SEMANTIC WEB SERVICE COMPOSITION FOR ERP BUSINESS PROCESS

*Anang Kunaefi, Riyanarto Sarno, Dwi Sunaryono, Imam Mukhlash

Institut Teknologi Sepuluh Nopember Surabaya
E-Mail: an_kunaefi@yahoo.co.id

Abstrak

Saat ini, ERP (Enterprise Resource Planning) bergerak menuju layanan SaaS (Software as a Service) dan Multi-Tenancy, di mana aplikasi ERP melayani beberapa penyewa dengan proses bisnis yang berbeda dalam lingkungan berbasis web service. Pada kasus Provider ERP, sangat penting untuk mencapai fleksibilitas proses bisnis penyewa sebagaimana didefinisikan dalam tingkat kematangan SaaS level 4, yaitu Configurable dan Scalable. Dengan cara ini, Penyedia dapat melayani proses bisnis penyewa secara dinamis. Penelitian ini menggunakan pendekatan komposisi semantik web service untuk menyelesaikan masalah fleksibilitas dalam proses bisnis. Ontologi digunakan sebagai representasi pengetahuan semantik pada domain pengetahuan ERP untuk proses pencarian dan komposisi web service. Selanjutnya, algoritma kemiripan berbasis fitur (Feature-based Similarity) dan kemiripan berbasis struktur (Structure-based Similarity) digunakan untuk melakukan pencarian kemiripan antara permintaan proses bisnis dari penyewa dan proses bisnis penyedia layanan ERP di Registry. Hasil penelitian menunjukkan bahwa metode yang diusulkan mampu memenuhi permintaan proses bisnis penyewa, baik workflow sederhana maupun workflow yang lebih kompleks dengan hasil yang baik.


Abstract

Nowadays, ERP (Enterprise Resource Planning) moves toward SaaS (Software as a Service) and Multi-Tenancy, where an ERP application serves multiple tenants with different business processes in a web-service based environment. In the case of ERP provider, it is very important to achieve business process flexibility among tenants as defined in SaaS Maturity Level 4, that is Configurable and Scalable. In this way, Provider can serve tenant’s business processes request dynamically. This research uses semantic Web Service Composition approach to address business process flexibility problem. Ontology is used as a semantic representation of ERP domain knowledge for web service discovery and composition. Afterwards, the combination of Feature-based Similarity and Structural-based Similarity algorithms are used to do the discovery and matchmaking process between tenant’s business process request and business process available in the ERP provider’s registry. The result show that the proposed method in this paper is able to fulfil tenant’s business process request both for simple workflow and complex workflow with a good result.

Keywords: Semantic Web Service, Semantic Composition, ERP Business Process, Feature-based Similarity, Structure-based Similarity.
INTRODUCTION

Nowadays, ERP moves toward SaaS (Software as a Service) and Multi-Tenancy, where an ERP application serves multiple tenants with different business processes. In order to achieve flexibility, SaaS ERP Provider needs a method to compose or configure business process at runtime. This requirement was defined by Qihongas SaaS Maturity Level 4 [1].

In this study, we use semantic web service composition to address business process flexibility problem. W3C defined Web Service Composition as a sequence of atomic services where one web service invoke another web service in order to make a new functional composite web service for its user.

In the previous studies about web service composition, some researchers used different methods and approaches to do the composition. Here are to name a few, Sirin used artificial intelligence approach to make service composition [2]. McIraith used logical programming to automatically compose web services [3], and Talantikite used ontology annotation to compose web service to meet user’s query [4]. Although many researchers have been addressed service composition problem with different techniques, composing a service from existing service is still very difficult [5]. This indicates that the world of service composition still in search for the best solution up until now.

Many studies used ontology to represent the metadata of the service, as ontology has been standardized by W3C [6,7,8]. However, for ontology approach alone, there are different methods used by researchers to do service composition. These methods are an extension of ontology specification designed for web service, for example, DAML as in [9,10], OWL-S as in [11,12,13], and WSMO as in [14,15].

Most researchers checked only the inputs and outputs of web services to make service composition [4,16,17,18]. By using this approach, the result of the composition can not be controlled by the user. Composition result can be very lengthy and inefficient, involving web services which should not be included.

In our study, we check the inputs and outputs (feature-based similarity), as well as the workflow structure (structure-based similarity) of the tenant’s request. We believe, combining these two similarity checking will lead to a better composition result.

In a usual scenario, user or tenant of ERP Provider will perform a query to discover whether their service request is available or not in ERP Provider’s registry before transaction take place.

In our scenario, however, tenant will perform a business process request using ontology annotation. Then, ERP Provider will perform similarity computation to search for similar composite service in the Registry. If, for example, the tenant’s request cannot be satisfied with existing composite service, ERP Provider will try to compose a new composite service based on a composite service that has the closest match with tenant’s request.

In order to accomplish this task, we explain our method in the next section.

DOMAIN ONTOLOGY

In heterogeneous environment, it is very important to use domain ontology as a reference for ontology processing. Domain ontology will help to solve heterogeneity problem in the case of concept misperception [19]. For example, tenant might use term “Order” to express purchase order process. On the other hand, in Provider’s side, terms “Order” might have at least two meanings, either “Purchase Order” or “Production Order”. Purchase Order and Production Order are different in the perspective of Provider’s Ontology, because they have their own characteristics. This condition is described in Figure. 1.

Figure 1. Example of ERP Domain Ontology.
Kunaefi et al., Semantic Web Service Composition For...

Figure 2. ProcessModel Representation in OWL-S.

Figure 3. Graph Structure Representation

Using these characteristics, the system is able to determine which class in the ontology that has similar meaning with user or tenant’s request, whether it is “Purchase Order” or “Production Order”.

We adopted layered ontology approach explain in [20], by separating ontology for ERP domain knowledge (Domain Ontology) and ontology for ERP application (Application Ontology). The benefit of this separation is to achieve comprehensive understanding about the meaning, purpose and usage of each concept in the ontology [20].

The ontology was made by following practical guide in [21] using OWL 1.2 specification with Protege 4.2.

SEMANTIC SIMILARITY

In this research, we employ two kinds of web service semantic similarity algorithms, which are, Feature-based Similarity and Structure-based Similarity.

Feature-based Similarity

In the case of web service, Feature is defined as Input, Output, Precondition and Effect (IOPE). Although the technique might different, many researchers used IOPE matchmaking as a basis to perform service discovery and service composition [16,17].

Therefore, in this study, we adopt Feature-based Similarity to semantically discover whether two services can be composed or not. In Feature-based Similarity, the more features in common between the two concepts, the greater the similarity value. To calculate the Feature-based Similarity, we used the formula from Ganjisaffar [22] as shown in Equation (1),

\[ \sigma(c, c') = \frac{2|F(c) \cap F(c')|}{|F(c)| + |F(c')|} \]

Where,

- \( F(c) \) : Features of Concept \( c \)
- \( F(c') \) : Features of Concept \( c' \)

Structure-Based Similarity

In the case of ERP, each tenant can have many different business processes. Each business process is implemented using a composite service contain a process workflow. In OWL-S, this process workflow called as Process Model [23]. In OWL-S, the Process Model is represented in XML as described in Figure 2.

In Figure 2, Process Model describes sequence of two web services i.e., createPurchaseRequest and createPurchaseOrder. This sequence can also be represented as a graph structure as described in Figure 3.

The purpose of using Structure-based similarity is to calculate the similarity of structure and sequence of the workflow process [24,25]. Graph-edit Distance Algorithm is used to compute the distance between the two graphs. Graph-edit Distance is defined as the minimum number of graph-edit operations necessary for the two graphs to be exactly the same. The definition of graph-edit operations include: node deletion and insertion; node substitution; edge deletion and insertion.

The detail process about how to compute Graph-edit Distance Algorithm has been discussed in [24].

WEB SERVICE COMPOSITION

In a service composition scenario using Input, Output, Precondition and Effect (IOPE), two services \( S_1 \) and \( S_2 \) can be composed, if the output of \( S_1 \) is used as an input of \( S_2 \) [26]. Furthermore, two services are compatible and composable, if the Effect value of \( S_1 \) is equal or similar to Precondition value of \( S_2 \).

In addition, the structure of the composition request is matched with the structure of composite service available in the Registry. The objective is, a new composition can be made
from existing composite service with minor addition.

![System Architecture](image)

**Figure 4. System Architecture.**

In our method as described in Figure 4, tenant’s business process request is handled by Service Composer. Service Composer utilizes Feature-based Similarity that is combined with Structure-based Similarity for matchmaking and composing workflow. For every similarity checking we use Domain Ontology to check the relatedness of every term of the service feature involved. Based on the result of matchmaking process, Service Composer return several Composition Plan that might be found to the Tenant.

**Composing a Simple Workflow**

Say, tenant define a request R as follows.

\[
R = \{ \text{input}, \text{purchaseOrderID}, \text{http://erp2011/domain.owl#purchaseOrderID} , \\
\text{output}, \text{stockStatus}, \text{http://erp2011/domain.owl#stockStatus} , \\
\text{precondition}, \text{itemOrdered}, \text{http://erp2011/domain.owl#itemOrdered} , \\
\text{effect}, \text{purchaseInvoiceCreated}, \text{http://erp2011/domain.owl#purchaseInvoiceCreated} \}
\]

**Extracting Feature**

Common web service features consist of Input, Output, Precondition and Effect (IOPE). For every feature in the above definition, there are two values separated with comma. The first value represent the name of the feature, the second value represent the annotation of the feature referencing to some ontology.

For example, input feature has value purchaseOrderID with characteristics similar to what is described in http://erp2011/domain.owl#purchaseOrderID.

From the domain ontology specification domain.owl#purchaseOrderID, we get the following description about input purchaseOrderID.

\[
F(\text{purchaseOrderID}) = \{(\text{hasInputName}, \text{purchaseOrderID}), \\
(\text{hasInputType}, \text{String})\}
\]

Using the same steps as input feature, we get all features of request R as follows.

\[
F(R) = \{(\text{hasInputName}, \text{purchaseOrderID}), \\
(\text{hasInputType}, \text{String}), \\
(\text{hasOutputName}, \text{stockStatus}), \\
(\text{hasOutputType}, \text{String}), \\
(\text{hasPrecondition}, \text{itemOrdered}), \\
(\text{hasEffect}, \text{purchaseInvoiceCreated})\}
\]

After getting the features, the next important task is getting the structure of the request. Again, we will use ontology description that is annotated by tenant in the request. The system read and extract Process Model of composite service in ontology file http://erp2011/receiving.owl. Figure 5 described an example of receiving ontology as a part of Application Ontology.

In the ontology, there are four classes, and five object properties. Receiving class has three processes, namely, createReceiving, checkQuantity and createPurchaseInvoice. Receiving class has hasInitialProcess property on createReceiving class, indicating that createReceiving class is the first process that should be executed in the workflow. On the other hand, it also has hasLastProcess property pointing on createPurchaseInvoice, indicating that createPurchaseInvoice is the last process of the workflow. It also relates with checkQuantity using hasProcess property.

![Example of Receiving.owl](image)

**Figure 5. Example of Receiving.owl.**
The position of checkQuantity process in the workflow can be determined with two properties, namely hasPreviousProcess and hasNextProcess. hasPreviousProcess means that checkQuantity can be executed after createReceiving is executed. On the contrary, hasNextProcess means that after executing checkQuantity, the system must execute createPurchaseInvoice.

In the bigger workflow involving more processes, there might be more than one hasPreviousProcess property relating with several classes. Each class should define its hasPreviousProcess and hasNextProcess property clearly.

The structure graph as a result of extraction process of ontology in Figure 5 is shown in Figure 6.

![Graph Structure Representation of Composition Request](image)

**Figure 6. Graph Structure Representation of Composition Request.**

**Service Discovery**

The next important task after getting request definition is service matchmaking process or service discovery [8, 18, 27]. In the matchmaking process, the system will search The Registry for composite service which has a similar graph structure with tenant’s request using similarity computation.

Say, for example, there is one composite service in the Registry as described in Figure 7. Similarity computation between graph structure in Figure 7 and Figure 8 begins with the similarity checking of nodes and edges. Dijkman in [24] used syntactic similarity on node labels using the Edit-Distance Algorithm.

In this study, we use Feature-based Similarity as seen in Equation (1) to compute similarity between nodes. This is because each node in the graph represents atomic web service, and each of them has features.

According to the result in Table 1, there are two similar nodes (indicated by the similarity value above threshold; threshold = 0.7) as follows:

1. createReceiving() similar with provideReceiving()
2. checkQuantity() similar with checkQuantityDifference()

Thus, according to Graph-Edit Distance Algorithm, we got the following computation.

\[
N1 = \text{All nodes of Graph}1 = 3 \\
N2 = \text{All nodes of Graph}2 = 2 \\
E1 = \text{All edges of Graph}1 = 4 \\
E2 = \text{All edges of Graph}2 = 3 \\
s_n = \text{All nodes} - \text{similar nodes} = 5 - 4 = 1 \\
s_e = \text{All edges} - \text{similar edges} = 7 - 4 = 3 \\
s_n = \frac{(|n| - s_n)}{|n|} = \frac{1}{3} = 0.33 \\
s_e = \frac{|e| - s_e}{|e|} = \frac{1}{2} = 0.5 \\
s_n = \frac{2\Sigma_{e\in E} \min(1,|n|) - |n|}{|E|} = \frac{2(1\times0.75 + 4\times0.5) - 4}{6} = \frac{542}{10} = 0.105 \\
\text{simgraph} = 1 - \arg(0.2 + 0.42 + 0.105) = 0.765
\]

The result of Structure-based Similarity score is equal to 0.765 (above threshold). The result indicates the two graphs are similar. However, if we checked on the IOPE parameter, the output part of the found composite service still did not match with tenant’s output request. This condition is described in Figure 8(a).

In order to fulfill the request, there should be another service to close the gap. Therefore, to complement the shortage, the algorithm will look for another appropriate service using Feature-based Similarity where the target output is the output request from tenants, and target input
istheoutputof the found composite service. The final result of the service composition is shown in Figure 8(b).

The result shown in Figure 8(b) might be one possibility of composition plan. Each tenant’s request can return several composition plans as a result of similarity computation. Each composition plan considered as a solution, if the average similarity score is above the threshold (threshold = 0.7).

Table 1. Feature-based Similarity Between Nodes.

<table>
<thead>
<tr>
<th>No</th>
<th>Node/Features</th>
<th>Feature/Similarity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>createReceiving()</td>
<td>provideReceiving()</td>
</tr>
<tr>
<td></td>
<td>(hasInputType, String), (hasOutputType, String), (hasInputName, purchaseOrderID), (hasOutputName, receivingID), (hasPrecondition, purchaseOrdered), (hasEffect, itemReceived)</td>
<td>(hasInputType, String), (hasOutputType, String), (hasInputName, purchaseOrderID), (hasOutputName, receivingID), (hasPrecondition, purchaseOrdered), (hasEffect, receivingCreated)</td>
</tr>
<tr>
<td>2</td>
<td>checkQuantity()</td>
<td>checkQuantityDifference()</td>
</tr>
<tr>
<td></td>
<td>(hasInputType, String), (hasOutputType, String), (hasInputName, purchaseOrderID), (hasOutputName, quantityDifference), (hasPrecondition, itemReceived), (hasEffect, quantityChecked)</td>
<td>(hasInputType, String), (hasOutputType, String), (hasInputName, purchaseOrderID), (hasOutputName, quantityDifference), (hasPrecondition, receivingCreated), (hasEffect, qtyChecked)</td>
</tr>
<tr>
<td>3</td>
<td>createPurchaseInvoice()</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(hasInputType, String), (hasOutputType, String), (hasInputName, purchaseOrderID), (hasOutputName, receivingID), (hasPrecondition, quantityChecked), (hasEffect, purchaseInvoiceCreated)</td>
<td>(hasInputType, String), (hasOutputType, String), (hasInputName, purchaseOrderID), (hasOutputName, receivingID), (hasPrecondition, quantityChecked), (hasEffect, purchaseInvoiceCreated)</td>
</tr>
</tbody>
</table>

Composing Complex Workflow

Example in the previous section shows only simple workflow consist of simple sequence. What if tenant, for example, request a more complex workflow consists not only sequence, but also conditional ifThenElse and parallel splitJoin execution. Figure 9 shows an example of a complex Order-to-Cash workflow in ERP.

As can be seen in Figure 9, the workflow begins when a customer order some products. Sales staff invoke salesOrder service to create sales order transaction. The process continue by invoking checkStock service for each product that is being ordered. After invoking checkStock service, there will be two possibilities. The product is in stock or out of stock. Therefore, the process continue with parallel execution. For products that is in stock, the system will invoke shipping service, and at the same time, for products that is out of stock, the system will invoke makeToOrder service. The next process is to determine, whether the sales is having return or not. If it is true, the process continue to invoke salesReturn, salesReturnInvoice and salesInvoice consecutively, otherwise it invokes salesInvoice service. The process ends by invoking cashbank service.

Unfortunately, to find a similar composite service having similar workflow with the one shown in Fig. 9 are very difficult. The system will end up with no result because the similarity score is low. However, the search will be much easier if the target composition is simple enough.
Therefore, we developed our algorithm to address this problem as follows.

(i) Get the workflow
(ii) Break the workflow into sequence compositions
(iii) For each new sequence composition, do the following
(iv) If new sequence composition found a match composite service, use the composite service,
(v) Otherwise, make a new composition from atomic services in the Registry.
(vi) If all the sequences have been satisfied, compose the sequence together, otherwise back to step iii.

In our method, we break the structure into several sequence compositions to get the following benefits: (a) to make the searching and composing process much more easier because the target composition is simple; and (b) to gain flexibility in the composition process because new services can be added in the middle of the structure.

As in Figure 9, the workflow request can be break into 6 sequence composition as follows.

By breaking the workflow into sequences, the composition process will be much easier and flexible. For example, sequence Seq-4 in Table 2 as in Figure 10, might have a match with one composite service in the Registry. Thus, the composite service can be reusable.

Service composition for each sequence in Table 2, follow the steps and guide in the previous section for simple workflow. Each of the sequence will have their own IOPE definition to help the composition process.

For example, Seq-1, will have Input and Precondition definition from salesOrder service; and Output and Effect definition from checkStock service.

After all sequences have been satisfied, the next step will be combining all the sequences with the control construct as define in the workflow according to its structure.

<table>
<thead>
<tr>
<th>Seq Composition</th>
<th>Services Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seq-1</td>
<td>salesOrder – checkStock</td>
</tr>
<tr>
<td>Seq-2</td>
<td>makeToOrder</td>
</tr>
<tr>
<td>Seq-3</td>
<td>shipping</td>
</tr>
<tr>
<td>Seq-4</td>
<td>salesReturn – salesReturnInvoice – salesInvoice</td>
</tr>
<tr>
<td>Seq-5</td>
<td>salesInvoice</td>
</tr>
<tr>
<td>Seq-6</td>
<td>Cashbank</td>
</tr>
</tbody>
</table>

Figure 10. Sequence Seq-4.

RESULT AND DISCUSSION

The method proposed is tested in a service-based ERP application environment with 60 web services. The result indicated that the algorithm was able to compose web services that are similar to tenant’s request specifications.

We use ROC (Receiver Operating Characteristics) classifier to classify the result of the composition discovery.

ROC pays attention on four possible conditions in the query result, namely, correct hit (True Positive/TP), correct rejection (True Negative/TN), incorrect hit (False Positive/FP) and incorrect rejection (False Negative/FN). The value of each condition can be obtain by manually observing the query result. For example, TP is the number of relevant result, while FP is not relevant. On the contrary, TN is the number of correct unreturned result, while FN is incorrect unreturned result.

Figure 11 describes ROC Curve, which is the plot of True Positive Rate (TPR) against False Positive Rate (FPR) of the result. TPR is the proportion of the relevant result returned by the system compare to all relevant result in the datasets.
Figure 11. ROC Curve.

FPR is irrelevant result detected as irrelevant compare to all irrelevant service in the dataset. TPR and FPR can be determined by the equation as shown in Equation (2) and Equation (3).

\[ TPR = \frac{TP}{TP + FN} \]  
\[ FPR = \frac{FP}{FP + TN} \]

The ROC Curve demonstrates several things. The closer the curve follows the left-hand border (above diagonal line), the more accurate the test. Additionally, the closer the curve comes to the 45-degree diagonal (under diagonal line), the less accurate the test. Based on curve in Figure 11, it can be seen that the matchmaking result for composite service is slightly better than atomic service.

The accuracy of the algorithm can be determined by formula as shown in Equation (4).

\[ Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \]

By observing the result, we got the accuracy of this method is between 72% and 97%. The accuracy of 97% can be reached if all features of the request service were determined by tenant, and there were services with similar features in Provider’s Registry. On the contrary, the accuracy of 72% was reached because tenant’s determined only two out of four features, which are input and output.

CONCLUSION AND FUTURE WORKS

In service-based application, there is a need, where tenants are able to construct their own business process according to their needs dynamically and semantically. This will make a new way for business to gain win-win solution between service provider and tenants.

In this paper, we proposed a method to make semantic composition using feature similarity and structural similarity.

Based on the composition formed, it can be concluded that the method used has been successfully compose web service that is similar to tenant’s specification.

REFERENCES


