Tag-based process recommendation for social business process modelling

Yanming Ye¹, ²*, Yueshen Xu², Zhilin Feng³

¹College of Information Technology, Hangzhou Dianzi University, 310018 Hangzhou, China
²College of Computer Science and Technology, Zhejiang University, 310027 Hangzhou, China
³College of Zhijiang, Zhejiang University of Technology, 310024 Hangzhou, China

Received 12 May 2014, www.tsi.lv

Abstract

Social BPM (Business Process Management) has become a new research hotspot in business process management field because of its capability of handling the flexibility and dynamics of process in social circumstance by means of integration of social software and BPM. The key technique is process modelling, and note worthy the process modelling is more complex in social BPM than in traditional BPM. This paper presents the definition of social business process model to identify the difference with the traditional business process model and gives a prototype of social BPM system. The modeller in the prototype system may be a common user without professional knowledge, so a tag-based process recommendation method is proposed to facilitate modelling. The experiment result shows that the method is valid and effective in computer-aided intelligent process modelling.

Keywords: process recommendation, process modelling, social business process model

1 Introduction

In traditional Business Process Management System (BPMS), process modelling is mainly conducted in a certain sequence via some modelling languages like BPMN (Business Process Modelling Notation). Further, those intermediate results are arranged and combined together by a process engine using a flow-oriented paradigm [1-3]. With popularity of Web 2.0 technologies, more and more business applications are deployed in form of Web applications, which lead to the socialization of business processes to better reach users and reduce the total cost. As a result, these business processes become more flexible and collaborative. Meanwhile, their activities become long running loose-coupled [4-7], which brings a challenge to traditional BPM systems that can only support the well-structured processes. Therefore, recently, people are making much endeavour to modify the traditional BPM to be adapted in new circumstances. Especially, social BPM techniques are getting more and more attentions for its combination of BPM and social software, while the latter can support social interaction and production, and raise the interaction level and scope, which are facilitated by computer networks [8-10].

In social BPM, social applications are integrated into business process management. Wherein, the participation of external stakeholders is strictly controlled in process design and implementation. To make social business more flexible and dynamic, process modelling may become more important and complex, so intelligent modelling methods are required by most of social BPM systems.

This paper focuses on the computer-aided process modelling of social business process. For clarity, JTangSBPMS that is the basic social BPM system of this paper is introduced firstly in Section 2. Then, in Section 3, we give some definitions with related basic instructions. Next, the process recommendation method is discussed in detail in Section 4 and experiments are demonstrated in Section 5. Finally, we show the conclusion and future work.

2 JTangSBPMS

Since social applications are prevailing in the whole world, the combination of microblog with existing BPMS is become necessary. Therefore, we build JTangSBPMS, a new social BPMS, whose prototype is shown in Figure 1.

As the same with traditional BPM, the lifecycle of social BPM can also be decomposed into the four following phases: design, configuration, enactment and diagnosis. In JTangSBPMS, process modellers and task performers are all microblog users. Besides, to complete the task assignment, a message containing the short URL of web application is sent to the target user, since the application owns the task widget. Then, the URL should be clicked to open the Web application, which will be deployed automatically. Users can send messages in the microblog about their feelings and experiences on the task or the business. The user collaboration can be

* Corresponding author e-mail: yeym@hdu.edu.cn
realized by the information exchanging through message sending.

In design phase, with the increasing of flexibility and collaboration of business process, process modelling also becomes more important. Process modelling is asic of the whole lifecycle. For the complexity, experts or professionals usually conduct the process modelling in traditional BPM systems. However, in social BPM systems or collaborative environment, the users or workers are peer to peer and the possible applied business areas are very wide, so the process modelling must be oriented to common users. Because the correctness and completeness of a process model is important for business process, it is necessary to provide intelligent modelling method to eliminate much of the hardship. This paper presents a tag-based process recommendation method to realize the intelligent modelling in JTangSBPMS.

The rest phases of JTangSBPMS are all based on the social business process model that is modelled in the design phase. For the configuration phase, the key steps are to configure Web applications carrying the real tasks according to activities of the process model. When an activity is on execution, the short URL of the Web application that carries the task of the activity will be sent to the performers in form of microblog message. In the enactment phase, any dynamic change on the business process will launch the process verification and recommendation for the robustness of the process. In each time when a process is finished, a social business process execution model will be produced, which can record the real running data. As for in the diagnosis phase, there are various social network analysis tools or methods that can be used to analyse the user’s behaviour and process structure as well as other social features of the process. Finally, a social business process can be improved based on these analyses.

3 Preliminaries

From the above discussion, the traditional BPM is flow-oriented, and business process tasks are scheduled in a specific sequence. Such an activity has a certain position in a process model, and the process is well-structured so that it can be executed in the pre-defined orders of activities in enactment phase. Meanwhile, the social BPM must deal with flexible and collaborative features of processes. For better comprehension, we will begin with the definitions of traditional business process model and social business process model.

**Definition 1. Traditional Business Process Model:** Let \((P, T, F)\) be a WF-net (as defined in [11]), and a traditional process model is a 5-tuple \((P, T, F, U, \Omega)\) where,

(i) \(U\) is a performers set,

(ii) \(\Omega: U \rightarrow P\) is the assignment function of performer (U) to a place (P) that includes certain task.

(iii) \(P, T,\) and \(F\) represent places, transformation, and flow structures respectively.

**Definition 2. Social Business Process Model:** A social business process model is a 5-tuple \((C, U, \xi, R, \Psi)\) where,
(i) C is an activity set in which each activity contains a specific task.
(ii) E is a user set as the subset of the users in Definition 3.
(iii) \( \xi : C \rightarrow \{ \text{subset}(U) \} \) is a function which maps activity to possible performers who are subset of the whole users.
(iv) R is a finite set containing five different kinds of relationships and \( R = \{ \text{before, after, and, or, undefined} \} \).
(v) \( \Psi : C \times C \rightarrow R \) is the relationship function between users.

From the above definitions, we can see that most of traditional process models are flow-oriented and the activity performer is usually appointed a forehand. But in the social process model, the order of activities may not be strict and we can determine the performers until running time.

Next, a formal definition of social network will be given as follows.

Definition 3 (Social Network): A social network SN is a graph denoted by a 5-tuple \( SN = (N, T, M, R, K) \) where,

(i) \( N \) is the finite set of nodes and each node stands for a user.
(ii) \( T \) is a set of social tags.
(iii) \( M : N \times T \rightarrow \{ \text{true, false} \} \) is the mapping function between users and tags.
(iv) \( R = \{ R_1, \ldots, R_n \} \) is a finite set of relationships.
(v) \( K : N \times N \rightarrow R \) is the relationship function.

In a social network, a user’s preference for a resource or a task can be inferred by tags, which can reflect the user interest and capability. Usually, the increasing of the number of tags means that a user is greatly interested in a certain resource or a user has many capabilities on a certain task.

Although the social business process model may be ill-defined, for example, the order of activities and the performers of activities are often unclear, any model execution will be conducted certainly on the unclear features in running time. Therefore, we give the definition of the process execution model as follows.

Definition 4. Social Business Process Execution Model: Let Y be a set of activity types and B be a finite set of tags for activities. A social business process execution model is a connected graph denoted by a 5-tuple \( P = (S, E, D, \alpha, \beta) \) where,

(i) S is a finite set containing activities.
(ii) E is the finite set of users.
(iii) \( D \subset S \times S \) is the finite set of edges.
(iv) \( \alpha : E \rightarrow S \) is the users assignment function for activities.
(v) \( \beta : S \rightarrow B \) is the activity tagging function.

In our system, the tags are divided into two types that are process tag and activity tag. There are three main components for any tagging system: a set of users, a set of tags created by the users and resources (URL, entities etc.). A social tagging system can be represented by a triple of users, tags and resources: \( (U, T, R) \), and the conceptual model of a tagging system is shown in Figure 2. [12-14]. Resources are restricted to be processes and activities depending on the specific purpose of our system in this paper. A tag is actually a word or a phrase that expresses certain meaning. However, people could use a wide variety of words to describe the same objects. Especially in the case of social tagging systems, users may apply “different terms as tags to describe the same resource” by using synonyms, homonyms and polysemy, leading to multiple and diverse descriptions for the same resource. The increasing number of vocabularies may imply that the connections between tags and resources will become less and less distinct, making information retrieval more difficult. To resolve the semantic problem, our system adopts an eclectic solution that provides a recommendation tag for each process model or activity based on the tags dictionary. Users can accept the recommendation default tag without any change or change it when they edit the related profiles or messages. If the users change a tag that is recommended, the tag will be added into the tags dictionary in proper catalogue unless it exists in the dictionary. The tags dictionary can
be trimmed from the WordNet Domains by means of some improvable methods [15–16]. While for simplicity, in our system the tags dictionary is in form of (tag, process-name/activity-name), of which the process tags are clustered by the probabilistic latent semantic analysis [17] on the set of the processes in the process repository. The activity tags are some fungible domain labels that are obtained from the WordNet Domains.

To model a complex social business process, it can be simplified as to confirm the next activity and its performer as clearly as possible, and then to specify the relationship between the activity and the existing activities of the modelling process fragment step by step. In general, two process models that belong to the same business area and resolve the similar business problem have the similar structure. A user may be more professional or skillful to perform a task if he or she has performed the similar task more times or more than other tasks. In a collaborative environment, a user mostly tends to work with the other certain user or users that he or she is familiar with. Therefore, we can recommend the next activities and their performers when a user is modelling a business process based on the historical process models and user behaviours. As what has been analysed, our recommendation method in fact is to find out process models whose structures are similar with each other and users whose behaviours are similar with each. Then, we will recommend the next activity that exists in most similar process models after the sequential activities that are contained in modelling process fragment, with the performers as users that have similar behaviour.

Definition 5 (Process Code): The process code of a process model (or process execution model) can be represented as the breadth first traversal order of tags on the process graph, which is a linear order, where the tags of the same hierarchies are by lexicographic order.

For example, the process code of the process execution model p, as show in Figure 3, is “ABCDEF” where the “A”, “B”, “C”, “D”, “E” and “F” are tags of activities of the process, not “ABDCEF” or others.

![Diagram](A-B-C-D-E-F)

**FIGURE 3** A sample of a social business process execution model - p.

The “A”, “B”, “C”, “D”, “E” and “F” can be seen as the tags of the activities that the users have labelled.

It is necessary to notice that the process code may be an array of chars or a string that can convert one to the other easily.

Definition 6 (Structure Similarity): the structure similarity between process execution model ep and ep2 with the process codes as cep1 and cep2 respectively can be calculated by the following formula,

\[
\text{Sim\_structure}(ep_1, ep_2) = \frac{|cep_1 \cap cep_2| + \sum_{t \in \text{synonym}(ti, tj)} \text{synonym}(ti, tj)}{\max(|cep_1|, |cep_2|)} \text{, where } |cep_i| = \text{length of } cep_i \text{ and } |cep_2| \text{ is the length of } cep_1 \text{ and } cep_2 \text{ respectively, and the synonym}(x, y) \text{ is the synonym function that returns 1 if the given tags } x \text{ and } y \text{ are synonyms and 0 otherwise.}
\]

Performers are important elements in a social business process, whereas many modellers have some doubts on how to assign more appropriate users to the tasks. The performer recommendation can facilitate the procedure, and the tag weight must be calculated in first.

**Definition 7 (User Behaviour Similarity):** The user behaviour similarity of a user on a tag can be calculated as the following formula,

\[
\text{Sim\_behavior}(u, t) = \frac{\text{tf}(u, t) \times \sum_{t \in \text{tf}(u, t)} \text{tf}(u, t)}{n} \text{, where tf}(u, t) \text{ is the tag frequency that represents the times that a tag } t \text{ has been tagged by a user } u \text{, and } n \text{ is the total number of tags that user } u \text{ has bookmarked.}
\]

### 4 Implementation

Tag-based process recommendation method turns to help process modellers fit processes to achieve a modelling intention with regard to the other modellers’ modelling behaviour and achievement.

The most important step of the method is to implement the algorithms of process structure similarity and user behaviour similarity. Next, we give the pseudo code of the structure similarity algorithms (shown in Figure 4).

**Input:** Process code: cep1, cep2  
**Output:** Process structure similarity: structuralSim  
**Declaration:**  
String[] intersect(String[] arr1, String[] arr2)  
String[] difference(String[] arr1, String[] arr2)  
int synonym(String arr1, String arr2)

```
1:  Initialize: structuralSim=0, int_syn=0
2:  int_lenofcep1 ← get the length of cep1
3:  int_lenofcep2 ← get the length of cep2
4:  arr_intercep ← intersect(cep1, cep2)
5:  int_lenofcep ← get the length of arr_intercep
6:  arr_diffcep1 ← difference(cep1, cep2)
7:  int_diffcep2 ← difference(cep2, cep2)
8:  for each l in arr_diffcep1 do
9:     int_tmpsyn=0
10:    for each m in arr_diffcep2 do
11:       int_tmpsyn ← synonym(l, m)
12:      if int_tmpsyn = 1 then break
13:     end if
14:    end for
15:   if int_tmpsyn= 1 then int_syn= int_syn+1
16: end if
17: structuralSim ← (int_lenofcep+int_syn)/max(int_lenofcep, p1, int_lenofcep2)
18: return structuralSim
```

**FIGURE 4** Pseudo code of process structure similarity algorithm
In the algorithm that is shown in Figure 4, intersect() and difference() are two functions that calculate

Intersection and difference set of two arrays of String type and synonym() is a function that determines whether two strings are synonyms.

For calculating the user behaviour similarity, we must use the social tagging system model. Further, as it has been discussed in the above section, a social tagging system can be represented by a triple of users, tags and resources: (U, T, R), while the resources refer in particular to the processes and activities in this paper. There is a database table named “tagging” with the columns of “user_id”, “tag_id”, “resource_id”, and “message_id”. A record “30015321, 00000271, 11120001, 30015321_00001032” represents a user “Jim” who is “30015321” has bookmarked a tag “consult” whose id is “00000271” on a activity whose id is “11120001” and the tag is published in a message whose id is “30015321_00001032”. The table “tagging” can create relationships with the tables of “tags”, “users”, “process” and “message” by the id of “tag_id”, “user_id”, “resource_id” and “message_id”. The performer recommendation is used more frequently than the activity recommendation, so the user behaviour similarity algorithm is implemented by the stored procedure and the pseudo code of the user behaviour similarity algorithm is shown in Figure 5.

```
DROP PROCEDURE IF EXISTS sim_behavior;
CREATE PROCEDURE sim_behavior(
    IN p_user_id VARCHAR(8),
    IN p_tag_id VARCHAR(8),
    OUT p_similarity INT)
BEGIN
    DECLARE m_tags INT DEFAULT 0;
    DECLARE m_tag INT DEFAULT 0;
    SET @COUNT_TAGS = CONCAT(SELECT COUNT(*) INTO @TAGS_TOTAL FROM tagging where user_id=’, p_user_id);
    SET @COUNT_TAG = CONCAT('SELECT COUNT(*)
                        INTO @TAG_TOTAL FROM tagging where user_id=’, p_user_id);
    PREPARE count_stmt FROM @COUNT.TAG;
    EXECUTE count_stmt;
    DEALLOCATE PREPARE count_stmt;
    SET m_tags = @TAGS_TOTAL;
    PREPARE count_stmt FROM @COUNT_TAGS;
    EXECUTE count_stmt;
    DEALLOCATE PREPARE count_stmt;
    SET m_tag = @TAG_TOTAL;
    SET p_similarity=m_tag / m_tags;
END;
```

5 Empirical Evaluation

In this section, we explain the experiment methodology that we followed to evaluate the described process recommendation method, and present the obtained results of that evaluation.

Our laboratory built a process model repository in which the process models are usually used in traditional BPM and collected from administrative examination and approval processes from administrative department of a local government (China). Firstly, we select 30 most frequently used process models as the test dataset and modify each process model by removing one or more activities and all the performers. Because the process models come from similar government administrative departments, so they can be placed in a category. After 30 minutes study of the modified dataset, 20 student volunteers can play the role of common knowledge workers that are not professional but experienced. Then the volunteers are requested to model the same business process with the similar topic on our system. The web application carrying the task of an activity is simplified by providing “accept”, “ignore”, “finish” and “recommend to others” buttons only. In our system, a social process model may be not well structured and some activities may be not strictly sequential (namely, each of the activities can be executed first) and most of performers are not clearly appointed to activities. In the experiment, the process models are put on running in turn after modelling and an activity is launched if the short URL of the web application carrying the task of the activity is sent to all users (an extreme case is that each user has the possibility to be the performer). The activities that have no strict order with each other (including implicit order) may be launched at the same time. The volunteers then are requested to select a role and accept any web application that his role is interested in or ignores it. Once a web application is accepted by a user, what the others do on it will be discarded unless the user turns to push “recommend to others” button. When the user finishes the task by pushing the “finish” button, the next activity may be launched. After all the process models are completely executed, there will be 20 social business process execution models according to the actual execution. Then five different volunteers (numbered 1, 2, 3, 4, 5) are selected to model the same business process using the commendation method after a brief introduction about the process. The modelling has a restriction on the finish time of 5 minute and the modellers can accept the recommendation or just follow their own judgments. The result is shown in Table 1.

<table>
<thead>
<tr>
<th>TABLE 1 Font styles for a reference to a journal article</th>
</tr>
</thead>
<tbody>
<tr>
<td>Volunteer Number</td>
</tr>
<tr>
<td>------------------</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>5</td>
</tr>
</tbody>
</table>

As it can be seen from Table 1, the models that the volunteer 2 and 4 modelled have 8 activities that all exist in the sample dataset. This may be caused by the
recommendation without any modification and the similar values of total time further verify the conclusion. The other volunteers spend more time on the modelling and the activities of their models are much more. The volunteer 1 spends more time than volunteer 5 while the amount of activities in their result models raise from 10 to 11. This may be because the latter makes a better use of the recommendation result than the former. It is necessary to notice that the modelling time may be related with the user’s experience and expertise. The average activities that are similar in the sample dataset exceeds 8 and this is a stronger evidence that the recommendation method is effective.

6 Conclusion and Future work

This paper presents a tag based process recommendation method for social business process modelling in the context of the social BPM. Social software is spreading quickly recently with various different application types and is affecting the modern people much more in life and work. More and more corporations have turned to develop their business or do some management by means of the tools of social software. Meanwhile, it is a challenge for traditional BPM systems that how to manage the business processes. Thus, some researchers attempt to combine the BPM and social software to realize so called social BPM. The JTangSBPMS that this paper discussed is a new try that integrates the microblog into traditional BPM system. The tag based process recommendation method is studied for the JTangSBPMS on intelligent process modelling. The experiment results show that the method is valid and effective in the system.

Still, much work has to be carried out in the future. For example, the paper does not take into account the time factor, while a new process model may be more useful for the modeller than a process model that is modelled years ago. Furthermore, how to use the social network analysis tools to improve the process modelling still leaves much to be developed. We hope that other researchers can be inspired by this paper to pay more attention to this field, and further can propose many more comprehensive methods.

Acknowledgments

This work is supported by the National Natural Science Foundation of China under Grant (No.61272129), National High-Tech Research Program of China (No. 2013AA01A213), New Century Excellent Talents Program by Ministry of Education of China (No.NCET-12-0491), Zhejiang Provincial Natural Science Foundation of China (LR13F020002) and Zhejiang Provincial Natural Science Foundation of China (LY13F020027).

References

[15] Locoro A 2011 Tagging ontologies with fuzzy wordnet domains In Fuzzy Logic and Applications Springer Berlin Heidelberg 107-14
### Authors

**Yanming Ye**, born on October 5, 1979, Longnan City, Gansu Province, P.R.China  
Current position, grades: Ph.D. candidate, Lecturer  
University studies: Artificial Intelligence  
Scientific interest: BPM, recommender system, social computing  
Publications: He has published more than 10 papers. Including:  
Experience: He is a Lecturer in Hangzhou Dianzi University and a Ph.D candidate at the College of Computer Science and Technology, Zhejiang University, China. He received his master degree in Zhejiang University, China in 2005.

**Yueshen Xu**, born on December 11, 1989, Weifang City, Shandong Province, P.R.China  
Current position, grades: Ph.D. candidate, grade 3  
University studies: Artificial Intelligence  
Scientific interest: service computing, recommender system, applied machine learning  
Publications: He has published more than 7 papers. Including:  
Experience: He is a full-time Ph.D candidate at the College of Computer Science and Technology, Zhejiang University, China. He received his bachelor degree in XiDian University, China in 2011.

**Zhilin Feng**, born on March 18, 1977, Hangzhou City, Zhejiang Province, P.R.China  
Current position, grades: Associate Professor  
University studies: Artificial Intelligence, Machine Learning  
Scientific interest: Data mining, Formal modelling  
Publications: He has published more than 15 papers. Including:  
Experience: He received his PhD degree from school of computer science and technology in Zhejiang University in 2005.