Metric Suite for Measuring Service Discoverability

1Research Scholar, Department of Banking Technology, Pondicherry University
2Assoc. Prof., Department of Banking Technology, Pondicherry University

1shundar_gss2004@yahoo.co.in, 2prasanna_v@yahoo.com

ABSTRACT

Service discoverability has greater significance in promoting reuse of services and achieving effective service composition. It stands as the prime quality attribute of Service Oriented Architecture (SOA). The stakeholders of SOA, service provider, service consumer and service broker are greatly influenced by service discoverability. Quantification of this quality attribute plays a important role in increasing the reusability of services and so on. From the various research works, we are concluding that the existing quantification techniques addressing service discoverability are in a primitive state which has to be explored more for enhancing discovery of services. In this paper, we propose a service discoverability metric suite. Through literature review, the constituents, aspects and sub-aspects corresponding to discoverability are mined and represented in form of taxonomy. The metric suite was proposed which covers entire aspects addressed in discoverability taxonomy. Practical applicability of proposed metric suite was demonstrated and its implications are discussed.

Keywords: Service Discoverability, SOA Quality, Discoverability Measures, Discoverability Aspects, SOA Quality Measures

1. INTRODUCTION

Qualities of SOA is an extensive research area, there are some of the researchers contributing on it. In particular, quantification techniques for measuring SOA qualities are highly needed but the contribution towards SOA quality measures are limited. From the existing works, the various quality attributes of SOA are reusability, interoperability, performance, security, reliability, availability, composability, discoverability, testability, modifiability, conformance, usability and accountability [1][2][3][4][5][6]. In various quality attributes of SOA, service reusability stands as one of the essential factor. Reusability being the core of SOA has helped in gaining its popularity and it also increased the adoption of SOA [5][8][9][10][38]. The key quality attribute of SOA used to realize reusability is service discoverability [3][38]. This quality attribute would play a prominent role in promoting the reuse of services and which has greatly influencing the various stakeholders of SOA.

Service discoverability comprises of various stakeholders of SOA such as service consumer, service provider and service broker [5][37][36]. As per definition of service discoverability is comprised of constituents and aspects. But none of the existing definitions are not exhibiting it aspects [3][5][10]. The rest of the paper is organized as follows, section 2 address the existing definitions and importance of service discoverability and brief survey about the existing works in measuring discoverability of SOA systems and to bring out the current state of art. Section 3 describes the proposed work, aspect mining and the proposed metric suite. Research experiment and its results are presented in section 4 and finally the conclusion and future directions are explained on section 5.

2. RELATED WORKS

In this section initially the various existing definition of service discoverability and importance of quantifying it has been presented.

2.1 Definitions for Service Discoverability

Discoverability has been recognized as an important quality aspect in the development of service centric systems [12]. It plays a major role in identifying the services which meet out the service consumer requirements. From the literature the following definitions of service discoverability are obtained:

• As per Thomas Erl [5] discoverability is defined as the process of searching for and finding solution logic within a specified environment.

• According to Zain Balfagih [3] it is defined as the degree which the individual service can be easily and accurately found consumers using UDDI consumers search their services through functional specification.

• Discoverability definition as per Si Won Choi [10] is a quality attribute measures the capability of the service to easily, accurately, and suitably found at both design time and runtime for the required service specification.

From the various definitions, it is clearly known that discoverability comprises of constituents and its corresponding aspects. But the definitions are not explicitly delivering it. Erl [5] has defined the discovering process alone; Zain Balfagih [3] has defined only functional specification of services. Similarly Si Won Choi [10] has talked about discovery process and more about functional requirements.

From, the research contribution of different researchers it has been known that service discoverability is not a single entity [2][13][17][36]. It contains the two constituents. The two discoverability constituents are Service Discovery and Service Interpretability. There are some researchers who have confusion with the use of these two terminologies (i.e. discoverability and discovery) [4][24][37][38]. From the definitions we know that discoverability is a
quality attribute and discovery is a constituent which enables discoverability. A detailed review of efforts on discoverability constituents are presented in our earlier contributions [35] [36]

Table 1: Literature review about Service Discoverability, their aspects and its corresponding metrics

<table>
<thead>
<tr>
<th>Period of Study</th>
<th>Contributors</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>2003 - 2011</td>
<td>T. Erl, [5]</td>
<td>• on fly they stated definitions about discoverability</td>
</tr>
<tr>
<td></td>
<td>Zain Balfagih and Mohd Fadzil Hassan [4]</td>
<td>• Discussed about the significance for measuring them.</td>
</tr>
<tr>
<td></td>
<td>Si Won Choi, et al., [8] [9][10]</td>
<td>• Some authors directly proposed measure to discoverability</td>
</tr>
<tr>
<td></td>
<td>Kozlenkov et al., [12]</td>
<td>• Defined the aspects of discoverability</td>
</tr>
<tr>
<td></td>
<td>Andreas Wombacher, [13]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ruben Lara et al., [14]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Benjamin Kanagwa and Agnes F. N. Lumaala [15]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Qi Yu Manjeet Rege et al., [17]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>R. Deepa and S. Swamynathan [16]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bensheng Yun [31]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hong-Linh Truong et al., [32]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Natalia Kokash [33]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ali ShaikhAli et al., [43]</td>
<td></td>
</tr>
</tbody>
</table>

From the literature review shown in Table 1, the issues that were not addressed are concluded and listed below

- Constituents of discoverability have not been addressed and measured exactly.
- The factor sufficient for the service discovery of discoverability has not been addressed. This is inferred from the analysis done to find out the factors that support discovery and interpretability.
- Interpretability and discovery metrics of services are not addressed correctly. The measurements are done based on the aspects of both constituents (discovery and interpretability) are in a primitive or early stage and needs more exploration.

In order to quantify discoverability, service discovery and interpretability has to be measured.

3. OUR PROPOSED WORK

From the above definitions stated in section 2, the consolidation of essential constituents of the service discoverability and adding missing concerns, formulated a new definition to service discoverability is attempted:

“Service discoverability is the quality attribute that quantifies the process of searching exact service for service consumer requirements based on exposed service description and interpretation of those services based on the functional and non-functional data used to describes its purpose and capabilities”

This definition exhibits that service discoverability contains two constituents, service discovery and service interpretability [5] [16] [32]. Service discovery deals with searching or finding the services and interpretability concerns with usage or invocation of those services discovered. Hence, the aspects related to these constituents have to be mined. The mined aspects are modeled, by proposing a taxonomy representing the constituents, their aspects and the sub aspects of service discoverability. Second part of this work concentrates on formulating measures for sub-aspects, aspects, constituents and the quality attribute. Metric of each aspect, sub-aspects and its constituents are formulated through bottom-up approach.

3.1 Aspects identification for Service Discoverability

In order to quantify discoverability it’s essential to identify the aspects or factors involved. Here the aspects addressing the discoverability are discussed in this section. Based on the identified aspects, taxonomy has been proposed for representing the major aspects and its sub aspects. For identification of aspects in service discoverability; top-down approach has been used.

3.1.1 Service Discovery Aspects

Actually, service discovery defines the process of identifying and invoking an appropriate service based on the supplied Meta data. It uses the central repository or registry where actually the services are registered by their providers. In order to measure service discovery, it is essential to measure
the following aspects which enable better search of services in the service registry [36].

- **Service Contract of required service in the Service Registry [34]**
  The service provider has to register their services with the service contract.

- **Meta Data of required service [24]**
  Service contract of the registered services is associated with the corresponding Meta data that describes the service functionalities.

- **Categorization of the required service in the Service Registry [23] [44]**
  In the registry, the services have to be positioned or located under the correct category to make the search easier.

### 3.1.2 Service Interpretation Aspects

**Interpretability of services** deals with the functional and quality of service data that to the invocation of appropriate services. To invoke or use the services efficiently the functional and non-functional aspects i.e. quality of service data of each registered service has to be defined or represented clearly [36] [37]. From the study we have found that the aspects which are listed below are essential for the invocation of the services.

#### 3.1.2.1 Functional Data

Functional data of service depict the purpose and capabilities of the services in the service registry [17]. The two sub aspects that are used to represent functional data are [28]

- **Well Described Semantic Elements** – The semantic elements are used to represent the purpose of the service (i.e. defines the scope of the services)
- **Well Defined Service Operation** - The syntax or interface, which depicts the operation or capabilities of services (i.e. it clearly represents what functionalities are offered by services)

#### 3.1.2.2 Quality of Service data

Quality of service data is used for finding the suitable service from the group of services that meet out consumer requirements. Various quality of service data are identified based on their usage by consumers for assessing or filtering their services are [5] [11] [21] [25] [41]:

- Compliance
- Response Time
- Throughput
- Latency
- Documents related to services
- Reliable messaging and best practices

Now, based on the above discussed factors, service registry has been designed which contains a complete list of attributes. The attributes, their value range relevant to discovery and interpretation of services are discussed below in research experiment section 4.

### 3.1.3 Taxonomy for service discoverability

The proposed taxonomy for service discoverability gives a clear picture about the different aspects that constitute the discoverability as shown in figure 1. This representation would help to define the measures for the constituents, aspects and sub-aspects completely. The levels are used to separate the discoverability, their constituents and aspects, addressing those constituents where level 0 indicates quality attribute, level 1 specifies the constituents, level 2 the aspects and level 3 the sub-aspects. A bottom up approach is to be used for proposing the measures. Lowest level i.e., sub-aspects are to be measured first, then the aspects that belong to the next level are to measured and finally the major aspects constituting the quality attribute are to be quantified.

![Figure 1: Taxonomy for Service Discoverability](image)

### 3.2 Metric suite for service discoverability

The proposed metrics are categorized in the same way as shown in the taxonomy stated above in figure1.

#### 3.2.1 Metric Template

Design template has been defined for metric formulation in order to represent the metrics as shown in figure 1.

<table>
<thead>
<tr>
<th>Metric Name</th>
<th>Metric Definition</th>
<th>Metric formulation</th>
<th>Metric Value</th>
<th>Metric Interpretation</th>
</tr>
</thead>
</table>

Design template elements such as definition describes the intent or purpose of each metric, formulation addresses the different factors used to define metric. Measures fall in ratio level scale. Hence the interval for each measure has been set from 0 to 1. 0 denotes lowest value and 1 is the highest value of each metric. Result of each metric can be interpreted based on the value between the intervals.

#### 3.2.2 Service Discovery Measures

Service discovery is a constituent of service discoverability as
defined in the level of taxonomy in section 3.1.3.

3.2.2.1 Service Contract
Metric Name: Checking Service Contract (CSC)
Definition: CSC is used to measure the existence of the required service contract. It is measured by using the number of attempts to get a hit of required service contract from possible number of attempts.

Metric formulation:
Let CSC be denoted by $K$ and it is calculated by Equation (3.1)

$$K = \frac{n_{hit}}{N_{hit}}$$  \hspace{1cm} (3.1)

Where, $0 \leq K \leq 1$ and $0 < n_{hit} < N_{hit}$

(i) When $n_{hit} = 0$, $K = 0$, it indicates that service contract of required service does not exist
(ii) When $K = 1$, $n = N_{hit}$, it indicates that all possible number of attempts are used to find the required service contract.
Here both the cases are impractical.

Metric value
Interval defined for CSC metric lies between 0 and 1.

Interpretation
Lower the value of $K$ indicates the higher existence of service contract. It implies that minimum numbers of attempts are used to get the required service contract and leads to better discovery.

3.2.2.2 Metadata
Metric Name: Check the Relevance of Metadata (RMD)
Definition: RMD metric is to check the relevance in the existing or supplied metadata. It is a ratio of relevant metadata to existing metadata.

Metric formulation
Let RMD be denoted by $P$ and it is calculated by Equation (3.2)

$$P = \frac{b_{md}}{B_{md}}$$  \hspace{1cm} (3.2)

And $0 \leq P \leq 1$ & $1 \leq b_{md} \leq B_{md}$

Metric value
RMD interval value falls from 0 to 1. Closer to 1 shows higher relevance and 0 indicates nonexistence of metadata.

Interpretation
Higher the value of $P$ indicates the higher relevance of metadata. It shows that existing metadata of service gives more relevance that leads to better discovery.

3.2.2.3 Exact Positioning
Metric Name: Check for Exact Positioning of Service (EPS)
Definition: EPS metric is to check whether the service is exactly positioned or not. The values of EPS are either 0 (position is not exact) or 1 (position is exact).

Metric formulation
Let EPS be denoted by $Q$ and it is calculated by Equation (3.3)

$$Q = \begin{cases} 
0, \text{Position is not exact} \\
1, \text{Position is exact}
\end{cases} \hspace{1cm} (3.3)$$

Metric value
EPS value should be either 0 or 1

Interpretation
Higher the value of $R$ indicates that the services are positioned properly. It seems that services are positioned in their exact category which in turn improves service discovery.

3.2.2.4 Service Discovery
Metric Name: Metric for Service Discovery (SD)
Definition: SD is measured by using the values of metrics of Checking Service Contract, Checking the Relevance of Metadata and Check for Exact Positioning of Service. SD is the linear combination of these metrics.

Metric formulation
Let SD be denoted by $D$ depends on $K$ & $P$ and the positioning of service ($R$) and it is calculated by Equation (3.4)

$$D = 1 - (C_1S + C_2Q)$$  \hspace{1cm} (3.4)

Here $C_1$ and $C_2$ are constants and

$$0 < D \leq 1 \ & C_1 + C_2 = 1$$

Metric value
Service discovery metric value is set to be range from 0 to 1. SD seems to be 0 when the service discovery fails.

Interpretation
Higher the value of service discovery metric gives the better Discovery of Services. It indicates that aspects of service discovery are addressed properly.

3.2.3 Service Interpretation Measures
The metrics addressing the aspects like functional data and quality of service data belongs to service interpretation (Level 1 of the taxonomy). The proposed sub-aspects (Level 2 of the taxonomy) metrics that constitute to major aspects metrics of service interpretation has discussed below.

3.2.3.1 Functional Data Measures
To measure functional data, the sub-aspects such as semantic elements and service operation are to be measured. Metrics
addressing functional data are illustrated in figure 7

a. Semantic Elements
Metric Name: checking for Matching Described Semantic Elements (MDSE)
Definition: MDSE is the ratio between matching and mismatching of well described semantic elements.
Metric formulation
Let MDSE be denoted by $M_{se}$ and it is calculated by Equation (3.5).
Let $U_{se}$ be the well described matching semantic elements and $V_{se}$ be the well described mismatching Semantic elements
Let $U_{se} \cdot V_{se}$ be the ratio between matching and mismatching well described semantic elements.

$$M_{se} = \frac{U_{se}}{U_{se} + V_{se}} \& 0 \leq M_{se} \leq 1$$  \hspace{1cm} (3.5)

Metric value
Value defined for MDSE lies from 0 to 1. Closer to 1 indicates the semantic elements are well described
Interpretation
Higher the value of M gives better matches for the Described Semantic elements of Services. It shows that purpose of the services is well described and would help for better invocation.

b. Service Operations
Metric Name: Check for well Defined Service Operations (DSO)
Definition: DSO is the ratio of the number of well defined service operations to total number of service operations along with a number of versions of the service.
Metric formulation
Let DSO be denoted by $G_{so}$ and it is calculated by Equation (3.6). Let $H_{i}$ be the number of well defined service operations with versions
Where, $i = 1, 2… w$
Let $\hat{\lambda}$ be the total number of service operations
Let $V_{j}$ be the number of versions of the service, $j = 1, 2… n$

$$G_{so} = \left( \frac{\sum_{i=1}^{w} H_{i}}{\frac{\lambda}{n} \sum_{j=1}^{V_{j}} V_{j}} \right)$$ \hspace{1cm} (3.6)

where, $0 < G_{so} \leq 1 \& 1 \leq \hat{\lambda} \leq 5$

Metric value
Value defined for DSO lies from 0 to 1. Closer to 1 indicates the service operations are well described.
Interpretation
Higher the value of DSO metric indicates operations of the service are well defined. It depicts that capabilities of service are well defined.

c. Functional Data
Metric Name: Metric for Functional Data (FD)
Definition: Functional data is denoted by FD and it is the ratio of functional relationship between metrics of Checking Matching Described Semantic Elements and Checking for well Defined Service Operations.

Metric formulation
Let FD be denoted by F and it is calculated by Equation (3.7).
Clearly F be the ratio of functional relationship between $M_{se}$ and $G_{so}$
It is defined as

$$F = C_{1} M_{se} + C_{2} G_{so}$$ \hspace{1cm} (3.7)

where $0 < F \leq 1 \& C_{1} + C_{2} = 1$

Metric value
Functional data metric value lies from 0 to 1.
Interpretation
Higher the value of FD metric indicates that purpose and functionalities of service are stated clearly.

3.2.3.2 Quality of Service Data Measures
The QoS attribute measures for each quality are described below. The expected minimum and maximum values for each quality attribute are arrived with the help of the metrics below.

a. Metric for Expected Minimum ($E_{\minQ}$)
Definition: $E_{\minQ}$ is the ratio of minimum value of respective QoS attribute to maximum value of respective QoS attribute
Metric formulation
Let $E_{\minQ}$ be denoted by $E_{1}$ and it is calculated by Equation (3.8).
Let $X$ be the minimum value range of each QoS data and let $Y$ be the maximum value range of each QoS data

$$E_{1} = \frac{X}{Y} \text{ where, } 0 \leq E_{1} \leq 1$$ \hspace{1cm} (3.8)

b. Metric for Expected Maximum ($E_{\maxQ}$)
Definition: $E_{\maxQ}$ is the maximum value for each quality of service data.
Metric formulation
Let $E_{\maxQ}$ be denoted by $E_{2}$ and it is calculated by Equation (3.9).

$$E_{2} = h \text{ where, } h = 1$$ \hspace{1cm} (3.9)

The numerator and denominator value of these metrics are taken from the service registry attributes which is defined in table 6.1. Expected minimum metric values are used only when the particular quality of service data is not available in the service registry.

c. Availability
Metric Name: Availability Metric
Definition: Availability Metric denoted by $R_{\text{avail}}$ is the maximum value between the ratio of agreed availability to maximum availability and expected minimum availability.
Metric formulation
Let Ratio of Availability be denoted by $R_{\text{Avail}}$ and it is calculated by Equation (3.10). Let $M_{\text{Avail}}$ be the maximum
Let \( E_{\text{Avail}} \) be the agreed availability, and let \( E_{\text{Avail}} \) be the expected minimum availability.

\[
R_{\text{Avail}} = \max \left( \frac{m_{\text{Avail}}}{M_{\text{Avail}}}, E_{\text{Avail}} \right) \text{ where, } E_{\text{Avail}} \leq R_{\text{Avail}} \leq 1 \quad (3.10)
\]

(i) If \( m_{\text{Avail}} = 0 \), \( R_{\text{Avail}} = E_{\text{Avail}} \)

When the agreed availability does not exist, then \( R_{\text{Avail}} \) would be the expected minimum value of availability.

(ii) If \( m_{\text{Avail}} \geq E_{\text{Avail}} \)

\[
R_{\text{Avail}} = \frac{m_{\text{Avail}}}{M_{\text{Avail}}} \text{ where } 0 < m_{\text{Avail}} \leq M_{\text{Avail}}
\]

**Metric value**

Value of Availability metric lies from 0 to 1. Closer to 1 indicates the services always exist.

**Interpretation**

Higher the value of this ratio indicates the high availability of services which in turn leads to better interpretation.

**d. Compliance**

**Metric Name:** Compliance Metric

**Definition:** Compliance Metric denoted by \( R_{\text{Comp}} \) is the maximum value between the ratio of agreed compliance to maximum compliance and expected minimum compliance.

**Metric formulation**

Let Ratio of Compliance be denoted by \( R_{\text{Comp}} \) and it is calculated by Equation (3.11)

\[
R_{\text{Comp}} = \max \left( \frac{m_{\text{Comp}}}{M_{\text{Comp}}}, E_{\text{Comp}} \right) \text{ where, } E_{\text{Comp}} \leq R_{\text{Comp}} \leq 1 \quad (3.11)
\]

(i) If \( m_{\text{Comp}} = 0 \), \( R_{\text{Comp}} = E_{\text{Comp}} \)

When the agreed compliance does not exist, then \( R_{\text{Comp}} \) would be the expected minimum value of compliance.

(ii) If \( m_{\text{Comp}} \geq E_{\text{Comp}} \)

\[
R_{\text{Comp}} = \frac{m_{\text{Comp}}}{M_{\text{Comp}}} \text{ where } 0 < m_{\text{Comp}} \leq M_{\text{Comp}}
\]

**Metric value**

Compliance metric value lies from 0 to 1.

**Interpretation**

Higher the value of this ratio indicates the services offers better compliance. This would in turn increase the service invocation.

**e. Response Time**

**Metric Name:** Response Time Metric

**Definition:** Response Time Metric is denoted by \( R_{\text{RT}} \). It is the minimum value between the ratio of agreed response time to maximum response time and expected maximum response time.

**Metric formulation**

Let Ratio of Response Time be denoted by \( R_{\text{RT}} \) and it is calculated by Equation (3.12)

\[
R_{\text{RT}} = \min \left( \frac{m_{\text{RT}}}{M_{\text{RT}}}, E_{\text{RT}} \right) \text{ where, } 0 < R_{\text{RT}} < 1 \quad (3.12)
\]

(i) If \( m_{\text{RT}} = 0 \), \( R_{\text{RT}} = 0 \)

When the agreed response time does not exist, \( R_{\text{RT}} \) be the expected maximum value of response time.

(ii) If \( m_{\text{RT}} < E_{\text{RT}} \)

\[
R_{\text{RT}} = \frac{m_{\text{RT}}}{M_{\text{RT}}} \text{ where } 0 < m_{\text{RT}} < M_{\text{RT}}
\]

**Metric value**

Response time metric lies from 0 to 1.

**Interpretation**

Lower the value of this ratio depicts better response from services. The value is normalized to 1 because all the sub-aspect metric ratios are in maximum value except for these two sub-aspects.

**f. Throughput**

**Metric Name:** Throughput Metric

**Definition:** Throughput Metric denoted by \( R_{\text{TP}} \) is the maximum value between the ratio of agreed throughput to maximum throughput and expected minimum throughput.

**Metric formulation**

Let Ratio of Throughput be denoted by \( R_{\text{TP}} \) and it is calculated by Equation (3.13)

\[
R_{\text{TP}} = \max \left( \frac{m_{\text{TP}}}{M_{\text{TP}}}, E_{\text{TP}} \right) \text{ where, } E_{\text{TP}} \leq R_{\text{TP}} \leq 1 \quad (3.13)
\]

(i) If \( m_{\text{TP}} = 0 \), \( R_{\text{TP}} = E_{\text{TP}} \)

When the agreed throughput does not exist, \( R_{\text{TP}} \) be the expected minimum value of throughput.

(ii) If \( m_{\text{TP}} \geq E_{\text{TP}} \)

\[
R_{\text{TP}} = \frac{m_{\text{TP}}}{M_{\text{TP}}} \text{ where } 0 < m_{\text{TP}} \leq M_{\text{TP}}
\]

**Metric value range**

Value range of throughput metric lies from 0 to 1.

**Interpretation**

Higher the value of this ratio indicates that the services handle more number of requests. This would enhance the usage of service.

**g. Latency**

**Metric Name:** Latency Metric

**Definition:** Latency Metric denoted by \( R_{\text{Lat}} \) is the minimum value between the ratio of agreed delay to maximum delay and expected maximum delay.
Metric formulation
Let Ratio of Latency be denoted by $R_{Lat}$ and it is calculated by Equation (3.14)
$$R_{Lat} = \min \left( \frac{m_{delay}}{M_{delay}} \right) \cdot E_{Lat} \quad \text{where, } 0 < R_{Lat} < 1 \quad (3.14)$$

(i) If $m_{delay} = 0$,
$R_{Lat} = 0$
When the agreed latency is not exist $R_{RT}$ be the expected maximum value of latency.
(ii) If $m_{delay} < E_{2Lat}$
$$R_{Lat} = \frac{m_{delay}}{M_{delay}}, \quad \text{where, } 0 < m_{delay} < M_{delay}$$

Metric value
Value of latency metric lies from 0 to 1. Closer to 0 indicates that delay is less.

Interpretation
Lower the value of this ratio indicates that the services offer less delay in processing requests. The value of latency are normalized to 1

h. Documents

Metric Name: Metric for Documents
Definition: Metric for Documents denoted by $R_{Doc}$ is the maximum value between the ratios of agreed documents to maximum documents supplied and expected Maximum doc.

Metric formulation
Let Ratio of Doc be denoted by $R_{Doc}$ and it is calculated by Equation (3.15) Let $M_{Doc}$ be the maximum documents supplied and $m_{Doc}$ be agreed documents, and Let $E_{1Doc}$ be the Expected Documents supplied
$$R_{Doc} = \max \left( \frac{m_{Doc}}{M_{Doc}} \right) \cdot E_{1Doc} \quad \text{where, } E_{1Doc} \leq R_{Doc} \leq 1 \quad (3.15)$$

(i) If $m_{Doc} = 0$,
$R_{Doc} = E_{1Doc}$
When the agreed documents related to service does not available, $R_{Doc}$ would be the expected minimum value of documents related to service.
(ii) If $m_{Doc} \geq E_{1Doc}$
$$R_{Doc} = \frac{m_{Doc}}{M_{Doc}}, \quad \text{where, } 0 < m_{Doc} \leq M_{Doc}$$

Metric value
Value defined documents metric lies from 0 to 1. Closer to 1 indicates that more documents related to service are supplied.

Interpretation
Higher the value of this ratio indicates the services offer more documents related to service which would enable better interpretation of service

i. Reliable Messaging

Metric Name: Reliable Messaging Metric
Definition: Metric for Reliable Messaging denoted by $R_{RM}$ is the maximum value between the ratios of agreed number of error messages handled to maximum number of error message handled and Expected number of error messages handled.

Metric formulation
Let Ratio of Reliable Messaging be denoted by $R_{RM}$ and it is calculated by Equation (3.16) Let $M_{RM}$ be the maximum number of error message handled and $m_{RM}$ be agreed number of error message handled, and Let $E_{1RM}$ be the Expected number of error messages handled
$$R_{RM} = \max \left( \frac{m_{RM}}{M_{RM}} \right) \cdot E_{1RM} \quad \text{where, } E_{1RM} \leq R_{RM} \leq 1 \quad (3.16)$$

(i) If $m_{RM} = 0$,
$R_{RM} = E_{1RM}$
When the agreed reliable messaging to service is not available $R_{RM}$ would be the expected minimum value of reliable messaging.
(ii) If $m_{RM} \geq E_{1RM}$
$$R_{RM} = \frac{m_{RM}}{M_{RM}}, \quad \text{where, } 0 < m_{RM} \leq M_{RM}$$

Metric value
Value of reliable messaging metric is lies from 0 to 1.

Interpretation
Higher the value of this ratio indicates that the services can handle more Error messages.

j. Best Practices

Metric Name: Metric for Best practices
Definition: Metric for Best Practices denoted by $R_{BP}$ is the maximum value between the ratios of agreed best practice to maximum best practice and expected minimum best practices.

Metric formulation
Let Ratio of Best Practices be denoted by $R_{BP}$ and it is calculated by Equation (3.17) Let $M_{BP}$ be the maximum best practices and $m_{BP}$ be agreed best practices, and Let $E_{1BP}$ be the Expected minimum best practices
$$R_{BP} = \max \left( \frac{m_{BP}}{M_{BP}} \right) \cdot E_{1BP} \quad \text{where, } E_{1BP} \leq R_{BP} \leq 1 \quad (3.17)$$

(i) If $m_{BP} = 0$,
$R_{BP} = E_{1BP}$
When the agreed practices does not exist $R_{BP}$ be the expected minimum value of best practices.
(ii) If $m_{BP} \geq E_{1BP}$
$$R_{BP} = \frac{m_{BP}}{M_{BP}}, \quad \text{where, } 0 < m_{BP} \leq M_{BP}$$

Metric value
Value of best practices metric lies from 0 to 1.

Interpretation
Higher the value of this ratio shows the services have adopted good practices which in turn improves the usage service

k. Quality of Service Data

Metric Name: Quality of Service Data Measure (QDM)
Definition: QDM is the ratio of value of each quality of service data to number of quality of service data.

Metric formulation
Let QDM be denoted by $Z$ and it is calculated by Equation (3.18). Let $n$ be number of Quality of Service Data and Let $Q_i$ be the values of each quality of service data obtained by using proposed metrics, where, $1 \leq i \leq n$

$$Z = \frac{\sum_{i=1}^{n} Q_i}{n}, \text{ where } 0 < z \leq 1$$ (3.18)

Metric value range
Value range of QDM lies from 0 to 1.

Interpretation
Higher the value of this metrics shows more quality of service data has been supplied.

3.2.5 Service Interpretation
Metric Name: Interpretation of Service Metrics (IoS)
Definition: IoS is the ratio of functional relationship between functional data and quality of service data measures.

Metric formulation
Let interpretability of Service be denoted by $I$ and it is calculated by Equation (3.19)

$$I = (C_1 \cdot F) + (C_2 \cdot Z)$$ (3.19)

where, $0 < 1 \leq 1 \& C_1 + C_2 = 1$

Where $C_1$ and $C_2$ are constants.

Metric value
Value of IoS is set from 0 to 1. Closer to 1 indicates that better interpretation of service.

Interpretation
Higher the value of these metrics shows functional and non-functional aspects are defined clearly.

3.2.6 Service Discoverability
Definition: Discoverability is measured by using its constituents and the corresponding aspects. Discoverability of service be denoted by DM and is the ratio of functional relationship between service discovery and interpretation of service.

Metric formulation
Let Discoverability of Service be denoted by DM and it is calculated by Equation (3.20)

$$DM = (C_1 \cdot D) + (C_2 \cdot I)$$ (3.20)

where, $0 < DM \leq 1 \& C_1 + C_2 = 1$

Where $C_1$ and $C_2$ are constants.

Metric value
Value of discoverability metric is from 0 to 1. Closer to 1 indicates that service discoverability would be better.

Interpretation
Higher Discoverability value gives better discovery of services. It shows that the discoverability constituents, the aspects and its sub-aspects have been defined completely.

3.2.7 Discoverability Metrics - Tree Structure

![Discoverability Metric Tree Structure](image)

This discoverability metric tree (figure 3) depicts how the metrics for sub-aspects, aspects, constituents and its quality attributes are arrived using bottom up approach.

4. RESEARCH EXPERIENCE

4.1 Experimentation
The main objective of this experimentation is to demonstrate the usability or practical applicability of the proposed metric suite of service discoverability. For designing and conducting the experiment, service registry and the corresponding information were designed based on the inspiration from the various research works from [8] [10] [14] [20] [21] [22] [23] [24]. The attributes chosen are based on the review of various works and the values for each attribute are defined based on the references [20] [21] [22] [24] [25] [34] [44] [45]. The chosen attributes describe the complete information on each registered service. Here the attributes are differentiated based on functional and quality of service data are shown in table 2. Based on those attributes three different service registries were designed and implemented. The service registry SR1 is designed with minimum or basic fields and SR2 with additional fields inclusive of SR1 and SR3 is the complete set which consists of all the fields. For our experimental purpose, the designed registries are with three different ranges of data (i.e. registry with 1000, 2000 and 3000 entries). From the related works, the complete list of attributes and the value ranges (i.e. from minimum to maximum for each attribute) were derived from the service registry.

The naming of the registry has been based on attributes chosen for the registry i.e. minimum set of attributes, next level or medium set of attributes and full set of attributes. The registries are named as SR1, SR2 and SR3.
• SR1 - Service Registry 1 is the basic registry which contains limited number of attributes with Service Name, Category, Service ID, Service Operation, Availability and Compliance.

• SR2 – Service Registry 2 extended version which contains additional attributes when compared to SR1. The additional attributes are Interface Name, Response Time and Throughput

• SR3 – Service Registry 3, registry which contains complete attributes Service Registry Attributes

In order to conduct the experiment, the queries were generated. The first query has been designed with the basic attribute. Further the queries have been designed with more number of attributes in addition to the basic attribute. For example, the first query selects service name where attribute i (condition operator) the next query selects service name where attribute i (condition operator) and attribute j (condition operator). The queries are designed as follows:

query no : 1 category filter
query no : 2 query no :1’s attribute and compliance filter
query no : 3 query no :2’s attributes and availability filter

In similar fashion, remaining queries are designed. In total, 11 queries are designed and it is applied on the various registries depending on the attribute that exists in each service registry. The experimentation is done on specific category (banking financial service and insurance services). Requirements for payment service were posed in the query and three services namely BFSI service 13, BFSI service 32, and BFSI service 91 were retrieved. These services provide the same payment functionality but each has been published through different service providers. Values obtained during each stage of experimentation are used by metrics aspects and sub-aspects of discoverability constituents. Metric results specific to constituents of discoverability such as service discovery and interpretability are presented.

4.2 Results and Discussion

4.2.1 Service Discovery

Output of the proposed metric suite has been represented in similar way as that of discoverability taxonomy. Each metric has been discussed with three service registries of various sample size (1000, 2000 and 3000 service entries). In BFSI services there exists innumerable services like authentication service, credit based services, payment services, cheque based services, loan services etc. For simplicity, it has been generally addressed as BFSI services and among those, services of the same functionality have been considered. For the discussion purpose only 3 services are taken which are BFSI service 13, BFSI service 32 and BFSI service 91. Service discovery (SD) has been measured by using the values of CSC, RMD, and EP metrics. The table 2 shows the discovery metrics values of registries with different sample size.

<table>
<thead>
<tr>
<th>Data Sets</th>
<th>Services</th>
<th>Service Discovery (SD) values</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000</td>
<td>BFSI Service 13</td>
<td>0.39</td>
</tr>
<tr>
<td></td>
<td>BFSI Service 32</td>
<td>0.41</td>
</tr>
<tr>
<td></td>
<td>BFSI Service 91</td>
<td>0.12</td>
</tr>
<tr>
<td>2000</td>
<td>BFSI Service 13</td>
<td>0.20</td>
</tr>
<tr>
<td></td>
<td>BFSI Service 32</td>
<td>0.42</td>
</tr>
<tr>
<td></td>
<td>BFSI Service 91</td>
<td>0.18</td>
</tr>
<tr>
<td>3000</td>
<td>BFSI Service 13</td>
<td>0.30</td>
</tr>
<tr>
<td></td>
<td>BFSI Service 32</td>
<td>0.41</td>
</tr>
<tr>
<td></td>
<td>BFSI Service 91</td>
<td>0.17</td>
</tr>
</tbody>
</table>

As shown in figure 4, it could be seen that the service discovery values of each service are varying. In case of BFSI service 13 of 1000 entries, SR2 shows 3% increase while compared with SR1 and SR3 shows 5% increase which indicates that service discovery is better in SR3 as it contains the exhaustive service attributes. It indicates that any service registry that has higher existence of service contract, more metadata information and exact positioning of services leads to better discovery of suitable services.

4.2.2 Interpretation of Service (IoS)

Table 3 depicts interpretation of service metric value. As mentioned in taxonomy of metric suite, interpretation deals with functional and quality of service data about service. FD
and QDM values are used to compute service interpretation. The graph shown in figure 5 illustrates that the interaction of services has been more in SR3 because it provides more functional relevance and quality information about services which are published in the registry.

While considering the BFSI service 13, the IoS value for three registries is 0.32, 0.42 and 0.64 respectively. It can be noticed that IoS value in SR3 is 22% more than SR2 and is 32% more than SR1 which has been due to inclusion of more functional and non-functional aspects of service. This leads to better interpretation in SR3 when compared to other registries. Interpretations of services are more when the services explore its functional and non-functional aspects clearly which has been observed from the experimentation.

Table 3: Interpretability Measure Values of Three Registries with Different Sample Size of Various Services

<table>
<thead>
<tr>
<th>Data Sets</th>
<th>Services</th>
<th>IoS values</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>SR1</td>
</tr>
<tr>
<td>1000</td>
<td>BFSI Service 13</td>
<td>0.32</td>
</tr>
<tr>
<td></td>
<td>BFSI Service 32</td>
<td>0.28</td>
</tr>
<tr>
<td></td>
<td>BFSI Service 91</td>
<td>0.32</td>
</tr>
<tr>
<td>2000</td>
<td>BFSI Service 13</td>
<td>0.32</td>
</tr>
<tr>
<td></td>
<td>BFSI Service 32</td>
<td>0.28</td>
</tr>
<tr>
<td></td>
<td>BFSI Service 91</td>
<td>0.32</td>
</tr>
<tr>
<td>3000</td>
<td>BFSI Service 13</td>
<td>0.32</td>
</tr>
<tr>
<td></td>
<td>BFSI Service 32</td>
<td>0.28</td>
</tr>
<tr>
<td></td>
<td>BFSI Service 91</td>
<td>0.32</td>
</tr>
</tbody>
</table>

Table 4: Discoverability (DM) Metric Value of Different Registries for Various BFSI Services

<table>
<thead>
<tr>
<th>Data Sets</th>
<th>Services</th>
<th>Discoverability values</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>SR1</td>
</tr>
<tr>
<td>1000</td>
<td>BFSI Service 13</td>
<td>0.36</td>
</tr>
<tr>
<td></td>
<td>BFSI Service 32</td>
<td>0.34</td>
</tr>
<tr>
<td></td>
<td>BFSI Service 91</td>
<td>0.22</td>
</tr>
<tr>
<td>2000</td>
<td>BFSI Service 13</td>
<td>0.26</td>
</tr>
<tr>
<td></td>
<td>BFSI Service 32</td>
<td>0.35</td>
</tr>
<tr>
<td></td>
<td>BFSI Service 91</td>
<td>0.25</td>
</tr>
<tr>
<td>3000</td>
<td>BFSI Service 13</td>
<td>0.31</td>
</tr>
<tr>
<td></td>
<td>BFSI Service 32</td>
<td>0.34</td>
</tr>
<tr>
<td></td>
<td>BFSI Service 91</td>
<td>0.25</td>
</tr>
</tbody>
</table>

Figure 5: Interpretation (IoS) Metric Values of Different Registries for Various BFSI Services

4.2.3 Service Discoverability

Quantification of discoverability uses the values of its constituents metric (SD and IoS). The table 4 depicts the outcome of the discoverability metric for experiment 1 with various BFSI services.

Figure 6: Discoverability (DM) Metric Values of BFSI Services with Different Registries

Result of discoverability metric (DM) values of each service in figure 6 shows huge variation. In the case BFSI service 13 of sample size 1000, in SR2 the DM value shows an increase of 10% when compared with SR1. Similarly in SR3 DM value shows an increase of 23% when compared with SR2. This variation in the value of DM metric is due to the impact on
providing relevant information about the service. SR3 offer high discoverability as depicted with a 23% increase because the information relevant to services is represented with the help of all its relevant attributes. Enhancing service discoverability would lead to an increase in usability of services. Quality attribute service discoverability has been quantified with measures of service discovery and interpretability constituents.

With help of proposed metric suite for service discoverability,

- Impact of each aspect in service discoverability has been quantified
- Effect on service discoverability with increase in number of attributes in service registry

5. CONCLUSION

Measures on service discoverability have greater significance in finding the exact services that fulfill the consumer requirements in order to promote the reuse of services. Research efforts are put in this direction and findings are summarized. Literature review was carried out on service discoverability and the need for its quantification. Service discoverability aspects were identified. The identified aspects were represented using our proposed taxonomy. A metric suite was proposed to measure service discoverability. This proposed metrics were applied in experiment. A detailed discussion on results reveals the significance of the metric suite. The future work concerned with how this quantification would enhance the other quality attributes such as service composability and service reusability.

REFERENCES


[16] Benjamin Kanagwa and Agnes F. N. Lumaala ,”Discovery of Services Based on WSDL Tag Level Combination of Distance Measures”, International Journal of Computing and ICT Research, Vol. 5, Special Issue, December 2011


[23] [http://www.infoq.com/articles/rest-discovery-dns - Retrieved on February 2012]


AUTHOR PROFILES

G. Shanmugasundaram obtained his B.Tech in Information Technology (2005) from BCET, Pondicherry University. He received his M.Tech in Computer Science and Engineering (2008) from SMVEC, Pondicherry University. Currently he is pursuing PhD in Department of Banking Technology, Pondicherry University. He is having 2 years of teaching experience and 1 year in software development. His research area includes Service Oriented Architecture and Web Technologies.

Dr. V. Prasanna Venkatesan is currently an Associate Professor, Department of Banking Technology, Pondicherry University. He earned his B.Sc in Physics (1986) from Arignar Anna Arts College, Karaikal. He received his M.C.A (1989) from Pondicherry Engineering College, M.Tech in Computer Science & Engineering (1995) from Pondicherry University and Ph.D in Computer Science & Engineering (2008) from Pondicherry University. He is having more than 20 years of teaching experience. He has published 3 books and papers in national and international journals/conferences. His research area includes Software Architecture, Banking Technology, Object Oriented Modeling and Design, Smart Banking.