Novel Approach for Computing Skyline Services with Fuzzy Consistent Model for QoS- based Service Composition

Fatma Rhimi, Saloua Ben Yahia and Samir Ben Ahmed
Faculty of Sciences LISI-INSAT, University of Carthage, Tunis, Tunisia

Keywords: Fuzzy Logic, Optimization, Skyline, Web Services Composition.

Abstract: Service composition is emerging as an effective solution to ensure the integration of multiple atomic web services in order to create value-added customized services. However, the exploding number of the deployed service candidates that is constantly increasing makes the process of choosing the best service candidates an important challenge. When there are multiple web services that offer the same functionalities, we need to select the best one according to its non-functional criteria (e.g. response time, price, reliability). Skyline is a technique that helps reducing the size of our search space and comes as a complementary approach to the optimization methods. In fact, Skyline consists in preselecting the best candidates in the search space according to their non-functional criteria. Those web services are considered optimal as they are not dominated by any other point in the search space. Therefore, we will eliminate all the irrelevant web services which will considerably reduce the complexity of the computation. Most of the current Skyline computation relies on a strict dominance relationship called Pareto-dominance. In this paper, we propose a new method to compute the Skyline points with a fuzzy approach which allows taking into consideration the users preferences. We will through this paper show how we could construct a consistent fuzzy model to overcome the shortcomings of web service composition computation. A detailed study of the approach will demonstrate the effectiveness and the efficiency of the proposed algorithm.

1 INTRODUCTION

Web services are software components designed to enhance the interoperability for machine-to-machine interaction among different applications and different platforms. This is why business structures are moving today towards the service-oriented architecture as web services seem to be the best solution to allow the exchanges between them. Service composition is a process that combines multiple atomic web services in order to create value-added web services. Hence, it is arising as an effective solution to deliver customised services to the different users.

However, today with the prevalence of paradigms such as Cloud Computing and XAAS (everything as a service) that provide services on demand, the number of available web services had exploded. This is why it has become difficult to choose the best candidates that would ensure an optimal composition.

Quality-of-Service (QoS) is widely employed to represent the non-functional characteristics of Web services and has been considered as the key factor in service selection. QoS is defined as a set of properties including response time, throughput, availability, reputation, etc. Hence, optimal composition can be defined as the composition that corresponds the most to the constraints provided by the end user in terms of non functional criteria.

The problem of QoS-based service composition becomes especially important as the number of candidate web services increases enormously every single day. Hence, performing an exhaustive search to find the best composition is not efficient in this case. In fact, even with hundreds of candidates the time execution of exhaustive algorithms is already very high and exceeds the time execution constraints as the number of possible combinations is very large. To tackle this problem, many researchers used methods such as Linear Programming methods which are proved to be very effective in a small space. However, today with the proliferation of the web technologies, there are multiple service providers who offer web services with the same functionality but with different QoS attributes. Those methods have an exponential cost in a context where the number of service candidates is large as

Rhimi F., Ben Yahia S. and Samir B..
Novel Approach for Computing Skyline Services with Fuzzy Consistent Model for QoS- based Service Composition.
DOI: 10.5220/0005499901350143
In *Proceedings of the 10th International Conference on Software Paradigm Trends* (ICSOFT-PT-2015), pages 135-143
ISBN: 978-989-758-115-1

the number of possible combinations grows exponentially.

Skyline is a technique that comes as a solution that helps reducing the search space based on a dominance relationship to preselect the best services and prune the others. Intuitively, a skyline query selects the “best” or most “interesting” points with respect to all dimensions. In this work, we define and exploit dominance relationships between services based on their QoS attributes. This is used to identify services in a service class that are dominated by other services in the same class. Most of the researchers addressed the Skyline with a Pareto dominance relationship: a service p dominates another service q if p is at least as good as q in all the dimensions and strictly better in at least one dimension. However, such strict dominance relationships privileges web services with some bad and some good attributes. Besides, in real world, the user’s preferences are usually complex and vague. It might be difficult to require a business user to express a crisp preference for an item or a feature of an item, and it is therefore difficult to represent the user‘s preferences with crisp numbers. In this study, fuzzy set techniques are used to describe user’s complex and vague preferences. We will through this paper address the Skyline based on a fuzzy dominance relationship which is a known to be a more flexible relationship.

1.1 Contributions

This paper aims to present a new approach for computing Skyline services in order to reduce the number of candidates. We suggest preselecting the best services based on fuzzy dominance relationships.

Fuzzy sets are more suited to the expression and the interpretation of the user’s preferences. Usually users use terms such as ‘rather

fast’, ‘not expensive’, ‘quite reliable’ to express their preferences. Besides, fuzzy sets can select service Skyline with a compromise between good and bad attributes as they use a more flexible dominance relationship.

However, unlike Pareto dominance relationship, dominance relationship of fuzzy sets does not preserve the transitivity property. Pruning services without checking this property can lead to erroneous results. Hence, constructing a consistent fuzzy model is crucial for the effectiveness of the computation. Furthermore, checking the dominance relationship between each pair of services is computationally expensive. So, using structures as R-tree may be very effective for reducing the cost of computation.

Considering all this, our main contributions may be summarized in the following:

- We will address the problem of computing service Skyline with a consideration of user’s preferences by making use of fuzzy preference relationships rather than Pareto dominance relationships.
- We introduce a novel approach to compute the service Skyline that consists in a two-phase algorithm: a transformation phase which constructs a consistent fuzzy model from the collected data and a computing phase which determines the dominance relationship with a branch-and-bound algorithm.
- We evaluate the efficiency and the effectiveness of the proposed method with a theoretical study and we will leave the experimental study for the future work.

1.2 Outline

Section 2 presents the related work of the web services composition problem, the Skyline techniques and fuzzy techniques. In sections 3 we will define the problem statement of our research and present the background of this work with a remainder of the concept of Skyline and fuzzy sets so we can advance the followed approach. Sections 4 and 5 will describe the different steps of the proposed approach. Section 6 contains an experimental study for the approach to evaluate the results. Finally, we will conclude the paper and give an overview on our future work.

2 RELATED WORK

The problem of QoS-based web service selection and composition has received a lot of attention during the last years. Local selection methods using techniques such as Simple Additive Weighting (SAW) were conducted to select services that ensure an optimal composition. However, local selection could not satisfy global constraints on the composition as it treats each service class individually. Zeng et al. (2003) tackled this problem using a global planning composition based on mixed integer Programming technique for dynamic and quality-driven selection. However, the costs of this approach are exponential in a large space. Linear programming methods are very effective when the size of the problem is small, but suffer from poor scalability due to the exponential time complexity of the applied search algorithms. In their work, Alrifai and Risse (2009) proposed a hybrid selection approach that combines local selection with global selection by decomposing global constraints into local constraints in order to find close-to optimal solutions. Canfora (2005) proposed a genetic algorithm to the QoS-based composition. Genetic algorithms are based on the evolution theory and in opposition to linear programming algorithms, the input data doesn't need to be linear. Besides, genetic algorithms are related to the number of service classes and not to the number of candidate web services, so they are more effective in a large space context. However, linear programming is proved to be faster than genetic algorithms and is preferred hence in a small space. Yu and Keiw-Jay (2004) proposed heuristic algorithms that can come as an alternative to exact solutions. The authors modelled the problem as combinatorial problem and proposed a heuristic Branch and Bound algorithm (WS HEU) and a heuristic graph model (MCSP-K). The two algorithms are proved to be more efficient than exact algorithms. Ardagna and Pernici (2007) tried to overcome the shortcomings of both local and global service composition by proposing an approach that addresses optimization problems under severe QoS constraints.

However, today, as we are moving from limited data systems to large scale systems, the methods proposed above are no longer practical. Cloud-based composition approaches were developed to deal with the problem of QoS-based web services composition in large scale systems. One can classify those approaches into five categories: classic approaches such as the work of Kofler, Haq and Schikuta (2010) where the authors tried To achieve a feasible concrete workflow for service composition with respect to the consumer QoS requirement, the problem is considered to be equivalent to a multi-dimensional multi-choice knapsack problem (MMKP) in which a parameter called happiness that is calculated based on QoS parameters is used as the utility function. Combinatorial approaches such as the works of Ludwig (2011) where in the service provider system an improved genetic algorithm is proposed; Yang, Mi and Sun (2012) and Ye where game theory is used to propose a service level agreement (SLA)-based service composition algorithm., Zhou and Bouguettaya (2011) where authors also applied a genetic algorithm to solve the composition problem in which a roulette wheel selection algorithm is used to select chromosomes to execute a crossover operation, framework based approaches like Pham et al. (2010) who proposed a new framework for service composition in which a composition agent is responsible for receiving the request and providing service management, machine based approaches with contributions such as the work of Baou and Dou (2012) researchers designed finite state machines to consider service correlations and finally structure based approaches such as the contribution of Wittern and Menzel(2012) where the composition problem is represented by a directed graph in which the nodes play service roles and the edges denote the relations between service.

Skyline technique is complementary to these solutions as it can be used as a pre-processing step to prune non-interesting candidate services and hence reduce the computation time of the applied selection algorithm. The analysis of the Skyline was originally considered as a mathematical problem. It was then introduced in the first place in the field of database by Borzsonyi, Kossmann and Stocker (2001). Given a set of points in d-dimensional space, the Skyline is defined as the subset containing the points which are not dominated by another point. Paradigms like Block Nested Loops (BNL) and Divide to Conquer are among the first attempts to solve the computing of Skyline.The index structures such as B-trees have also been utilized to improve the performance of analyzing the Skyline. Nearest Neighbour (NN) and Branch and Bound Skyline (BBS) are two representative algorithms that can progressively address the Skyline based on R-tree structure.

In recent works, many researchers focused on computing skyline services in the context of service composition. However, the majority of these works relied on Pareto dominance relationship for this purpose Alrifai, Skoutas and Risse (2010), Chen (2014), Abourezk and Idrissi (2014). Pareto dominance has the shortcoming of neglecting the smoothness and fuzziness of human preferences. A definition and an example of Pareto dominance is given further in our work. Bouguettaya et al. (2013) addressed the problem of uncertainty in service composition and defined a concept called p-dominant service skyline.

Fuzzy logic was addressed in the optimization techniques for service composition in many contributions such as those of Almulla, Almatori and Yahyaoui (2010), Torres, Astudillo and Salas (2011), Ping et al. (2006), Xuan and Tsuji (2008). However, in all these works, fuzzy techniques were used to find global optimization solutions for web services selection. Only few works used fuzzy logic for Skyline computation. To the best of our knowledge, the work that follows a similar line is the work of Benouaret et al. (2011) where the authors defined a concept called the α-dominant service skyline. However, to overcome the shortcomings of fuzzy relationships, they proceed in double-checking the services which is very time-consuming and complex especially when the search space is very large. We suggest in this work another approach for computing skyline with fuzzy dominance relationship by constructing a consistent fuzzy model. The consistent fuzzy model allows a direct pruning of the irrelevant services in the search space, which will enormously reduce the complexity of computation. A detailed study is given in the rest of the paper.

3 BACKGROUND

3.1 Skyline Computation

Skyline can be formally defined as follows:
Given a set of points S in a space with D dimensions, Skyline points are the points who are not dominated by any other point in the search space according to those dimensions. A definition of the dominance concept is then crucial to the understanding of the skyline concept.

Pareto dominance

Definition

Given d the number of dimensions in the space and s_i, s_j two web services in the space, we say that s_i dominates s_j denoted by $s_i \prec s_j$ *iff* s_i is at least as good as s_j in all the dimensions and strictly better in at least one dimension.

In the literature, many researchers made use of this concept to compute the service Skyline. In their work, Alrifai et al. (2010) used the Pareto dominance to reduce the size of the search space. They extended their work to compute the set of representative skyline in the case where the size of the initial set is still too large. Yu and Bouguettaya (2011) also used Pareto dominance to introduce a concept of C-Sky which computes the overall Skyline of the composition by a combination of the Skyline sets of individual service classes. Although those contributions are efficient and effective, their reliance on Pareto dominance presents some shortcomings. Pareto dominance is a strict dominance relationship that privileges services with good and bad attributes. Besides, it might neglect user's preferences. The example below illustrates this.

Illustrative Example

Let's consider the common example in the literature that selects the set of interesting hotels in a reservation service represented in Fig.2. The hotels are represented by two criteria: their prices and their distances from the beach. It is obvious that a hotel with a low price and a small distance from the beach is preferred in this case. According to this, the Skyline points are a,l and m as they are the only points that are not dominated by any other point in the search space.

One can notice that selecting Skyline points with a Pareto dominance relationship is strict and can discard potentially good candidates. Let's consider for example the point h (4, 3): According to the Pareto dominance, this point is dominated by other points in the space so it is discarded from the Skyline set. However, some users who accord an importance to the price would prefer the point h over the point a which has a small distance from the beach but has a high price (1, 9).

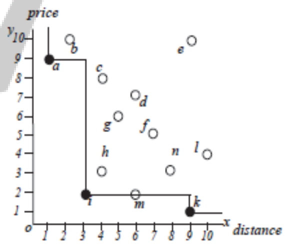

Figure 1: Example of Skyline dataset.

Fuzzy dominance relationships can overcome this problem as they express the user's preferences gradually. The next section will demonstrate this affirmation.

Fuzzy dominance

Given two points in a space with d dimensions, we can define the dominance relationship as follows:

$$\deg_{\mu_{\varepsilon,\lambda}}(s_i \prec s_j) = \frac{\sum_{m=1}^{d} (\mu_{\varepsilon,\lambda}(q_m(s_i), q_m(s_j))}{d} \quad (1)$$

With:
d : The space dimensions (i.e. the QoS attributes in our context)
s_i, s_j : two points in the search space
$q_m(s_i), q_m(s_j)$: The values of the m^{th} attribute for s_i

and s_j respectively.

$\mu_{\varepsilon,\lambda}$: A fuzzy membership function that is defined as follows:

$$\mu_{\varepsilon,\lambda}(x,y) = \begin{Bmatrix} 0 \text{ if x-y} \le \varepsilon \\ 1 \text{ if x-y} \ge \lambda+\varepsilon \\ \dfrac{x-y-\varepsilon}{\lambda} \textit{otherwise} \end{Bmatrix} \qquad (2)$$

With $\varepsilon \ge 0$, $\lambda > 0$.

Let's for example consider $\varepsilon = 0$ and $\lambda = 0.3$. This would result in $\mu_{0,0.3}(\text{a},\text{h}) = 0.5$ and $\mu_{0,0.3}(\text{h},\text{a}) = 0.5$. Hence, point h would not be discarded which proves that fuzzy dominance privileges points with good compromise.

In this paper, we will propose an approach for computing Skyline with fuzzy dominance. The next section will be a remainder of the basics of fuzzy logic and fuzzy sets that are crucial to the understanding of the rest of the paper.

3.2 Fuzzy Sets

Fuzzy sets were introduced by Lotfi Zadeh A. (1965) as an extension of the classical notion of sets. Fuzzy sets are described by means of a membership of a real value in the interval [0, 1] function. Fuzzy sets generalize classical sets, since the indicator functions of classical sets are special cases of the membership functions of fuzzy sets, if they do not take the values 0 or 1. The theory of fuzzy sets can be used in a wide range of areas in which information is incomplete or inaccurate. Furthermore, they are well suited for expressing non exact linguistic terms such as 'rather reliable', 'cheap' and 'not expensive' that are widely used by web services users.
Binary fuzzy relation between two non-empty fuzzy sets is a fuzzy subset of the Cartesian product $X \times Y$ namely:

$R = \{\langle\langle x,y\rangle, U_r(x,y)\rangle : x \in X, y \in Y\}$

Where $U_r : X \times Y \to [0,1]$ is a membership function which assigns to each pair $x \in X, y \in Y$ the membership degree $U_r(x,y)$, interpreted as the degree of the relation between $x \in X, y \in Y$. $U_r(x,y) = 1$ means that the two components x and y are fully connected. $U_r(x,y) = 0$ means that the two elements are completely independent.

Definitions

Fuzzy preference relationships

A fuzzy preference relation P on a set of alternatives X is a fuzzy set on the product X × X that is characterized by a membership function $u_p : X \times X \to [0,1]$. The fuzzy relation can be represented by the matrix $N \times N$ $p = (p_{ij})$ with $p_{ij} = u_p(x_i, x_j)$ $i, j = 1..n$.

P_{ij} is interpreted as the degree of preference of x_i over x_j: $P_{ij} = \frac{1}{2}$ indicates indifference of preference between the alternatives, $P_{ij} = 1$ is interpreted as x_i is totally preferred to x_j and $P_{ij} > \frac{1}{2}$ is interpreted as x_i is preferred to x_j. Hence we have $P_{ii} = 0.5$.

Fuzzy preference relationships are assumed to be additively reciprocal which implies that $p_{ij} + p_{ji} = 1$.

Additive transitivity

Additive transitivity for fuzzy preference relations can be seen as a property to characterize consistency in the case of fuzzy preference relations. The mathematical formulation of the additive transitivity was given by Tanino (1988).

$$(p_{ij} - 0.5) + (p_{jk} - 0.5) = (p_{ik} - 0.5) \forall i,j,k \in \{1,...,n\} \qquad (3)$$

This equation can be written as follows:

$$P_{ij} + P_{jk} + P_{ik} = \frac{3}{2} \quad \forall i < j < k \qquad (4)$$

4 CONSISTENCY OF FUZZY MODELS

In decision making, the study of consistency when the decision makers express their opinions by means of preference relations becomes a very important aspect in order to avoid misleading solutions. In decision making problems based on fuzzy preference relations the study of consistency is associated with the study of the transitivity property. The decision-making process with preferences is based on fuzzy preference relationships, where the process is related to a degree of preference of an alternative over another. Therefore, establishing properties to check for preference relations is very important for the design of valid models for the decision making process. One of these properties is called the consistency property. The lack of consistency in decision making can lead to incoherent conclusions; this is why it is important, if not essential, to study the conditions under which consistency is satisfied. Transitivity is one of the most important properties concerning preferences. In a fuzzy context, where a user expresses his opinions using fuzzy preference relations, a traditional requirement to characterize consistency is using transitivity, in the sense that if an alternative x_i is preferred to alternative x_j and this one to x_k then alternative x_i should be preferred to x_k.

Our approach consists in generating a preference fuzzy model comprising n-1 preference values collected from users and generated from a membership function. This model will respect the properties of reciprocity and transitivity. Our idea is that by constructing a consistent model from the start we can reduce the time cost of the checking process in the Branch and Bound Skyline with fuzzy dominance relationship. In fact, without preservation of consistency, the pruning of irrelevant services can be direct as a service can be in the same time dominating and dominated by a service. To overcome this problem, we suggest injecting the transitivity property from the beginning, when collecting the preference values from the users, by applying a set of transformations and operations on fuzzy sets.

Herrera et al. (2004) chose the additive transitivity for the construction of the consistent model. A fuzzy preference model is consistent if and only if it fulfills Eq. (3).

This leads to establish the following result:

$$p_{i(i+1)} + p_{(i+1)(i+2)} + \dots + p_{(j-1)i} + p_{ji} = \frac{j-i+1}{2} \quad (5)$$

The proof of this affirmation is found in their work. For sake of simplicity, we will only use the result of this proof in the rest of the paper. This property allows us to construct the $n-1$ preference values $p_{12}, p_{23}, \dots, p_{(n-1)n}$ collected from the users with a fuzzy relationship function.

According to the definitions above, these two matrices are additively transitive; hence the fuzzy model is consistent. It is worth to notice that in certain cases, we would have obtained a matrix P with entries not in the interval [0, 1], but in an interval [−a, 1 + a], being a > 0. In such a case, we would need to transform the obtained values using a transformation function which preserves reciprocity and additive consistency. Herrera et al. (2004) proposed the following function for normalizing the values:

$$f(x) = \frac{x+a}{1+2a} \quad (6)$$

Where Eq. (6) is a function verifying the following properties:

$$f(-\mathrm{a}) = 0$$

$$f(1+\mathrm{a}) = \mathrm{a}$$

$$f(x) + f(1-x) = 1 \forall x \in [-\mathrm{a}, 1+\mathrm{a}]$$

$$f(x) + f(y) + f(z) = \frac{3}{2}, \forall x, y, z \in [-\mathrm{a}, 1+\mathrm{a}]$$

$$\textit{such that } x + y + z = \frac{3}{2}$$

Hence, we can summarize the method to construct a consistent fuzzy model from the $n-1$ preference values collected in the following steps:

1. Compute the set of preference values A with (1) such as:

$$A = \{p_{ij}, p_{ij} = \frac{j-i+1}{2} - p_{ii+1} - p_{i+1i+2} \dots - p_{j-1j} \forall i < j\}$$

2. Compute the overall preference relationship matrix P. As P is reciprocal we will have:

$$p = A \cup \neg A \cup \{p_{12}, p_{23} \dots, p_{n-1n}\} \cup \{1-p_{12}, 1-p_{23} \dots, 1-p_{n-1n}\}$$

3. Determine the range for the normalization :

$$\mathrm{a} = |min\{B \cup \{p_{12}, p_{23} \dots, p_{n-1n}\}\}|$$

1. Compute the normalized preference relationship matrix such as:

$$p' = f(p) \textit{ with}$$

$$f: [-\mathrm{a}, 1+\mathrm{a}] \rightarrow [0,1]$$

$$f(x) = \frac{x+a}{1+2a}$$

5 SKYLINE USING A BRANCH AND BOUND ALGORITHM WITH A CONSISTENT FUZZY MODEL

5.1 Branch and Bound Skyline using Pareto Dominance Relationship

Branch and Bound Skyline was first introduced by Papadias (2003) for the computation of Skyline points with a Pareto dominance relationship. Branch and Bound Skyline is an algorithm based on R-Tree structure known for its efficiency and effectiveness in large spaces. It is widely used to reduce the search space. An example of R-Tree data is illustrated in Fig.2: Data points are regrouped in nodes according to their distance from the origin. An entry is the Minimum Bound Rectangle of a node and a leaf entry is a data point. Papadias used a Pareto dominance relationship to determine the Skyline points in the search space. In Pareto dominance, the property of transitivity is verified. However, this property is not verified by the fuzzy dominance relationship. When transitivity is not preserved, discarding some points can lead to erroneous results. Thus, constructing a consistent fuzzy model with transitivity property is crucial for establishing correct results. Benouaret, Benslimane and Hadjali (2011) used the Branch and Bound Skyline with a fuzzy dominance membership function. In order to address the above problem, they developed an optimization technique called a - *dominant* service Skyline. This new concept is proved to be effective and efficient. However, the authors only considered the lack of antisymmetry in the fuzzy relationships. They double-check every point in the heap for dominance before inserting it in the Skyline. Our proposed algorithm comes as a solution to this problem. We suggest constructing form the collected data a consistent fuzzy model by applying fuzzy transformations. This model will preserve the transitivity property. Hence, the pruning process can be direct and we will be sure that the results are not erroneous.

5.2 Branch and Bound Algorithm with Fuzzy Dominance Relationship

The proposed algorithm is a two-phase algorithm. The first phase consists in transforming the data

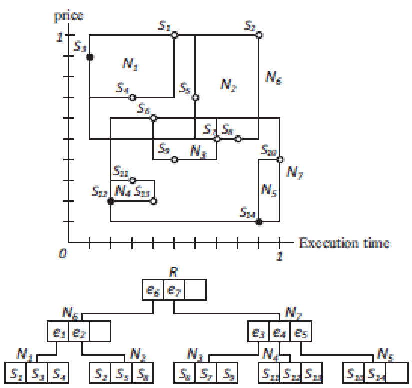

Figure 2: Example of R-Tree.

points of the search space into a consistent fuzzy model. The second step is determining the Skyline points with a Branch and Bound algorithm according to the fuzzy dominance relationship.

Algorithm 1: Fuzzy Consistent Branch and Bound Skyline

```
Input: service R-Tree, fuzzy membership function
Output: set of Skyline points
   Begin
   1. Heap H=Ø, Skyline S=Ø
   2. For all the services in the search space inserted in the R-Tree:
   3. Compute the dominance degree for all the QoS criteria of each pair of services inserted in the R-tree according to Eq. (1) and store them in the R-Tree
   4. Transform the computed values according to Eq. (5).
   5. Determine the value of the normalization range
   6. Compute the overall normalized values according to Eq. (6).
   7.For all entries in the Root:
      8. Insert all entries in the heap
   9. While heap not empty
    s 10. Remove the entry e with the min distance
        11. If e is fuzzy-dominated by a point in S discard e
        12. Else
          13. if e is an intermediate entry
            14. For each child c of e
               15.If c is not fuzzy-dominated by some point in S Insert c into heap
               16. Else Insert c into S
   17. End while
   End FCBB
```

6 EXPERIMENTAL EVALUATION

In this section we verify the effectiveness and efficiency of our proposed algorithm referred to as *CFBBS* (Consistent Fuzzy Branch and Bound Skyline). We conduct a set of experiments by comparing our algorithm to a Branch and Bound Skyline algorithm with Pareto dominance relationship referred to as *PDBBS*. In this experiment we will focus on the size of the service

Skyline provided by both algorithms in order to study how the fuzzy dominance affects the selected Skyline points. Then, to study the scalability of our algorithm, we developed a native approach for skyline computation referred to as *NA*. We took in this study into consideration the effects of the number of services. Other parameters will be studied on future work. The parameters are summarized in Table1. The algorithms were implemented in Java. The experiments were conducted on a 2.00 GHz Intel core I7 CPU and 8 GB of RAM, running Windows.

6.1 Size of the Skyline

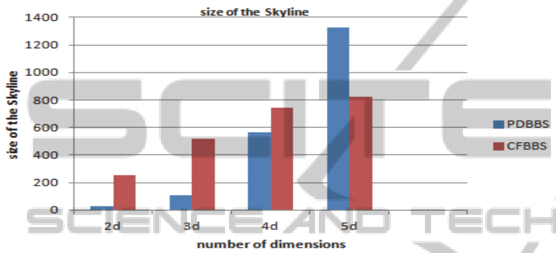

Figure 3: Effect of the number of dimensions on the Skyline size.

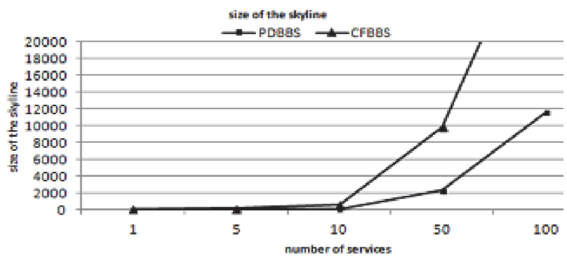

Figure 4: Effect of the number of services on the Skyline size.

Figure 3 shows that the size of the Skyline increases as the number for dimensions for both algorithms. However, this increase is more accentuated for the Pareto dominance skyline. Figure 4 shows that the size of the fuzzy dominance skyline is larger than the size provided by the Pareto dominance skyline. Both algorithms show that the size of the skyline increases as the number of services increases. This result is expectable since the fuzzy dominance would privilege services with compromises between good and bad values.

6.2 Execution Time and Scalability

Figure 5 and Figure 6 show that the CFFBS is more scalable than the basic approaches for skyline

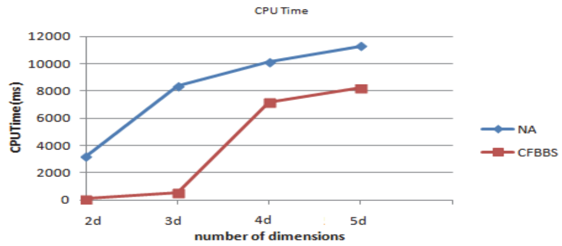

Figure 5: Effect of the number of services on the CPU Time.

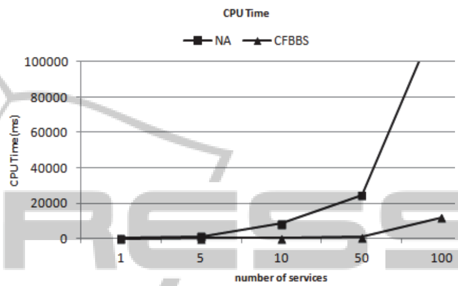

Figure 6: Effect of the number of services on the CPU Time.

computation. This result is significant as the number of services increases which is due to the pruning process of the R-Tree structures. Besides, constructing a consistent fuzzy model allow us to discard directly dominated points without double-checking the search space. However, we can notice that when the number of dimensions increases, the two algorithms have almost the same performances. This is because R-Tress structures perform poorly when the number of dimensions is high.

Table1: Parameters and considered values.

Parameters	Values
Number of services	1K, 5K, 10K, 100K
Number of dimensions	2d, 3d, 4d, 5d
ε, λ	0.1, 0.2

7 CONCLUSIONS

We have addressed in this paper the problem of QoS-based web service composition. We advanced that exhaustive search in the context of large-scale systems is not a practical solution. Hence, Skyline technique, as it allows reducing the search space can improve the performance of the composition methods. We have tackled in this paper the skyline with fuzzy preference values to take into consideration the possible compromises between the QoS values. So we have proposed a method to construct from the collected fuzzy values a consistent fuzzy model that would reduce the time cost of the double-checking process in the Branch and Bound Skyline with fuzzy dominance. We have proved the efficiency and the effectiveness of our proposed algorithm with an experimental study.

In our future work we will focus on extending this concept to cover the whole QoS-based composition process.

REFERENCES

M. Abourezk, A. Idrissi, 2014. Introduction of an outranking method in the Cloud computing research and Selection System based on the Skyline. *Research Challenges in Information Science (RCIS), 2014 IEEE Eighth International Conference on*. 1-12.

D. Ardagna, B.Pernici, 2007.Adaptive Service Composition in Flexible Processes. *IEEE Trans. Software Eng*. 369-384.

M. Almulla, K. Almatori and H. Yahyaou , 2011. A QoS-Based Fuzzy Model for Ranking Real World Web Services. *Web Services (ICWS), 2011 IEEE International Conference on*. 203 – 210.

M. Alrifai and T. Risse, 2009. Combining global optimization with local selection for efficient qos-aware service composition. *In International World Wide Web Conference* .881–890.

M. Alrifai, D. Skoutas, and T. Risse , 2010. Selecting skyline services for qos-based web service composition. *In WWW*.11-20.

H. Bao and W. Dou, 2012. A QoS-Aware Service Selection Method for CloudService Composition. *Parallel and Distributed Processing Symposium Workshops & PhD Forum (IPDPSW)*. 2254 – 2261.

K. Benouaret, D. Benslimane, A. Hadjali, 2011. On the Use of Fuzzy Dominance for Computing Service Skyline Based on QoS. *In the 9th International Conference on Web Services (IEEE ICWS 2011).*

S. Borzsonyi, D. Kossmann, and K. Stocker. The skyline operator, 2001. *In International Conference on Data Engineering* .421–430.

G. Canfora, 2005.An approach for QoS-aware service composition based on genetic algorithms*. Proceedings of the 2005 conference on Genetic and evolutionary computation*. 1069-1075.

L. Chen, 2014. Ensuring reliability and qos optimizing for web service composition. *Computational intelligence and security (cis)*. 510-513.

E. Herrera-Viedma, F. Herrera, F. Chiclana, M. Luque, 2004. Some issues on consistency of fuzzy preference relations, *European Journal of Operational Research.*

K. Kofler, I. U. Haq, and E. Schikuta.2010. User-centric, heuristic optimization of service composition in clouds, *EuroPar'10 Proceedings of the 16th international Euro-Par conference on Parallel processing*.405- 417.

S. A. Ludwig, 2001. Clonal selection based genetic algorithm for workflow service.

D. Papadias,Y. Tao, G. Fu, 2005. Progressive skyline computation in database systems.*journal ACM Transactions on Database Systems (TODS).* 41-82.

T. V. Pham, H.Jamjoom, K. Jordan and Z.-Y Shae, 2010. A service composition framework for market-oriented high performance computing cloud.*In Proceedings of the 19th ACM International Symposium on High Performance Distributed Computing*. 284–287.

T. Tanino, Fuzzy preference relations in group decision making, 1988. J. Kacprzyk, M. Roubens, (Eds.).

T. Vuong Xuan, H. Tsuji,2008. QoS Based Ranking for Web Services: Fuzzy Approaches. Next Generation *Web Services Practices, NWESP '08. 4th International Conference on*, 77 – 82.

P. Wang, K. M. Chao, C. C. Lo; C. L. Huang; Y. Li, 2006. A Fuzzy Model for Selection of QoS-Aware Web Services. *e-Business Engineering. ICEBE '06. IEEE International Conference on* . 585-593.

E. Wittern, J. Kuhlenkamp, and Menzel, 2012. Cloud service selection based on variability modelling.*LNCS*. 127–141.

Y. Yang, Z. Mi, and J. Sun, 2012. Game theory based IaaS services composition in cloud computing environment. *Advances in Information Sciences and Service Sciences*. 238–246.

Z. Ye, X. Zhou and A. Bouguettaya, 2011. Genetic algorithm based QoS-aware service compositions in cloud computing. *In J. Yu, M. Kim, & R. Unland (Eds.).Database Systems for Advanced Applications*. 321–334.

Q. Yu and A. Bouguettaya, 2011. Computing service skyline from uncertain qows. *IEEE T. Services Computing,* 16-29.

T. Yu and L. Kwei-Jay,Service selection algorithms for Services web withend-to-end QoS constraints , 2004. *Lin e-Commerce Technology Proceedings IEEE International Conference* .129 – 136.

L. A Zadeh, 1971. A similarity relations and fuzzy orderings. ch.1.

L. Zeng, B. Benatallah, M. Dumas, J. Kalagnanam, and Q. Sheng, 2003. Quality-driven Web Service Composition. *In WWW* 411-421.