Dear Author,

Please check your proof carefully and mark all corrections at the appropriate place in the proof (e.g., by using on-screen annotation in the PDF file) or compile them in a separate list. Note: if you opt to annotate the file with software other than Adobe Reader then please also highlight the appropriate place in the PDF file. To ensure fast publication of your paper please return your corrections within 48 hours.

Your article is registered as a regular item and is being processed for inclusion in a regular issue of the journal. If this is NOT correct and your article belongs to a Special Issue/Collection please contact j.aranha@elsevier.com immediately prior to returning your corrections.

For correction or revision of any artwork, please consult http://www.elsevier.com/artworkinstructions

Any queries or remarks that have arisen during the processing of your manuscript are listed below and highlighted by flags in the proof. Click on the ‘Q’ link to go to the location in the proof.

<table>
<thead>
<tr>
<th>Location in article</th>
<th>Query / Remark: click on the Q link to go please insert your reply or correction at the corresponding line in the proof</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1</td>
<td>AU: Please confirm that given names and surnames have been identified correctly.</td>
</tr>
<tr>
<td>Q2</td>
<td>AU: Please check the telephone number of the corresponding author for correctness.</td>
</tr>
<tr>
<td>Q3</td>
<td>AU: Please provide page range for references [119] and [127].</td>
</tr>
</tbody>
</table>

Please check this box or indicate your approval if you have no corrections to make to the PDF file.

Thank you for your assistance.
Towards automated composition of convergent services: A survey

Armando Ordóñez a,∗, Vidal Alcazar b, Oscar Mauricio Caicedo Rendon c, d, Paolo Facarind, e
Juan C. Corrales f, Lisandro Zambenedetti Granville c

a Intelligent Management Systems Group, University Foundation of Popayán, Calle 5 No. 8-58, Popayán, CA, Colombia
b Planning and Learning Group, Universidad Carlos III, Av. Universidad, 30, 28911 Leganés, MA, Spain
c Institute of Informatics, Federal University of Rio Grande do Sul, Av. Bento Gonçalves, 9500, Porto Alegre, RS, Brazil
d School of Architecture, Computing and Engineering, University of East London, 4-6 University Way, E16 2RD, London, UK
e Grupo de Ingeniería Telemática, Universidad del Cauca, Calle 5 No. 4-70, Popayán, CA, Colombia

A convergent service is defined as a service that exploits the convergence of communication networks and at the same time takes advantage of features of the Web. Nowadays, building up a convergent service is not trivial, because although there are significant approaches that aim to automate the service composition at different levels in the Web and Telecom domains, selecting the most appropriate approach for specific case studies is complex due to the big amount of involved information and the lack of technical considerations. Thus, in this paper, we identify the relevant phases for convergent service composition and explore the existing approaches and their associated technologies for automating each phase. For each technology, the maturity and results are analysed, as well as the elements that must be considered prior to their application in real scenarios. Furthermore, we provide research directions related to the convergent service composition phases.

© 2015 Published by Elsevier B.V.
operators from vigilance over routine and tedious tasks [8]. Several approaches have appeared to fully automate the composition of services available in the Web from both academia [9–14] and industry [15,16]. Similarly, some research has revolved around fully automation of communication service composition, coming mainly from European Projects [17,18]. Considering the evolution of service composition in the Telecom and Web domains, we raise the first research question addressed in this paper: Which is the level of automation that has been achieved to date from previous service composition approaches?

In addition to the aforementioned automation, it is important to consider that the service composition carried out in the Telecom fundamentally different of the conducted in the Web, which means that some technical issues must be contemplated when combining services from both domains [19,20]. Moreover, several other characteristics pertaining to specific application domains, like user preferences, response time, adaptability, and scalability must be taken into account. This is because the fact that convergent service composition often involves the challenge of merging two areas of knowledge with their respective concepts, technologies, and architectures. For this reason, a systematic way of analysing the strengths and weaknesses of each existing solution for a particular application domain is needed. In this way, we raise the second research question addressed in this paper: which elements must be considered for applying automated services composition in different specific domains?

Convergent service composition is also a complex process that often involves several elaborated steps or phases, from processing the input to selection of services to its execution. As these phases often focus on one or more aspects of the service composition process, it is needed to classify them. In this way, we raise the third research question addressed in this paper: which are the phases for convergent service composition and the approaches useful in each one of them?

The above raised research questions indicate that building up a convergent service is not trivial, because although there are significant approaches that aim to automate the service composition at different levels in the Web, telco networks, and computer networks, selecting the most appropriate approach(es) for specific case studies is complex due to the big amount of involved information and the lack of technical considerations. Thus, in this paper, we identify the relevant phases for convergent service composition and explore the existing approaches and their associated technologies for automating each phase. For each technology, the maturity and results are analysed, as well as the elements that must be considered prior to their application in real scenarios. Furthermore, we provide research directions in the convergent service composition phases.

The research conducted to answer the raised questions led to the following contributions:

- Identify the relevant phases for convergent service composition.
- Explore and analyse the existing approaches and their associated technologies for automating each phase.
- Provide research directions in each one of convergent service composition phases.

The remainder of this paper is organised as follows. In Section 2, we present the methodology followed to build up this survey. In Section 3, we review foundations about service composition. In Section 4, we introduce the phases for convergent service composition. In Sections 5–7, we describe and discuss each phase and its associated approaches. In Section 8, we present research directions. Finally, in Section 9, we provide some conclusions that help to select approaches for convergent service composition.

2. Methodology

To solve the research questions, we follow the next steps [21]: (i) formalize the service composition process by defining different phases, (ii) classify the existing approaches under the umbrella of requirements and criteria for their application in convergent composition; and (iii) define the guidelines for employing diverse approaches and their technologies in convergent service composition.

In this paper, we carry out a comparative literature review [22], aiming to identify the existing approaches in service composition in recent years through the lens of its applicability in convergent services. In order to select the works presenting the approaches useful for our study, we introduced search keywords on three electronic databases: ACM, IEEE, and Science Direct. As convergent composition is a relatively new field, we do not expect to find many publications using the exact keyword. Therefore, we used the following keyword phrases: “convergent service composition”, “automated service composition”, “Telecom 2.0 service composition”, and “mashup composition”. We found 250 works. Afterwards, we looked through the titles of such works to eliminate those not related to the field; for the rest, we browsed the abstracts to estimate their relevance and whether they have experimental results. We selected both recent works from recognized conferences and relevant ones from high impact journals. As a result, 65 works were selected for a full-text review.

Tables 1 and 2 present, in no order of relevance, the works selected for reviewing. These works are analysed according to: the origin (i.e., academy, industry, and from both in alliance), the domain (i.e., Web, Telecom, and convergent); deals with methodologies for convergent composition), the level of automation provided to users (i.e., high, medium, low; when users participate little in the process or needs to know little about technical issues, higher is the automation), the use of semantic in any of the phases above mentioned, and the result of the work (i.e., prototype, model, product, and standard). The detail of the literature review as well as the rest of
the analysis carried out in this paper are presented in the following chapters.

3. Foundations on service composition

In this section, we introduce the concepts of service composition and convergent service composition. Furthermore, we present the differences among the services offered in the domains of Telecom and Web as well as current approaches for describing such services.

3.1. What is a service?

The term “service” is a commonplace in many fields and therefore different definitions can be found in the literature: in the workflow area, the concept of e-service is commonly associated with the completion of a task [79]; meanwhile, in the Web, a Web service can be seen as a way to expose the functionality of an information system and make it available through standard Web technologies [80]. More precisely, the World Wide Web Consortium (W3C)—an organisation for developing Web standards—defines a Web service as a software system identified by a URI, whose public interfaces and bindings are defined and described using the eXtensible Markup Language (XML) or the JavaScript Object Notation (JSON). Such a service can be discovered by other software systems which, in turn, may then interact with the Web service in the manner prescribed by its definition, using XML or JSON based messages conveyed by Internet protocols [81]. Usually, in computer science, most authors refer to both “e-services” and “Web services” by simply “services” [82].

On the other hand, the term “service” in the Telecom domain can be understood as the capability to exchange information through a telecommunication medium, provided to a user by a service provider [83]. Different standardisation associations in this domain have

### Table 2

Works selected for reviewing.

<table>
<thead>
<tr>
<th>Work</th>
<th>Phase</th>
<th>Origin</th>
<th>Domain</th>
<th>Automation</th>
<th>Semantic</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>[37]</td>
<td>Creation</td>
<td>Both</td>
<td>Convergent</td>
<td>Medium</td>
<td>✓</td>
<td>Prototype</td>
</tr>
<tr>
<td>[38]</td>
<td>All except adaptation</td>
<td>Academic</td>
<td>Convergent</td>
<td>High</td>
<td>✓</td>
<td>Prototype</td>
</tr>
<tr>
<td>[39]</td>
<td>All except adaptation</td>
<td>Academic</td>
<td>Convergent</td>
<td>High</td>
<td>✓</td>
<td>Prototype</td>
</tr>
<tr>
<td>[40]</td>
<td>Creation</td>
<td>Industry</td>
<td>Convergent</td>
<td>Low</td>
<td>✓</td>
<td>Standard</td>
</tr>
<tr>
<td>[41]</td>
<td>Creation</td>
<td>Industry</td>
<td>Web</td>
<td>Low</td>
<td>✓</td>
<td>Standard</td>
</tr>
<tr>
<td>[42]</td>
<td>Creation</td>
<td>Industry</td>
<td>Web</td>
<td>Low</td>
<td>✓</td>
<td>Standard</td>
</tr>
<tr>
<td>[43]</td>
<td>Creation</td>
<td>Industry</td>
<td>Telecom</td>
<td>Medium</td>
<td>✓</td>
<td>Prototype</td>
</tr>
<tr>
<td>[44]</td>
<td>Creation</td>
<td>Both</td>
<td>Convergent</td>
<td>Medium</td>
<td>✓</td>
<td>Prototype</td>
</tr>
<tr>
<td>[45]</td>
<td>Execution</td>
<td>Academic</td>
<td>Convergent</td>
<td>n/a</td>
<td>✓</td>
<td>Prototype</td>
</tr>
<tr>
<td>[46]</td>
<td>Execution</td>
<td>Academic</td>
<td>Convergent</td>
<td>n/a</td>
<td>✓</td>
<td>Prototype</td>
</tr>
<tr>
<td>[47]</td>
<td>Creation</td>
<td>Industry</td>
<td>Web</td>
<td>n/a</td>
<td>✓</td>
<td>Standard</td>
</tr>
<tr>
<td>[48]</td>
<td>Execution</td>
<td>Academic</td>
<td>Convergent</td>
<td>n/a</td>
<td>✓</td>
<td>Prototype</td>
</tr>
<tr>
<td>[49]</td>
<td>Adaptation</td>
<td>Academic</td>
<td>Convergent</td>
<td>High</td>
<td>✓</td>
<td>Prototype</td>
</tr>
<tr>
<td>[50]</td>
<td>Adaptation</td>
<td>Academic</td>
<td>Convergent</td>
<td>High</td>
<td>✓</td>
<td>Prototype</td>
</tr>
<tr>
<td>[51]</td>
<td>Creation</td>
<td>Industry</td>
<td>Convergent</td>
<td>Medium</td>
<td>✓</td>
<td>Prototype</td>
</tr>
<tr>
<td>[52]</td>
<td>Adaptation</td>
<td>Academic</td>
<td>Convergent</td>
<td>High</td>
<td>✓</td>
<td>Prototype</td>
</tr>
<tr>
<td>[53]</td>
<td>Adaptation</td>
<td>Academic</td>
<td>Convergent</td>
<td>High</td>
<td>✓</td>
<td>Prototype</td>
</tr>
<tr>
<td>[54]</td>
<td>Creation</td>
<td>Academic</td>
<td>Convergent</td>
<td>Medium</td>
<td>✓</td>
<td>Prototype</td>
</tr>
<tr>
<td>[55]</td>
<td>Execution</td>
<td>Industry</td>
<td>Convergent</td>
<td>n/a</td>
<td>✓</td>
<td>Prototype</td>
</tr>
<tr>
<td>[56]</td>
<td>Execution</td>
<td>Industry</td>
<td>Convergent</td>
<td>n/a</td>
<td>✓</td>
<td>Prototype</td>
</tr>
<tr>
<td>[57]</td>
<td>Execution</td>
<td>Industry</td>
<td>Convergent</td>
<td>n/a</td>
<td>✓</td>
<td>Prototype</td>
</tr>
<tr>
<td>[58]</td>
<td>Execution</td>
<td>Industry</td>
<td>Telecom</td>
<td>n/a</td>
<td>✓</td>
<td>Prototype</td>
</tr>
<tr>
<td>[59]</td>
<td>Execution</td>
<td>Industry</td>
<td>Convergent</td>
<td>n/a</td>
<td>✓</td>
<td>Prototype</td>
</tr>
<tr>
<td>[60]</td>
<td>Execution</td>
<td>Industry</td>
<td>Convergent</td>
<td>n/a</td>
<td>✓</td>
<td>Prototype</td>
</tr>
<tr>
<td>[61]</td>
<td>Execution</td>
<td>Industry</td>
<td>Communication</td>
<td>n/a</td>
<td>✓</td>
<td>Standard</td>
</tr>
<tr>
<td>[62]</td>
<td>Execution</td>
<td>Industry</td>
<td>Convergent</td>
<td>n/a</td>
<td>✓</td>
<td>Standard</td>
</tr>
<tr>
<td>[63]</td>
<td>Execution</td>
<td>Industry</td>
<td>Convergent</td>
<td>n/a</td>
<td>✓</td>
<td>Standard</td>
</tr>
<tr>
<td>[64]</td>
<td>Generation</td>
<td>Academic</td>
<td>Web</td>
<td>High</td>
<td>✓</td>
<td>Prototype</td>
</tr>
<tr>
<td>[65]</td>
<td>Generation</td>
<td>Academic</td>
<td>Web</td>
<td>High</td>
<td>✓</td>
<td>Prototype</td>
</tr>
<tr>
<td>[66]</td>
<td>Generation</td>
<td>Academic</td>
<td>Web</td>
<td>High</td>
<td>✓</td>
<td>Prototype</td>
</tr>
<tr>
<td>[67]</td>
<td>Generation</td>
<td>Academic</td>
<td>Web</td>
<td>High</td>
<td>✓</td>
<td>Prototype</td>
</tr>
<tr>
<td>[68]</td>
<td>Generation</td>
<td>Academic</td>
<td>Telecom</td>
<td>High</td>
<td>✓</td>
<td>Prototype</td>
</tr>
<tr>
<td>[69]</td>
<td>Generation</td>
<td>Academic</td>
<td>Web</td>
<td>High</td>
<td>✓</td>
<td>Prototype</td>
</tr>
<tr>
<td>[70]</td>
<td>Generation</td>
<td>Academic</td>
<td>Web</td>
<td>High</td>
<td>✓</td>
<td>Prototype</td>
</tr>
<tr>
<td>[71]</td>
<td>Generation</td>
<td>Academic</td>
<td>Web</td>
<td>High</td>
<td>✓</td>
<td>Prototype</td>
</tr>
<tr>
<td>[72]</td>
<td>Creation, generation</td>
<td>Academic</td>
<td>Web</td>
<td>High</td>
<td>✓</td>
<td>Prototype</td>
</tr>
<tr>
<td>[73]</td>
<td>Generation</td>
<td>Academic</td>
<td>Web</td>
<td>High</td>
<td>✓</td>
<td>Prototype</td>
</tr>
<tr>
<td>[74]</td>
<td>Creation, execution</td>
<td>Academic</td>
<td>Convergent</td>
<td>Medium</td>
<td>✓</td>
<td>Prototype</td>
</tr>
<tr>
<td>[75]</td>
<td>Creation, execution</td>
<td>Academic</td>
<td>Convergent</td>
<td>Medium</td>
<td>✓</td>
<td>Prototype</td>
</tr>
<tr>
<td>[76]</td>
<td>Creation, execution</td>
<td>Academic</td>
<td>Web</td>
<td>Medium</td>
<td>✓</td>
<td>Prototype</td>
</tr>
<tr>
<td>[77]</td>
<td>Creation, execution</td>
<td>Academic</td>
<td>Web</td>
<td>Medium</td>
<td>✓</td>
<td>Prototype</td>
</tr>
<tr>
<td>[78]</td>
<td>Creation, execution</td>
<td>Academic</td>
<td>Web</td>
<td>Medium</td>
<td>✓</td>
<td>Prototype</td>
</tr>
</tbody>
</table>
created similar definitions. One of the most important associations for standardisation in Telecom is the Tele-Management Forum (TMF) that describes a service as a set of independent functions being an integral part of one or more business processes. This functional set consists of the hardware and software components as well as the underlying telecommunication medium. The user sees all these components as a fusion; a service can be a service component of another service [84]. Other association of standards for mobile technologies is the 3rd Generation Partnership Project (3GPP). The 3GPP defines a service a component of the portfolio of choices offered by service providers to a user, in the sense of functionalities provided to a user [15].

Basic Telecom services (i.e., networking services) provide connectivity to a network, like voice over the Internet Protocol (IP) and basic signalling, for example. Added-value services, a broader concept of service, provide other functionalities such as: voice/video connectivity, community tools, presence, conferencing, gaming, and TV broadcasting. The added-value services are based on NGN [82], whose purpose is to deploy next generation architectures that allow the access to independent services over converged fixed and mobile networks.

To sum up, the fundamental objective of services is to have a collection of network-resident software services accessible via standardized protocols, whose functionality can be discovered, used individually or composed to form more complex services [79].

### 3.2. How can a service be described?

As stated before, an atomic service consists of a set of network-accessible functionalities described using a machine understandable description. The set of methods that allow accessing these functionalities is known as the interface of the service. The description of service interfaces will be used afterwards for service discovery, compatibility, verification, and composition. The first step to take when describing a service is to analyse its inputs and outputs; Fig. 1 shows diagram of the input and output of a service that performs the sum of two integers. Currently, such description depends on the architectural style (i.e., the Service Oriented Architecture—SOA—and the Representational State Transfer—REST) and the technologies used for implementing the service.

The Web Services Description Language (WSDL) [85] is the fact standard for describing services based on SOA that are usually implemented using the Simple Object Access Protocol (SOAP) and XML. WSDL [86] provides a computer-understandable description of Web services and describes Web services as collections of communication endpoints that can exchange certain messages. This description language includes atomic method calls (also called operations) in terms of input and output messages where each operation can be a one-way operation, request-response, solicit-response or notification.

Listing 1 depicts an example of WSDL. The code starts with the definition of the namespaces and the definition of messages to exchange (i.e., SayHelloRequest and SayHelloResponse). The portType includes the allowed operations (i.e., SayHello). The Binding (i.e., greeting_binding) specifies a particular protocol and the encoding style for provision of the service. Finally the service element (i.e., greeting_service) includes information about the concrete address of the service.

The services based on SOAP and XML, there is still not a widely accepted standard for describing services based on the architectural style REST that are usually implemented using the HyperText Transfer Protocol (HTTP) and JSON or XML. Relevant efforts for describing REST-based services are: (i) the Web Applications Description Language (WADL) [87] that offers a machine-processable protocol description format for services using HTTP and XML, (ii) the WSDL 2.0 [88] that provides a machine-readable description for SOAP and HTTP based services by accepting the binding to all the HTTP request methods (not only GET and POST); and (iii) the HTML for REST (hREST) [89] that offers a XHTML-based microformat for representing and describing Application Program Interfaces (APIs) and services based on REST.

The second level of service description is concerned with pre- and post-conditions. As aforementioned, description solutions, such as WSDL (1.1 or 2.0), WADL, and hREST, describe a single interaction with a service in terms of request and response messages. However, to assess whether a service can be used by other services, it does not suffice to know the type of data generated or received by a service, but also the change in the world that it produces. Pre-conditions must be true in the input in order to the service to be executed; at the same time, post-conditions are conditions that become true after the service is successfully executed. For example, a service called send_message may require as input a well-formed email address and may have the effect (post-condition) of informing someone of something, which represents the change in the world.

Fig. 2 depicts the service pay that has as input the payment value and the payment data (e.g., card number and customer) as well as an additional pre-condition. Such precondition ensures that the available credit is higher than the amount to pay. A post-condition for the service pay may be to change the state of the transaction to payed, for instance. Some languages, such as the Web Ontology Language for Services (OWL-S), the Semantic Annotations for REST (SA-REST), and the micro Web Service Modelling Ontology (microWSMO) [89,90], allow the annotation of services with pre- and post-conditions apart from the inputs and outputs. It is important to highlight that currently any of these languages is dominant in the second level of service description.

### 3.3. What are the differences between services available in the Web and Telecom domains?

Telecom services are fundamentally different from Web services based on SOA or REST [17]. Next the most remarkable differences are analysed.

**Standard language for service description.** In the Web domain, for services based on SOAP, there is a large consensus on the main standard for service description: WSDL. Unlike, for services based on REST, there is not a widely accepted/standard accepted for describing them. In the same direction of REST-based services and contrary to SOAP-based services, in spite of some initiatives such as ONEAPI [61] and Twilio [58] for standardising service interfaces, there is no standard language for service description in the Telecom domain. A review of proposals shows that service description languages in Telecom are often protocol specific, not extensible nor reusable [53].

**Level of detail of services.** In the Web, the details of atomic services are mostly of the time open to the world by means of REST and SOAP APIs, widgets, and Web front-ends. That is to say that the users are often aware of the parameters of the demanded services. Two prominent examples of highly parameterized Web services whose APIs are public for example Google Translator and Google Maps. On the other hand, Telecom operators offer services without authorising the user for configuring them in deep or even note that he/she is working with them. For example, services such as call forwarding or billing processes are often completely opaque to users.

---

Please cite this article as: A. Ordóñez et al., Towards automated composition of convergent services: A survey, Computer Communications (2015), http://dx.doi.org/10.1016/j.comcom.2015.07.025
Availability of services. Availability refers to the ability to access a system or service. In the case of Web services provided through the Internet, the availability depends highly from the underlying networks. Conversely, the percentage of availability of Telecom services is very high, and terms such as $5 \times 9$ (i.e., 99.999% availability) and fault tolerance are commonly used [91].

Time constraints. In the Telecom domain, there are tight real-time requirements in protocols and user requirements (e.g., post-dial delay is typically bounded [19]). In contrast, a best-effort response time is typically offered by services available along the Web.

Number of protocols. In the Web most of the communications are based in synchronous protocols, mostly in HTTP. Conversely, in Telecom, there is a broad range of protocols, each one specific for some services (e.g., voice transfer and video) or networks (e.g., Wi-Fi, mobile, and fixed). In spite of the efforts for integrate all the protocols over the Session Initiation Protocol (SIP), that allows phone calls to be made over the Internet [92], current technologies require the implementation of diverse protocols in order to be able to be integrated with legacy hardware and software.

3.4. What is service composition?

In a broad sense, the process of creating a composite service by combining atomic services is known as “service composition”. This composition is required when the request of a client cannot be satisfied by any single service. Due to the differences in standards, service descriptions and formalisation of the involved processes, service composition is subsequently described separately for the Web and Telecom domains.

Service composition in the Web domain. According to the previously given description, service composition in the Web consists of joining the individual services and defining the control flow and the data flow, similar to workflow approaches. The resulting composition can be viewed from two perspectives: orchestration and choreography. Orchestration describes the executable process that can interact through messages with both internal and external Web services.
Orchestration includes business logic and task-execution orderings from one partner perspective.

Choreography tracks the sequences of messages among multiple partners rather than in a specific process executed by a single partner. Choreography is more collaborative that orchestration, and allows to describe the public message exchange between services. Also, choreography is useful for establishing the rules for the exchange of messages in order to guarantee interoperability among different agents. In summary, orchestration describes an executable process while choreography specifies the protocol of peer-to-peer interactions. Fig. 3 illustrates the relationship among both concepts.

The main language for describing the composition of SOAP-based services is the Web Services-Business Process Execution Language (WS-BPEL) [47], often shortened to BPEL in many works. Business process is a term used to describe the interaction between services in terms of work and data flow. BPEL is an orchestration language that models business processes that interact with external parties. In BPEL each part is a service described using WSDL. BPEL includes representation in XML of concepts such as data, control flow, data flow, error handling and partners among others.

Listing 2 presents an example of a whole process in BPEL. In this example, a process called AmericanAirline defines the partnerLinks (services involved in the process) as well as the variables of the process. The BPEL-based workflow of process is a sequence of tasks that includes: receive, assign, and invoke (make a call-back). It is important to note that the element invoke includes parameters related to the service description defined in the WSDL such as operation and portType.

BPEL does not include any graphical representation for business processes, so other standards have appeared to provide such functionality. The most used notation for graphical representation of business processes in service composition is the Business Process Model and Notation (BPMN) [93]. BPMN is a standard that provides a graphical notation for specifying business processes and is very similar to the activity diagrams of the Unified Modelling Language (UML) [94]. The goal of BPMN is offering a representation tool for both technical and business users. To do so, BPMN provides a notation that is intuitive to business users, yet able to represent complex process semantics. The BPMN specification also provides a mapping to particular languages such as BPEL. Fig. 4 depicts an example of modelling with BPMN.

Unlike the SOAP-based services, for composing REST-based services there is still no a language standardized or broadly accepted. Significant researches are: (i) BPEL for REST [95] proposes an extension of BPEL that covers the composition of Web Services based on SOAP and REST by adding specialized activities as < get >, < post >, and < resource >, (ii) Bite [96] presents a lightweight and extensible composition language to create workflows where the main entities of composing are REST-based services; and (iii) The Lightweight Process Modelling Language (LPML) [97] allows to model a business process, that uses Web 2.0 solutions, by means a subset of BPMN.

On the other hand, regardless of the method used to represent the composite service, service composition in the Web is commonly performed at design time. In this sense, recently, non-technical users have begun to play an active role in service composition, especially by creating their own services, called mashups, in the Web 2.0. Mashups are composite Web applications focused on end-users and created by combining different resources available along the Web [98,99]. In this sense, MDEs like Yahoo Pipes (see Fig. 5) allow users to create their own services by assembling service building blocks.

MDEs work with a rather limited set of services in comparison with all ones available on the Web, but at least offer some degree of personalisation/customisation to users. Still, service composition made by users has some problematic issues such as the heterogeneous, dynamic, and spontaneous nature of their needs [37], which current development environments directed towards them cannot address yet.

Service composition in the Telecom domain. In Telecom, service composition is known as feature application composition [19] that is as a process inherent to the establishment of any communication. For instance, a basic call service between two subscribers involves multiple features, such as voicemail, call forwarding, and call waiting. All these features are invoked in a sequence even if the user is never aware of the feature composition.

Similarly to the individual service description, in the Telecom domain, there is a shortcoming related to the little consensus about standard descriptions for service composition. In recent years, the TelcoML [40] and DiaGen [100] were proposed for coping with such shortcoming. TelcoML has been proposed by the Object Management Group (OMG) [40], offering a UML Profile for designing advanced and integrated telecom services. In turn, DiaGen is a framework based on the implementation of a declarative language over Java that provides service discovery and high-level communication mechanisms. Such a framework also generates the class skeletons of composite services.

On the other hand, some standards, such as the IP Multimedia Subsystem (IMS) [62], Parlay X [101], and OneAPI [61], aim to carry out composition of individual Telecom services into new more
complex services, leveraging the reuse and standardisation of sub-

cent technologies [19]. In general terms: (i) IMS can be seen as an
architectural framework for delivering IP multimedia services by per-
mitting the integration of services based among others on SOAP, REST,
and SIP, (ii) Parlay X introduces the concept of Telco Web services to
allow the creation of composite services; and (iii) ONEAPI exposes

network facilities as services, aiming to facilitate the combination of
these facilities with services from the Web domain.

Like service composition in the Web, the service composition in
the Telecom domain may also be carried out by using SCEs [4] and
MDEs [5,6]. These environments allow the user to combine services
functionalities using drag-and-drop and wire tools that hide the tech-
nical details of technologies like IMS, Parlay X, and ONEAPI. Fig. 6 de-
picts a screenshot of Radisys [43] that is an example of a typical SCE.
In the user interface of Radisys, the functionalities to be composed
graphically in the central canvas are located at the left pane.

3.5. What is convergent service composition?

The term “convergent” has been used widely in telecom and com-
puter networks with different meanings: to describe the fusion of
fixed and mobile services [2], or as a synonymous of the NGN that
represents the integration of the Internet and the Public Switched
Telephone Network (PSTN) [3]. As previously stated, we use the term
“convergent” as efficient coexistence of services available on the Tele-
com and Web domains to provide innovative and more complex ser-
vices to the user.

The rise of composition of services appears in the midst of a grow-
ing expectation of traffic volume by telecom operators. This is due
to the increasing offer of added-value services to customers in addi-
tion to traditional voice and data communication [102]. Novel added-
value services combine features from the Telecom with technologies
of the Web; this combination is known as “convergent composition”
[23] or “unified composition” [19].

To illustrate what the process of convergent composition is, we
present an example in the Early Warning Management domain in
which an environmental manager must make decisions regarding the
environmental alarms of a region. He/she can request information
from sensor networks deployed at several spots and can also
use services available on a telecommunication network and the Web
to process basic data and send information to both farmers and

Fig. 4. Simplified example of a process modelled using BPMN.
Fig. 5. Service composition using Yahoo Pipes.

Fig. 6. The Radisys SCE [43].

Please cite this article as: A. Ordóñez et al., Towards automated composition of convergent services: A survey, Computer Communications (2015), http://dx.doi.org/10.1016/j.comcom.2015.07.025
Fig. 7. Architecture for management of environmental early warnings.

As the environmental manager may come from fields like biology or agriculture, his/her technological background may be low. Also, electronic devices in the area may be scarce and not reliable or obsolete. Furthermore, the data network coverage may be deficient too. Thus, the preferred way to enter information to the system is by voice and in natural language. This way, users do not have to know how the system internally works and can make requests from regular mobile devices or landline phones. Commonly, the user expresses his/her request informally. Here are two examples of requests in natural language: (i) “I need to compute the hydrological balance of zone 1”; (ii) “If the river flow of zone 2 is greater than 15% of the safety limit on average, send a warning to every farmer within a radius of 2 miles from the river and create an action map including emergency and rescue groups”.

In the first request, a service composition could be: gathering information from the sensors in zone 1, using hydrological services from the Internet to process the sensed data, obtaining the hydrological balance map from Google maps, composing the final image, and sending it to the user by the Multimedia Message Service (MMS).

In the second request, a service composition could be: sensor data are evaluated and, if necessary, an emergency map is created. The map is created drawing a radius of 2 miles around each sensor reporting danger. To do so the system may search for geographical services available in the Web. Finally, the system informs about the alarm to the farmers inside the emergency area. The most convenient way to send the information is selected: SMS, Cell Phone call, fixed telephone call, or voice message.

In both request examples, services from the Web and Teleco are used for attending user request. Additionally, the context of the user for both farmers and managers must be considered in order to get a service composition that fulfills the needs of the user adequately. For instance, if the user is currently driving then a service could automatically provide information through voice, such as SIRI [16] and Vlingo [103], instead of text. Another way of taking into account the context of the user is allowing services that can be triggered under specific conditions; for example, it might be possible to determine whether a rescue vehicle is close to an affected area, in which case an emergency notification would be triggered [102].

4. Convergent service composition

In a general sense, the process for carrying out convergent service composition can be decomposed in several relatively independent steps, beginning with the user request and ending with the execution of the composite service. The whole process (see Fig. 8) can thus be divided into four phases: creation, generation, adaptation, and execution.

An important consideration to be noted is that the automation of each phase is independent of the automation of the rest of the other phases. This means that the automation of the whole service composition can be described as the sum of the automation of its parts. Next, each phase is described as well as the automation level expected.

4.1. Creation phase

In this phase, the user (advanced or novice) specifies the user request (i.e., ru) in a natural way in order to be solved by the system (see Fig. 8, phase a). The user specifies ru through a device or computer that hides any technical (platform or language) detail and allows him to clearly state his necessities with his own words. Such ru needs to be translated to a machine understandable representation (i.e., rm). An example of the input and output of the creation phase may be: ru: "I want to receive the route to the closest ATM in my cell phone" and rm: get("calMap","myPosition","ATM").

In the creation phase, the automation is associated with the capability of the system to extract rm from ru through a device or computer that hides any technical (platform or language) detail and allows him to clearly state his necessities with his own words. Such ru needs to be translated to a machine understandable representation (i.e., rm). An example of the input and output of the creation phase may be: ru: "I want to receive the route to the closest ATM in my cell phone" and rm: get(“calMap”, “myPosition”, “ATM”).

4.2. Generation phase

The next step in the process for convergent service composition consists on translating the machine understandable request rm into
4.4. Execution and reconfiguration phases

The execution phase (see Fig. 8, phase d) is responsible for managing the lifecycle of CS. Therefore, in this phase, the convergent service is deployed into an execution environment (e.g., a server) to be subsequently executed. The execution phase is inherently automatic: once the user deploys the service, it runs in the execution environment without human intervention.

The reconfiguration phase is required, for instance, when a malfunction (e.g., server failure) or an error (e.g., a Web service becomes offline) happens during the execution of CS. The usual procedure for reconfiguring a convergent service consists in notifying the developer/user, who must go to the corresponding SCE/MDE for re-examining the whole composition process.

Complete automation of this phase requires the composite service to be monitored and reconfigured automatically in case of failure [105].

5. Service creation

As afore defined, service creation is the process of translating the user request to a language or descriptor understandable by the computer. For instance, the translation may be done from the request expressed in natural language to BPEL/Bite language or to executable code. This phase depends largely on the level of expertise and the device used by the user. If the user is an expert and uses a desktop personal computer, a graphical interface for creating workflows is very suitable. On the contrary, the same interface is not suitable for a non-expert user, or for another user who may be driving when needs to compose a service for satisfying his/her needs.

In terms of academic research, a clear separation between programmers/developers (a.k.a, professional users) and end-user needs to be done according to the conclusion of European SPICE [31] and OPUCE [27] projects. In this way, in the following subsections, the main approaches for service creation will be exposed considering focus on professional users (i.e., SCE-based approaches) and end-users (i.e., approaches using natural language processing—NLP—and mashup-based approaches). Most of them comes from academic initiatives.

5.1. User-driven service creation (U-D)

Users can request the composition of an abstract service from a pool of atomic services. The concept of user-driven service creation (U-D), as shown in [105], is an interruption of the whole lifecycle of composition: the service execution is the only phase that is carried out in a self-automated manner.

5.2. Template-based approach

This is a simple and relatively easy approach in which an atomic service is composed to another atomic service using a template. In this approach, the template contains predefined services as well as the information describing how these services need to be composed.

5.3. Mashup-based approach

This approach is similar to template-based approach but in a more general form. In this case, the template is created by the user (or experts) and represents a business process, including all the atomic services to be included in a composite service.

Listing 3. Example of Java executable composite service.

public class CompositeService{
    public void Request()
        { //calculate coordinates of cash machine and my position
            here = Locator.getActualPosition();
            cash_machine = Locator.getPosition(Gazeter.get(here,
                cash_machine));
            Location my_position = Locator.getLocation(here);
            Location destination = Locator.getLocation(cash_machine);
            //Generate the map in an image
            MapImage map = MapService.render(my_position, destination);
            //Send the message to the user
            TelcoApi.sendSMS(user.getPhoneNumber, map, "result");
        }
}
5.1. SCE-based approaches

SCEs ease the service creation by offering mechanisms for integrating services and functionalities graphically. This idea was initially suggested by researchers that proposed the use of Integrated Development Environments (IDE) able to visually integrate atomic components representing Telecom and Web 2.0 services [30].

The SCE proposed by the European OPUCe project offers intuitive tools for graphical creation of composite services [26, 27, 30]. This SCE allows the integration of Web and Telecom services taking into account some user profile criteria. Furthermore, the SCE of OPUCe includes a semantic browser and a translator to BPEL.

HSCE is other research work that carries out SCE [28]. This work apart from providing a low latency execution environment for convergent services also offers a template-based service creation platform.

An important shortcoming of HSCE is that it is based on BPEL templates where most of the design tasks must be done manually, which makes it ill-suited for automatic service composition.

Another SCE-based approach was proposed in the field of the Artificial Intelligence (AI) [11]. Such approach aims at automating the Web Services composition under semantic awareness by AI planning techniques and a graphical environment in which even non-expert users (i.e., end-users) can create composite services. One of the most striking characteristics of this approach is that the expertise required by non-experts for carrying out composition activities is essentially similar to discover and invoke an atomic Web Service.

In the TeamCom project arose other SCE-based approach that describes the development of added-value services through a tool based on BPEL [53]. This tool additionally allows to translate to Java and deploy on a Telecom execution server the developed added-value services.

Finally, regarding SCE-based approaches, it is important to highlight that in recent years, the concept of Service Delivery Frameworks has also been proposed to provide service creation in convergent scenarios [25]. We consider these frameworks just like other form to refer a solution using SCEs.

5.2. Mashup-based approaches

As an alternative to SCEs, an important trend of research has emerged around the mashup technology that allows creating solutions, named mashups. These mashups are composite Web applications centred in end-users and created by combining different resources available along the Web [98, 99]. End-user centric means that mashups may be developed by users who usually do not have advanced programming skills [37], leveraging MDEs that tend to be user friendly and hide technical details [34]. Furthermore, mashups encourage the cooperation among end-users and the reuse of existing Web applications [106, 107].

The proponents of the use of mashups for creation of convergent services promote the generation of mechanisms for service composition from both Web and Telecom using friendly interfaces. This is done following the Web mashups philosophy based on the fact that a large number of services are accessible over the Web through open interfaces.

An example of mashup-based approach for service composition is a framework that introduces the concept of mobile mashups [33]. This framework proposes a device-independent architecture applicable to both the server and the device side that allows integrating device capabilities (e.g., camera and Bluetooth) with Telecom features and Web accessible services.

MARIO [51] is a tool that implements a “pursuit of desires” pattern, in which users set their preferences by selecting tags describing services to which they want to subscribe. The internal composer of MARIO makes the correspondence between the tags and the existing services. Such correspondence is in turn refined by the user until he/she gets the final composition.

In the domain of added-value convergent services, a reference architecture was proposed to facilitate the provisioning of telco mashups for end-users [20]. A telco mashup is a composite service that combines functionalities from telecommunication networks (e.g., streaming, quality of service, and billing) and device capabilities (e.g., geographic location and health information) with services (e.g., instant messaging and IP telephony) accessible via Web-based protocols. In the literature, we did not find a prototype of a mashup system or a telco mashup that implements this reference architecture.

In the network management domain, a mashup-based approach was proposed to deal with the security problem of botnets, in an attempt at providing an easily extensible, and usable way [74]. This approach was carried out in a mashup application. The mashup built (called botnet mashup) integrates dynamically botnet information collected from existing mitigation tools, such as sandboxes and antiviruses, with geographic location retrieved from online mapping and geolocation APIs.

A mashup system was also introduced to evaluate qualitatively the feasibility of using Web 2.0 technologies on network management [108]. Using the proposed system prototype, a mashup (called BGP mashup) was built in order to monitor the traffic of the Border Gateway Protocol (BGP) among two autonomous systems by integrating traffic router information. This integrated information was collected by using the Simple Network Management Protocol (SNMP) and presented by combining images retrieved from a generator of traffic graphs and maps generated from a mapping service.

As a continuation of the two works above cited, a generic architecture was proposed to support the composing of network management applications [5]. This architecture was carried out in a MDE that allows network administrators to design their own management applications through the composition of network management and services available in the Web. The qualitative evaluation of the proposed architecture was conducted by analysing the development of the BGP mashup and the botnet mashup previously referred.

As an evolution on using mashups for network management arose the concept of mashments [13, 14]. Mashments are solutions based on the Situation Management (SM) discipline and the mashup technology that aim to facilitate the work of network administrators when facing unexpected, heterogeneous, and dynamic situations. To achieve such aim, mashments combine diverse types of resources (data, application logic, and user interfaces) from multiple providers (from the Web and the network management) and automate the investigative and control aspects of SM.

An important shortcoming of the mashup-based approaches above cited is that the selection of atomic services used for creating mashups is performed manually by using mainly searches based on keywords. In this sense and with the goal of overcoming such shortcoming, in the literature of service composition arose diverse and relevant semantic mashup approaches. For instance, SensorMashup is a semantic-based platform aimed to facilitate the creation of mashups by following the linked open data principle [76] that enables the dynamic discovery, integration, and querying of heterogeneous data sources. This platform provides a visual composer that allows end-users to easily publish sensors as data sources identified by URIs and to rapidly combine/link such sensors for building up mashups.

In service composition, other groundbreaking approach was proposed for improving the process of developing mashups by adding semantic service discovery capabilities [77]. In a general way, these capabilities were added by combining semantic Web Services discovery frameworks, Universal Description Discovery and Integration (UDDI) repositories, and existing mashup tools.

In the Cloud Computing domain, a semantic-based approach was proposed to make the engineering process of creating mashups more agile and flexible [78]. In particular, this approach uses the Cloud...
Computing technology to increase the number of public REST APIs and further introduces the Semantic Annotation for Web Application Description Language (SAWADL) to carry out the *semantisation* of such APIs. SAWADL is a WADL extension that enables the annotation, combination, and deployment of REST APIs, aiming to facilitate the development of mashups in a cloud environment.

Instead of relying on the expertise of the user as the above mentioned mashup-based approaches, other ones focus on capturing the user request, defined in a specific/proprietary description language of composition, and translating it into composite services. In the approaches using these specific or proprietary languages, the composite service description made by the user is translated into a process definition language (e.g., BPEL and/or BPEL for REST) before being executed. For instance, in the SOA4All project [97], non-expert users interact with a graphical process editor called SOA4All Studio to create new composite services. This editor directly interacts with a reasoning engine that creates a composed service by inferring in a semantic way the most suitable connection among parameters of the services exposed on it.

### 5.3. NLP-based approaches

Recent approaches have also explored the application of NLP techniques to automate the composition of services. These approaches have been used mostly in intelligent home environments [35,36] and offer mechanisms to map words to basic functionalities of the system.

The goal of a NLP-based system is to extract the workflow of the composite service from the user request expressed in natural language.

In the domain of environmental early warnings, a NLP-based system named AUTO was proposed for composite service creation [38,39]. In AUTO, the user defines his/her demand and context information using natural language. This demand is handled by a request processing module that is responsible for transforming natural language and context information into a plan of composite service. This plan combines services from the Web and Telecom domains.

In the SPICE project, a specialized language called SPATEL [31] was proposed to describe service compositions formed by RESTful services (i.e., RESTful is a Java implementation of the REST architectural style) and Telecom features. These descriptions are interpreted by a translator module [32] that is responsible for generating code for specific execution technologies. SPATEL and the translator module are the core of the generation and execution of composite services offered by the SPICE platform.

### 5.4. Comparison

After surveying the previous approaches, it can be stated that the works in the service creation phase can be classified in three groups: SCE-based, mashup-based, and NLP-based. Each one of them has different particular elements that make them suitable for distinct environments. Table 3 summarizes the differences between the three approaches.

SCE-based approaches offer an incomparable technical detail for convergent service composition. When employing SCEs, the limit on the intricacy of the service composition depends only on the number of available services and the knowledge of the user. However, this obviously requires the user to be an expert in the specific application domain. Furthermore, because the complexity and nature of SCEs, usually they must be installed in a desktop personal computer, which greatly restricts the mobility of the user.

SCE allows to create convergent services with a remarkable level of detail, these interfaces allow to customize different aspects of the services and the subjacent protocols. Unlike Web services, the protocol in telecom environments is very important for the provision of the service and may directly benefit the quality of services. Therefore SCE might be a very good alternative for technical and experts when defining the template of convergent services that could be used later in further phases of automated composition.

Mashup-based approaches, in turn, offer more alternatives by enabling service creation from Web graphical interfaces that do not need to be installed in any desktop personal computer. Additionally, the usability as well as the range of audience for these interfaces is bigger than for SCEs. The main disadvantage over SCEs is the loss of expressive power when creating complex compositions, which may be too simple for the needs of some users.

Some approaches create Mashups automatically from natural language request and allow to refine some details at a later time. On the other hand, Mashups are generally open an may be shared and reused which may lead to a collaborative creation of convergent services.

The approaches based on SCE and mashups allow creating composite services that may be defined using different languages, such as BPEL, BPEL, and BPEL for REST. In these approaches, the core of the service composition task is completely carried out in a manual way. In this way, the novelty of the application of NLP-based approaches is that, in some scenarios, the user just requests his/her desire or goal in natural language and does not define the workflow of the composite service. This workflow is obtained automatically from the user request by the NLP system. Furthermore, by using voice as the primary input medium, a NLP system may be used with any kind of calling device.

Existing research results have shown that creation phase have been automated significantly when natural language is used. However, several important issues are open:

1. To map user request and context to a service is a big challenge, especially when the service does not have a standard interface description like in telecommunications domain, many semantic approaches translate requests to formal representations using natural language processing techniques and ontologies. These approaches have good performance in specific domains [109]. The main idea is to extract concepts from the user request and match them with service concepts [110].

2. There exist some differences between finding services on Telecom and on the Web. The main challenge of finding a service in a telecom environment is the lack of standard descriptions for the functional properties of services. Some approaches for automated creation use traditional Web representations for Telecom environments such as WSDL, however the trend of moving telecom facilities to the Web such as Skype or Twilio allow to glimpse the use of semantic representations of Services.

### 6. Service generation

As above described, the service generation is related to the creation/composition of a chain or a linked set of services that can be represented as a business process encoded, for instance, in BPEL or...
BPEL for REST. There exists a broad range of research approaches around the topic of automated/semi-automated service composition in the Web domain, most of which is applicable to convergent composition. Such approaches can be divided into two categories: workflow-based approaches and approaches that employ AI techniques.

### 6.1. Workflow-based generation

The approaches on service generation using workflow-based techniques assume that a composite process can be perceived as a workflow. The workflow of a composite service includes a set of atomic services along with a data and control flow. Workflow-based approaches can be subdivided into two groups: static, template, and dynamic composition.

The static composition (see Fig. 10) means that user builds an abstract process model or template at design time, and then at execution time atomic services are selected automatically from a set of candidate services. Graphically the composite service is depicted as a directed acyclic graph, with control and dataflow. This graph defines the order of execution of the nodes of the process, in which the nodes are automatically bound with particular services. This automation focuses mainly on the automation of the instantiation process.

The use of template-based composition is located at the middle of static and dynamic composition. Using templates, the user retains control over the process and automatic techniques can be employed to customize and implement the process, achieving very good performance and quality. The template-based composition depends on whether standard procedures are available for the particular domain of application, such as in medical protocols. The templates can be parameterized with respect to some variables to allow customization based on user needs and preferences.

A concept that uses templates is the Hierarchical Task Networks (HTN) [64]. HTN planning is especially promising for service composition because its concept of task decomposition is very similar to the concept of composite process decomposition. In HTN planning, a task represents the activities to perform. A task may be either primitive or compound (which in our case represents an individual and composite service respectively). A compound task is one that needs to be decomposed into smaller tasks using a method. Methods indicate how to decompose a compound task into a partially ordered set of subtasks, each of which can be again compound or primitive. The lowest level task that cannot be decomposed further is called a planning operator. In service composition the highest level task or goal is extracted from the user request and the lowest level planning operator represents a primitive service directly associated with implementation services.

The dynamic composition proposes the generation of the service abstract model automatically. In this way, individual services are selected in an automatic way (see Fig. 11). Some relevant approaches on dynamic composition are, first, a rule-based solution for semi-automated composition that offers the user the ability to guide the composition process. In this solution, the user defines the flow and the rules that allow automatically building the composition. Second, a mechanism that works by synthesising its workflow based on the semantics of the components and the context of users [72]. This mechanism composes applications requested in natural language considering in addition to the context the preferences of the user. These preferences are extracted from user-specified rules and also via learning.

Important shortcomings of the workflow-based composition are that it usually requires extensive domain knowledge and at least some degree of manual implementation by developers, making it difficult, time consuming, and error-prone. In order to overcome such shortcomings arose the service composition based on AI.

### 6.2. AI-based generation

One important group of approaches is focused on semantics. Semantic composition has been used widely for services composition. The main idea is to add semantic information to methods and I/O parameters in order to judge the matching degree between services by calculating semantic similarity of concepts in ontology tree [114]. Furthermore, some authors extract concepts from user goal description and matched them with ontology concepts so as to guide the service composition [115]. For example, Slaimi et al. present an approach called OVWSC (Ontology based Web Service Composition) that includes two phases: first the domain service ontology is filtered for reducing the services research’s space. Second, an abstract workflow representing the composite service is generated.

In spite of the high precision that might be achieved with semantic techniques for services composition, some limitations exist: i) this process is time-consuming due to the inference required during search. ii) The lack of unified ontologies in different domains prevents its spread. iii) It requires users to have enough domain knowledge to perform semantic annotation and searching, which is not feasible in real application. For improving performances of semantic search, some approaches present structures such as quick Extended Quick Service Query List (EQSQL); these structures associate service and request concepts at design time in order to decrease the search time. In the author presents a platform for context-aware service composition based on user-defined rules. The platform interprets ontology descriptions of services and creates user-defined rules based on the ontology description of services. The authors present a review of the recent research methods on semantic Web Service composition.

Nowadays, many approaches for service composition integrate diverse AI techniques: such as rule-engines and logic reasoning, for example. Tang et al. present a method for automatic Web Service Composition based on logical inference of Horn clauses and Petri nets. The problem of service composition is transformed into the logical inference problem of Horn clauses by exploring the dependency relations among services. Besides, the authors present an algorithm for determining if the composite service exists in linear time. Li et al. present a novel composition mechanism by extending OWL-S model. This mechanism considers user preferences and combines the strong knowledge representation and reasoning capabilities of Description logic with the modelling capability of action state.
transformation in AI planning, and finally testifies its availability in DL. Finally, Priya and Chandramathi [122] present an approach for QoS based web service selection. The authors address the composition of Web services using a Fuzzy rule based system. This composition allows imprecise information to be considered and help to model highly complex problems that have multi dimensional data.

Different techniques from AI have been applied in automated composition, the most successful up to date being automated planning [10,121]. Fig. 12 depicts the general automated service composition in which the process model (i.e., service model) is created automatically.

Automated planning can be defined as the task of coming up automatically with a sequence of actions that achieve a goal [123]. In planning there are three types of elements: states, goals, and actions. Given an initial state, the goals and the set of possible actions, a planning system must find a sequence of actions (called plan) that leads to some state in which the goals are true.

An action in automated planning is specified in terms of preconditions and effects. The preconditions are the facts that must hold true before the action can be executed, while the effects depict how the state changes when the action is executed. In this way, a service can be seen as an action and the composite service as a plan [73].

This kind of composition is automatic and does not require knowledge about some predefined workflow.

Several approaches have emerged from AI strengthened by the fact that OWL-S [90], the language of choice for the semantic description of Web services based on SOAP, maintains a similar structure to the description languages used in automated planning [10]. In this sense, for instance, an adaptation and extension of Golog was proposed to allow automated construction of composite services [66]. Similarly, an approach that combines HTN and automated planning was introduced for generating service composition [69]. This approach allows the implementation of preferences during service composition.

Other relevant AI-based approaches are, first, a proposal that includes a semantic description of the composition process and of the discovery of the atomic services [11]. Second, a framework for composition based on automated planning that uses an algorithm for service discovery for consider QoS requirements as well as transactional constraints during the service generation process [65]. Third, a solution that proposes to employ planning in NGN by modelling Telecom services using situation calculus and reducing the composition problem to a planning task [68].

7. Service adaptation and execution

In this section, we present approaches for service adaptation. The adaptation phase translates the composite service into an executable component runnable on a Telecom application. In this sense, the service composition problem can be reduced in some cases to the Boolean satisfiability problem [125], a well-known NP-complete problem. In this sense, algorithms that carry out exhaustive search may work for small scale compositions, but for larger problems approximate algorithms that find sub-optimal solutions are more desirable [104].

To finish this comparison, it is also important to highlight that for service generation, approaches based on AI and workflow can be combined (i.e., these approaches can be complementary). For instance, a framework was introduced for intelligently navigating service composition [70]. This framework consists of two primary processing modules: the Planning Module and the Constraint Satisfaction Problems (CSP) Solving Module. Taking advantage of the functional attributes of the services, the Planning Module aims to produce referred composite plans that function as general patterns. The CSP Solving Module focuses on selecting qualified services to implement the composite plans, based on their non-functional attributes evaluation.

7.1. Approaches for service adaptation

The adaptation phase translates the composite service into an executable component runnable on a Telecom application. In this sense, the automation of the adaptation phase consists of translating the composite service into an executable component without human intervention.

The most remarkable approach for automated service adaptation is presented within the scope of the TeamCom project [49,50,52] (see Fig. 13), which deals with cases in which the developer describes the business process using BPEL but the resulting BPEL process is not necessarily executable on a BPEL-engine. The TeamCom platform translates the resulting BPEL process in XML; next, these files are in turn transformed into Java code that is deployable on a JAIN Service Logic Execution Environment (JSLEE) server [53].

7.2. Industrial approaches for service execution

Regarding composition of services in the Web, the predominant vision for representing composite services is BPEL and its extensions like BPEL for REST (see Fig. 14). Notwithstanding such predominance for representing business process, BPEL/BPEL4REST have not been developed for control of specific communication services [50], particularly in real time like usually Telecom services require. Consequently, BPEL and its associated execution servers are not appropriate in the sense that they cannot provide the availability requirements necessary for convergent service execution [48]. In this way, we consider that the best option for executing convergent services must be found in the range of approaches coming from the Telecom domain. The
following paragraphs present a summary of the Telecom technologies that provide alternatives for service composition execution.

From the Telecom industrial perspective, IMS [62] is the most important global architecture for service execution. IMS is an NGN architectural framework for delivering multimedia services based on IP and SIP. In its origins IMS represented an approach to delivering “Internet services” over data services for 2G and 3G mobile networks. The purpose of IMS was later updated for supporting also Wireless Local Area Networks (WLAN), mobile protocols, and fixed lines. An important characteristic of IMS is that it uses standard technologies for its communications in order to provide convergence of services. Despite the potential that IMS has, the migration to IMS-based platforms has been overall slow due to the high cost of updating the equipment and software for Telecom management necessary to adopt IMS.

Fig. 15 presents the layers of IMS. The Access layer manages the connection of users with IMS terminals, such as phones, personal digital assistants (PDAs), and computers, through different protocols. The only requirement is that these devices must support IP and SIP. Most of the protocols of mobile Telecom networks (e.g., the general packet radio service—GPRS) and wireless networks (e.g., WLAN) are supported. Other non-IMS-compatible phone systems like traditional telephone devices (i.e., PSTN terminals) are supported through gateways.

The Control layer manages the authentication and routing; this layer also provides the capability of associating or composing services from the Services layer. The main element in this layer is the Call Session Control Function which orchestrates the interaction between the services in the service layer. IMS uses a simple orchestration mechanism to select the services that are needed within a Telecom session. However, this method is not dynamic, so any service that a user might need is placed in the chain of services, regardless of their actual usefulness for a given session.

The Services layer holds all of the actual set of services. This set includes traditional Telecom services, such as call forwarding and voicemail, as well as services accessible via Web-based protocols. The IMS services may be installed in application servers or in media servers. In the Service layer, most of the researchers [49,50,52,126–128] have focused on the use of JSLEE as the subjacent application server for executing convergent services.

JSLEE (see Fig. 16) is a high throughput and low latency event processing application environment. It provides a robust platform for rapid development and deployment of convergent services. The access to network resources and services available in the Web is offered by the Resource Adapters (RAs). Furthermore, the use of RAs allows equipping the JSLEE platform with capabilities such as Authentication, Authorisation, and Accounting (AAA) among many others [129].

JSLEE includes a component model to structure the logic of convergent applications where the basic services (from the Telecom and/or the Web domain) are represented as Service Building Blocks (SBBs). JSLEE also enables the composition of SBBs by combining two or more basic services. Recent studies have demonstrated that JAIN SLEE is suitable for provisioning of convergent services in production environments by using commercial solutions, such as the Mobicents JAIN SLEE [131] and the Rhino Application Server [132]. In this sense, for instance, some works [128,130] have demonstrated that even for complex services, the throughput obtained is comparable with the traffic handled by a large gateway node of a VoIP operator.

The other type of Java application servers that can be used in the Services layer of IMS is based on the SIP Servlet specification [63]. SIP Servlet contains and manages SIP applications and provides access to SIP mechanisms via a Java API. The SIP Servlet API defines a component called Application Router (AR) which is responsible for the selection and composition of services. In this sense, the Ericsson composition platform [59] is a SOA-based solution that allows the composition of convergent applications using services from Web, enterprise and Telecom domains, for both circuit-switched and IP-based services. This solution selects the services to be used and informs IMS about the SIP service to be executed. Other approaches [133,134] have explored the integration of SIP-based servers (e.g., Sailfin) with Java Enterprise Edition Servers (e.g., GlassFish) to conduct IMS service composition.

Parlay X [101] is an important alternative to provide convergent services in NGN-based networks, permitting the use of Web services in the Telecom domain. Parlay X [60,135] is an standard that comprises a set of services (from the Web and Telecom domains) offered like APIs that facilitate the composition of more complex functionalities to be provided in Telecom networks. An example of a commercial server supporting Parlay X is the Unified Service Platform ZXUP10 of ZTE Inc [136].

OneAPI [61] appears as an evolution of Parlay X. OneAPI is also a set of supported APIs that exposes network capabilities, such as messaging, payments, and location, on the Internet. Using such guidelines any mobile operator or service provider may implement or use OneAPI. As an example, any messaging service could be used by any other application for sending SMS messages. Regarding the OneAPI it is also important to note that it already provides simple interfaces for different Telecom and Web protocols, HTTP among them. Supporting
HTTP, ONEAPI can execute convergent applications formed by SOAP-based services, Telecom features, and RESTful services. In this vein, some companies provide Telecom services under subscription over the internet using SIP protocol, among them: Hoiio [54], Tropo [55], Twilio [58], CallFire [56], and Skype [57]. Each one of these companies provides their own languages to describe the offered functionalities.

7.3. Academic approaches for service execution

Recently, some academic alternatives have reaffirmed the preponderance of JSLEE [137]. In this sense, a service platform was proposed in order to fulfilling the Telecom requirements (i.e., low latency, real time, and fine-grained events) when convergent services are requested. This platform is event-based architecture, provides asynchronous iteration, and supports JSLEE [45].

Other existing approaches [29,48,128,135] mix JSLEE with Enterprise Service Bus (ESB) to support convergence service composition over heterogeneous networks. ESB (see Fig. 17) is an architectural model that provides a common connectivity layer between services of the Web. Services supported by ESBs are not restricted by a protocol like HTTP nor have to be described by a specific descriptor such as WSDL. In the ESB world, a service is a functional component that is used as a meta-data descriptor understandable by computers. Such descriptor is published so it can be used without needing the implementation of the service. All clients communicate with the ESB in the same way; in turn the ESB is on charge of the translation of messages (format and structure) between client and provider services [138].

On the other hand, some approaches [139,140] propose the integration of SIP and BPEL for the execution of convergent services. These approaches present, for instance, the abstract design of a hybrid multimedia conference framework with low latency services, which allows the execution of both primitive and composite conference services.

Finally, HSCEE [28] is an approach that employs ESB, JSLEE, and BPEL to carry out template-based service composition. HSCEE was employed to build a prototype of a typical communication multimedia conferencing system based on Web services. This system can concurrently run multiple instances of conferences in common public networks, being able to support multi-user conferences employing various network terminals from anywhere. Authors also present promising performance results in the presented prototype.

7.4. Comparison

To select a technology for Execution of convergent services, different criteria must be considered. Previous works have focused on the comparison of these technologies [91] and most of them agree that the best selection depends exclusively on of the needs of the user. As aforementioned, IMS is an architectural framework that defines a series of guidelines for future implementation of convergent scenarios. In the Service layer of IMS both SIPServlets and JSLEE are applicable. Thus, our analysis will focus on these two technologies.

JSLEE and SIPServlets are both standards regarding the execution of convergent applications. Both standards define a container-based model for applications, which describes how applications may be built and deployed in their respective containers and how services and features can be supplied by the containers. JSLEE is a more complex specification than SIPServlets. JSLEE includes mechanisms for fault tolerance and high availability, whereas SIPServlets does not. Essentially SIPServlets remain only as a programming model whereas JSLEE aims to be a robust application environment.

In this paper, we are interested in providing guidelines about the convergent service composition. Such labour establishes a set of basic requirements that are the basis of our comparison (see Table 4) in the execution phase: support for different protocols, support for reconfiguration, and model for representing the composition.

Support for different protocols. Whereas SIPServlet provides support for SIP and HTTP, JSLEE is protocol agnostic; that is to say that JSLEE can be extended to support additional protocols (adding RAs) and external resources with a consistent event model, regardless of protocol/resource. As convergent services are formed by atomic services (using a lot of different protocols) available in the Telecom and Web domains, IMS application servers must support many protocols and APIs (not just SIP).

Support for reconfiguration. Consumers of convergent services will expect high quality of service, so an execution environment must be a carrier-grade platform that addresses corresponding requirements of continuous availability and performance. To provide an acceptable degree of availability a mechanism for monitoring and reconfiguration is mandatory due to the inherent changing nature of services available in the Web domain.

About reconfiguration is important to highlight: (i) SIPServlet has no standard management mechanisms defined until now, JSLEE counts with standard management interfaces based on the Java for Management Extension (JMX) independent of the subjacent management protocol. The main supported functions are life cycles, upgrades, profiles, and tracing, (ii) regarding the utilities for applications that can be useful for monitoring and reconfiguration, SIPServlet has the timer function, while JSLEE has: timer, trace, alarm, and statistics among others, (iii) JSLEE has a container-managed state and wide support for transactions, while SIPServlet has no such functionality. (iv) SIPServlet cannot define a failure model while JSLEE provides a robust failure model via transactions; and (v) while SIPServlets are stateless and the shared state may be stored in a separate session object as a name-value (String, Object) visible within the session, the SBBs may be stateful or stateless. SBB state is private, type safe, transacted and a property of the SBB itself (i.e. an SBB is an object); the shared state may be stored in a separate context via a type safe interface and the access to the shared state may be specified at deployment time.

Table 4

Comparison of approaches for convergent service execution.

<table>
<thead>
<tr>
<th></th>
<th>Standard</th>
<th>Protocols</th>
<th>Reconfiguration</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>SipServlet</td>
<td>HTTP and SIP</td>
<td>Agnostic</td>
<td>Timer</td>
<td>Servlet-based</td>
</tr>
<tr>
<td>JSLEE</td>
<td></td>
<td></td>
<td>Timer, transactions, failure management, and JMX</td>
<td>Object-based</td>
</tr>
</tbody>
</table>

Please cite this article as: A. Ordóñez et al., Towards automated composition of convergent services: A survey, Computer Communications (2015), http://dx.doi.org/10.1016/j.comcom.2015.07.025
Model for representing the composition. As the convergent service composition must be the most automated possible, the model of the subjacent technology must easily support the automation of the composition. In this sense, it is important to noteworthy: (i) Services in SIPServlet are represented as entities based in HTTP Java Servlets without a standard model for composition; and (ii) JSLEE represents services using an Object Oriented model that see the services as SBB, providing support for composition and reuse.

From the previous discussion it is clear that SIPServlet provides support for the adoption of the SIP protocol and the development of simple SIP services. However, convergent services have requirements that exceed the capabilities of SIPServlet. In contrast, JSLEE has been developed to satisfy the requirements of Telecom services in tier-one service providers [91]. This leads us to declare JSLEE as the better option for executing convergent services.

Regarding the automation of the execution phase, JSLEE has still relevant shortcomings. As mentioned previously, the execution phase is inherently automatic, as once the user deploys the service it runs in the execution environment without human intervention. The problem arises when a malfunction or an error appears. In this vein, the role of automated reconfiguration in the convergent service composition is very important. However little has been said about this topic in spite of the existing tools provided by the JSLEE server. Some approaches [105,141,142] offer alternatives for JSLEE service reconfiguration based on advanced selection algorithms of services in case of failure. Nonetheless, we think is important to analyse the feasibility of using current approaches of Web service recovery [143] for reconfiguring convergent compositions.

8. Research directions

From the above review, we can see that many advances have been made in various research aspects, including automated generation, graphical creation, and execution of convergent services in telecom environments. However, there are still many open research issues that need to be solved before the automated composition can be of practical use. This section presents the research directions and how existing approaches can be used or contribute to address the challenges of convergent service composition.

Availability of services. Whereas in the Telecom domain the availability of services is close to 100%, in the Web domain, services can raise, change, or disappear continually. Recent approaches are focusing its efforts on defining approaches for reconfiguring compositions that involve services from both domains. These approaches tackle the ongoing changing in the service libraries and, consider failover mechanisms (e.g., to replace services in failure) that allow the reconfiguration of composite services. These approaches can be divided into two groups.

One group of approaches for reconfiguration aims for solving convergent reconfiguration adapting mechanisms from Web services reconfiguration in convergent domains. These approaches mostly deal with replacement of failing services, and consequently, most of the existing related approaches are focused on service selection of potential replacement services. In [144] the authors present architectures for automated reconfiguration of convergent services in Telecommunication environments. In order to reduce reconfiguration times while holding the initial constraints, the authors present an iterative algorithm which does not replace individual services but whole regions of services, specified with Hierarchical Tasks Networks (HTNs).

Another group of approaches for convergent reconfiguration adapts services composition according to the changing situation and environment of the user, most of the proposed techniques may be used in the convergent composition: Chen and Clarke [145] present a model for service composition in high mobility environments. The model for decentralized service composition adapts itself by composing fragments on-demand, as appropriate to the constraints of services consumers and providers. [145], Neiat et al. [146] present a failure-proof composition model for Sensor-Cloud services based on dynamic features such as spatio-temporal aspects. The authors present a failure-proof composition algorithm based on D+ Lite to handle QoS changes of Sensor-Cloud services at run-time.

Number of services. In the Telecom domain, the number of services is relatively low (e.g., messaging, phone calls, location, and video call). In contrast, the number and nature of services in the Web domain is essentially high and changing (e.g., shopping, translator, and social networks). This high number of services turns complex the use of traditional methods, such as SCEs and MDEs, for creating convergent services. Therefore, some of the research approaches, analysed here especially the group of approaches based on automated planning may be used for discovering convergent services in Telecom infrastructure; for example, Ordóñez et al. [109] present a framework for including automated composition approaches based on automated planning in Telecommunications scenarios. On the other hand, recent algorithms such as [147,148], uses mechanisms for reducing the services search space by pruning candidates by dominance relationships and constraint validations at candidate level. Finally, some approaches are focused on the use of data mining in services composition. Bayati et al. [149,150] use data mining techniques on Web services in order to find the best composition in a set of possible compositions. These works use techniques such as association, classification and clustering are for improving performance of the Web Service Composition.

With the increasing presence of web services on the Internet, quality of service (QoS) is becoming important for describing non-functional characteristics of services, and is often employed in Web Service Composition [147,151]. In [147] a service composition algorithm using partial selection technique for QoS-Aware Web Service Composition, this algorithm reduces the search space due to the fact that it prunes candidates by dominance relationships and constraint validations at candidate level [147]. Equally Moustaaid and Zhang [148] propose two algorithms based on reinforcement learning for composing and adapting Web services in open and dynamic environments: the first algorithm introduces an offline learning scheme for service composition where one online learning task is transformed into a series of supervised learning steps. The second algorithm presents a coordination mechanism in order to enable multiple agents to learn the service composition task cooperatively. In [152] the authors present a novel multi-objective that addresses QoS-aware Web Service Composition with conflicting objectives and various restrictions on quality matrices. The approach uses reinforcement learning to deal with the uncertainty inherent in open and decentralized environments. Experimental results reveal the ability of the proposed approach to find a set of Pareto optimal solutions, which have the equivalent quality to satisfy multiple QoS-objectives with different user preferences.

The aspect of services discovery is related to the description of services that will be addressed below.

Services description. Though standardized or broadly accepted language do not yet exist for describing convergent services, the trend indicates that, as in Social Web and Telecom network management, REST-based services have turned de facto standard to publish functionalities, we consider that existing languages (e.g., WSDL, WADL, and hREST) may be adapted to specific domains. Existing languages such as: BPEL, BPEL for REST, BPEL and LPMI offer the base for unifying description of service compositions, especially when these compositions include services based on protocols such as SOAP/HTTP, JSON/XML, SOAP, Atom, and SIP. Regarding Telecommunication providers, different service providers based on Cloud computing, such as Skype or Twilio, may offer its own standard or language. In this scenario the option of Enterprise Service Buses (ESB) acquires high relevance due to its potential for translating different standards. Despite improvements, research for describing Telecom services

Please cite this article as: A. Ordóñez et al., Towards automated composition of convergent services: A survey, Computer Communications (2015), http://dx.doi.org/10.1016/j.comcom.2015.07.025
using well established techniques from the Web services must be expanded. These extensions and improvements will be fundamental to facilitate the evolution of convergent service composition.

**Interfaces of user.** In the Web domain, the user interacts with the system using mostly interfaces controlled with a keyboard and a pointing device or a tactile screen. In contrast, in the Telecom domain, the voice is the primary element to interact with the phone, especially when the user is moving or driving. Therefore, existing academic and industrial approaches based on natural language (both commercial and open source) can be used for providing and consuming convergent services with richer user interfaces. Here, rich user interfaces must be understood as those in which the user can access the convergent applications using keyboards, tactile screens, and voice depending of the user context. For example [153], natural language processing is used to identify services required to accomplish the user request. This approach represents the identified tasks in model and from this model a user interface (UI) is generated. The authors claim that the system can extract the tasks from how-to instructions Web pages with high precision. The work of Ordóñez et al. [110] presents the use of natural language processing in convergent domains by translating user request and context into automated planning representations.

**Context of user.** In the Web domain, the context and preferences of the user considered during service composition rarely change in real time. Conversely, in the Telecom domain, the user is often in constant movement and the type and capabilities of the device cannot be known beforehand in many cases, which means that the user context must be considered, modelled, and included in the composition process. In this regard, many research works can be applied for particular solutions of convergent composition. For example [154] presents an automated composition system that incorporates SOA paradigm and devices implemented by device profile for web service (DPWS). This approach presents a composition engine that considers context information, composition templates, and predefined policies. Equally NGSON [155] is an IEEE sponsored effort to standardise an open service ecosystem that allows composing and delivering services dynamically across different service providers and network operators domains. NGSON uses context information to drive the interactions of the NGSON overlay nodes with service and network operator infrastructures.

**Time constraints and number of protocols.** During development of convergent composition solutions, the heterogeneity of protocols, networks and time restrictions may be tackled by selecting robust environments for service execution. JSLEE offers a high reliable environment for supporting Telecommunications constraints. JSLEE provides independence of protocol, and interfaces for diverse networks; equally JSLEE considers time constrains proper of telecom domains.

9. Conclusions

Selection of technologies for specific convergent composition solutions is complex due to the big amount of considerations. In this paper, we have identified the relevant phases for convergent service composition and explored the existing approaches for automating each one of these phases. It has been analysed that all phases of convergent composition have reached high level of automation using different approaches. Following, as conclusion, we provide general guidelines that help to select the most suitable approach for each domain or problem.

**Domain knowledge.** If the user possesses domain knowledge, he will be able to create a full specification that details the specific tasks and the data/control flow among them, as well as other non-functional requirements. Conversely, it cannot be expected of a non-expert user to create a full workflow. Thus, for non-expert users, a specification stated in terms of preconditions and effects should be created and used to perform the composition of services. Such specification should be done with domain-independent planning techniques that do not require knowledge of the process model.

**Level of automation required.** When the number of services and the complexity of services composition increase, automatic composition becomes more desirable. There are different variations; for example, a semi-automated approach can generate basic templates automatically that are later presented to the user so he/she chooses one of them and refines it manually.

**Context of device and user.** If the user requires mobility and does not have advanced knowledge on programming, he/she may use Web-enabled graphical interfaces provided by mashup-based approaches or alternatively even consider voice-based interfaces.

Conversely, if the domain is oriented to expert users that do not have such constraints, SCE-based approaches are the best choice.

**Dynamic domains.** In convergent composition, services may become unavailable or be redefined by their provider. This consideration justifies the need for continuous monitoring and reconfiguration during the execution. In workflow-based approaches, this means that the user must supervise the execution process and perform the necessary changes; in automatic approaches the system must react and adapt itself to the circumstances with as little human intervention as possible. For convergent domains, JSLEE offers a set of features that could ease reconfiguration implementation: standard management interfaces based on JMX, independence of management protocol, and interfaces for managing applications. Also, it is important to highlight that some AI-based planning platforms include execution and replanning, which may alleviate the reconfiguration problem to some extent. These AI-based approaches are an interesting alternative, but they need to consider the complexity and performance requirements of the Telecom domain.

Acknowledgements

The research of PhD(c) Caicedo is supported by the PECPG (Agreement Program Students Graduate) of the CAPES (Brazil) and the University of Cauca (Colombia).

References


Please cite this article as: A. Ordóñez et al., Towards automated composition of convergent services: A survey, Computer Communications (2015), http://dx.doi.org/10.1016/j.comcom.2015.07.025

Please cite this article as: A. Ordóñez et al., Towards automated composition of convergent services: A survey, Computer Communications (2015), http://dx.doi.org/10.1016/j.comcom.2015.07.025.

ARTICLE IN PRESS

Please cite this article as: A. Ordóñez et al., Towards automated composition of convergent services: A survey, Computer Communications (2015), http://dx.doi.org/10.1016/j.comcom.2015.07.025