A cloud computing platform for ERP applications

Chin-Sheng Chen, Wen-Yau Liang, Hui-Yu Hsu

A R T I C L E   I N F O

Article history:
Received 13 March 2013
Received in revised form 15 October 2014
Accepted 11 November 2014
Available online 21 November 2014

Keywords:
Web services composition
ERP
Cloud computing
Genetic algorithms
Rough set theory

A B S T R A C T

Cloud computing enables many applications of Web services and rekindles the interest of providing ERP services via the Internet. It has the potentials to reshape the way IT services are consumed. Recent research indicates that ERP delivered thru SaaS will outperform the traditional IT offers. However, distributing a service compared to distributing a product is more complicated because of the immateriality, the integration and the one-shot-principle referring to services. This paper defines a CloudERP platform on which enterprise customers can select web services and customize a unique ERP system to meet their specific needs. The CloudERP aims to provide enterprise users with the flexibility of renting an entire ERP service through multiple vendors. This paper also addresses the challenge of composing web services and proposes a web-based solution for automating the ERP service customization process. The proposed service composition method builds on the genetic algorithm concept and incorporates with knowledge of web services extracted from the web service platform with the rough set theory. A system prototype was built on the Google App Engine platform to verify the proposed composition process. Based on experimental results from running the prototype, the composition method works effectively and has great potential for supporting a fully functional CloudERP platform.

1. Introduction

Traditional business applications such as computer aided design (CAD), product data management (PDM), Computer aided manufacturing (CAM), enterprise resources planning (ERP) and manufacturing execution systems (MES) all rely on a central server and procedural software. These systems are not autonomous or flexible enough to support a dynamic business environment [34]. With the advance of Internet technology and globalization, these enterprise applications, especially ERP systems have been web-enabled, providing access to information and communications via the Internet as a part of global business strategy [16]. Along with the emerging demand for mobility and on-demand services, the development of web-based ERP systems becomes an urgent research and development issue [64].

The subscription to web services for ERP applications has two essential advantages: ease of integration and reduction in costs through the hosted application model [57]. Wu et al. [67] presented a framework for measuring the scalability of service based applications in a Cloud Computing environment and propose an assignment strategy to improve the scalability of composite Web services in terms of services productivity. Recent research indicates that ERP delivered thru SaaS will outperform the traditional IT offers as a consequence of the current economic crisis and will help the economies to recover [21]. Although ERP is lagging behind other applications in terms of SaaS based applications there seems to be a general consensus that ERP in SaaS is gaining momentum. To grab this momentum, the four big players in the ERP systems market SAP, Oracle, Sage and Microsoft are positioning their ERP offers in SaaS model [24]. However, distributing a service compared to distributing a product is more complicated because of the immateriality, the integration and the one-shot-principle referring to services [23]. Also, the process of analyzing and selecting services in the Web services composition process is more complex than the one of analyzing and selecting parts for a product design [37]. It is further complicated by the customer’s request in terms of the scope of application. One specific need is the development of efficient composition methods which evaluate and optimally integrate these possibly heterogeneous services on the Web, especially in the ERP application domain, in response to an enterprise customer’s request.

Therefore, this paper proposes a CloudERP platform on which enterprise customers can select web services and customize a unique ERP system to meet their specific needs. The CloudERP aims to provide enterprise users with the flexibility of renting an entire ERP service through multiple vendors. This paper also addresses the
challenge of composing web services and proposes a web-based solution for automating the ERP service customization process. This study proposes a method that makes use of the genetic algorithm (GA) concept and the rough set theory to solve the Web services composition problem. The genetic algorithm incorporates with rough set theory to solve the web services composition problem has been discussed and applied [5,36,37]. However, these all focus on how to use rough set theory to extract rules and ignore the feature of the application domain. The novelty of the proposed method lies in the application domain (Cloud ERP).

This remainder of this paper is organized as follows: Section 2 reviews the concepts of Web-based ERP and Web services composition. In Section 3, a novel ERP platform called CloudERP is proposed. In Section 4, the proposed composition method for Web ERP services is presented. In Section 5, a system prototype is presented along with experimental data analysis and then followed by Section 6 which provides concluding remarks and summary of future research directions.

2. Literature review

2.1. Web-based ERP

ERP systems are one of the most adopted information technology (IT) solutions in organizations [2]. Because of their scale and substantial resources consumption, it is not surprising that ERP systems have been a center of focus by both researchers and practitioners [11]. The key competitive edge for every enterprise in the 21st century is in its ability to prescribe, standardize, and adapt its business activities and collaborations with customers, suppliers, partners and competitors [34]. Most ERP vendors today recognize this interoperability issue as significant and have built up Internet-enabled supply chain/logistics modules to facilitate integration with the back-end systems of supply chain partners relying on a diverse set of legacy databases, IT infrastructure and applications [32]. For example, Gollakota [15] reported a company which created kiosks with Internet and computer access and operated a web portal serving the needs of the farming industry. The portal provided information relating to farming techniques, farm business information, and general information such as weather and climate, and access to the firm’s ERP system. Separately, Tarantilis et al. [57] presented a Web-based ERP system developed to address business problems and manage real-world business processes ranging from a simple office automation procedure to complex supply chain planning. Zhang et al. [75] explored the IT service innovation in textile industrial clusters from a service system perspective. They argued that the IT enabled producer service could be used to ensure the structural upgrading of the textile industrial clusters. Mital et al. [44] also developed an integrative framework to identify the determinants of choice of SaaS in the specific context of SaaS based e-procurement and ERP.

In summary, one of the most important trends in the recent years is cloud computing. It has the potentials to reshape the way IT services are consumed. More recently, some ERP vendors have moved some of their offerings to the cloud e.g., SAP By Design. However, there is still a lot to be done in order for the customers to see more and more services and suites moving to the cloud. Therefore, more research efforts are still needed in order to elucidate knowledge on the marriage of the two [11].

2.2. Web services composition

The capability of composition is an important strength of any Web services provider. Rather than accessing a single service, composing services is essential as it adds better benefits to its users [13]. By ensuring high-level interoperability, Web services offer have the capability of composing compatible processes referred to as composite Web services, independent of specific platforms and computing paradigms [42]. While elementary Web services do not rely on other Web services to fulfil external requests, composite services integrate multiple service components to fulfil a customer’s request [43].

Several approaches and applications have been proposed to exploit the concept of Web services composition. One research direction looked into the role of policies and context for framing the composition of Web services. Policies are to govern the behavior of Web services engaged in composition; and context is to support the development of adaptable Web services [42]. Yu et al. [73] designed a broker-based architecture for selecting Quality of Service (QoS)-based services. The objective of service selection is to maximize an application-specific utility function under the end-to-end QoS constraints. Park [46] presented a decentralized protocol design called the Web services co-allocation protocol, aiming to facilitate the execution of composite Web services, while Lee et al. [32] proposed a Web services-based Multidisciplinary Design Optimization (MDO) framework that synthesizes both disciplinary and cross-disciplinary resources available for MDO. Taking advantage of Web services, Zhao et al. [77] built a biomedical digital library infrastructure called the Living Human Digital Library (LHDL) that allows clinicians and researchers to preserve, trace, and share data resources, as well as to collaborate at the data-processing level. Recently, Yahyaoui et al. [70] proposed a novel matchmaking approach between fuzzy user queries and real world Web services. The matchmaking spans over a domain dependent classification step that produces fuzzy classification rules for Web services. Furthermore, these rules are leveraged to classify Web services into categories, which allow reducing the matchmaking space. One study developed an efficient approach for automatic composition of Web services using the state-of-the-art Artificial Intelligence (AI) planners [79].

Rajeswari et al. [51] revealed various challenges in the QoS parameter for Web service composition because it is difficult to recognize. In summary, Web services composition is a complex issue. The complexity initially arises from the diversity and compatibility of the composition components of Web services. It is further complicated by the customer’s request in terms of the scope of application. In theory, service components are developed by different organizations and offered by different providers at different rates. There is a general need for developing principles and methodologies for managing composite Web services. One specific need is the development of efficient composition methods which evaluate and optimally integrate these possibly heterogeneous services on the Web, especially in the Cloud ERP application domain, in response to an enterprise customer’s request.

3. CloudERP

Cloud computing is defined as both the applications delivered as services and the hardware and systems software in the data centers that provide those services [3]. Kim [29] anticipated that Cloud computing would become a key computing paradigm for the next 5–10 years. Cloud services can be viewed as a cluster of service solutions based on cloud computing, which involves making computing, data storage, and software services available via the Internet. Generally, cloud services can be divided into three categories [69]:

(1) Software as a service (SaaS): Applications services delivered over the network. SaaS simplifies the utilization of a large
amount of software applications remotely, elastically and seamlessly [65].

(2) Platform as a service (PaaS): A software development framework and components all delivered on the network. Offered as on-demand, pay for usage model. A PaaS model packages a computing platform including operating system, programming language execution environment, database, and web server. A PaaS client is able to develop and run its applications at the software layer [65].

(3) Infrastructure as a service (IaaS): An integrated environment of computing resources, storage, and network fabric delivered over the network. Offered as an on-demand, pay for usage model.

Among them, SaaS is regarded as a potential segment and the utilization of SaaS solutions can lead to many benefits for enterprise users with profound consequences in improving IT performance [6]. Service providers can greatly simplify software installation and maintenance and centralizes the control of versioning. End users on the other hand can access the service “anytime, anywhere,” share data and collaborate with partners readily, while keeping their data stored safely in the infrastructure. As a result, an enterprise customer does not have to acquire the whole enterprise software suite, and yet is able to choose each module from different vendors, creating a unique, cost-efficient and customized enterprise solution [57]. Within the hype of cloud services, ERP systems delivered as Software as a Service (SaaS) is receiving more focus from ERP vendors. ERP vendors have for many years developed and sold ERP as ‘standard software’ that fits the needs of many firms, and now SaaS as a new approach to deliver software has emerged [24]. The proposed composition method can be implemented as a SaaS, running on a PaaS by a Cloud services provider. Depending on its architecture, cloud computing can be categorized into three types: external/public clouds—resources dynamically provided on a self-service basis over the internet via web services from an off-site third-party provider; internal/private clouds—data and processes managed within an organization without the restriction of network bandwidth or security exposures; and hybrid clouds—the environment consisting of multiple internal and external cloud computing solutions [53]. This paper propose a hybrid clouds computing which consisting of multiple internal and external cloud computing solutions. Fig. 1 depicts a CloudERP platform that supports interoperable service-to-service interaction over the Cloud. The CloudERP aims to provide enterprise users with the flexibility of renting an entire ERP service through multiple vendors. The platform has three major players:

1. The Cloud services provider, which enables communications among ERP providers and enterprise customers;
2. The ERP providers, which provide an XML format, computer-readable description of Web services for execution of various application functionalities;
3. The enterprise customers, which select, compose, and lease the Web services to meet their ERP objectives.

To satisfy the need of an enterprise customer for ERP application, the following platform steps have to exist and occur in the sequence as outlined in Fig. 1.

**Step 1 (Submit and Assess)**
1.1 ERP providers submit Web services to the platform.
1.2 The platform checks the compatibility. If not, return to ERP providers.
1.3 Experts start to assess each Web services.

**Step 2 (Publish)**
2.1 Publish the Web service on this platform and notify users.

**Step 3 (Select)**
3.1 Users input requirements and constraints.
3.2 Users select composite method.

**Step 4 (Implement)**
4.1 The platform composites Web services and configure into the user’s virtual Cloud.
4.2 The platform notifies users.

**Step 5 (Access)**
5.1 Users access the composited Web service through the virtual Cloud.
5.2 Users evaluate the process and send feedbacks to the platform.

This paper focuses on the selection and composition process in step 3 and proposes a composition method with use of the GA concept and the rough set theory to select and compose web services for CloudERP users.

### 4. The proposed genetic algorithm

This section details the proposed genetic algorithm imbedded with rough set theory. It was developed and coded with basic units of Web services. The `reduct` rules generated by the rough set approach were used to reduce the basic units’ domain range of the initial population and validate the feasibility of offspring when a crossover is processed to achieve optimization of the objective function. It was designed to improve the effectiveness of the GA’s evolution and achieve rapid convergence of the search process.

#### 4.1. Genetic algorithm

Genetic algorithms are used to perform elegant and robust search to improve a known solution. They allow application of optimization methods to find answers in complex or poorly understood search spaces [1]. Genetic Algorithms are a powerful tool to solve combinatorial optimizing problems [41] and are a suitable choice.
for discredited optimization problems [28], Vallada and Ruiz [63] proposed three genetic algorithms for the permutation flow shop scheduling problem with total tardiness minimization criterion. The results showed that the proposed algorithms were effective, outperforming the existing methods used for comparison in the study. Udhayakumar et al. [62] applied GA to solving the P-model of chance constrained data envelopment analysis (CCDEA) in a case of Indian banking sector.

GA techniques have been also applied in others fields. For example, Liang and Huang [35] used GA to solve the problem of product synthesis, where there is a conflict between performance and cost. One study proposed genetic algorithms for match-up rescheduling with non-reschedule and reschedule strategies which accommodate new orders by manipulating the available idle times on machines and by resequencing operations, respectively [74]. Sadrzahe and Khalilia [54] presented a genetic algorithm-based meta-heuristic to solve the facility layout problem (FLP) in a manufacturing system, where the material flow pattern of the multi-line layout is considered with the multi-products. Kuo and Lin [31] proposed an evolutionary-based clustering algorithm based on a hybrid of genetic algorithms (GA) and particle swarm optimization algorithm (PSOA) for order clustering in order to reduce surface mount technology (SMT) setup time. More recently, Toledo et al. [59] applied a genetic algorithm with hierarchically structured population to solve unconstrained optimization problems. Meanwhile, one research considered the discrete optimization via simulation problem with a single stochastic constraint and presented two genetic-algorithm-based algorithms that adopt different sampling rules and searching mechanisms, and thus deliver different statistical guarantees [60]. An advanced novel heuristic algorithm is presented, the hybrid genetic algorithm with neural networks (HGA-NN), which is used to identify an optimum feature subset and to increase the classification accuracy and scalability in credit risk assessment [45].

However, typical GA techniques are viewed with shortcomings, such as poor local searching, premature converging, and slow convergence speed [25]. The next subsection thus is devoted to the genetic algorithm imbedded with the rough set theory to overcome the above search problems.

4.2. Rough set theory

Rough set theory (RST) was devised by Pawlak in 1982 as a discovery tool that can be used to induce logical patterns that are hidden in massive data. A rough set approach can capture useful information from a set of mixed data and output this information in the form of decision rules [48]. The RS approach uses a decision table with rows containing objects and columns containing criteria or features to derive decision rules through an inductive process. Rough set theory has become a well-established theory for uncertainty management in a wide variety of applications related to pattern recognition, image processing, feature selection, neural computing, conflict analysis, decision support, data mining and knowledge discovery [49]. In rough set theory, a reduct is defined as a minimal sufficient subset of a set of attributes, which has the same ability to discern concepts as when the full set of attributes is used [78]. Basically, the reducts represent necessary condition attributes in decision making. Further, a subset of the attributes can have more than one reduct, so simplifying the decision rules will not yield unique results. To implement the rough set theory, a procedure for determining the reducts is needed, such as generating reducts and identifying the decision rule.

The rough set theory has been applied to various areas: Tseng and Huang [61] derived the decision rules from historical data for identifying features that contribute to CRM; Chu et al. [9] proposed a expert systems for assisting mapping from performance space to design space (ESMPD); Herawan et al. [18] proposed a new technique called maximum dependency attributes (MDA) for selecting clustering attribute. The proposed approach is based on rough set theory by taking into account the dependency of attributes of the database; Shyng et al. [55] used two processes (pre process and post process) to select suitable rules and to explore the relationship among attributes; A systematic approach to analyze existing patient information based on rough set theory with the consideration of resource allocation is developed [19], a case study is presented to demonstrate the contribution of the proposed approach which assists on decision-making in patent reform or invention with constraint resource; Kaniwai and Kudo [26] proposed a method for mining such local patterns from sequences and described an algorithm for generating decision rules that take into account local patterns for arriving at a particular decision; Wu [68] attempted to segment the ERP users into two subgroups according to the notion of Herzberg's Motivation-Hygiene theory, and further, to uncover imperative perceived benefits for distinct subgroups of ERP users employing the rough set theory. The previous literatures have shown that the rough set theory is very useful to extract knowledge and help the decision making.

Meanwhile, in the past decade, several extensions of the rough set model have been proposed to improve the effectiveness, such as the variable precision rough set model [76], the Bayesian rough set model [33], the Dominance-based rough set model [12], the fuzzy rough set model [39]. In this study, the difference of these models will not be addressed. Generally, the generated rules from each method can be applied to the proposed approach.

4.3. The evaluation schema

Many methods have been considered for ERP selection, including scoring, mathematical programming, analytic hierarchy process (AHP), and multi-criteria decision analysis [27]. Other methods such as zero–one goal programming [27], fuzzy method [66], Data envelopment analysis (DEA) [66], artificial neural network (ANN) [71], and analytic network process (ANP) [38] have also been proposed for selecting a suitable ERP.

Deriving from expert opinion and earlier studies, Karsak and Özogul [27] cited the use of criteria for ERP system selection such as total cost of ownership, functional fit of the system, user friendliness, flexibility, and vendor’s reputation. Weight associated with each selected criterion is also an important factor in the decision-making process. The proposed composition method ranks composited Web services by the ranking model. Given the objective function, it considers normalizing different criteria and allowing for different weight for each criterion. The notations and computations for criterion value are specified as follows:

\[
\alpha, \beta, \gamma, \ldots \text{ represent all kinds of Web services} \\
i \text{the number of Web services composition} \\
W \text{ weight} \\
C \text{ cost of ownership} \\
FF \text{ functional fit of the ERP system} \\
UF \text{ user friendliness} \\
F \text{ flexibility} \\
VR \text{ vendor’s reputation} \\
TI \text{ total index}
\]

- Cost of ownership \((C): C(\alpha), C(\beta), C(\gamma)\) represents the cost of ownership. \(C(\alpha, \beta, \gamma, \ldots)\) represents an integral cost of the Web service \(\alpha, \beta, \gamma, \ldots\)

\[
C(\alpha, \beta, \gamma) = C(\alpha) + C(\beta) + C(\gamma) + \ldots
\]

(1)

- Functional fit (FF): A single Web service may influence integral composition Functional fit. \(FF(\alpha), FF(\beta), FF(\gamma)\) represents the
Functional fit of the Web service and $FF(\alpha, \beta, \gamma, \ldots)$ represent the integral Functional fit of the Web service $\alpha, \beta, \gamma, \ldots$ composition.

$$FF(\alpha, \beta, \gamma) = FF(\beta) \times FF(\gamma) \times \ldots$$  \hspace{1cm} (2)

- User friendliness (UF): $UF(\alpha), UF(\beta), UF(\gamma)$ represents the User friendliness of the Web service. $UF(\alpha, \beta, \gamma, \ldots)$ represents an average User friendliness of the Web service $\alpha, \beta, \gamma, \ldots$

$$UF(\alpha, \beta, \gamma) = (UF(\alpha) + UF(\beta) + UF(\gamma) + \ldots)/n$$  \hspace{1cm} (3)

- Flexibility (F): A single Web service may influence integral composition flexibility. $F(\alpha), F(\beta), F(\gamma)$ represents the flexibility of the Web service and $F(\alpha, \beta, \gamma, \ldots)$ represents the integral flexibility of the Web service $\alpha, \beta, \gamma, \ldots$ composition.

$$F(\alpha, \beta, \gamma) = F(\beta) \times F(\gamma) \times \ldots$$  \hspace{1cm} (4)

- Vendor’s reputation (VR): $VR(\alpha), VR(\beta), VR(\gamma)$ represents the VR Vendor’s reputation of the Web service. $VR(\alpha, \beta, \gamma, \ldots)$ represents an average Vendor’s reputation of the Web service $\alpha, \beta, \gamma, \ldots$

$$VR(\alpha, \beta, \gamma) = (VR(\alpha) + VR(\beta) + VR(\gamma) + \ldots)/n$$  \hspace{1cm} (5)

- Total Index (TI): Through the predefined weight (W1~W5, for each criterion) and normalization, the total index can be obtained using the following equation (i means the number of Web services composition).

$$TI = \left( \frac{C^{\text{max}} - C}{C^{\text{max}} - C^{\text{min}}} \right) \times W_i + \left( \frac{FF - FF^{\text{min}}}{FF^{\text{max}} - FF^{\text{min}}} \right) \times W_2 + \left( \frac{UF - UF^{\text{min}}}{UF^{\text{max}} - UF^{\text{min}}} \right) \times W_3 +$$

$$\frac{F^{\text{max}} - F^{\text{min}}}{F^{\text{max}} - F^{\text{min}}} \times W_4 + \left( \frac{VR - VR^{\text{min}}}{VR^{\text{max}} - VR^{\text{min}}} \times W_5 \right), W_1 - W_5 \in [0, 1], \sum_{j=1}^{5} W_j = 1$$

4.4. The proposed composition process

The GA-based method utilizes relevant knowledge extracted using the rough set theory to improve the search performance by reducing the domain range of the initial population. GA algorithms are efficient search methods based on the principles of natural selection and population genetics in which random operators in a population of candidate solutions are employed to generate new points in the search space [8]. The increase in data and information, however, often hinders the performance and capacity of the GA, raising the cost of finding a solution by using it. To overcome the problem, Passone et al. [47] combined the GA with guidance provided by domain-specific knowledge. The proposed method makes use of the rough set theory introduced by Pawlak as a mathematical method [50] to improve the GA performance. With the rough set method, the minimal attribute sets can be extracted without deterioration in the quality of approximation and minimal length decision rules corresponding to lower or upper approximation [22]. The rough set routine performs off-line, and only when the database is updated. The reduce rules are applied in the GA’s evolution process in real time while composing Web services, in attempt to increase the effectiveness of the GA’s evolution, and more rapidly achieve convergence. The proposed composition process consists of two phases. The first phase is called the predisse of rough set. It is a pre-composition activity and is performed off-line periodically. The second phase is the rough set theory imbedded GA for web service composition. Both phases are outlined in Fig. 2 and detailed as follows:

**Phase I: Predisse of rough set:** This phase focuses on using the database to find relevant rules. The rule identification algorithm developed by Huang and Tseng [20] is used to identify and compose candidate reduct rules. The algorithm is outlined as the following five steps:
Table 1
A partial table of the Web services component database.

<table>
<thead>
<tr>
<th>Functionality</th>
<th>Vendor</th>
<th>Flexibility (range 1–100)</th>
<th>Reputation (range 1–10)</th>
<th>Cost (range 1–10)</th>
<th>Functional fit (range 1–100)</th>
<th>User friendliness (range 1–10)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Finance</td>
<td>SAP</td>
<td>98</td>
<td>9</td>
<td>8</td>
<td>89</td>
<td>7</td>
</tr>
<tr>
<td>Distribution</td>
<td>SAP</td>
<td>69</td>
<td>6</td>
<td>7</td>
<td>93</td>
<td>9</td>
</tr>
<tr>
<td>Human Resource</td>
<td>Salesforces</td>
<td>95</td>
<td>7</td>
<td>8</td>
<td>93</td>
<td>7</td>
</tr>
<tr>
<td>Manufacturing</td>
<td>SAP</td>
<td>55</td>
<td>9</td>
<td>10</td>
<td>97</td>
<td>7</td>
</tr>
<tr>
<td>Finance</td>
<td>Microsoft</td>
<td>62</td>
<td>8</td>
<td>7</td>
<td>97</td>
<td>8</td>
</tr>
<tr>
<td>Procurement</td>
<td>Microsoft</td>
<td>65</td>
<td>10</td>
<td>8</td>
<td>87</td>
<td>6</td>
</tr>
<tr>
<td>Distribution</td>
<td>Oracle</td>
<td>94</td>
<td>6</td>
<td>9</td>
<td>90</td>
<td>7</td>
</tr>
</tbody>
</table>

2.1 Set chromosome p to 1.
2.2 Set gene q to 1.
2.3 Select one Web service that satisfies the rule from the database.
2.4 Set q = q + 1.
2.5 If q < n, then go to Step 2.3.
2.6 Set p = p + 1.
2.7 If p = m, then go to Step 2.2; otherwise, go to Step 3.3.

Step 3 (Evolution): the GA evolution occurs in this step.
3.1 Calculate each chromosome using the objective function.
3.2 Check termination condition. If satisfied, go to Step 4.
3.3 Use a suitable selection strategy to select the new population.
3.4 Crossover: perform crossover according to the crossover rate.
3.5 Mutation: perform the mutation process according to the mutation rate.
3.6 Set generation k = k + 1, and go back to Step 3.1.

Step 4 (Detailed design)
4.1 Allow the user to evaluate the result and modulate the parameters.

5. System prototyping and experimental analysis

Many well known IT companies such as Google, Amazon, Yahoo, Microsoft, IBM, and SAP are active participants in cloud computing, on both SaaS and PaaS. Google operates a cloud computing business platform called Google App Engine (GAE), which is viewed as currently leading and more mature cloud computing platform. Software developers are able to write applications on the platform, and enterprise customers are able to customize network services [72].

The prototyping system was coded in JAVA and JSP based on the GAE for testing and validation of the proposed method. The implementation scenario of CloudERP is shown in Fig. 3. The information of Web services components was published and stored in the system database. Table 1 represents a partial table of Web services components (The table is set randomly). The value of each criterion is predefined by experts.

5.1. Example illustration

In the IT industry, product life cycle is extremely short. Companies need to deliver new products while they have market value. In some IT industry segments, original equipment manufacturing (OEM) and original design manufacturing (ODM) are the main business, of which companies are not involved in their customers’ sales and marketing activities. The cross-functional cooperation of information systems in such an IT industry segment is believed more important than the industry segments with a longer product life cycle, in order to cope with the rapid changes in customer needs and the extremely short product life cycles [56]. A typical ERP system for such an application consists of financial, human resources, manufacturing, procurement, and distribution modules [7]. Thus the scenario is an enterprise customer would like to select these five functional modules from a CloudERP platform to customize its own ERP system. This system prototype is intended to use roulette wheel selection and crossover rate/mutation rate to generate the offspring. Since this enterprise customer wants to pay more attention to flexibility, thus the weight for flexibility is set at 0.6. The weights for the other four criteria are set at 0.1. The proposed method works with the example as follows:

**Phase 1: Predispose of rough set**

Step 1: Create basic units and enter them in the database
Following the approach proposed in Liang and Huang [37], used supplier, type, and implementation language were selected as the conditional attributes of the rough set. Also considered were network-related parameters such as network domain, download frequency, and reputation as argued by Tian et al. [58], Hansen et al. [17], and Ko et al. [30]. Consequently supplier, implementation language, network domain and number of downloads were adopted as the four conditional attributes. On the other hand, high performance, high reputation, and low cost were used as the three decision variables. Each attribute in Table 2 has its own code name. Thirty records were randomly selected from the dataset for this example. Based on Table 2, a decision table was generated as in Table 3 and then organized into an elementary set by their attributive values as shown in Table 4.

Step 2: Calculate the lower and upper approximations for basic units
The elementary sets’ upper and lower approximations were computed in this step. And decision attribute’s elementary set was completely classified as shown in Table 5.

Step 3: Find the core and reduct of attributes
This step created a discernibility matrix to obtain the core and reducts, using the absorption law to calculate the reduct result. In this example, the reduct (C3: No. of download; C4: Implemented Language) was as shown in Table 6.

Table 2
Description of all attributes.

<table>
<thead>
<tr>
<th>Conditional attributes</th>
<th>Decision attribute</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1: Supplier</td>
<td></td>
</tr>
<tr>
<td>C2: Network domain</td>
<td></td>
</tr>
<tr>
<td>C3: No. of download</td>
<td></td>
</tr>
<tr>
<td>C4: Implemented Language</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Step 4: Find the core and reduct of attributive values
The unnecessary values in the condition attributes in the decision table were eliminated in this step, according to the reduct \( \{C_3, C_4\} \) and the discernibility matrix. After calculation, the final decision table version was generated as shown in Table 7.

Step 5: Find the relevant rules

Table 3
Decision Table.

<table>
<thead>
<tr>
<th>Object No.</th>
<th>C1</th>
<th>C2</th>
<th>C3</th>
<th>C4</th>
<th>Decision attributes</th>
</tr>
</thead>
<tbody>
<tr>
<td>X1</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>X2</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>X3</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>X4</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>X27</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>X28</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>X29</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>X30</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 4
Elementary set.

<table>
<thead>
<tr>
<th>Cluster No.</th>
<th>Objects</th>
<th>Decision attributes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( {x_1,x_3,x_7,x_{10},x_{11},x_{13},x_{17},x_{20},x_{21},x_{23},x_{27},x_{30}} )</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>( {x_2,x_4,x_{12},x_{14},x_{18},x_{22},x_{24},x_{28}} )</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>( {x_5,x_6,x_{15},x_{16},x_{19},x_{25},x_{26},x_{29}} )</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 5
Lower and upper approximations.

<table>
<thead>
<tr>
<th>Cluster No.</th>
<th>Number of objects</th>
<th>Lower approximations</th>
<th>Upper approximations</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>1.0</td>
</tr>
<tr>
<td>2</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>1.0</td>
</tr>
<tr>
<td>3</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Table 6
Reduct of attributes.

<table>
<thead>
<tr>
<th>Reduct</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( {C_3, C_4} )</td>
</tr>
</tbody>
</table>

Table 7
Final decision table.

<table>
<thead>
<tr>
<th>C3</th>
<th>C4</th>
<th>O</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>*</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>*</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>*</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>*</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 8
Decision rule.

<table>
<thead>
<tr>
<th>Reduct No.</th>
<th>C3</th>
<th>C4</th>
<th>O</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>*</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>*</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 9
The explanation of decision rules.

<table>
<thead>
<tr>
<th>Rule No.</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>IF the No. of download is high and implemented language is Java, then the Web service has high performance.</td>
</tr>
<tr>
<td>2</td>
<td>IF implemented language is VB, then the Web service has low cost.</td>
</tr>
<tr>
<td>3</td>
<td>IF the No. of download is high, then the Web service has high reputation.</td>
</tr>
</tbody>
</table>

In this step, a set of decision rules was summarized as shown in Table 8, in accordance with Table 7. The meaning of the decision rules is interpreted in Table 9.

Phase II: Genetic Algorithm for Web services composition

Step 1 (Define the parameters)

1.1 Set the initial population size \( m \) to 300.
1.2 Set the generation \( k \) to 1.
1.3 Set the number of chromosome \( n \) as 5 (number of composited Web services)
1.4 Set the mutation rate to 0.001, crossover rate as 0.8 and termination condition as 500 generations.
1.5 Set the weights (Cost: 0.1; Functional Fit: 0.1; User Friendliness: 0.1; Flexibility: 0.6; Reputation: 0.1).

1.6 Select the rule: the final reduct rule in this example was applied to the genetic algorithm to reduce the domain range of the initial population. Since this company assigns a heavier weight to flexibility (0.6), the reduct rule 1 in Table 9 which says “IF the No. of downloads is high and implemented Language is Java, then the Web service has high performance.” was selected since the extent of flexibility can be measured by one of its metrics: efficiency [14].

1.7 Set the percentage of rule matching to 90%: To avoid finding a local optimal solution, weight analysis was used to make the rules fuzzy rather than absolute. Using the weight, the reduct rule 1 in Table 9 implies that 90% of the initial population selected was considered satisfactory.

Step 2 (Initialization)
This step aims to produce 500 populations, composed of five categories (A: Finance, B: Human Resource, C: Manufacturing, D: Procurement, and E: Distribution), each of which satisfies the reduct rule. Table 10 is a partial population that was obtained.

Step 3 (Evolution)
This step focuses on the GA evolution of the 500 generations. In this case, the proposed GA converges at the 70th generation, at which all the population is the same and the fitness functions are all equal to 1. The best, worst and average fitness (objective) converged at 1 as shown in Figure 4. The evolution process therefore was terminated. This example was converged at \(\{36,66,36,92,96\}\).

Step 4 (Detailed design)
The final result for this case study is summarized in Table 11. The enterprise user now can adjust the parameters to test and verify.

5.2. Experimental analysis

5.2.1. Experiment I – small sample
To verify the proposed method, the results were compared with those obtained by standard GA and exhaustive enumerations. The exhaustive enumerations are executed at local machine due to the limitation of GAE (A request handler has a limited amount of time to generate and return a response to a request, typically around 60 s. Once the deadline has been reached, the request handler is interrupted). Although the exhaustive enumerations can find all feasible solutions and the global optimum, they are computationally costly. In this case, experimental parameters are identical to those used in the previous example except that the number of candidate Web services component in each category is set to 50. The exhaustive enumerations in this study reached the global optimal composition in 222 s. Accordingly, the two methods (i.e., the standard GA approach and the proposed approach) were each run 1000 times. Table 12 shows the experimental results. The hit rate is the percentage of the global optimum obtained from the exhaustive enumerations. The proposed method has a much higher high rate.

5.2.2. Experiment II – large sample
In this case, experimental parameters are identical to those used in the previous example except that the number of candidate Web services components in each category is set to 100. Table 13 shows the proposed method has a much higher hit rate.
### 5.2.3. Experiment III – sensitivity analysis

Sensitivity analysis is often used to analyze how sensitive a system is with respect to the change of parameters [40]. A sensitivity analysis of the proposed GA method was carried out with three parameters, which are population size, crossover rate and mutation rate. These parameters largely determine the success and efficiency of a GA routine for solving a specific problem. 36 sets of parameters in total combining different crossover rates (0.7, 0.8, 0.85, 0.9), mutation rates (0.001, 0.0015, 0.005) and population numbers (100, 200, 300) were used to run the experiments and to obtain the hit rate. The findings are summarized below:

- The greater the population, the higher the hit rate as shown in Table 14.
- There is no significant difference in hit rate among different crossover rates as shown in Table 15.
- There is no significant difference in hit rate among different mutation rates as shown in Table 16.

The sensitivity analysis on the parameters shows that the user can recognize which parameter to focus on for the problem of interest. For example, the nature of the hit rate in this case study highly relates to the population number. Eiben et al. [10] conducted a comprehensive review and classification of parameter control methods for evolutionary algorithms. Yet, the selection of these control parameters is rather complex and needs further research, beyond the scope of this study.

### 5.3. Discussion

The experimental results show that the proposed composition method is significantly better than traditional GA as summarized in Tables 12 and 13. The execution time for the exhaustive enumeration method is much longer than the one for the proposed method. It indicates that application of constraints in the search process improves the solution quality. Specifically when the number of candidate Web services increased from 100 to 300, the time required for the exhaustive enumerations to find the global optimum rose exponentially to 22 days. Assuming the trend continues, exhaustive enumeration methods will not be able to find the optimal solution for a much larger number of Web services. The proposed composition method on the contrary is not subject to this limitation.

Johansson and Ruivo [24] explore vendor’s perspective on what factors affect adopting ERP as SaaS and found 10 factors: Costs, Security, Availability, Usability, Implementation, Ubiquity, Flexibility, Compatibility, Analytics and Best-practices. Costs, data security and system availability were perceived by the experts as the most important factors in customer perspective for adopting ERP systems in a SaaS delivery model. A SaaS provider should realize that a successful market establishment of its offer lays not so much on the product itself but on the kind of support given in the SaaS model and the customer experience with provided service. That is, the paradigm changes from product feature to service trust. Given the emerging trend of ERP delivered thru CloudERP platform, researchers can investigate other issues like privacy, customer-centric and type of firm capabilities. The real strength of cloud computing is that it is a catalyst for more innovation. In fact, as cloud computing continues to become cheaper and more ubiquitous, the opportunities for combinatorial innovation will only grow. It is true that this inevitably requires more creativity and skill from IT and business executives. In the end, this not something to be avoided. It should be welcomed and embraced [4].

### 6. Conclusion

This paper proposed a CloudERP platform and outlined a method for composing web services for ERP providers and enterprise users. This paper zooms in the selection process to propose a Web services composition method for Cloud platform providers in order to automatically customize an EPR service in response to an enterprise customer’s need. The proposed composition method makes use of the GA concepts and employs rules generated by the rough set. The modified GA-based composition method appears to operate effectively and promote both convergence and hit rate, according to the experiments. Toward developing a fully functional CloudERP platform, more research and development efforts are needed to refine the proposed composition process, especially with respect to the submission, assessment, publishing, and implementation activities.

### Acknowledgment

This work was partially supported by funding from the Ministry of Science and Technology of the Republic of China (NSC 99-2410-H-018-016-MY3; NSC 100-2918-I-018-001; MOST 103-2410-H-018-015-MY2).

### References


