Journal of Theoretical and Applied Information Technology
31st January 2016. Vol.83. No.3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

AUTOMATED SERVICE IDENTIFICATION FRAMEWORK (ASIF)

[1]ALI TAEI ZADEH, [2] MURIATI MUKHTAR, [3] SHAHNORBANUN SAHRAN, [4] ZAHRA LOTFI

[1,4] Dr., Department of Industrial Computing, UKM, Kuala lumpur, Malaysia

[2, 3] Assoc. Prof., Department of Industrial Computing, UKM, Kuala lampur Malaysia

E-mail: [1]alitaee@gmail.com, [2,3]mm,shah@ftsm.ukm.my, [4] lotfinasri@yahoo.com

ABSTRACT

Service oriented architecture (SOA) has been used to achieve business-IT alignment in industries for recent years. Service identification represents the first phase in service modeling, a necessary step in SOA. Reviewing existing service identification methods (SIMs) reveals the lack of business alignment, automation, and quality of services as challenging issues. Hence, the main aim of this research is to propose a service identification framework (ASIF) that results in the automated identification of business aligned services which satisfy service quality factors such as cohesion, coupling and reusability. ASIF relies on business aligned entities based on business processes and business goals as unique sources of inputs to guarantee delivery of business aligned services. The proposed ASIF will serve as a comprehensive guideline for service identification that spans from scope determination to refinement of service candidates. In order to achieve this goal, it is necessary to firstly identify and select suitable SIM input types and secondly, to determine appropriate techniques to extract services automatically. Subsequently, the proposed framework is then validated and finally an ASIF tool to support automation is developed.
Keywords: SOA, service identification, AHP, Business-IT alignment, automated tool, BPMN, KAOS.

1. INTRODUCTION

Agility supports the issue of business-IT alignment by backing the change frequency of business processes. Agility is the product of service orientated architecture (SOA) flexibility. SOA provides a flexible set of applications that can be changed with less cost and efforts in comparison with previous software methodologies. SOA has gained considerable attention in solving today's business challenges including agility and flexibility related issues as well as addressing well-known software challenges such as lack of integration and reusability. Two different aspects of SOA including business and software are discussed. The software aspect only concentrates on using pre-existing technology in relation to services, while the business aspect focuses on achieving services which are business-aligned to support business-IT alignment. However, implementation difficulties such as lack of comprehensive guidelines and technical skill requirement for SOA implementation undermine realization of its benefits [1].

Service identification is an important activity in the service modeling lifecycle and should be supported by comprehensive methodologies in order to achieve an appropriate set of services [2]. It has attracted many authors with academic and industrial backgrounds in recent years [3, 4]. Service identification throws light on building blocks of SOA which are so critical and any fault in this phase can lead to defects in all SOA layers. Therefore, faults in basic phases of such architecture are costly and difficult to correct which indicates the importance of service identification methods (SIMs) [5, 6]. In addition, a method which can systematically introduce an appropriate set of services, rooted in business entities, is still missed [1, 2].

According to different views regarding SOA, the existing SIMs differ significantly in terms of delivery strategy, business-oriented (top-down) or technical-oriented (bottom-up) [7]. This difference is related to the origins of service portfolio that varies from business domain to technical domain. In fact, the reason for selecting different service

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

identification configurations such as input types and service identification techniques by SIMs is whether they adopt technical-oriented or business-oriented perspective. Services in the business-oriented methods correspond to business domain entities including business processes (as-is) and business goals (to-be), while technical-oriented SIMs merely rely on applications' assets as their inputs such as database, interface or other technical entities. It is critical to take into consideration as-is and to-be situations. Likewise, Frey et al. [8] emphasize that "typical challenges of identifying the right services, which are aligned with current and future business needs several years ahead." Nevertheless, most service orientation researches have focused on technical domain instead of business domain. Establishing a link between business entities such as business processes and web services will guarantee business-IT alignment [9].

Hence, business-oriented methods are more effective, but due to descriptive nature of business domain entities, automization of these methods is not fully developed and consequently they are underutilized [10]. Automation is a challenging subject in existing SIMs due to lack of an automated business-oriented method. However, among existing SIMs few of them aim to propose automated method. The reason behind lack of automated techniques is the complexity and fuzzy nature of top-down SIMs. Besides, the majority of automation techniques in SIMs have been proposed within software-oriented (bottom-up) SIMs which essentially rely on legacy systems' transformation. Furthermore, business entities such as business goals and business processes have descriptive nature, which make their processing costly and time consuming. Therefore, they are needed to be quantified to be involved in automating SIMs. Also, there is no comprehensive framework encompassing all the steps of service identification with detailed guidelines, and the existing SIMs have not been validated via case studies. In addition, the aim of service identification is to come up with qualified services based on service quality factors such as reusability, cohesion, coupling, and granularity [3].

The objective of the current study is to propose a comprehensive framework for service identification that embeds techniques to identify services automatically. Therefore, this paper addresses the mentioned objectives by providing i) a comprehensive SIM according to techniques and phases of existing SIMs, namely scope determination, service identification (SI) input, SI techniques, SIM issues and service refinement which should be covered in ii) quantification of service identification metrics iii) automating as much as possible of service identification activities from inputs' preparation to service refinement iv) evaluating applicability of ASIF via detailed case study.

The rest of this paper is outlined as follows. Section 2 presents some related work focusing on issues of service identification methods; then, the results and discussion including the quantification of the service quality factors (reusability, cohesion and coupling), clustering algorithm and tool development are discussed in Section 3. Section 4 deliberates the conclusion and closing remarks of the study.

2. LITERATURE REVIEW

Service orientated architecture (SOA) has gained considerable attention in solving today's business challenges including agility and flexibility as well as addressing well-known software challenges such as integration and reusability. Service identification throws light on building blocks of SOA which are so critical and any fault in this phase can lead to defects in all SOA layers. Services that correspond to meaningful business entities, tend to use domain decomposition and analysis from enterprise's business side to reach the services which have been categorized as top-down SIMs. However, implementation difficulties such as lack of comprehensive guidelines and technical skills requirement for SOA implementation undermine realization of its benefits [1]. On the other hand, bottom-up SIMs are extracted based on information collected from source code or other specifications of information systems such as database, interface or other technical entities. Hence, bottom-up methods try to extract services from software orientation. In addition, some of SIMs use both top-down and bottom-up approaches which is named meet-in-middle [11].

To propose a comprehensive SIM, all necessary activities and techniques of service identification that reveal the service portfolio should be considered. Therefore, the required steps in service identification were investigated based on the necessity and importance stated and emphasized by existing SIMs. The identified steps are scope determination, service identification technique, and

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

service refinement. In addition, the input type selection has been highlighted as important decision that affects the selected technique as well as the satisfaction degree of service quality factors [2]. Furthermore, the business-IT alignment is linked to input types such as business processes and business goals representing most important business domain assets [12]. Therefore, the top-down SIMs which aim to achieve business-IT alignment highly depends on business processes or business goals as their input types. Few SIMs have considered both business processes and business goals in their methods with details of their involvement in their proposed method [13-15]. The GSM method for goal modeling is proposed by [13]. However, it has not provided a measurable way to assess the alignment degree of each identified service with business goals as well as it lacks guidelines or examples to increase the applicability. Likewise, Fareghzadeh [14] that stresses on business goals and business processes has not presented how to satisfy the alignment of business processes and goals with identified services. Therefore, these two input types highly affect the business-IT alignment and consideration of them is a necessity in a top-down method.

Moreover, there is a clear emphasis on automation of SIMs to decrease required expertise as well as the human involvement [16]. Besides, in recent years, the top-down SIMs have increasing trends towards automated techniques. However, the automation including input type's transformation activities, these activities in top-down SIMs remains manual although the standards and great volume of the input types imposes much cost. Utilizing automated tools will result in decrease in the complexity and implementation costs of service modeling [16]. In addition, effective implementation of a SIM depends on the involvement of automated tools [17]. The most automation is realized in bottom-up methods since their input types can be executed and taken effortlessly. However, regarding the importance of business processes and necessity of involving them in the business-aligned SIM, their automated transformation to services through top-down SIMs has attracted recent researchers [3, 7, 10, 18].

In the case of service quality factors, majority of SIMs state their critical roles in service identification; nevertheless, a minority of them have quantified their calculation. Likewise, some of SIMs that provide quantitative methods to assess the quality of their services present automated tools. In order to realize a top-down SIM, it is important to calculate the service quality factors based on business-aligned input types such as business process models, while existing SIMs calculate the cohesion and coupling according to technical oriented input types or atomic elements of business aligned input types such as CRUD matrixes extracted from high abstract input types. However, the atomic operations such as read, write, update and delete did not represent the relationship aggregations in high level input types. Therefore, some methods depend on descriptive reasoning of quality factor satisfaction, while other SIMs formulate the situation of those factors to achieve clear numerical-based values for each factor as their weights [10, 13]. The evaluation of service quality factors based on formal methods is reasonable because they provide numerical measurement of the situation of each quality factor[10, 19]. The possibility of utilizing numerical metrics of quality factors depends on the techniques selected in a SIM. As discussed, due to descriptive specification of input types in top-down methods, quantification of service identification process as well as measurement of service quality factors have not been fully developed yet.

The importance of service refinement phase is emphasized; however, methods that evaluate the business-IT alignment in this phase are rarely mentioned. Van Nuffel refines the candidate services on the basis of the cohesion and coupling situation of each candidate service [20]. Yousef et al. [21] adopt the QoSOnt ontology and Arsanjani et al. [13] apply the refinement step based on Litmus Test [13]. They use Litmus Test; however, there is no details about litmus Test questions nor automated method for addressing the Litmus Test questions. Although there is stress on the importance of service refinement but there is no SIM that refer to automated method for service refinement. Also, existing SIMs do not provide detailed guidelines nor validation such as case studies or examples.

3. RESULT AND DISCUSSION

The contribution of this paper is an automated service identification framework (ASIF) which contains procedure of service identification is discussed at length. ASIF methodology consists of five distinguished phases, namely 'scope determination', 'goal modelling and business process modelling', 'weighting', 'clustering' and 'service refinement' (Figure 1). Whenever there is a need for identification of a service list or an

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

existing service list needs to be updated, ASIF, as a rotational framework, can be applied.

ASIF has tried to prepare clear and detailed guidelines in a form of recursive steps for identifying services as basic building blocks of SOA. Over the time, the intermediate goals change due to new customers' requirements or new conditions, consequently the related business processes reveal that goal should change as well.

3.1 Preparation phases

The first phase is scope determination. One of the common causes of failure in projects is an undetermined scope that increases the ambiguity and decreases the project transparency [22]. The subject of scope determination has ties with comparing the priority of the scope of different candidates based on the multiple criteria. Sikdar and Das propose some factors that influence scope determination in SOA such as organization structure, competency, technology, resources and business processes of the firm or its stakeholders [23].

AHP was proposed in 1980 in order to decompose the decision problem into a multi-criteria hierarchical process that utilize pair-wise comparison to weight according to relative comparisons scores [24]. Thus, AHP is widely used in evaluating situations where direct establishment of weights and ratings is difficult. According to AHP the major business entities should be placed as measures to be ranked by stakeholders. The effective factors those could be used as criteria within prioritization process based on literature are enterprise's goals, risks, costs, key business requirements, alternative modules, solution constraints and prioritized business processes [1, 25]. According to AHP technique, the priority and importance of each criterion in comparison with other criteria should be calculated by, pair wise comparisons between them. After this step the candidates' scopes can be ordered by their importance based on set of criteria.

The second phase deals with as-is and to-be information elicitation. The as-is situation is achieved via BPMN or UML AD, which are standard business process models to show the current business activities. In addition, the to-be situation is indicated based on KAOS [26] goal model. Goal orientation in SOA refers to the consideration of the business goals in SOA especially during the service identification process. Each goal is divided according to its scale into coarse-grained and fine-grained goals [27]. Coarse-grained goals can be refined to fine-grained ones and ultimately a set of them can satisfy coarse-grained goals. Thus, goal refinement process offers traceability in coarse-grained and fine-grained goals or from actionable entities to coarse-grained goals and vice versa [28].

The third phase which copes with weighting quality factors quantifies the business goals affinity, reusability, cohesion, and coupling to provide the prerequisites of identifying qualified services. The list of validated business processes which each includes sets of tasks is placed in the bottom of the goal model representing the operational layer. Goal-graph leafs, as a result of goals' decomposition process, often are fine grained. Then, by assessing the operationalization of goal-graph leafs based on HOW questions, each one of those business processes can be related to one or more goals of the KAOS goal model. Therefore, by presenting both non-functional requirements and operational layers in integrated model and then relating each goal to one or more related task (and vice versa) the graph formation step will be complete. The mapping between the identified goals which have been validated on one side and the real and updated list of business processes on the other side will display the importance of each business process based on its role in satisfying business goals. The goal-decomposition process which relates high level goals to operational business processes represents justifications and reasons of performing each process and also the value that each business process adds to the enterprise.

Basically, we consider the degree of goals and business processes relation as a way of measuring each process which is represented by goal factor (GF). GF of each operational entity is calculated by adding up the number of edges between each process and goals or requirements. The number of relations of process i is ($0\leq j \leq n$) where j is the relation number of process i that is 0 when a process does not support any goals and n when it supports and is related to n goals. In addition, GF should distinguish between AND and OR relation types, where AND relation between a goal and a process means that it is mandatory to perform that process to achieve the goal, while in an OR relation case, satisfying any business process among the ones related to the goal is enough. Hence, a sub goal that is related to the parent goal via AND relation partially contributes to the parent goal since without its contribution the parent goal cannot be satisfied. In addition, if the sub goal is related via

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

OR relation, it will contribute to the parent; however, if it is denied then the parent goal can be satisfied by another sub goal. Obviously, the importance of AND relations are more than OR relations which indicate the importance role of the sub goal or operation which participate in the AND relation to satisfy a parent goal. Thus, in order to represent the superior position of AND relations, we sum up the AND relations of each process and multiply it by two. It indicates the importance of AND in comparison with OR in the goal model.

Clearly, the AND edges of the entity i to entitiy j is called AND-edges and is represented by (α_{ij}) and the OR-edges of the entity i to entity j is represented by (β_{ij}). We assign value 1 as the weight to each OR-edge (β_i), and value 2 to each AND-edge (α_i). Obviously, if a requirement or goal has only one relation with lower levels in the goal graph, it is a compulsory relation and has the same value with AND relations. Therefore, the equation for weight calculation of each task is proposed as:

Goal Factor of the process i

$$(GF_i) = \sum_{j=1}^{n} (2 * \alpha_{ij} + \beta_{ij}) \qquad (1)$$

The j variable represents other goal model entities which have directly or indirectly relation with entity i. The GF illustrates the effectiveness of each process in meeting the business goals. Assigning relation weight that is a critical concept in analysing goal affinity is included in GF calculation; for example, in goal graph each relation has been graphically weighted and represented in Figure 2. This example is composed of four goals in two levels, four requirements and four processes in operational layer.

According to Figure 2, to calculate the GF for the 'survey and analysis service problem' and 'assigning IP address to all application devices' as business processes, the number and type of the corresponding relations should be identified for those tasks. For example, GF calculation of Survey and analysis service problem: (GF)= (2*3+0)= 6. The above integrated model which weighting each business process based on its support degree of business goals will be used in service refinement.

In addition, creating automated tool depends on availability of numerical data. Extracting the concepts inside the business process models and transforming them to quantitative domain are identified as milestones of automatic service identification. Therefore, descriptive BPMN diagrams should be quantified by converting the relations between tasks of the business processes to numbers in a matrix format. The matrix of relationships between tasks is a squared matrix that represents the relations between tasks of one or more business processes. When a task does not have any relation to other tasks, the value is zero and when one or more relations exist, the cell value is one.

$$\forall\, Task_i,\ Task_j \in Task\ Set\ :\ if\ Task_i \rightarrow Task_j \Rightarrow Tasks\ Relation\ Matrix[i,j] = 1 \qquad (1)$$

Figure 3 illustrates a tasks-relation matrix which includes tasks and displays their relations by placing '1' in related cell.

The matrix has a vital role because it demonstrates a) The role of each task in the business process b) Calling frequency of each task by others (Reusability) c) Degree of relativity and dependency among a set of tasks (Cohesion) d) Degree of relativity and dependency between two set of tasks (Coupling).

Accordingly, the service quality factors namely, reusability, coupling and cohesion should be quantified. Since the aim of ASIF is to achieve business-aligned services, the service quality factors need to be extracted from the tasks relation matrix. Quantifying the value of each quality factor paves the ground for clear measurement and also the result of these quantified quality factors can be an indicator of SIM's quality.

Reusability of a task indicates the dependency of other tasks on it, which in turn will be prioritized compared to others. To calculate the reusability value of each task in the tasks relation matrix, relations in BPMN diagrams that have ended in a task such as $TASK_i$ should be counted to show its reusability weight. Thus, if we assume 'x' as the value of a cell in the relation matrix, it is defined as:

Sum of calling frequency= x, $x \geq 0$ (2)

To calculate the reusability value of each cell, each relation from $TASK_i$ to $TASK_j$ will increase the cell value of $TASK_j$ in the matrix by one.

Based on tasks relations matrix, to calculate the cohesion value of a cluster of tasks, the cohesion can be quantified by calculating the ratio of relations inside a cluster divided by maximum possible relations inside a cluster – without considering the recursive calls- that is

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

$(n_i * (n_i - 1))$, where n_i indicates the number of nodes in cluster i.

The maximum cohesion value in a cluster is 1, when the graph of tasks and their relations are a complete graph, and it is 0, when there are no relations and dependencies between cluster nodes that means there is no consistency between tasks in the cluster. We customized the Bunch cohesion C_i in cluster i which consists of n_i components in which the number of relations between tasks inside the cluster i is β_i as:

Cohesion $$C_i = \frac{\beta_i}{n_i * (n_i - 1)} \quad (3)$$

Tightly-coupled sets of tasks are a reason for inappropriate clustering of the tasks which need to call other tasks in other clusters frequently. Thus, the objective is to achieve a cluster set in which each cluster is independent. To this end, dependency of a task on other clusters should be minimized to increase the flexibility of each service, and as a result any change in tasks of a cluster will not force change in other clusters. We define the coupling A_{ij} between cluster i and cluster j where the number of nodes in cluster i is n_i and n_j is the number of nodes in cluster j and also α_{ij} is number of relations between cluster i and cluster j as:

Coupling

$$A_{ij} = \begin{cases} 0, & \text{If } i = j \\ \frac{\alpha_{ij}}{2 n_i n_j} & \text{If } i \neq j \quad n_i, n_j \neq 0 \end{cases} \quad (4)$$

3.2 Service Identification

The current challenging issue is how to find best composition of tasks in order to provide a service list. Clustering algorithms are considered as the best tool to find the best composition of a set of elements. In addition, clustering algorithms rely on objective function which has the responsibility to calculate the factors of each composition of elements. Clustering of tasks into high-cohesive and low-coupled services are decisive parameters in proposing sets of tasks as services. Hence, the weights of tasks will be updated based on cohesion and coupling measures. Then, the objective function as the engine of clustering algorithm is proposed to make a balance between various factors. Next, the clustering algorithm is applied to the weighted matrix according to the objective function that guarantees the best values of the objective function in output candidate services. The Bunch clustering algorithm [29] is selected to be applied as a basis for clustering. Bunch is originally proposed for clustering software modules, due to specific situation of the software domain, it is necessary to reconfigure the algorithm. The goal of clustering step is to categorize the tasks in sets that result in satisfaction of quality factors.

The core of the clustering algorithm is adjusting and applying the objective function that tries to meet all qualification requirements. We name the value of objective function as Service quality (SQ) that represents the status of the quality factors of clusters. Objective function factors may have conflicts with each other; hence, the main goal of the clustering is finding a trade-off between those factors. Regarding the clustering process, the main activity is to find compositions of tasks that gain the maximum value of the objective function among other possible compositions of tasks. This is why objective functions are considered as the main characteristics of clustering algorithms.

SQ value quantifies the quality of the service candidates based on mentioned quality factors. The best situation of a cluster set is when tasks inside each service have maximum intra-relations and minimum relations. This can satisfy the reusability factor by composing a reusable task with its related tasks in a cluster to maximize its intra-relations in the service, and also minimize its inter-relations. The additional benefit of such a service organizing method is managing the appropriate service granularity through which the size of each service is reasonably determined. Therefore, the effective factors in forming the service candidate portfolio which are directly managed by SQ are cohesion, coupling, and reusability, while granularity is managed indirectly.

In order to calculate the SQ for a service portfolio, with k clusters, the cohesion of each cluster is rewarded while the coupling degree between candidate services is fined. Consequently, the service quality first calculates the average cohesion value, and then subtracts it from average coupling value of all clusters. Cohesion average of entire clusters is achieved by dividing them by the cluster numbers (k) while the coupling refers to distinct number of pairs $(\frac{k(k-1)}{2})$. The coupling degree is fined by subtracting the cohesion average

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

from coupling average. Thus, the SQ value is defined as:

$$SQ(\text{Service Quality}) = \frac{1}{k}\sum_{i=1}^{k} C_i - \frac{1}{\frac{k(k-1)}{2}}\sum_{i=1}^{k}\sum_{j=i+1}^{k} A_{ij} \quad (5)$$

In which C_i indicates the inner relations of clusters that means the cohesion value in cluster 1 to k. In addition, A_{ij}represents the relation degree between clusters or coupling value that is calculated by the coupling equation which was illustrated formerly. Next step is to calculate the SQ through the tool that uses the cohesion and coupling factors to find the best combination of tasks in each cluster.

After applying the clustering algorithm, the service portfolio is exposed that includes a service list which determines the tasks inside each cluster. Specifications which are exposed after executing the clustering algorithm include service candidate number, name of related business process, and tasks inside the service candidate.

After clustering phase, a service list is made; however, it may need to be refined due to a number of reasons such as high number of exposed services, limitation in implementation time or budget and the tendency for incremental implementation of SOA. The Litmus Test is used as a tool for selecting services with higher priority to be realized from a set of candidate services based on a set of questions. Likewise, answers to such questions will prioritize the services. It is used when there are resource constraints such as budget, time, and implementation difficulties [30]. The Litmus Test results determine the most appropriate services to be realized. Litmus Test questions can be the degree of relation of the services with the business goals, required time as well as the budget to implement each service.

The motivation factor in the decision to preserve a service in the identified service list in the final refining process is the strength of its support degree of the business goals. Besides, as discussed earlier, the importance of each task is tied to the intensity of its relationship with high-level business goals. Goal-affinity factor in this research is defined as a quantified measure used to calculate the support degree of services to goals. It also can be used as a determinative factor in service refining stage which based on this factor each service can be assessed in terms of the degree of its relation with business goals. The degree of Goal-Affinity factor is placed as a question into the Litmus test that uses the questions list. Therefore, service refining can be performed through ignoring the services with low priority in satisfying business goals. While the majority of services consist of more than one task, each task undergoes the goal-affinity calculation. Therefore, in order to obtain the goals-affinity of each service the average of goal-affinity of tasks should be calculated. The goal affinity of each service is represented as GAS (goals affinity of service) and also goals-affinity of task i as GAT_i. Formally the GAS is proposed as: $GAS = \frac{\sum_{i=1}^{n} GAT_i}{n}$ (6)

For example, the “Backup service” in Figure 4 consists of three tasks, therefore, the calculation of the goal-affinity of the service will be: GAS$= \frac{6+8+3}{3} = 5.6$

3.3 SIMs’ Comparison

The main objective of this research is to achieve business aligned and automated SIM. The current SIMs’ challenges can be summarized in necessity for a comprehensive method, satisfying the service quality factors, realizing business-IT aligned services, automation, and validation [1, 11]. This section presents a comparison between ASIF and most approximate SIMs to ASIF objectives. The criteria set for comparison are firstly, comprehensiveness of a SIM which refers to inclusion of all necessary service identification activities that have been investigated in chapter two namely, scope determination, and service refinement as complementary phases of a SIM. Secondly, service quality factors that include cohesion, coupling, reusability and granularity, thirdly, supporting the business dimension within service identification to evaluate if a SIM satisfies business-IT alignment via supporting the business processes and business goals. In addition, we consider existing of automation tool as criteria due to automation importance of service identification activities. Each feature is assessed by two conditions: existence and guidelines. Existence, indicates if the feature is considered in the SIM or not, and guidelines field shows existence of the accessible guidelines regarding to each feature. The guidelines can be seen as spectrum of methods that only introduced an activity without providing details, to methods that provide a lot of details such as heuristics, examples, and case studies to increase the clarity and support of its applicability. To illustrate the guidelines situation, three types of indicators have been used namely, minus (-)

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

indicates absence of guidelines, plus (+) shows that partially supporting details are available, and double plus (++) standing for sufficient details and guidelines through systematic methods, algorithms, case studies and examples.

The comparison which presented in Table 1, was performed by a coder, and also based on literature review results. The results of the comparison between coder results and literature review were matched which validate the comparison results.

We have selected SIMs whose objectives mostly approximate the objectives of the current research. The aim was to select SIMs which emphasize on identifying business aligned services and automation. Therefore, the most appropriate SIMs which were involved in this comparison include methods proposed by [3, 10, 13, 18, 31].

Regarding to scope determination phase, only [13] put forth the idea of scope determination and stress that it should be consistent with business goals; however, they did not present details to address this issue. On the other hand, ASIF provides clear scope determination method based on AHP which supports its applicability via guidelines.

In addition, according to [18], the service refinement is addressed based on behavioural constraints of each identified service that facilitates the service refinement using human based activity, but they do not provide illustrations about the service prioritization phase. In addition, [3] stress on cohesion and coupling status of service candidates in order to refine service candidates automatically. Similarly, [18] conduct the refinement step to decide between realizing the candidate service as service or implementing it based on previous techniques such as Sun Enterprise JavaBeans that relies on experts' experience. They utilize Litmus Test that consists of a set of questions used for service refinement. However, the questions and the process of refinement have not been declared. In addition, the answering of questions is based on experts' knowledge and it has not been facilitated to address the questions quantitatively to make the answering process clear. Alternatively, ASIF utilized the goals affinity factor of each candidate service which was calculated automatically to present a business-aligned indicator for service refinement. Besides the goals affinity factor, a set of questions are proposed to help the service refinement based on Litmus Test.

Regarding the service quality factors, [3] calculate the cohesion value of each candidate service and the coupling value for service refinement without addressing the reusability and granularity factors. [13] utilize the service granularity as criteria in service refinement to be recognized by experts as human based activity. Likewise, [10] propose a set of quantitative equations for service quality factors calculation based on the low-level entities in their CRUD matrix. ASIF proposes quantitative calculation of the cohesion, coupling and reusability based on relations between business processes' tasks as business aligned entities. The granularity can be satisfied via a trade-off between cohesion and coupling values.

The business-alignment of services is assessed based on the degree of relationship of a service with business processes and business goals. However, the business processes have been selected by all mentioned SIMs through different business process models. [10, 13, 18] utilize general process models from the business domain, due to informality of those process models, these SIMs have lack of guidelines in how the process models can be utilized. In addition, [31] used UML as a business process model and [3] selected the BPMN as specific business process model supported with clear guidelines. In addition, the business goals involvement only mentioned by [13] supports the business-IT alignment based on business goals. In contrary, ASIF utilizes KAOS based goal modelling with detailed guidelines to assess the identified services based on goals affinity factor of each candidate service.

Furthermore, the SIMs' automation tools lack comprehensive coverage of the service identification activities. Besides, the service identification preparation phase that should undertake the input types' transformation to an automatable domain is not automated. Obviously, by considering the volume of input types, remaining the preparation phase un-automated, imposes the expertise involvement. In addition, quantifying the service quality factors to enable the automation is not supported by majority of SIMs' tools. Furthermore, the service refinement in all five automated tools is based on experts' involvement, while ASIF Tool covers major service identification phases. It supports the preparation phase by establishment of tasks' relation matrix automatically. Also, the service identification phase is supported by service quality factors quantification and utilizing the clustering

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

algorithm. Eventually, ASIF supports the service refinement by automatic calculation of goals affinity degree of each service candidate. Hence, the comparison indicates the superiority of ASIF based on selected criteria.

3.4 ASIF Tool Development

The automated tool is developed to integrate automatically the service identification process from BPMN diagrams to service candidates. The ASIF Tool has consists of two parts, the convertor from tasks-relation matrix to clustering algorithm format (Figure 5). The second part includes clustering algorithm that clusters the tasks inside a business processes according to cohesion, coupling and reusability weights of each composition of tasks.

The customized objective function that calculates the SQ is developed in the tool to come up with a list of service candidates based on clustering algorithm calculation of best SQ value. In addition, visualizing the output clusters in the form of service candidates is necessary to increase the clarity and understandability by presenting inter and intra relations of candidate services. We use Graphviz as a known tool for visualizing data in mathematical or graph forms that can be easily configured [32]. Figure 6 illustrates the execution of the developed tool.

4. CONCLUSION

ASIF pays attention to service identification preparation phases. Business-aligned input types and scope determination based on AHP has been presented to help the firm in selecting right scope according to SOA requirements. ASIF has tried to present an automatic method with a comprehensive guideline as the key to addressing ambiguities and increasing applicability. It focuses on proposing a top-down method and so all selected inputs such as BPMN and business goals have selected from business side that enrich with business concepts. Furthermore, ASIF gets success in presenting an automation method which considerably reduces the necessity of experts' knowledge and also decreases the required time to process a large number of business processes and other inputs in order to identify services.

ASIF emphasizes on quantification of inputs though their attributions are not near to numerical domain to enable the application of mathematical and software tools within identification process. Thus, relationships between a business process's entities have converted to a matrix form and quality factors such as reusability, cohesion and coupling, which are decisive in the services ASIF has quantified through mathematical equations. In addition, first version of ASIF tool has presented which integrates all ASIF steps and ease ASIF usability.

Creating a balance between service quality factors is considered as a challenging task like cohesion and coupling of service candidates. Cohesive services encapsulate most related tasks in one cluster which acts to reasonably ease achieving appropriate granularity level. Consequently, a cluster set that is cohesive and loosely coupled, indirectly manages granularity of each service to make a trade-off in service size based on its relations.

This paper has presented ASIF to support service identification with regard to their goal affinity so as we can visually recognize the importance of services and also the probable gap in identifying services to be consider as required services. Future research objective can aim involving of more quality factors besides, ASIF tool can cover service specification phase to expose services in WSDL form.

REFRENCES:

[1] Kohlborn, T., Korthaus, A, Chan, T, Rosemann, M (2009). Identification and Analysis of Business and Software Services-A Consolidated Approach., IEEE Transactions on Services Computing 2(1): 50-64.

[2] Börner, R., M. Goeken, Goeken, M, Rabhi, F (2012). SOA development and service identification: A case study on method use, context and success factors,Working Paper Series, Frankfurt School of Finance & Management.

[3] Bianchini, D., Cappiello, C, De Antonellis, V, Pernici, B (2013). Service Identification in Inter-Organizational Process Design, IEEE Transactions on Services Computing, doi: 10.1109/TSC.2013.26.

[4] Simon, K (2013). U.S. Patent No. 8,359, 224. Washington, DC: U.S. Patent and Trademark Office.

[5] Boerner, R. M. Goeken (2009).Service identification in SOA Governance literature review and implications for a new method. 3rd IEEE International Conference onDigital Ecosystems and Technologies, DEST '09.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[6] Dijkman, R. M, Dumas, M., Ouyang, C (2008). Semantics and analysis of business process models in BPMN.Information and Software Technology50 (12): 1281-1294.

[7] Gu, Q, Lago, P (2010). Service identification methods: a systematic literature review. In Towards a Service-Based Internet (pp. 37-50). Springer Berlin Heidelberg.

[8] Frey, F. J., Hentrich, C., Zdun, U (2013). Capability-based Service Identification in Service-Oriented Legacy Modernization. 18th European Conference on Pattern Languages of Programs:1-12.

[9] Cherbakov, L., Galambos, G., Harishankar, R., Kalyana, S., Rackham, G (2005). Impact of service orientation at the business level. IBM Systems Journal,44(4): 653-668.

[10] Jamshidi, P., Mansour, S., Sedighiani, K., Jamshidi, S., Shams, F (2012). An Automated Service Identification Method (p.2). Technical Report, TR-ASER-2012-01, Automated Software Engineering Research Group, Shahid Beheshti University.

[11] Borner, R., M. Goeken, M., Kohlborn, T. Korthaus, A (2011). Fragment selection and context factors in situational methods for service identification, Research Challenges in Information Science (RCIS), 2011 Fifth International Conference. 1-8.

[12] Brown, P.C. (2009). service Specification and Business Process Design. From www. soasymposium.com/ home2011/pdf_brazil/Paul_C_Brown_Service_ Specification.pdf [May 2011].

[13] Arsanjani, A, Ghosh, S, Allam, A, Abdollah, T, Ganapathy, S, Holley, K (2008). SOMA: A method for developing service-oriented solutions. IBM systems Journal, 47(3), 377-396.

[14] Fareghzadeh, N. (2008). Service identification approach to SOA development. Proceedings of World Academy of Science, Engineering and Technology 35. 258-266.

[15] Ramadour, P, Fakhri, M. Cauvet, C. (2010). A goal-driven method for business services engineering. 4th International Conference on New Trends in Information Science and Service Science (NISS). 554-558.

[16] Baghdadi, Y. (2006). Reverse engineering relational databases to identify and specify basic Web services with respect to service oriented computing. Information Systems Frontiers 8(5): 395-410.

[17] Jamshidi, P., Khoshnevis, S., Teimourzadegan, R., Nikravesh, A., Khoshkbarforoushha, A. Shams, F (2009). ASSM: Toward an automated method for service specification. Proceding of Services Computing Conference APSCC 2009. IEEE Asia-Pacific. 451-456.

[18] Leopold, H., Mendling, J (2012). Automatic derivation of service candidates from business process model repositories. In Business Information Systems : 84-95. Springer Berlin Heidelberg.

[19] Qian, M., Zhou, N., Zhu, Y. Wang, H (2009). Evaluating Service Identification with Design Metrics on Business Process Decomposition. IEEE International Conference on Services computing.160-167.

[20] Van Nuffel, D. 2007. Towards a service-oriented methodology: business-driven guidelines for service identification, In On the Move to Meaningful Internet Systems 2007: OTM 2007 Workshops. 294-303.

[21] Yousef, R., Odeh, M., Coward, D., Sharieh, A (2009). BPAOntoSOA: A generic framework to derive software service oriented models from business process architectures. ICADIWT'09. Second International Conference on the Applications of Digital Information and Web Technologies. 50-55.

[22] Berntsson-Svensson, R. Aurum, A (2006). Successful software project and products: An empirical investigation, ACM.

[23] Sikdar, S, O. Das (2009). Goal based project scope determination approach. IEEE Toronto International Conference on Science and Technology for Humanity (TIC-STH).

[24] Mohamad, R., Aziz, M. W., Jawawi, D. N. A., Ghazali, M., Arbaie, M. Z., Ibrahim, N (2012). Service identification guideline for developing distributed embedded real-time systems. Software, IET. 6(1): 74-82.

[25] Hallikainen, P., Kimpimäki, H., Kivijärvi, H. (2006). Supporting the module sequencing decision in the ERP implementation process. In Hawaii International Conference On System Sciences IEEE. : 181.

[26] Van Lamsweerde, A, Dardenne, A, Delcourt, B, Dubisy, F (1991). The KAOS project: Knowledge acquisition in automated specification of software. In Proceedings of the AAAI Spring Symposium Series.

[27] Van Lamsweerde, A. (2009). Requirements engineering: from system goals to UML models to software specifications, Wiley.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[28] Van Lamsweerde, A., Darimont, R., Massonet, P. (1995). Goal-directed elaboration of requirements for a meeting scheduler: problems and lessons learnt, In Requirements Engineering, 1995., Proceedings of the Second IEEE International Symposium on :194-203.

[29] Mitchell, B. S. (2002). A heuristic search approach to solving the software clustering problem. (Order No. 3039424, Drexel University). ProQuest Dissertations and Theses, , 262-262 p. Retrieved from http://search.proquest.com/docview/304803712?accountid=41453. (304803712)

[30] Ackerman, L., Y. Chaix. P Wik. (2012). Top 3 Tips for Using Automation to Bring Governance to Service-Oriented Analysis and Design.

[31] Dwivedi, V. Kulkarni. (2008). A model driven service identification approach for process centric systems, Congress on Services. IEEE. Los Alamitos: 65-72.

[32] Ellson, J., Gansner, E., Koutsofios, L., North, S. C. Woodhull, G. (2002). Graphviz—open source graph drawing tools. In Graph Drawing. Springer Berlin Heidelberg 483-484

[33] T.S. Bhatti, R.C. Bansal, and D.P. Kothari, "Reactive Power Control of Isolated Hybrid Power Systems", *Proceedings of International Conference on Computer Application in Electrical Engineering Recent Advances (CERA)*, Indian Institute of Technology Roorkee (India), February 21-23, 2002, pp. 626-632.

[34] B.N. Singh, Bhim Singh, Ambrish Chandra, and Kamal Al-Haddad, "Digital Implementation of an Advanced Static VAR Compensator for Voltage Profile Improvement, Power Factor Correction and Balancing of Unbalanced Reactive Loads", *Electric Power Energy Research*, Vol. 54, No. 2, 2000, pp. 101-111.

[35] J.B. Ekanayake and N. Jenkins, "A Three-Level Advanced Static VAR Compensator", *IEEE Transactions on Power Systems*, Vol. 11, No. 1, January 1996, pp. 540-545.

Journal of Theoretical and Applied Information Technology
31st January 2016. Vol.83. No.3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

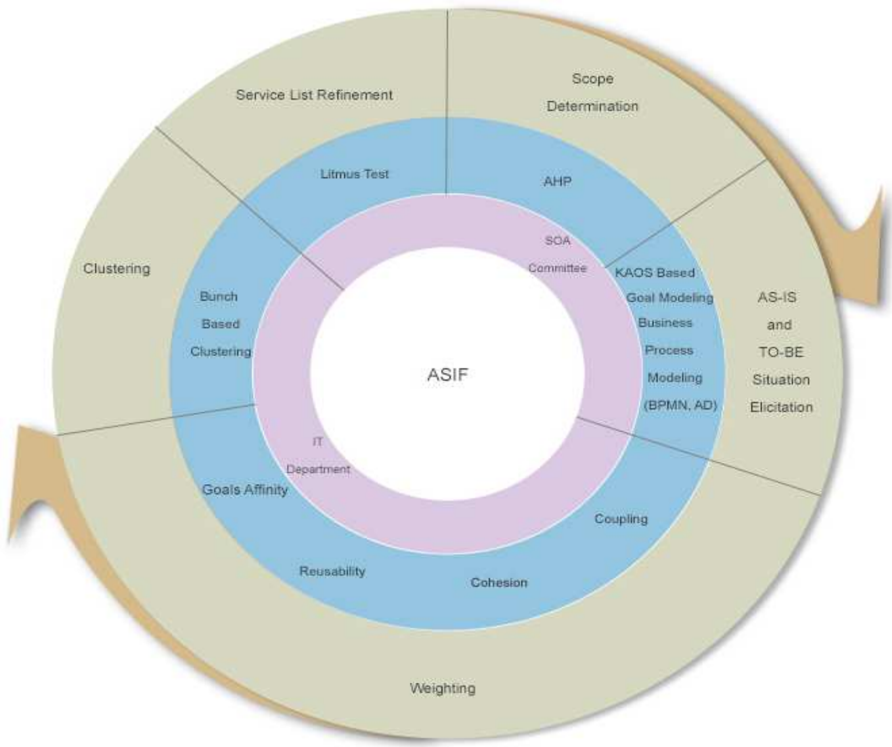

Fig 1 Automated Service Identification Framework (Asif)

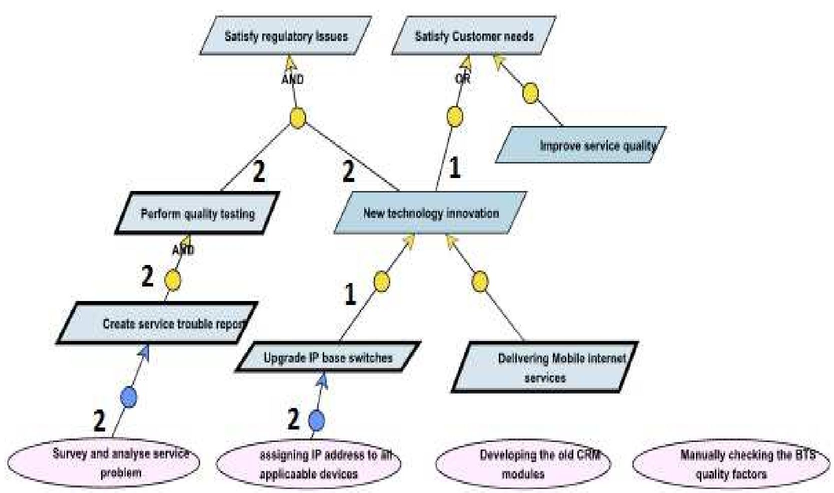

Fig 2 Example Of Relation Weighting Through A Goal Graph Based On KAOS

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

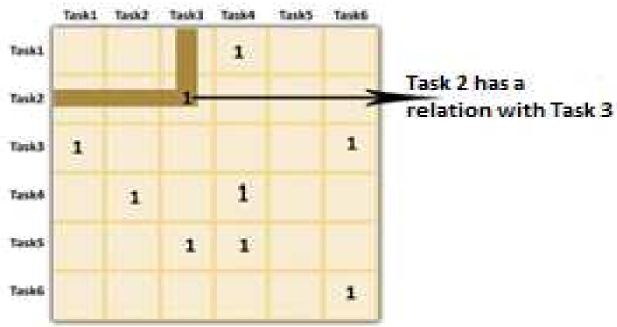

Fig 3 Example Of Tasks-Relation Matrix

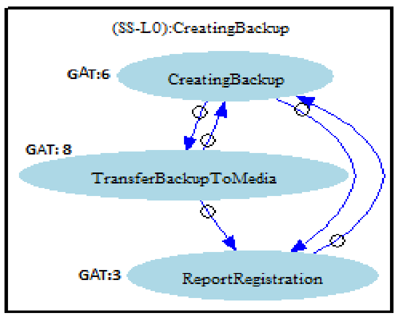

Fig 4 Backup Service That Consist Of Three Tasks

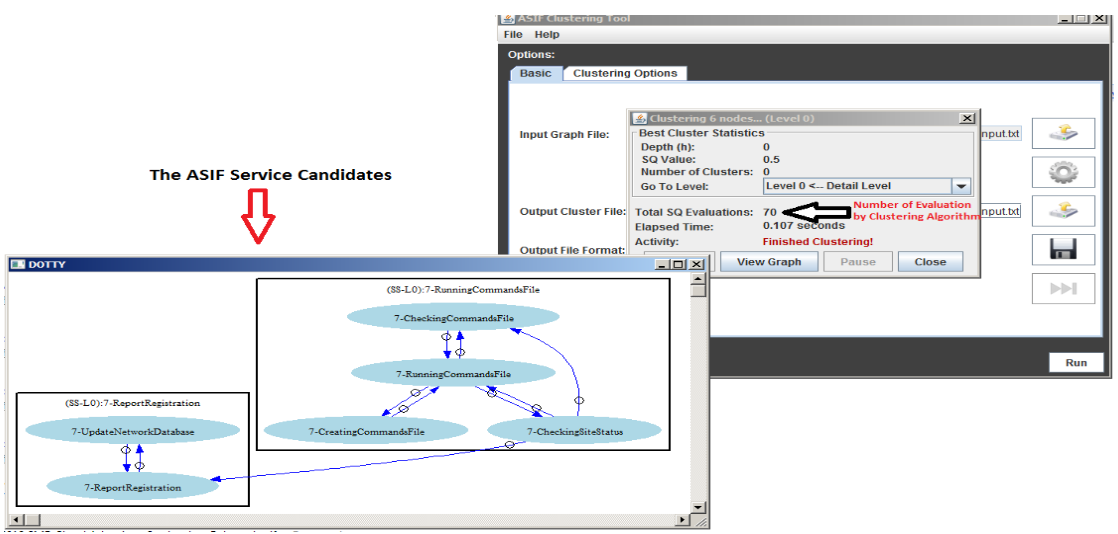

Fig 6 Screenshot Of The Supporting Tool That Shows The Identified Services Candidates Graphically

Journal of Theoretical and Applied Information Technology
31st January 2016. Vol.83. No.3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 1: Comparison Of Service Identification Methods With ASIF Regarding To Key Features

Factors	Complementary Phases				Service Quality Factors								Business-Alignment				Automation Tool	
	Scope Determination		Service Refinement		Cohesion		Coupling		Reusability		Granularity		Business Process		Goals			
Source	Existence	Guidelines	Existence	Guidelines	Existence	Guidelines	Existence	Guidelines	Existence	Guidelines	Existence	Guidelines	Existence	Guidelines	Existence	Guidelines	Existence	Coverage*
[18]	-	-	√		-		-	-	-	-	-	-	√ ++		-	-	√	T
[31]	-	-	-	-	-	-	-	-	-	-	-	-	√	-	-	-	√	T
[3]	-	-	√	++	√	++	√	++	-	-	-	-	√	+	-	-	√	T,R
[13]	√	+	√	+	√	-	√	-	√	-	-	-	√	-	√ +		-	-
[10]	-	-	-	-	√	++	√	++	√	-	√	+	√	+	-	-	√	T,S
ASIF	√	++	√	++	√	++	√	++	√	++	√	-	√	++	√	++	√	P,T,S,R

**Automated Parts In SIM: Preparation (P), Technique (T), Service Quality Factors (S), Service Refinement (R)*

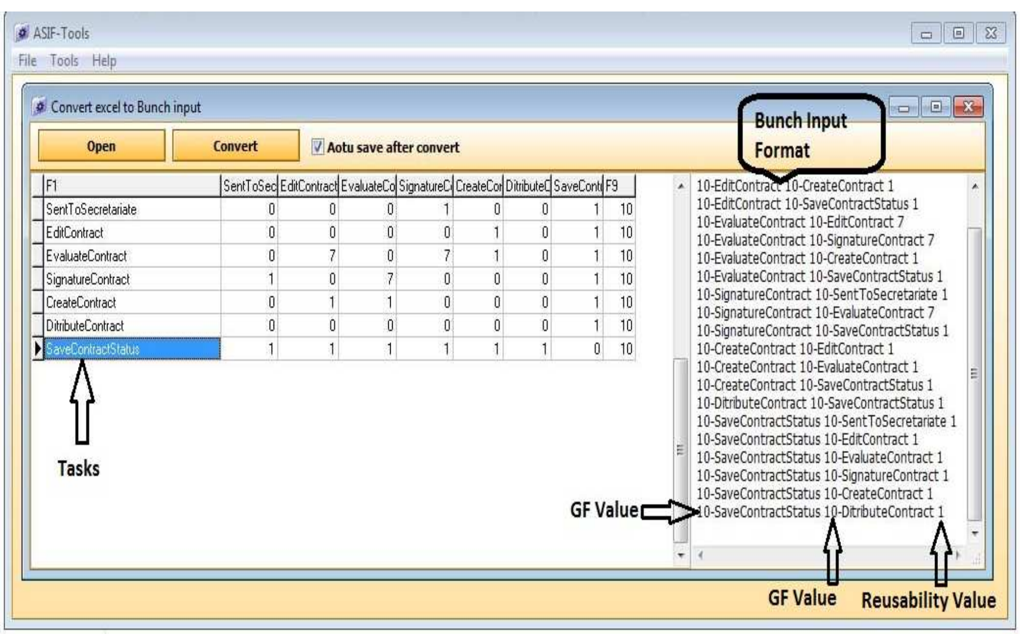

Fig 5 Screenshot Of The Supporting Tool That Shows Converting 'Create Site' Relation Matrix In Excel Format To Bunch Input