,n}) = \frac{P_u(r_{m,n}) \times \sum_{k=1}^{K} Com\_Sim(U_m, U_k) + P_s(r_{m,n}) \times \sum_{l=1}^{L} Com\_Sim(S_n, S_l)}{\sum_{k=1}^{K} Com\_Sim(U_m, U_k) + \sum_{l=1}^{L} Com\_Sim(S_n, S_l)} \tag{15}$$

where $P_{com}(r_{m,n})$, $P_u(r_{m,n})$ and $P_s(r_{m,n})$ denote a vector of predicted QoS values of the entry $r_{m,n}$ in the $M \times N$ user-service rating matrix using the combined filtering method, user-based filtering method and service-based filtering method, respectively; K denotes the number of the nearest neighbours with the highest similarities for the active service user U_m; L denotes the number of the nearest neighbours with the highest similarities for the target manufacturing service S_n; $Com\_Sim(U_m, U_k)$ denotes the combined similarity between the service user U_m and its similar service user U_k; and $Com\_Sim(S_n, S_l)$ denotes the combined similarity between the manufacturing service S_n and its similar manufacturing service S_l.

If the accumulated combined similarity between the service user U_m and all its similar service users is higher than that between the manufacturing service S_n and all its similar manufacturing services, i.e., $\sum_{k=1}^{K} Com\_Sim(U_m, U_k) > \sum_{l=1}^{L} Com\_Sim(S_n, S_l)$, the predicted QoS value using the user-based filtering method, i.e. $P_u(r_{m,n})$ has more importance than that using the service-based filtering method, i.e. $P_s(r_{m,n})$ in its contribution to the combined predicted QoS value, and vice versa.

Based on the above collaborative filtering method, if an active service user has used, tagged or rated more manufacturing services in his/her historical service retrieval, the offline similarity computation will be more accurate, which will consequently improve the online prediction accuracy of the missing QoS values in the online service recommendation. This causal relationship will encourage the service users to tag or rate more manufacturing services after using them.

5.2 *Service recommendation*

After the missing QoS values of diverse manufacturing services for the active service user have been predicted, the personalised manufacturing service recommendation can be generated to the active service user through single QoS comparison or multi-objective decision-making among multiple QoSs. The single QoS comparison will recommend the active service user, the top k optimal manufacturing services with higher QoS values. The multi-objective decision-making among multiple QoSs will recommend the active service user the top k optimal manufacturing services with higher weighted sum of multiple QoS values.

6. Case study and experimental evaluation

This section demonstrates how the proposed social network-enhanced collaborative filtering method achieves the practicality and accuracy to personalised manufacturing service recommendation in the real application for cross-enterprise collaboration. A Java-based object-oriented software prototype is implemented. Related manufacturing ontology and manufacturing service ontology for semantic similarity calculation of manufacturing concepts and manufacturing service concept, which have been developed in our previous work (Cai, Zhang, and Zhang 2011; Zhang et al. 2013) is loaded into the system through Protégé-2000 (Gennari et al. 2003), a widely accepted ontology editor.

The manufacturing service repository of the current prototype system under evaluation contains 867 manufacturing services. There are 145 registered users, who have provided preference data, tagging data and QoS data to some manufacturing services in their historical service retrievals. QoS ratings vary from '0' to '10', with '0' indicating the worst and '10' indicating the best.

6.1 *A case study of personalised manufacturing service recommendation*

This sub-section illustrates an example of practical personalised manufacturing service recommendation. An active service user U_1 is searching for a manufacturing service with the following semantic properties and QoS properties:

(1) Manufacturing type: *Mechanical processing*;
(2) Manufacturing operation: *Finish turning* with *Surface roughness* <2 μm, *Dimension accuracy* <5 μm and *Linearity* <3 μm;
(3) Manufacturing object: *Stainless steel shaft* with *semi-diameter* <30 cm, *length* <80 cm and *weight* <150 kg;
(4) Manufacturing feature: *Hole*, *Cylindrical Surface* and *Conic surface*.
(5) Performance QoS: *Best*.

Figure 2 shows the user interface for personalised manufacturing service recommendation in the prototype system. The active service user U_1 can develop the service query by clicking at the 'Build query tree', which makes it easy to formulate queries. The query conditions can be input in the right window, and the query tree can be updated and shown in the left window. Figure 2 is showing the query conditions of semantic properties in the right window when the 'Semantic properties' tab is selected, while showing both semantic properties and performance QoS property in the developed query tree in the left window.

After the 'Search' button at the bottom of query tree, the personalised manufacturing service recommendation process proceeds with the following eight steps:

(1) The semantics-based manufacturing service matchmaking is firstly adopted to find some manufacturing services with semantically similar capabilities through semantic similarity calculation, as can be referred to our previous work (Cai, Zhang, and Zhang 2011; Zhang et al. 2013), and will not be repeated here for conciseness. In this case, six lathe finish services (S_1 to S_6) that satisfy the semantics-based functional requirement of service user U_1 are found.
(2) The active service user U_1 now needs to select one of six semantics-matching lathe finish service with the best performance QoS. Suppose there are five historical service users (U_1 to U_5) including the active service user U_1, have provided preference data, tagging data or performance QoS data to these six lathe finish services in their historical service retrievals. Table 1 shows a user-service rating matrix, from which, we find the service user U_1 and manufacturing service S_6 are an extreme cold start user and extreme cold start service, respectively, because they have only one performance rating. The missing performance QoS values of all lathe finish services except S_5 for the active service user U_1 need to be predicted, so that the active service user U_1 can select a lathe finish service with the best performance QoS.
(3) *Preference similarity computation*. Based on the preference data of previous retrievals from past service users (U_1 to U_5), the user–service preference matrix is shown in Table 2. The preference similarity matrices between service users and between manufacturing services are computed using Equations (1) and (2), and shown in Tables 3 and 4, respectively.

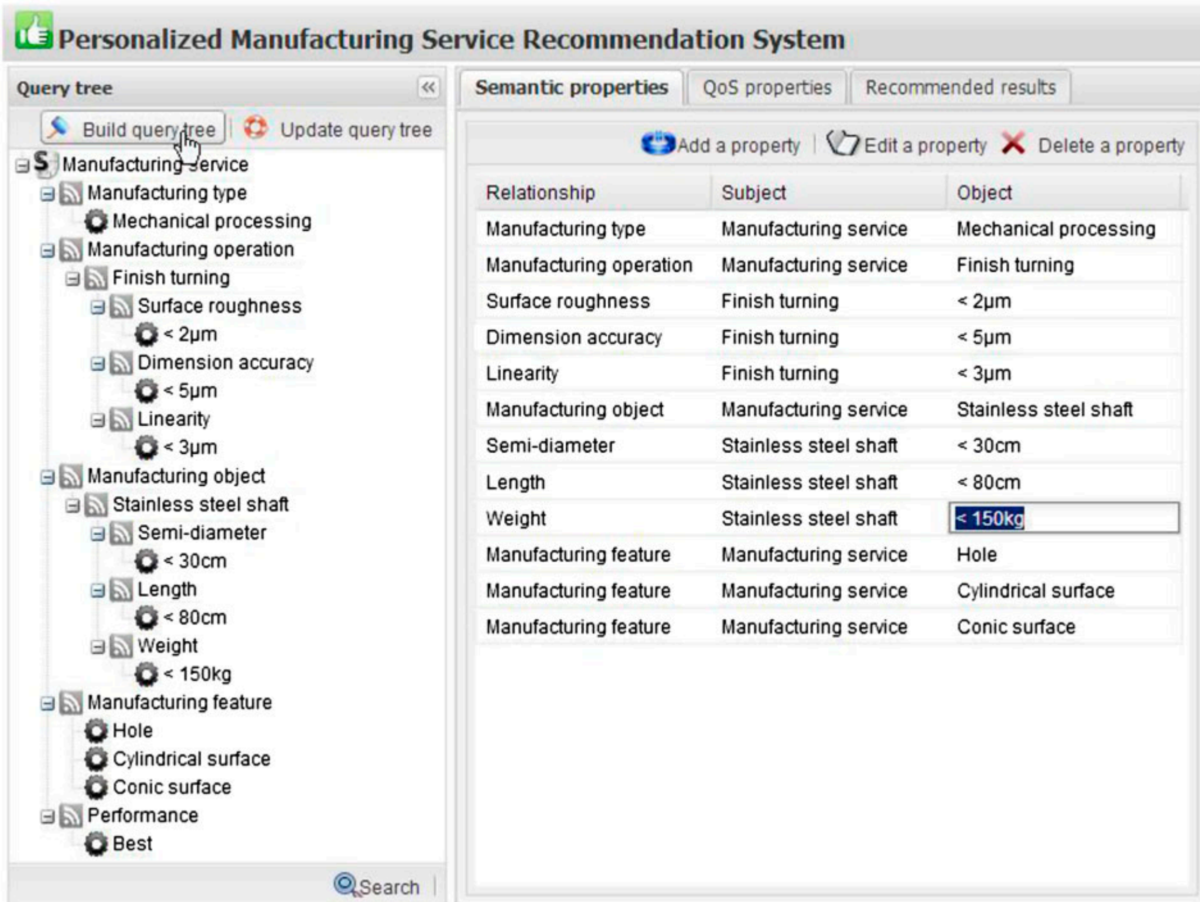

Figure 2. Graphic user interface for manufacturing service recommendation in the prototype system.

Table 1. The user-service rating matrix.

		Manufacturing services					
Performance QoS rating		S_1	S_2	S_3	S_4	S_5	S_6
Service users	U_1	*Null*	*Null*	*Null*	*Null*	9	*Null*
	U_2	6	5	9	*Null*	*Null*	*Null*
	U_3	3	4	*Null*	4	*Null*	*Null*
	U_4	2	*Null*	8	*Null*	7	*Null*
	U_5	*Null*	5	*Null*	9	*Null*	3

(4) *Tagging similarity computation*. Based on the tagging data of previous retrievals from past service users (U_1 to U_5), the user–service tagging matrix is shown in Table 5. The tagging similarity matrices between service users and between manufacturing services are computed using Equations (3) and (6), and shown in Tables 6 and 7, respectively.

(5) *Rating similarity computation*. According to the user-service rating matrix shown in Table 1, the rating similarity matrices between service users and between manufacturing services are computed using Equations (9) and (10), and shown in Tables 8 and 9, respectively.

(6) *Combined similarity computation*. The above preference, tagging and rating similarities between service users and between manufacturing services are combined to produce the combined similarity matrices between service users and between manufacturing services using Equations (11) and (12), and shown in Tables 10 and 11, respectively. In this case study, let three partial similarities have equal weights that sum to 1.

Table 2. The user-service preference matrix.

The number of usage times		Manufacturing services S_1	S_2	S_3	S_4	S_5	S_6
Service users	U_1	2	4	5	0	5	1
	U_2	3	6	9	0	0	0
	U_3	6	7	8	4	6	6
	U_4	7	0	7	0	2	5
	U_5	0	1	6	1	2	3

Table 3. The preference similarity matrix between service users.

Preference similarity		Service users U_1	U_2	U_3	U_4	U_5
Service users	U_1	1	0.592	0.786	0.129	0.445
	U_2	0.592	1	0.833	0.338	0.518
	U_3	0.786	0.833	1	0.483	0.622
	U_4	0.129	0.338	0.483	1	0.443
	U_5	0.445	0.518	0.622	0.443	1

Table 4. The preference similarity matrix between manufacturing services.

Preference similarity		Manufacturing services S_1	S_2	S_3	S_4	S_5	S_6
Manufacturing services	S_1	1	0.091	0.439	0.301	0.213	0.647
	S_2	0.091	1	0.518	0.521	0.335	−0.161
	S_3	0.439	0.518	1	0.274	−0.387	0.062
	S_4	0.301	0.521	0.274	1	0.648	0.679
	S_5	0.213	0.335	−0.387	0.648	1	0.48
	S_6	0.647	−0.161	0.062	0.679	0.48	1

(7) *Performance QoS prediction*. Based on the above combined similarities between service users and between manufacturing services, the missing performance QoS values of manufacturing services for the active service user U_1 can be predicted using Equation (15). In this case study, let the number of the nearest neighbours to be 3. As shown in Table 8, the active service user U_1 has no qualified neighbour, and is an extreme cold start user in terms of the original collaborative filtering method. However, as shown in Table 10, the active service user U_1 has now three qualified neighbours, i.e. the service users U_2, U_3 and U_5, in terms of the proposed social network-enhanced collaborative filtering method. Clearly, the proposed approach has alleviated the user-based cold start problem. Similarly, it is easy to find the user-based data sparsity problem, and service-based data sparsity and cold start problems can be alleviated as well. The predicted performance QoS values of manufacturing services for the active service user U_1 can be found in Table 12.

(8) *Manufacturing service recommendation*. Based on the predicted performance QoS values of diverse manufacturing services for the active service user U_1, the lathe finish service S_3 whose name is 'Lathe finish service-Zhenyuan' is recommended because it has the highest performance QoS value 10 for the active service user U_1. Figure 3 shows the recommended manufacturing services in the ranking order of performance QoS value, which are shown through the 'Recommended results' tab.

The above personalised manufacturing service discovery only takes a non-expert 10 s to obtain the solution, but it takes an experienced human planner at least 1 h to obtain the similar solution when many manufacturing services have been published in the Internet.

Table 5. The user-service tagging matrix.

Tags		Manufacturing services					
		S_1	S_2	S_3	S_4	S_5	S_6
Service users	U_1	Good service; Fair price; Good machine	Good reliability; Good price;	Null	Null	Good service; Good availability; Fast delivery	Easy communication; New machine
	U_2	Good equipment; Friendly service; Cheap cost	Good equipment; Fast response; Good price	Good equipment; Good communication; Cost effective	Null	Null	Null
	U_3	Speedy response	Good facility; Cheap service; Friendly communication	Null	Helpful staff; Reliable service; Easy communication	Null	Null
	U_4	Good machine; Fast response	Null	New machine; Fast delivery; Easy communication	Null	Good price; Easy communication	Null
	U_5	Null	Good price; Reliable service	Null	Good sales staff; Right service; Reliable equipment	Good equipment; Helpful staff	Fast service

Table 6. The tagging similarity matrix between service users.

Tagging similarity		Service users				
		U_1	U_2	U_3	U_4	U_5
Service users	U_1	1	0.683	0.267	0.317	0.439
	U_2	0.683	1	0.333	0.517	0.5
	U_3	0.267	0.333	1	0.45	0.467
	U_4	0.317	0.517	0.45	1	0.35
	U_5	0.439	0.5	0.467	0.35	1

Table 7. The tagging similarity matrix between manufacturing services.

Tagging similarity		Manufacturing services					
		S_1	S_2	S_3	S_4	S_5	S_6
Manufacturing services	S_1	1	0.455	0.567	0.167	0.309	0.367
	S_2	0.455	1	0.6	0.367	0.259	0.2
	S_3	0.567	0.6	1	*Null*	0.333	*Null*
	S_4	0.167	0.367	*Null*	1	0.5	0.167
	S_5	0.309	0.259	0.333	0.5	1	0.1
	S_6	0.367	0.2	*Null*	0.167	0.1	1

6.2 *Evaluating the accuracy of personalised manufacturing service recommendation*

In this section, we will use the evaluation metric of Mean Absolute Error (MAE) to evaluate the accuracy of QoS prediction of our proposed approach in comparison with other approaches. MAE is a widely used metric in personalised recommendation system (Herlocker, Konstan, and Riedl 2002).

$$\mathrm{MAE} = \frac{\sum_{s=1}^{S} |r_s - \hat{r}_s|}{S} \tag{16}$$

Table 8. The rating similarity matrix between service users.

Rating similarity		Service users				
		U_1	U_2	U_3	U_4	U_5
Service users	U_1	1	*Null*	*Null*	*Null*	*Null*
	U_2	*Null*	1	−0.083	0.748	1
	U_3	*Null*	−0.083	1	1	0.555
	U_4	*Null*	0.748	1	1	*Null*
	U_5	*Null*	1	0.555	*Null*	1

Table 9. The rating similarity matrix between manufacturing services.

Rating similarity		Manufacturing services					
		S_1	S_2	S_3	S_4	S_5	S_6
Manufacturing services	S_1	1	0.676	0.986	1	1	*Null*
	S_2	0.676	1	1	0.949	*Null*	*Null*
	S_3	0.986	1	1	*Null*	1	*Null*
	S_4	1	0.949	*Null*	1	*Null*	*Null*
	S_5	1	*Null*	1	*Null*	1	*Null*
	S_6	*Null*	*Null*	*Null*	*Null*	*Null*	1

Table 10. The combined similarity matrix between service users.

Combined similarity		Service users				
		U_1	U_2	U_3	U_4	U_5
Service users	U_1	1	0.638	0.527	0.223	0.442
	U_2	0.638	1	0.361	0.534	0.673
	U_3	0.527	0.361	1	0.644	0.548
	U_4	0.223	0.534	0.644	1	0.397
	U_5	0.442	0.673	0.548	0.397	1

Table 11. The combined similarity matrix between manufacturing services.

Combined similarity		Manufacturing services					
		S_1	S_2	S_3	S_4	S_5	S_6
Manufacturing services	S_1	1	0.407	0.664	0.489	0.507	0.507
	S_2	0.407	1	0.706	0.612	0.297	0.02
	S_3	0.664	0.706	1	0.274	0.315	0.062
	S_4	0.489	0.612	0.274	1	0.574	0.423
	S_5	0.507	0.297	0.315	0.574	1	0.29
	S_6	0.507	0.02	0.062	0.423	0.29	1

where S denotes the number of the total predictions, r_s denotes the predicted QoS value and $\hat{r}_s$ denotes the actual QoS value. The lower the MAE value, the higher the prediction accuracy.

About 145 registered service users are divided into training users and active service users whose proportion is 4:1. The actual QoS values rated by the training users are used as training sets to calculate the predicated QoS values for the active service users, and the actual QoS values rated by the active service users are used as the test sets for comparison.

We compare our proposed social network-enhanced collaborative filtering method with the traditional combined user- and service-based collaborative filtering method. In our proposed approach, we let three partial similarities have

Table 12. The predicted missing performance QoS values for the active service user U_1 in the user-service rating matrix.

Performance QoS rating		Manufacturing services S_1	S_2	S_3	S_4	S_5	S_6
Service users	U_1	**7.221**	**8.264**	**10**	**9.51**	9	**5.409**

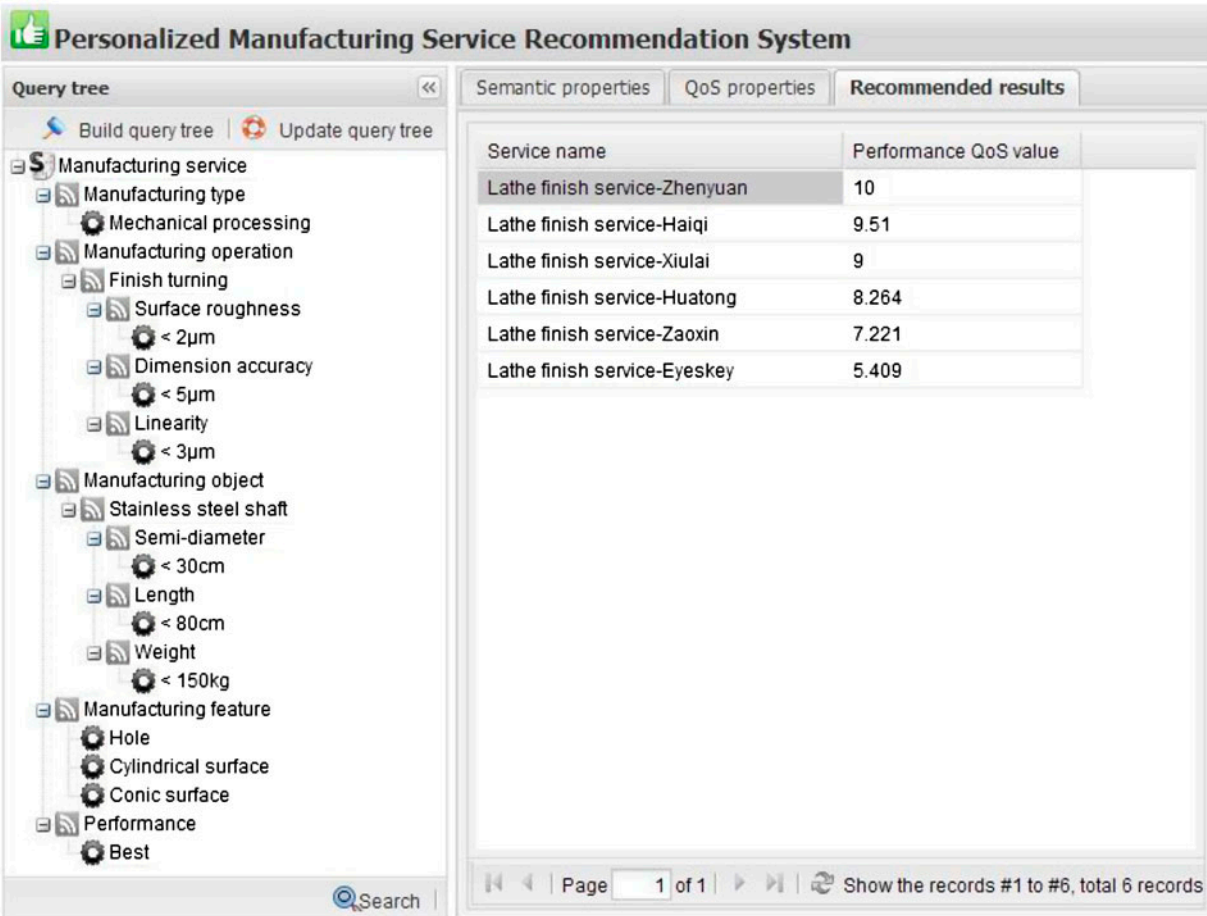

Figure 3. The recommended results in the illustrative example.

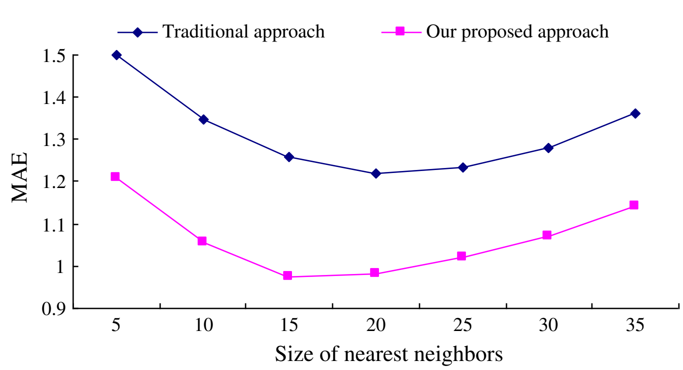

Figure 4. Comparison of accuracy of QoS prediction (MAE) between the traditional approach and our proposed approach with different size of nearest neighbours.

equal weights that sum to 1. Figure 4 shows the MAE values obtained from the traditional approach and our proposed approach with different size of nearest neighbours. The result demonstrates the accuracy of QoS prediction of our proposed approach outperforms the traditional approach (Figure 4).

7. Conclusion

This paper presents a social network-enhanced collaborative filtering technique to predict the missing QoS values of manufacturing services for the active service users, improving the effectiveness of personalszed QoS-aware service recommendation. As the implementation of personalised service recommendation systems in e-manufacturing is still in its infancy, the outcome of this research will be of great value for personalised cross-enterprise collaboration.

The social network explores the usage of preference and tagging relationships among service users and manufacturing services in making personalised recommendation, alleviating the data sparsity and cold-start problems, the traditional collaborative filtering techniques have suffered from, while achieving the higher prediction accuracy than the traditional collaborative filtering techniques.

However, the proposed approach is still in an early stage of development and has a limitation for large-scale real world applications. In our future work, we are going to improve our personalised manufacturing service recommendation system by taking into account the more comprehensive social network information, for example, the tracing records of posts, replies, home page visit and interactive activities, for more accurate prediction of QoS values in personalised manufacturing service recommendation.

Acknowledgements

The work has been supported by China National Natural Science Foundation (Nos. 51375429, 51175462, 71071144), Zhejiang Natural Science Foundation of China (No. LY13E050010, LY12G02006).

References

Adomavicius, G., and A. Tuzhilin. 2005. "Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-art and Possible Extensions." *IEEE Transactions on Knowledge and Data Engineering* 17 (6): 734–749.

Ameri, F., and D. Dutta. 2008. "A Matchmaking Methodology for Supply Chain Deployment in Distributed Manufacturing Environments." *Journal of Computing and Information Science in Engineering* 8 (1): 1–9.

Berners-Lee, T., J. Hendler, and O. Lassila. 2001. "The Semantic Web." *Scientific American* 284: 34–43.

Bouzakis, K.-D., G. Andreadis, A. Vakali, and M. Sarigiannidou. 2009. "Automating the Manufacturing Process under a Web Based Framework." *Advances in Engineering Software* 40 (9): 956–964.

Cai, M., W. Y. Zhang, and K. Zhang. 2011. "ManuHub: A Semantic Web System for Ontology-based Service Management in Distributed Manufacturing Environments." *IEEE Transactions on Systems, Man, and Cybernetics Part A: Systems and Humans* 41 (3): 574–582.

Chan, N. N., W. Gaaloul, and S. Tata. 2012. "A Recommender System Based on Historical Usage Data for Web Service Discovery." *Service Oriented Computing and Applications* 6: 51–63.

Dou, W. C., C. Lv, X. Y. Zhang, and J. J. Chen. 2012. "A Collaborative QoS-aware Service Evaluation Method Among Multi-users for a Shared Service." *International Journal of Web Services Research* 9 (1): 30–50.

Duke, A., J. Davies, and M. Michardson. 2005. "Enabling a Scalable Service-oriented Architecture with Semantic Web service." *BT Technology Journal* 23 (3): 191–201.

Gennari, J. H., M. A. Musen, R. W. Fergerson, W. E. Grosso, M. Crubezy, H. Eriksson, N. F. Noy, and S. W. Tu. 2003. "The Evolution of Protégé-2000: An Environment for Knowledge-based Systems Development." *International Journal of Human-Computer Studies* 58: 89–123.

Georgios, M., and B. Nick. 2010. "Structural and Role-oriented Web Service Discovery with Taxonomies in OWL-S." *IEEE Transactions on Knowledge and Data Engineering* 22 (2): 278–290.

Herlocker, J., J. A. Konstan, and J. Riedl. 2002. "An Empirical Analysis of Design Choices in Neighbourhood-based Collaborative Filtering Algorithms." *Information Retrieval* 5 (4): 287–310.

Jiang, J. J., and D. W. Conrath. 1997. "Semantic Similarity Based on Corpus Statistics and Lexical Taxonomy." Proceedings of the International Conference on Research in Computational Linguistics, Taiwan, 19–33, August 22–24.

Li, Y. H., Z. A. Bandar, and D. McLean. 2003. "An Approach for Measuring Semantic Similarity Between Words Using Multiple Information Sources." *IEEE Transactions on Knowledge and Data Engineering* 15 (4): 871–882.

Li, M. S., B. Yu, O. Rana, and Z. D. Wang. 2008. "Grid Service Discovery with Rough Sets." *IEEE Transactions on Knowledge and Data Engineering* 20 (6): 851–862.

Mathew, G. E., J. Shields, and V. Verma. 2004. "QoS Based Pricing for Web Services." Proceedings of the 5th International Conference on Web Information Systems Engineering (WISE 2004), Lecture Notes in Computer Science, 3307, 264–275, 2004.

McGuinness, D. L., and F. V. Harmelen. 2007. "OWL Web Ontology Language Overview, 2004." Accessed March. http://www.w3.org/TR/2004/REC-owl-features-20040210/.

Omrana, H., F.-Z. Belouadha, and Q. Roudies. 2012. "Template-based Matching Algorithm for Dynamic Web Services Discovery." *International Journal of Information and Communication Technology* 4 (2–4): 198–209.

Rodgers, J. L., and W. A. Nicewander. 1988. "Thirteen Ways to Look at the Correlation Coefficient." *The American Statistician* 42 (1): 59–66.

Shambour, Q., and J. Lu. 2011. "A hybrid Trust-enhanced Collaborative Filtering Recommendation Approach for Personalized Government-to-Business e-services." *International Journal of Intelligent Systems* 26 (9): 814–843.

Shambour, Q., and J. Lu. 2012. "A Trust-semantic Fusion-based Recommendation Approach for E-business Applications." *Decision Support Systems* 54: 768–780.

Wang, P., K. M. Chao, and C.-C. Lo. 2010. "On Optimal Decision for QoS-aware Composite Service Selection." *Expert Systems with Applications* 37 (1): 440–449.

Wang, H. C., C. S. Lee, and T. H. Ho. 2007. "Combining Subjective and Objective QoS Factors for Personalized Web Service Selection." *Expert Systems with Applications* 32 (2): 571–584.

Wu, J., L. Chen, Y. P. Feng, Z. B. Zheng, M. C. Zhou, and Z. H. Wu. 2013. "Predicting Quality of Service for Selection by Neighbourhood-based Collaborative Filtering." *IEEE Transactions on Systems, Man, and Cybernetics Part A: Systems and Humans* 43 (2): 428–439. doi:10.1109/TSMCA.2012.2210409.

Xu, Y. H., X. T. Guo, J. X. Hao, J. Ma, R. Y. K. Lau, and W. Xu. 2012. "Combining Social Network and Semantic Concept Analysis for Personalized Academic Researcher Recommendation." *Decision Support Systems* 54: 564–573.

Zhang, W. Y., and J. W. Yin. 2009. "Weaving a Semantic Grid for Multidisciplinary Collaborative Design." *International Journal of Production Research* 47 (11): 2511–2527.

Zhang, W. Y., J. W. Yin, M. Cai, J. Wu, and L. F. Lin. 2012. "Manufacturing Deep Web Service Management: Exploring Semantic Web technologies." *IEEE Industrial Electronics Magazine* 6 (2): 38–51.

Zhang, W. Y., S. Zhang, M. Cai, and J. Wu. forthcoming. "An Agent-based Peer-to-peer Architecture for Semantic Discovery of Manufacturing Services Across Virtual Enterprises." *Enterprise Information Systems*. doi:10.1080/17517575.2012.747002.

Zhang, S., W. Y. Zhang, J. Liu, and J. Wu. 2013. "A Time-aware Bayesian Approach for Optimal Manufacturing Service Recommendation in Distributed Manufacturing Environments." *Journal of Manufacturing Systems* 32 (1): 189–196.

Zheng, Z. B., H. Ma, M. R. Lyu, and I. King. 2011. "QoS-aware Web Service Recommendation by Collaborative Filtering." *IEEE Transactions on Service Computing* 4 (2): 140–152.

Zhong, J., A. Li, P. Deng, and L. Su. 2012. "Web Service Recommendation Based on Latent Features." *Journal of Information and Computational Science* 9 (1): 121–134.