A QSQL-based Efficient Planning Algorithm for Fully-automated Service Composition in Dynamic Service Environments

Kaijun Ren¹,², Xiao Liu², Jinjun Chen², Nong Xiao¹, Junqiang Song¹, Weimin Zhang¹
¹College of Computer, National University of Defense Technology, Changsha, Hunan 410073, P.R. China
renkaijun@nudt.edu.cn
²Centre for Information Technology Research, Swinburne University of Technology, Melbourne 3122, Australia
jchen@swin.edu.au

Abstract

Web service composition is emerging as a promising technology for supporting large-scale, sophisticated business process integration in a variety of complex e-science or e-business domains. Particularly, semantics have been proposed as a key to automatically solving the discovery and composition problem. However, most of semantic composition approaches still remain at a stage of low efficiency because of the performance issues brought by the involved ontology reasoning and manual processing. To address this problem, in this paper, we present a QSQL-based service composition algorithm towards a fully-automated service composition. QSQL (Quick Service Query List) is an efficient service query index list which can achieve about the same semantic service discovery effects as other existing semantic composition methods, but with much less reasoning. With our proposed QSQL-based service composition algorithm, composition plans can be created to meet a user’s query in an automatic, efficient and semantic manner. In particular, with our algorithm, most existing composition plans in QSQL can be founded and ranked by exploiting a weighted Petri net representation; which will facilitate the execution verification. The final experiment is conducted to further demonstrate the feasibility of our proposed composition approach and its efficiency.

1. Introduction

Web service composition, by binding two or more existing services into a new one, is emerging as a promising technology for supporting large-scale, sophisticated business process integration in a variety of complex e-science or e-business domains. With the increase in importance of business process automation and highly dynamic nature of the internet, automated composition approaches have been the most promising methods to be applied in real world contexts. Currently, there exists many different research efforts on automating service composition; such as Petri net-based composition approaches [1, 2], AI planning-based approaches [3, 4], context and non-functional properties-based approaches [5-7]. But most importantly, semantic service compositions that take semantics of services into account to automatically solve the discovery and composition problem, have been a recent active research field [8-10]. However, despite the merits and the importance of semantic information contained by services, most existing algorithms for semantic service composition have taken direct ontology reasoning for generating process plans, which have caused performance issues. For example, [11] provided a hybrid semantic service discovery method based on direct ontology reasoning on top of OWL-S described services. The provided examples contain 582 services, 29 query requests and the average response time for each query is about 8 seconds when being simulated in the computer with a 2.4 GHz CPU, 1024MB memory. Therefore, if there is no single web service satisfying the requests, a process plan by composing multiple services has to be generated, thus the response time will be much longer. As such, the traditional semantic composition approaches involving much reasoning are unable to offer the capabilities of quick response for user’ query when he demands a quick response in highly dynamic internet service environments. Additionally, with the growing service resources, manual processes in some service composition approaches can be another
With this approach, an explicit goal definition needs to be used to determine the combination of actions [3, 4].

Problems as AI planning problems, where a planner is responsible for representing process constructs. Section 5 presents a formal and weighted Petri net description of our concrete composition algorithm and its complex analysis. Section 6 presents the experiment comparison. Our final experiment demonstrates the feasibility of our proposed composition method and the corresponding efficiency.

The remainder of this paper is structured as follows: Section 2 gives an analysis of related work. Section 3 gives a summary of QSQL principles. Section 4 provides a formal and weighted Petri net description for representing process constructs. Section 5 presents our concrete composition algorithm and its complex analysis. Section 6 presents the experiment comparison. The final section gives a conclusion and our future work.

2. Related Work and Problem Analysis

Over the last several years, different representative composition strategies have been proposed in the literatures. To the best of our knowledge, we categorize and analyze them below.

- Composition based on Petri nets:
  [1] Proposes a goal-driven service composition approach based on a Petri nets modeling technique. With this method, goals are recursively decomposed to be sub-goals until they become executable elementary goals. Because Petri net is a suitable tool for formal analysis and verification of correctness, this approach is of some significance. However, the author didn’t give the efficiency comparison of goal decomposition. In actuality, goal decomposition may be time-consuming tasks because of similar complex AI searching.

- Composition based on AI planning:
  Many research efforts regard the service composition problems as AI planning problems, where a planner is used to determine the combination of actions [3, 4]. With this approach, an explicit goal definition needs to be provided. However, such explicit goal is usually not available. Therefore, most of these approaches are restricted to sequential compositions. Additionally, performance issue can also present another problem when applying sophisticated AI algorithms for automated plan generation.

- Composition based on formal models: (Process algebra, Finite-State Machines)
  Some composition approaches are based on formal models such as process algebra [12], Finite-State Machine [13]. The above formal composition methods are very useful to check the existence of a composition and return a composition plan if one exists. However, they fail to automatically find most of such existing composition plans and rank them. In dynamic network environments, alternate execution plans are very important when exceptions occur.

- Composition based context and non-functional properties:
  Some composition approaches depend on context and non-functional properties of services. [5] proposes a dynamic and graph theory-based service composition approach aiming at capturing the useful characteristics of resources in a pervasive computing environment. However, in this method, the basic matching is still on the direct semantic reasoning level. Therefore, the performance problems for producing composition plans should also be overcome. Actually, most of these approaches did not pay much attention to the efficiency of automatically generating process flow.

- Composition based on semantic approaches
  Semantics have been proposed as a key to increasing automation in applying Web services and managing Web processes within and across enterprises, and the World Wide Web Consortium (W3C) has recently finished an important standard—Semantic Annotations for WSDL and XML Schema (SAWSDL) for Semantic Web services(SWSs)[14]. Currently, there are many existing semantic service composition approaches[8-10] and semantic composition projects such as METEOR-S[15], IRS[16] SHOP2[3], IBM semantic tools[17]. However, these methods still either stay in a semi-automated state, or in low efficiency of producing a composition plan due to a direct reasoning style.

By comparison, our proposed QSQL-based service composition approach can overcome the performance issue by avoiding large scale of ontology reasoning so that a given query can be achieved a quick response during the creation of composition plans. Particularly, most of existing plans can also be founded and ranked by exploiting a formal and weighted Petri net representation for facilitating the execution verification.
3. Summary of Quick Service Query List (QSQL)

In order to overcome the low discovery efficiency brought by the traditional semantic service discovery algorithm based on direct reasoning, we have proposed a pre-reasoning based service discovery method in [18]. In this method, ontology reasoning can be done in advance when services are published, and the reasoning results can be recorded in specially designed data structures which are based on graph storage theory to form quick service query index called QSQL. The basic data structures of QSQL mainly include two parts. One part is the domain of link which is mainly used to avoid repeated reasoning. The other part is the domain of data, which is primarily used to record service information in the corresponding INPUT/OUTPUT data vectors such as Exact_vector, Plugin_vector, Sib_vector, Gr apar_vector, Grachd_vector according to their corresponding semantic relationships. In order to clearly understand QSQL, let’s assume there are some web service models which need to be published to form QSQL in Table 1 (These examples will also be used to exemplify the later composition algorithm). Particularly, the parameters of service models have been simplified to directly match the corresponding ontology concepts. Some semantic relationships between ontology concepts with respect to the relevant service models in Table 1 are as follows.

- $USprice \supseteq USdollar$ datetime $\supseteq$ somedatetime
- $somedatetime \supseteq$ givendatetime cityid = citycode
- $areacode \supseteq$ cityid location $\equiv$ cityname

Table 1. Abstract web service models

<table>
<thead>
<tr>
<th>ID</th>
<th>Operation</th>
<th>Input1</th>
<th>Input2</th>
<th>Input3</th>
<th>Output1</th>
<th>Output2</th>
</tr>
</thead>
<tbody>
<tr>
<td>w1</td>
<td>UStoRMB</td>
<td>USprice</td>
<td></td>
<td></td>
<td>RMBPrice</td>
<td></td>
</tr>
<tr>
<td>w2</td>
<td>GetstockPrice</td>
<td>cityID</td>
<td>some</td>
<td>stockID</td>
<td>USdollar</td>
<td></td>
</tr>
<tr>
<td>w3</td>
<td>citytoCode</td>
<td>cityID</td>
<td>name</td>
<td></td>
<td>stockID</td>
<td></td>
</tr>
<tr>
<td>w4</td>
<td>getStocked</td>
<td>stock name</td>
<td></td>
<td></td>
<td>stockID</td>
<td></td>
</tr>
<tr>
<td>w5</td>
<td>getregionCode</td>
<td>location</td>
<td></td>
<td></td>
<td>areacode</td>
<td></td>
</tr>
</tbody>
</table>

According to the introduced service publication algorithm in [18], part publication records in QSQL for services in Table 1 are given in Table 2.

Table 2. Part records of QSQL for models in Table 1

<table>
<thead>
<tr>
<th>ID</th>
<th>ontology</th>
<th>domain</th>
<th>Iexact</th>
<th>Iplugin</th>
<th>Isab</th>
<th>Kgrapar</th>
<th>Kgrachd</th>
<th>Oexact</th>
<th>Oplugin</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>subprice</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td></td>
<td></td>
<td>…</td>
<td>…</td>
</tr>
<tr>
<td>2</td>
<td>USprice</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td></td>
<td></td>
<td>…</td>
<td>…</td>
</tr>
<tr>
<td>3</td>
<td>givendatetime</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td></td>
<td></td>
<td>…</td>
<td>…</td>
</tr>
<tr>
<td>4</td>
<td>USdollar</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td></td>
<td></td>
<td>…</td>
<td>…</td>
</tr>
<tr>
<td>5</td>
<td>somedatetime</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td></td>
<td></td>
<td>…</td>
<td>…</td>
</tr>
<tr>
<td>6</td>
<td>cityid</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td></td>
<td></td>
<td>…</td>
<td>…</td>
</tr>
<tr>
<td>7</td>
<td>cityname</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td></td>
<td></td>
<td>…</td>
<td>…</td>
</tr>
<tr>
<td>8</td>
<td>stockid</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td></td>
<td></td>
<td>…</td>
<td>…</td>
</tr>
<tr>
<td>9</td>
<td>citycode</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td></td>
<td></td>
<td>…</td>
<td>…</td>
</tr>
<tr>
<td>10</td>
<td>location</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td></td>
<td></td>
<td>…</td>
<td>…</td>
</tr>
<tr>
<td>11</td>
<td>stockname</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td></td>
<td></td>
<td>…</td>
<td>…</td>
</tr>
<tr>
<td>12</td>
<td>datetime</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td></td>
<td></td>
<td>…</td>
<td>…</td>
</tr>
<tr>
<td>13</td>
<td>area code</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td></td>
<td></td>
<td>…</td>
<td>…</td>
</tr>
</tbody>
</table>

For example, when publishing

\[ w5(\text{getregionCode}(\text{input}(\text{location})), \text{output}(\text{areacode})) \]

an output parameter of ws5, the publication algorithm first inserted ontology concept $\text{areacode}$ into QSQL, then added unique ID of ws5 to its Output Exact _ vector domain (ID 13 of Table 2, o.exact means Output Exact _ vector, similar for others); further the algorithm depended on ontology reasoning to find out all other related ontology classes such as equivalent, super, sub classes etc for $\text{areacode}$, then respectively added ID of ws5 to their corresponding output vectors. For instance, in the above semantic relationships, because of $\text{areacode} \supseteq \text{cityid}$, $\text{cityid} \equiv \text{citycode}$, the algorithm inserted ontology concepts $\text{cityid}$ and $\text{citycode}$ into QSQL; then added ID of ws5 to their Output Plugin _ vector domains (ID 6, 9 of Table 2). Repeatedly, the publication algorithm processes each parameters of ws5. As such, all ontology concepts which have semantic relationships with WS5 could be found to record the information of ws5 in their corresponding vectors by algorithm. Therefore, QSQL can make sure a large number of ontology reasoning can be processed at service publication stage so that semantic service discovery and composition can be achieved a quick response. More details can be found in [18].

4. Extended Petri net Representation for Process Model

A Web service behavior is basically a partially ordered set of operations, and it is straightforward to map it into a Petri net in [1, 2]. In our methods, we extend the definition of Petri net and composition operators in [1, 2] so that our composition plans can be directly mapped by the weighted Petri net for dynamic creation of process flow.

Definition 4.1 (Extended Basic Petri net (EBPN))
An extended basic Petri net is a weighted service net SN with a 7 tuple $\langle P, T, R, I, O, W, S \rangle$ where:

- (1) $P$ is a finite set of places;
- (2) $T$ is a finite set of transitions representing the operations of the service;
- (3) $R$ is a finite set of arcs where $R \subseteq \{P \times T\} \cup \{T \times P\}$;
- (4) $I$ is the input place; $O$ is the output place;
- (6) $W : R \rightarrow \{1, 2, 3, 4, 5\}$ is a weight function which will represent a corresponding matching degree according to the definition in [18]. Namely, 1 denotes $\text{Grachd};$ 2 denotes $\text{Grapar};$ 3 denotes $\text{Sib};$ 4 denotes $\text{Plugin};$ 5 denotes $\text{Exact}$;
- (7) $S : T \rightarrow A \cup \{\Theta\}$ is a labeling function where $A$ is a set of operations of all WS (1, 0, .) $\Theta$ denotes a null operation.

A composition process flow is composed of weighted basic service nets by effective composition operators. These composition operators are used to perform algebraic operations over services to obtain
composite services[1]. In this paper, we only give three basic composition operators: sequence, choice, and parallelism operator.

**Definition 4.2 Sequence Operators** \( \triangleright \)

\( \triangleright \) denotes an operator of sequence. \( w_s \triangleright w_s \) represents two adjacent services in a sequential relationship. Figure 1 shows the graphical representation of \( w_s \triangleright w_s \).

![Figure 1. Sequence services \( w_s \triangleright w_s \)](image)

The corresponding net SN of \( w_s \triangleright w_s \) is \( \{P,T,R,I,O,W,S\} \) where

- \( P = P_1 \cup P_2; \ T = T_1 \cup T_2 \cup \{t_i, t_j\}; \ R = R_1 \cup R_2 \cup \{(o_i, t_i), (o_j, t_j)\} \)
- \( I = i_1; \ O = o_1; \ W(o_i, t_i) = W(t_i, o_i); \ S = S_1 \cup S_2 \cup \{t_i, t_j\} \)

**Definition 4.3 Choice Operators** \( \otimes \)

\( \otimes \) is a choice operator which allows building a composite service out of two services. The component services linked by the choice operator are executed alternatively. Figure 1 shows the graphical representation of \( w_s \otimes w_s \).

![Figure 2. Choice services \( w_s \otimes w_s \)](image)

The corresponding SN of \( w_s \otimes w_s \) is \( \{P,T,R,I,O,W,S\} \) where

- \( P = P_1 \cup P_2 \cup \{i, o\}; \ T = T_1 \cup T_2 \cup \{t_i, t_j\} \)
- \( R = R_1 \cup R_2 \cup \{(i, t_i), (i, t_j), (o_i, t_j), (o_j, t_j), (o_i, t_i), (o_j, t_i)\} \)
- \( I = i_1 \cup i_2; \ O = o_1 \cup o_2; \ W(o_i, t_i) = W(t_i, o_i); \ S = S_1 \cup S_2 \cup \{t_i, t_j\} \)
- \( W(i, t_i) = W(t_i, o_i); W(t_i, o_i) = W(o_i, t_j); W(o_i, t_j) \)

**Definition 4.4 Parallelism Operators** \( \oplus \)

The component services linked by the parallelism operator \( \oplus \) are executed concurrently. Figure 3 shows the graphical representation of \( w_s \oplus w_s \).

![Figure 3. Parallelism services \( w_s \oplus w_s \)](image)

The corresponding SN of \( w_s \oplus w_s \) is \( \{P,T,R,I,O,W,S\} \) where

- \( P = P_1 \cup P_2 \cup \{i, o\}; \ T = T_1 \cup T_2 \cup \{t_i, t_j\} \)
- \( R = R_1 \cup R_2 \cup \{(i, t_i), (i, t_j), (o_i, t_j), (o_j, t_j), (o_i, t_i), (o_j, t_i), (t_i, o)\} \)
- \( I = i_1 \cup i_2; \ O = o_1 \cup o_2; \ S = S_1 \cup S_2 \cup \{t_i, t_j\} \)
- \( W(i, t_i) = \text{sum}(W(t_i, o_i), W(t_i, o_j)), W(t_i, o) = \text{sum}(W(o_i, t_i), W(o_j, t_j)) \)

5. QSQL-based Planning Algorithm

5.1. Algorithm Design

In the previous paper [18], we have proposed service publication rules and algorithm to build up QSQL. Simultaneously, we also have presented a service discovery algorithm for searching a single service from QSQL to meet the user’s requirements. However, in real application, service composition is very useful when there is no single service to satisfy user’s request. In this section, we give an efficient composition algorithm to automatically combine multiple services from QSQL to meet user’s requirements.

We take service models in Table 1 as a composition example. For a given query \( \text{wsr}(\text{cityname, stockname, datetime}, \text{o(RMBprice)}) \), there will be no solution for single service meeting user’s requirements. But intuitively, as shown in Figure 4, \( \text{ws}3, \text{ws}4, \text{ws}2 \) and \( \text{ws}1 \) can be composed in some orders to meet the query. Alternatively, \( \text{ws}5, \text{ws}4, \text{ws}2 \) and \( \text{ws}1 \) can also be another plan to meet the query with a different matching effect. In order to automatically offer such composition plans meeting user’s query, we designed and implemented a QSQL-based composition algorithm. In our composition method, the only thing users need to do is to specify a query including the inputs and outputs. Particularly, these inputs and outputs will be mapped to the corresponding ontology concepts. Actually, many semantic tools and methods have been proposed to help annotate semantic information to services[17, 19]. Therefore, for a given user’s query, we can assume that the specified inputs and outputs in queries have been mapped by ontology concepts.

![Figure 4. Intuitive composition plans for query](image)
IC' semantically contains the sub goal (step 4). If it is true for IC' ⊇ subgoal, it means the sub goal can be matched directly by the user’s inputs, and the weighted are between some of query’s inputs and sub goal.

**Table 3:** The definition of symbols in algorithm

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>I'</td>
<td>user’s query</td>
</tr>
<tr>
<td>O'</td>
<td>inputs of service model</td>
</tr>
<tr>
<td>V'</td>
<td>outputs of service model</td>
</tr>
<tr>
<td>V'</td>
<td>Output-Exact vector</td>
</tr>
<tr>
<td>V'</td>
<td>Output-Plugin_vector</td>
</tr>
<tr>
<td>V'</td>
<td>Output-Sib_vector</td>
</tr>
<tr>
<td>V'</td>
<td>Output-Grapar_vector</td>
</tr>
<tr>
<td>I'</td>
<td>outputs of service models</td>
</tr>
<tr>
<td>O'</td>
<td>specified inputs by user</td>
</tr>
<tr>
<td>I'</td>
<td>specified outputs by user</td>
</tr>
<tr>
<td>V'</td>
<td>Output-Exact_vector</td>
</tr>
<tr>
<td>V'</td>
<td>Input-Exact_vector</td>
</tr>
<tr>
<td>V'</td>
<td>Input-Plugin_vector</td>
</tr>
<tr>
<td>V'</td>
<td>Input-Sib_vector</td>
</tr>
<tr>
<td>V'</td>
<td>Input-Grapar_vector</td>
</tr>
</tbody>
</table>

should be a part of a successful searching path. Therefore, the algorithm inserts the path (I'(i), subgoal, matchingweight) to the composition plan (step 6). Then the algorithm begins to pop and match the next sub goal (step 7) from the stack. However, if IC' doesn’t semantically contain subgoal, it means this subgoal can’t be met by the user’s inputs. Hence the algorithm begins to search QSQL and retrieve the related service models to meet subgoal (step 9). For this step, because all published service models have been semantically recorded in output vectors V', x=1,2,3,4,5 (defined in Table 3) of each ontology concept in QSQL, the algorithm can easily retrieve all service models to match subgoal by finding its corresponding ontology concept from QSQL. If ∃x, subgoal, V' x is null, it means this subgoal is unreachable; further all related paths with this sub goal should be deleted from the composition plans (step 11). Contrarily, for any V' x, if they are not null, all service models WS, which belong to subgoal, V' x (x=1,2,3,4,5) can sequentially match the sub goal, so the algorithm adds the path (WS, subgoal, x) to the composition path with the weight value x (step 14). Notably, x simultaneously represents the meaning of the matching degree according to the definition of Table 3. For example, if WS, (I’, O’, UID ∈ subgoal, V’), then (WS, subgoal, 5) shows WS meets subgoal with the matching degree EXACT. Now back to the algorithm, if ∃i, j, x, WS, (I’, O’, UID ∈ subgoal, V’), WS, (I’, O’, UID ∈ subgoal, V’), then the relationship between WS and WS is WS ⊗ WS, which means they can be the alternate path to meet this subgoal (step 15). In particular, for each WS, (I’, O’, UID ∈ subgoal, V’), the inputs I’ needs to be concurrently met by either user inputs, or the outputs of other service models so that WS can be successfully executed to meet subgoal.

Therefore, the inputs I' of WS should be pushed into the stack to be next searching goals (step 16). Considering the concurrent, these new sub goals should be added to the path of composition plans in parallelism order (step 17). Similarly, all sub goals should be popped from the stack and processed by step 3 to 17 until the stack is null. When the stack is null, either all composition plans in QSQL will be generated, or there is no such plans to meet the user’s query. Because the algorithm has taken the extended and weighted Petri net model, all final composition plans can be represented by Petri nets.

**Algorithm 1: FindAllProcessPlans**

```
Inputs: WSRI(V, I, O, UID); QSQL;
Outputs: composition plans
1 For I’, IC’ ← extendingInputsource(I’); //Forming extended input source I’ of I; by user reasoning
2 push (IC’); //For each output o, pushing o to stack so that o will be a search goal
3 subgoal ← pop(); //Retrieving a subgoal from target stack;
4 relationJudging(IC', subgoal); //Judging the semantic relations if IC’ ⊇ subgoal
5 Case (true): IC’ ⊇ subgoal
6 addpath(I’, IC’, subgoal, matchingweight); //Subgoal has been founded, adding path from one input of I’ to subgoal, inserting the matching weight value in the path
7 GO TO 3; (Popping next subgoal);
EndCase 1
8 Case (false): IC’ ⊌ subgoal
9 V’ ← retrieveAllO_vectorsfromQSQL(subgoal), x = 1, 2, 3, 4, 5; //Retrieving all output vectors v of subgoal from QSQL
10 If ∀∅(V’ ∧ subgoal) V’ = ∅ Then deleteLocalPath(subgoal); //It is impossible for subgoal to be arrived. Deleting the paths which are related with subgoal
11 GO TO 1;
12 Else For ∀V’ j, j ∈ WSRI(I’, O’, UID = ∅), V’; WSRI(I’, O’, UID = ∅) Do
13 insertSequenceOperatorPath(WS, subgoal, VS, j); //WS can match subgoal with matching degree x
14 insertChoiceOperatorPath(WS, VS, j, I’); //Because of meeting the same subgoal; existing relationship ws, ws, push (I’); //For a x, (I’, O’), v ∈ V’; making j’ become subgoals, pushing j’ to stack
15 insertParallelismOperatorPath(WS, I’, j’); //For inputs I’ of WS, they need to be matched simultaneously to receive ws’
16 EndFor
EndCase 2
19 Repeat 3 and 17 until stack is null
```

According to algorithm 1, our proposed composition algorithm mainly achieved the following distinguished advantages:

- For a given query, the algorithm can automatically find all composition plans from QSQL;
- All composition plans can be ranked by their weighted matching degree;
- All generated plans can be built efficiently without user interaction and much ontology reasoning;
- All generated plans can be represented by extended Petri nets which are formally introduced in section 4.

**5.2. Complexity of Algorithm**

As shown in algorithm 1, the computing complexity mainly depends on the scale of searching goals in stack and the cycle numbers between step 13 and 18. The number of searching goals will be changed
dynamically. Let’s consider the worst case, assume the total number of published service models is N, the maximum number of input parameters for each service model is µ, then the upper limit of goals in stack is \( \mu N + |\alpha'| \cdot (|\alpha'|) \) requested output number for query). Additionally the maximum scanning number between step 13 and 18 is 5N. Consequently, the total cycle numbers is \( (\mu N + |\alpha'|) + 5N \). Normally, \( \mu \) and \( |\alpha'| \) are constants. Therefore, the total complexity of our composition algorithm is \( O(N^2) \) under polynomial time. In particular, with our algorithm, the reasoning time happens only at the beginning when forming the extended input set \( IC' \). Nevertheless, in the cycle body, there is not any reasoning happening because QSQL has stored semantic relationships between concepts and services in advance. By comparison, semantic reasoning is normally a key factor for performance issue to generate composition plans in the traditional semantic service composition approaches. We will further illustrate this in our experiment analysis.

5.3. Composition Example

For a given composition example in Table 1 and QSQL records in Table 2, we use our composition algorithm to find all composition plans to meet the specified query \( wsr(i(cityname,stockname,datetime),o(RMBprice)) \). The algorithm first formed the extended input source \( IC' \) for query’s inputs \( (cityname,stockname,datetime) \) by racer reasoning such as \( datatime \subset \text{somedatet ime} \) (step 1), then pushed the query’s output \( RMBprice \) into the stack to be searching goal (step 2). The algorithm began to pop a search goal \( RMBprice \) from the stack (step 3), then judged if \( IC' \supseteq \text{RMBprice} \) (step 4). Because \( IC' \) didn’t contain \( RMBprice \), step 9 was executed. The algorithm found the matched ontology concept \( rmbprice \) in QSQL, then retrieved the related service model from the output vectors of \( rmbprice \). As we know from QSQL in Table 2, \( rmbprice.V_1 \) = \{ws \} (ID 1 of Table 2), so \( ws \) matched the goal \( rmbprice \) with matching degree 5 (EXACT). The algorithm then added the path \( (ws,G: \text{RMBprice}, 5) \) (step 14) to the composition plans (as shown in Figure 5).

Further, for the execution of \( ws \), the inputs \( ( \text{givenDatatime} , \text{USprice} ) \) of \( ws \) should be met concurrently by either query’s inputs or other service model’s outputs. Therefore, the algorithm pushed inputs \( \text{givenDatatime} \) and \( \text{USprice} \) into the stack to be next sub searching goals (step 16). Simultaneously, according to step 17, the path \( (G_{32} : \text{USprice}, ws, 5) \) and \( (G_{32} : \text{givenDatatime}, ws, 5) \) should be added to the composition plan, and the operation relationship between \( G_{31} \) and \( G_{32} \) is \( G_{31} \oplus G_{32} \). Then the algorithm went back to step 3 to pop the next sub goal \( USprice \). Similarly, because \( USprice \subset IC' \), \( USprice.V_1 = ws, t \) (ID 2 of Table 2), \( (ws,G_o : \text{USprice}, 5) \) should be added to the composition plan. (shown in Figure 5).

![Figure 5. Composition example](image)

Figure 5. Composition example plans, and the inputs \( \{ \text{cityID}, \text{stockID}, \text{someDatatime} \} \) of \( ws \) were pushed into the stack to be next searching goals. Repeatedly, the searching goals in the stack were processed by the algorithm. Finally, when the stack is null, all weighted composition plans are generated and ranked. With this example, Figure 5 gives the whole processes for a given query \( wsr(i(cityname,stockname,datetime),o(RMBprice)) \). As shown in Figure 5, the final generated composition plans are \( (ws_1 \oplus (ws_2 \oplus ws_3)) \supset ws_1 \supset ws_2 \supset ws_3 \). Therefore, there exists exactly two plans \( (ws_1 \oplus ws_2) \supset ws_1 \supset ws_2 \), and \( (ws_2 \oplus ws_3) \supset ws_2 \supset ws_3 \) to meet the query. At the same time, according to the former introduced weight computing method of section 4, the total weight value of \( (ws_1 \oplus ws_2) \supset ws_1 \supset ws_2 \) is \( 36(5+5+4+5+5+5+2+5) \), while \( (ws_1 \oplus ws_2) \supset ws_1 \supset ws_2 \) is \( 35(5+5+4+5+5+5+2+5) \).
Therefore, \((\text{ws}_1 \oplus \text{ws}_2) \rightarrow \text{ws}_3 \rightarrow \text{ws}_4\) is better than \((\text{ws}_1 \oplus \text{ws}_2) \rightarrow \text{ws}_3 \rightarrow \text{ws}_5\). However, if \((\text{ws}_1 \oplus \text{ws}_2) \rightarrow \text{ws}_3 \rightarrow \text{ws}_5\) fails to meet the user’s requirements, the plan \((\text{ws}_1 \oplus \text{ws}_3) \rightarrow \text{ws}_2 \rightarrow \text{ws}_4\) will become an alternate one.

6. Experiment and Evaluation

The simulation was performed in our real-world grid workflow management system called SwinDeW-G [21] to demonstrate the efficiency of the above algorithm. Particularly, we mainly focus on the comparison and analysis of the performance for producing composition plans without including the performance comparison of running actual composition plans by taking separating techniques between abstract composition plans and concrete composition execution plans. Currently, we have produced and published 3000 service models to QSQL by using six selected ontology domains[22]. In addition, we also produced 30 composition queries. Query client, QSQL database, and service repository are distributed in different nodes. In our experiment, we mainly compared the performance including composition time, memory consuming between QSQL-based composition methods and the traditional direct reasoning-based composition methods.

**Figure 6.** The comparison for composition query time

**Composition time:** Figure 6 shows the execution time for each query by using two different composition methods. As shown in Figure 6, the following distinctions can be easily achieved. First, the average execution time (11544ms) of each query by direct reasoning-based composition methods is higher than the time by QSQL-based composition method (2297ms). Secondly, the fluctuation of each query by direct reasoning is obvious. Thirdly, the execution time by direct reasoning-based composition method appears to an increasing trend. For the first distinction, the reason is that the direct reasoning-based composition methods need to involve much ontology reasoning when creating the process flow to meet the user’s query. By comparison, because the large number of ontology reasoning has been processed previously in QSQL, our algorithm can offer a quick response for each query. For the second distinction, different queries contained different numbers of ontology reasoning by direct reasoning-based methods, which directly caused to the fluctuation of composition time. While in our composition algorithm, ontology reasoning only happened in the beginning when forming the extended user’s input collection, thus the composition time is relatively stable. For the third distinction, next section can give a good explanation.

**Figure 7.** QSQL-based service composition

**Figure 8.** Direct reasoning-based approach

**Memory Consumption:** In our experiments, the memory consumption has been monitored and analyzed. Figure 7 and Figure 8 give the comparison results. Apparently, when the traditional direct reasoning composition algorithm continuously processed queries, more ontology will be loaded and classified which means more memory space is needed. When RAM is limited, the garbage collection engine in JVM had to work to release more memory which caused higher latency (Figure 8). This also proved the above third distinction. Therefore, from both sides of memory consumption and response time, the direct reasoning-based composition method has become unbearable. By comparison, our QSQL-based composition method can continually process queries in more stable way without the need of much reasoning.
Because the publishing time is not the critical time for application, QSQL-based composition methods are more efficient and meaningful.

7. Conclusions and Future

In this paper, we have presented a new efficient QSQL-based composition algorithm for automatic service composition. For a given user query, the proposed algorithm can not only ensure all existing composition plans in QSQL can be efficiently found in an automatic and semantic manner, but also can rank them by exploiting a weighted Petri net representation for further verification. Our experiments and valuations have demonstrated that our proposed QSQL-based composition algorithm is more efficient than other typical ones. In future, we will focus on how to dynamically bind service instances to abstract composition plans for making resource’s sharing more flexible.

8. Acknowledgement

We are grateful for the English proofreading by D. May, and the foundation support by the National “973” Research Plan Foundation of China under Grant No. 2003CB317008, by National Nature Science Foundation of China under Grant No. 60573135, 40505023 and 60736013, and by Swinburne Dean’s Collaborative Grants Scheme 2007-2008, and by Swinburne Research Development Scheme 2008.

9. References