www.s2is.org
International Journal on Smart Sensing and Intelligent Systems
ISSN 1178-5608

SURVEY OF SEMANTIC SIMILARITY MEASURES IN PERVASIVE COMPUTING

Djamel Guessoum[1], Moeiz Miraoui[2], Chakib Tadj[1]

[1]Ecole de Technologie Superieure, Universite du Quebec, Canada
djamel.guessoum.1@ens.etsmtl.ca

[2]Al-Leith computer college, Umm Al-Qura university, Saudi Arabia
mfmiraoui@uqu.edu.sa

Submitted: Oct. 6, 2014 ***Accepted: Jan. 8, 2015*** ***Published: Mar. 1, 2015***

Abstract- ***Semantic similarity measures usage is prevalent in pervasive computing with the following aims: 1) to compare the components of an application; 2) to recommend and rank services by degree of relevance; 3) to identify services by matching the description of a query with the available services; 5) to compare the current context with already known contexts. The existing works that apply semantic similarity measures to pervasive computing focus on one particular issue. Furthermore, surveys in this domain are limited to the recommendation or discovery of context-aware services. In this article, we therefore present a survey of context-aware semantic similarity measures used in various areas of pervasive computing.***

Index terms***:*** **pervasive computing, semantic similarity, context-aware, service discovery, service recommendation**

I. INTRODUCTION

Similarity involves the assessment of intrinsic common characteristics between two or more concepts. A characteristic is intrinsic to an object when it defines the nature of the object itself and cannot be separated from it. In information systems in particular, similarity relates to the assessment of likeness in an analyzed date set in order to quantify these similarities in the interval [0,1]. As a result, it is possible to order and prioritize them or extract invariants. Generally, the similarity evaluation involves three types of data processing, namely classification, identification, and characterization (Bisson 2000). Classification aims to structure data in a heterogeneous group according to similarity, while identification endeavors to recognize the class to which an unknown object is likely to belong. Finally, the characterization process allows the explicit representation of information that is common to a set of data.

Semantic similarity measures are referenced to the similarity measure based on human judgment. This latter notion was first introduced in the study of Ruinstein and Goodenough in 1965, in which two groups of 51 people evaluated the synonyms of 65 pairs of names. In 1991, Miller and Charles repeated the original experiments of Rubinstein and Goodenough using 30 pairs of names taken from the original list of 65: 10 pairs had a high level of synonymy, 10 an average level, and 10 a lower level (Saruladha et al. 2010).

In pervasive computing, semantic similarity measures were implemented as a mechanism to properly adapt the applications and services between the user and environment. Semantic similarities measures in a pervasive computing system (PCS) are thus applied in order to select the modules of a context-aware application that are appropriate to the user's current context, to choose the best advertised service by matching the user's query to the available service description and classifying the selected services according to relevance, and finally, to identify the current context by comparing information collected from the environment with a set of predefined situations.

a. Dynamic adaptation of services in pervasive computing

In pervasive computing, the adaptation of services is a dynamic process wherein services are offered reactively to a user in response to a change in context or proactively by predicting a change in context and reacting accordingly (Germán 2010). Several definitions are proposed in the literature, although the most generic is given by Efstratiou (2004) who generalizes the concept of adaptation for mobile equipment and context-aware applications in a PCS by assuming that an application or system is adaptive when it changes its behavior in response to a change in context (this change occurs in either the context or equipment resources). Zouari (2011) recently defined the dynamic adaptation of a context-aware application according to its ability to change its behavior during the execution phase in line with fluctuations in the environment or changes in user requirements.

Another approach has been adopted in other studies, such as that used by Simonin and Carbonell (2007), which categorizes the dynamic adaptation of services according to the purpose of adaptation, thus distinguishing two types of adaptation: adaptation to the user profile and to the environment. This approach requires the user context and environment to be sources of information for an appropriate adaptation of services. The following works, however, are more comprehensive in terms of services. For example, Nicklas et al. (2008) categorize the adaptation of context-aware applications into four classes: 1) the selection of information and services; 2) the presentation of information and services; 3) the automatic execution of a service for a user; 4) the marking of context with information for later retrieval. Benazzouz (2012) classifies the adaptation into three classes, notably the personalization, recommendation, and reconfiguration of services. According to this classification, the personalization of services is linked directly to the user’s preferences, deriving its contextual information from the user environment (e.g., ambient temperature, geographical location). Recommendation is a particular form of personalization that draws from user-stored preferences (history) to recommend the most adequate services. Finally, reconfiguration takes into account the system environment (e.g., releasing memory space for an application). Note that reconfiguration does not consider the user’s environment.

In what follows, Section 2 of this survey discusses the concept of semantic similarity in general. Section 3 introduces the various applications of semantic similarities measures in the field of ubiquitous computing; semantic similarity measures are discussed between contexts, for the recommendation of services, in context-aware applications, and for the service discovery.

II. NOTION OF SEMANTIC SIMILARITY

a. Introduction

In pervasive computing, where the notion of context plays a very important role, the semantic similarity measure is a tool to evaluate the resemblance between instances of a context. It allows services to be chosen and classified according to their relevance to a given query, and a user's profile and preferences to be compared to those of other users in order to recommend similar services. Finally, semantic similarity aims to evaluate the similarity between application components in order to propose the most relevant one in a current context.

Harispe et al. (2013) classify semantic similarity measures according to the type of elements to be compared (i.e., words, sentences, paragraphs, and documents, concepts or groups of concepts, semantically related instances) and the semantic proxies used to extract the required semantics from the measure. In terms of the latter, the semantic proxies are of two types (Mihalcea et al. 2006): corpus-based proxies in which the similarity between two concepts is determined based on the information extracted from a large corpora, and knowledge-based proxies in which the similarity between two concepts is evaluated using information derived from the semantic networks (e.g., ontologies, WordNet).

b. Notion of distance and similarity

b.i Distance

Distance is associated with all quantifiable (scalar or vector) or measurable information that describes a context, such as temperature, noise, time, and geographical position (Lavirotte et al. 2005). Thus, for a space E comprising contexts $E_1, E_2, \ldots \ldots E_n$ as described by *m*-dimensional vector entities, X $= (x_1, x_2, \ldots x_m)$, $Y = (y_1, y_2, \ldots y_m), \ldots \ldots$, the function d: $E \ x \ E \rightarrow \mathbb{R}^+$ associated with *X* and *Y* has the following properties:

$$\begin{cases} d\,(X,Y) \geq 0 \\ d\,(X,Y) = 0 \ \Leftrightarrow X = Y \ \ (separation) \\ d\,(X,Y) = d(Y,X)\ (symetry) \\ d\,(X,Y) \leq d(X,Y) + d(Y,Z)\ (triangular\ inequality) \end{cases} \qquad (1)$$

This is known as the *distance* or dissimilarity.

b.ii Similarity

Definition: Semantic similarity measures are mathematical tools used to quantitatively or qualitatively estimate the robustness of semantic relations between units of language, concepts, or instances of concepts through a numeric or symbolic description obtained from a semantic support, such as a text or knowledge representation supporting its meaning or describing its nature (Harispe et al. 2013).

The function *s* that defines semantic similarity must have the following properties:

For a set of concepts in a domain X, the function s: X×X→$\mathbb{R}^+$ is called "similarity" in X, if $\forall\, x, y \in X$:

$$\begin{cases} s(x,y) = s(y,x) \quad (symetry) \\ s(x,x) = 1 \quad (a\ concept\ is\ similar\ to\ itself) \\ and \quad \forall\, y \in X\ and\ x \neq y : s(x,x) \geq s(x,y) \end{cases} \tag{2}$$

The transformations most frequently used to obtain the distance or dissimilarity d from similarity s bounded by 1 are as follows (Michel and Deza 2007):

$$d = 1 - s, = \frac{1-s}{s}, d = \sqrt{1-s}, d = \sqrt{2(1-s^2)}, d = arccos\,(s), d = -\ln(s)$$

c. Semantic similarity measures applied to ontologies

With the advent of the internet and need for information and knowledge sharing on a semantic level, the use of ontologies has become necessary, and as a result, they have considerably developed. The advantages in adopting ontologies as a tool for knowledge representation in a PCS are summarized by Viterbo et al. (2008) as follows:

(1) Ontologies are semantically richer than taxonomies or object-orientated models.
(2) Knowledge is described through accurate representations.
(3) Ontologies are formal; those in web ontology language (OWL-DL) map directly onto the DL (first-order logic).
(4) Formal ontologies in OWL-DL can be verified or classified through inference mechanisms (e.g., RACER, FaCT): verification of the consistency, classification, and discovery of new information.

(5) Ontologies in OWL use XML/RDF syntax, which allows them to be automatically manipulated and understood by most internet resources.
(6) Ontologies capture and represent knowledge in detail.
(7) Ontologies can be used to reduce ambiguity by providing a model for sharing information.
(8) Ontologies are modular, reusable, and independent of the application's code.
(9) Ontologies can be combined with the emerging rule-based languages like semantic web rule language.

The similarity measures between ontologies occur on two levels—lexical and conceptual—which include concepts with semantic relationships (Maedche and Staab 2002). In the case of PCS, the semantic similarity measure must take context into account so that the results are relevant and up-to-date.

Ehrig et al. (2005) classify the "contextual" semantic similarity measures between ontologies and intra-ontologies into three layers (Figure 1). First, in the data layer (representation layer), similarity measures are only simple measures between the values of the entities (i.e., integers or characters). Second, in the ontology layer (layer of meaning), the similarity between two concepts is based on the ontological structure and semantic relations represented by the ontology. Third, in the context layer, the external factor of the measure is considered, namely the context in which the ontology develops. Note that the semantic similarity measures between concepts made through the comparison of their common characteristics are also an integral part of the data layer. For example, the concept jaguar (car) and jaguar (animal) are syntactically similar, but very different when described according to their characteristics: vehicle or wheels versus animal or feline. For this reason, the data layer is divided into syntactic and semantic similarities.

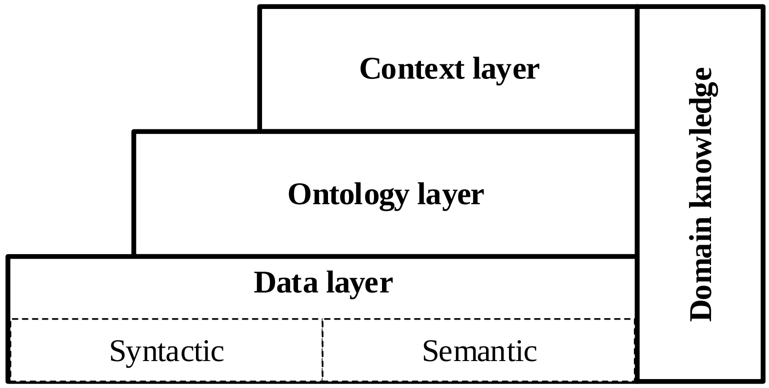

Figure 1. Layered model of semantic similarity measures between ontologies and intra-ontologies (inspired by Ehrig et al. 2005)

Recently, Sanchez et al. (2012) and Saruladha (2011) made the most developed semantic similarities measures to date (Table 1) based on the ontological representation of knowledge (ontology layer), especially in its taxonomic form. The measures are as follows:

- Edge counting-based measures;
- Informational content-based measures;
- Feature-based measures;
- Hybrid measures.

Edge counting measures were first introduced by Rada et al. (1989). These apply to ontologies with relations between concepts of the taxonomic type (is-a). The basic idea of these measures is the fewer number of edges between two concepts, the more similar they are. The semantic similarity between two concepts, C1 and C2, in this case is given as:

$$dis(C1, C2) = \min(path(C1, C2)) \qquad (3)$$

Wu and Palmer (1994) considered the depth of the ontology in the measure, because the more specific two concepts are (in the lower ontological levels), the more similar they will be, and vice versa. This measure is given as:

$$sim(C1, C2) = \frac{2 \times P}{N_1 + N_2 + 2 \times P} \qquad (4)$$

Where N_1 is the number of (is-a) edges between the concept C1 and the least common subsumer (LCS) of (C1,C2), N_2 is the number of (is-a) edges between the concept C2 and the LCS of (C1,C2), and P is the number of edges (is-a) between the LCS and ontology root.

Several other measures were subsequently introduced by Leacock and Chodorow (1998) and Li et al. (2003), as the authors attempted to make adjustments for a particular aspect of Wu and Palmer's measure. This type of measure is simple to implement, but it is limited to ontologies with taxonomic relations (is-a). Furthermore, it does not allow for the context and can give incorrect semantic similarity measures.

Semantic similarity measures based on the informational content of the common notion underlying two concepts were first introduced by Resnik (1995). The informational content of a concept is its probability of occurring in a corpus such as WordNet: the higher the occurrence of the concept, the less the informational content. The informational content is given as:

$$IC(C) = -\log P(C) \qquad (5)$$

Several other measures inspired by Resnik were subsequently proposed. For Lin (1998) and Jiang and Conrath (1997), for example, the informational content of the concepts C1 and C2 is considered when evaluating the shared information more accurately.

Among the limitations of these measures is their dependence on the corpus, as the concepts may be sometimes ambiguous or even not present. They also give the same result for any pair of concepts with the same LCS (Sánchez et al. 2012). Their dependency on the design of the ontology and their lack of consideration for the context are also limitations.

Finally, the semantic similarity measures based on the features of the concepts are based on Tversky's model of similarity (1977), whereby two concepts are more similar if they have more common characteristics and less non-common characteristics.

Let $\emptyset(C1)$ and $\emptyset(C2)$ be the characteristics of C1 and C2. $\emptyset(C1) \cap \emptyset(C2)$ are the shared characteristics of C1 and C2. $\emptyset(C1) \mid \emptyset(C2)$, while the non-common characteristics of C1 and C2 are $\emptyset(C2) \mid \emptyset(C1)$. The semantic similarity between C1 and C2 is thus given as:

$$Sim(C1, C2) = \alpha . F(\emptyset(a) \cap \emptyset(b) - \beta . F(\emptyset(C1)|\emptyset(C2)) - \gamma . F(\emptyset(C2)|\emptyset(C1))) \quad (6)$$

Where F reflects the important characteristics of C1 and C2, and α, β, γ are the weighting parameters. Note that the characteristics depend on the context of their definition.

The determination of the weighting parameters represents the major disadvantage of this type of semantic similarity measure.

Table 1: Semantic similarity measures inspired by Saruladha (2011), Gomaa et al. (2013), and Meng et al. (2013)

1. Semantic similarity measures based on the ontological representation of knowledge		**Studies**	**Specificity**
Inter-ontologies	**Path-based**	Al Mubaid and Nguyen (2009)	- Concept specificity, shortest path, concept depth, is-a relation
	Feature-based	Rodriguez and Egenhofer (2003)	Three independent similarity assessments: 1) similarity of synonym sets, 2) a feature similarity 3) types of semantic relations
Intra-ontologies	**Path-based**	Rada et al. (1989)	- Simple, is-a relation, number of edges in a taxonomy - Two pairs with the shortest path of equal length will have the same similarity
		Hirst and St Onge (1998)	- Relatedness measure with different semantic relations, shortest path, automatic detection and correction of malapropisms
		Bulskov (2002)	- Is-a relation, path length, weighted paths, information retrieval
	Depth-based	Wu and Palmer (1994)	-Simple, is-a relation, number of edges, taxonomy depth ……
		Sussna (1993)	- Based on all possible links, weighted relations, measure between two adjacent concepts - Sensitive to: 1) shortest path between concepts, 2) density of concepts along this path, 3) shortest path from the root to the LCS
		Leacock and Chodorow (1998)	-Simple, Is-a relation,Similarity value using a logarithmic function, -Shortest path in taxonomy, -Maximum depth…..
Informational content-based measures (corpus-based)		Resnik (1995)	- Simple, is-a relation, information content in LCS - Coarse measure less likely to suffer from zero counts - Two pairs with the same LCS will have the same similarity
		Lin (1998)	- Same as Resnik's measure plus commonalities and distinct features of a concept considered
		Jiang and Conrath (1997)	- Is-a relation, shortest path and edges weighted by IC in a taxonomy
Hybrid measures		Li et al. (2003)	-Simple, shortest path, depth of LCS, local semantic density of concepts, multiple corpora used
Feature-based measures		Tversky (1977)	- Features common and distinct between concepts, aymmetrical measure, computational complexity
		Pirró Measure (2010)	- Common features among concepts and different features among concepts, defined in terms of information theoretic domain, corpus independent
2. Corpus-based semantic text similarities		HAL (Hyperspace Analogue to Language) (1995)	- Co-occurrence of words in a corpus (matrix of words appearing next to each other, similarity by cosine of vectors) - Only information found in the corpus used - No human bias or influence
		LSA (Latent Semantic Analysis) (1997)	- Use of singular value decomposition (SVD), method for dimensionality reduction, information retrieval and pattern recognition - Solves polysemy, synonymy, and term dependence - Low efficiency and high data storage
		DISCO (Extracting DIStributionally similar words using CO-occurrences) (2009)	- Distributional similarity - Words with similar meaning occur in similar context - Use a context window of size ±3 words for counting co-occurrences
3. Logic-based representation of semantic measures		D'Amato (2007), D'Amato et al. (2009)	- Similarity value between objects is the result of the common and different features - Similarity between individuals and between a concept description and an individual - Clustering and retrieval on DL knowledge bases - Weakness in cases involving individuals

III. APPLICATION OF SEMANTIC SIMILARITY IN PERVASIVE COMPUTING

In a typical PCS, a context-aware application interacts with the physical environment and the user's system in order to provide appropriate services. This interaction may be a response to the user's request for a specific service or to the current context information with the aim of providing services that are relevant to the user. In such an environment, semantic similarity measures have been applied at several levels (Figure 2): the comparison of an application's components with respect to their appropriateness in a current context; the recommendation of services and collaborative filtering when comparing the preferences of multiple users with the ranking of services according to their relevance during the recommendation process; service discovery by the matching the description of a request with available services; lastly, the comparison of the current context with already known contexts or the detection of current situations. These applications are detailed in the forthcoming sections.

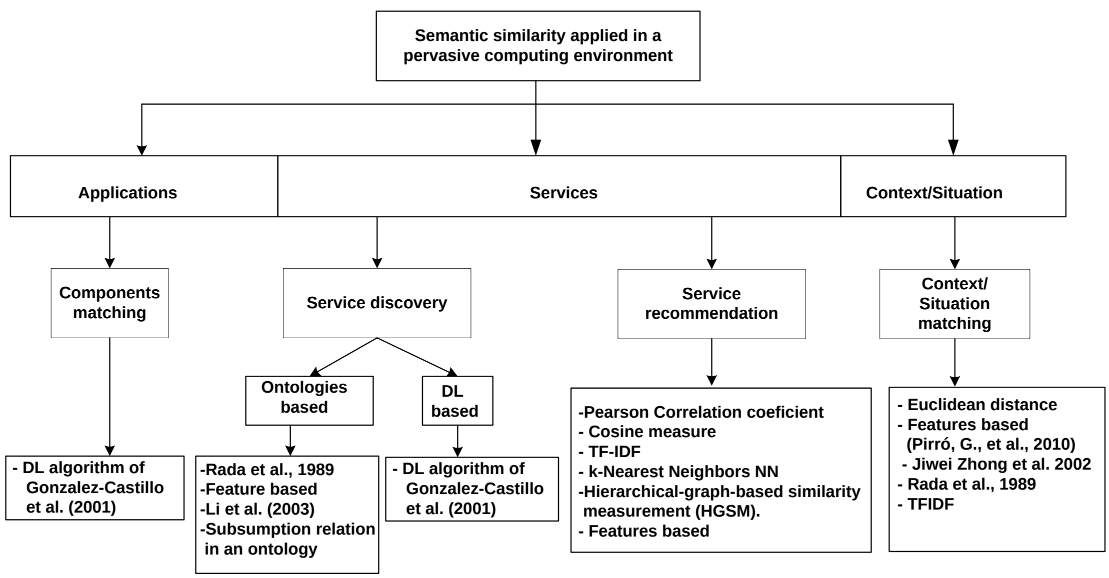

Figure 2. Application of semantic similarity measures in pervasive computing systems

a. Semantic similarity measures and context

The definition of context according to Petit (2005) along with the majority of researchers is based on the four following axes:

(1) There is no context without context: the concept of context must be defined in terms of a purpose. For example, the aim may be to adapt the interactive capabilities of a system dynamically.

(2) Context is an information space that serves the interpretation: context capture is not an end in itself, but captured data must serve an objective.

(3) Context is an information space shared by several actors: the user and the system.

(4) Context is an infinite and dynamic space of information: context is not permanently fixed, but is constructed over time.

The following definitions of context should be in accordance with the aforementioned axes. First, Brezillon et al. (1999) defined two concepts relating to context: 1) the set of contextual knowledge (e.g., time, location) to be used in a decision problem, which is latent and cannot be used without an emergent objective; 2) the context as the product of the emergent objective or intention that uses a large part of contextual knowledge.

In 1994, Schilit and Adams categorized context according to six areas. The first three relate to the human factor: user information (e.g., clothes, biophysical conditions), social environment (e.g., proximity to other people), and user tasks (e.g., active user tasks). The other three areas concern the physical environment: location, infrastructure (e.g., resources, communication), and physical conditions (e.g., noise, brightness, weather conditions).

The definition of Dey et al. (2001) is the most cited: “context is any information that can be used to characterize the situation of an entity. An entity is a person, or object that is considered relevant to the interaction between a user and an application, including the user and the application themselves” (p. 5). This definition is evidently similar to Schilit’s because context is defined as a set of information collected from the user environment (person), physical environment (physical object), or system environment, with the objective of collection being the characterization of these environments.

Given the preceding definitions, we may say that context is definitely a set of information characterizing an environment, whether the user, physical, or system environment, and that the collection of this information must serve for an objective.

a.i Impact of context

Keßler (2007) defines context relative to the similarity measure in the following terms: "A similarity measurement's context is any information that helps to specify the similarity of two entities more precisely concerning the current situation. This information must be represented in the same way as the knowledge base under consideration, and it must be capturable at maintainable cost" (p. 4). This definition gives rise to the following questions regarding the choice of contextual information to be included in the similarity measure between two concepts:

(1) Impact: does the chosen contextual information improve the accuracy of the semantic similarity?
(2) Representation: can this contextual information be represented in the knowledge base?
(3) Acquisition: can this contextual information be acquired at a reasonable cost?

Formally, for a contextual information $\mathrm{c_n}$ of a context C to be considered in a calculation of semantic similarity between contexts, its impact should be calculated by measuring the semantic similarity that includes and excludes this information. The impact must be greater than a minimum threshold δ:

$$Imp(c_n) = \frac{\sum\left|sim_{(C_n \in C)}(a,b) - sim_{(C_n \notin C)}(a,b)\right|}{|C|} \qquad (7)$$

Where $C = \{c | imp(c) > \delta\}$ is the final context including all relevant contextual information.

Most semantic similarity measures are between concepts without taking account of the context of the measure, which sometimes leads to implausible results. "Tablet" and "smart phone" are two similar concepts in terms of "information processing," but completely different in terms of "telephony." The limiting factor according to Janowicz (2008) in the collection of contextual information does not concern how much information can be collected, but rather whether this information can be incorporated into the similarity measure (e.g., through weights) and whether it plays a significant role (i.e., an impact on the result of the similarity measurement).

In pervasive computing, the introduction of context has improved existing semantic similarity measures by introducing weights to the characteristics and semantic links. Furthermore, it has facilitated the application of semantic similarity measures in the calculation of contextual similarities between situations, contexts, concepts, or instances of concepts.

a.ii Semantic similarity between contexts

In a PCS, the services provided to a user relate to the user context (environmental, system-based). The identification of context is thus an essential task. The question that arises is therefore, “What services must an intelligent device in a PCS provide to a user when the current context is identified?” The identification of the current context is defined by the contextual information related to the triggering of a service as well as a situation or “current context” in the set of current contextual information, similar to a known situation or context (Benazzouz 2012), with each identified situation being linked to one or more of the services to be provided. This identification forms the basis of the rule-based adaptation mechanism, which is a set of conditional rules with the form: if (contextual information I) then (service S).

A situation is “a snapshot of the environment at a given point in time” (Ramparany et al. 2011). Identifying a situation is based on data mining techniques. Once identified, semantic similarity measures are applied in order to compare it with situations with known services. In Dietze et al. (2008), semantic similarity is measured against the Euclidean distance between the contextual data vectored in mobile situation spaces. Gicquel (2012) modeled the spatio-temporal context of a museum visitor in an ontological form, with the semantic similarity measures being used to recommend artwork similar to the interests of the user by comparing the properties of two concepts in the knowledge base. The similarity measure is a modified version of the similarity proposed by Pirró and Euzenat (2010), which combines the similarity calculation based on Tversky’s model with that of informational content.

Benazzouz (2012) and Ramparany et al. (2011) applied semantic similarity measures to group data and “pure” contexts in order to build relevant situations for the adaptation of services declared within the ontology of context. First, syntactical and conceptual (semantic) similarity measures between contextual data are applied based on the measures of Zhong et al. (2002) and Rada et al. (1989). Second, conceptual and relational similarity measures between “pure” contexts are used based on the quantification of information common to two graphs and on the statistical technique known as TF-IDF (term frequency–inverse document frequency).

Ontological representation is also used to model a set of situations that occur frequently, such as the locations “at home” or “at work.” Semantic similarities are made between contextual variables representing the current situation and the “frequent” situations, while the services provided are a set of appropriate notifications (Meissen et al. 2005).

A similar approach was proposed by Kirsch-Pinheiro et al. (2006) for the adaptation of content found in an intelligent device with a PCS. The authors used semantic similarity measures to assess the degree of matching between the predefined profiles of situations and the current context of the user with the aim of prioritizing them. Using a graph-modeled context, it estimates the proportion of elements in the graph defined by the user's current context, with the graphical elements defined by each user profile. This measure is determined as follows:

$$sim(C_u, C_p) = x, \quad x \in [0,1]$$

$$\text{With: } \begin{cases} x = 1 \ \textit{if each element of } C_u \text{ has an equal element in } C_p \\ \qquad\qquad Else \\ x = \frac{|X|}{|C_u|} \ \textit{with} \quad X = \{x | x \ \textit{equal to} \ y, x \in C_u, y \in C_p\} \end{cases} \tag{8}$$

Semantic similarities between contexts in a PCS are thus based on the collection of one or several elements of contextual data that are relevant to one or several services. The description and semantic relations of these services are described in an ontological form, thus allowing the application of known semantic similarity measures.

b. Recommendation of services in a PCS

In a PCS, the recommendation of services must consider the context as well as the user's preferences (Figure 3). The context and user preferences can be used to limit the number of recommended services or rank them according to their relevance to the user (Van Setten et al. 2004), while the contextual information can also serve to reduce the issue of limited data (Liu et al. 2010).

Formally, if C is a set of users, S a set of products (services) to be recommended (e.g., books, movies), and u the utility function represented by the rating of how much a user c has appreciated the service s, then the measure of the relevancy of a product or service s $\in$ S to the user c $\in$ C is $\mathrm{u}: \mathrm{C} \times \mathrm{S} \rightarrow \mathrm{R}$, where R is a bounded set of integers or reals. For each user c $\in$ C, we want to select the product or service $\mathrm{s}' \in \mathrm{S}$ that maximizes the utility function, where $\forall\, \mathrm{c} \in \mathrm{C}$, $\mathrm{s}'_{\mathrm{c}} = \arg\max_{\mathrm{s} \in \mathrm{S}} \mathrm{u}(\mathrm{c}, \mathrm{s})$.

The most popular types of service recommendations found in the literature are the following:

(1) Collaborative filtering: The user's ratings of a product/service are collected, and services recommended to the user based on the ratings of other similar users. The two most popular approaches for measuring similarities between users are those of Adomavicius and Tuzhilin (2005) and Liu et al. (2010), notably the correlation and cosine approaches, defined as follows:

- The *correlation approach* uses the Pearson correlation coefficient:

$$sim(x,y) = \frac{\sum_{s\in S_{xy}}(r_{x,s}-\bar{r}_x)(r_{y,s}-\bar{r}_y)}{\sum_{s\in S_{xy}}(r_{x,s}-\bar{r}_x)^2.\sum_{s\in S_{xy}}(r_{y,s}-\bar{r}_y)^2} \qquad (9)$$

  Where x, y are two users rating the same services, S_{xy} is the set of services rated by users x and y, ($r_{x,s}$, $r_{y,s}$) are the ratings of service s by the users x and y, and ($\bar{r}_x$, $\bar{r}_y$) are the average ratings of x and y.

- The *cosine approach* considers users as a set of vectors in a space with the dimension of the set of services S_{xy}:

$$sim(x,y) = \cos(\vec{x},\vec{y}) = \frac{\vec{x}.\vec{y}}{\|\vec{x}\|_2 \times \|\vec{y}\|_2} = \frac{\sum_{s\in S_{xy}} r_{x,s}r_{y,s}}{\sqrt{\sum_{s\in S_{xy}} r^2{}_{x,s.}}\sqrt{\sum_{s\in S_{xy}} r^2{}_{y,s}}} \qquad (10)$$

  Where $\vec{x}.\vec{y}$ is the dot product between vectors $\vec{x}$ and $\vec{y}$.

(2) Content filtering: The services are recommended to a user based on their description and the user profile and preferences (Adomavicius and Tuzhilin 2005; Henricksen et al. 2006; Sharma and Gera 2013). The utility function is represented by:

$$u(c,s) = \cos(\overrightarrow{w_c},\overrightarrow{w_s}) = \frac{\sum_i^K w_{i,c}.w_{i,s}}{\sqrt{\sum_{i=1}^K w^2{}_{i,c}}.\sqrt{\sum_{i=1}^K w^2{}_{i,s}}} \qquad (11)$$

Where $\overrightarrow{w_c}$ and $\overrightarrow{w_s}$ are the TF-IDF vectors of the keyword weights (keyword describing the content of an item), $u(c,s)$ is the utility function, and K is the total number of keywords in the system.

Case-based reasoning (CBR) is another technique used in context-aware systems for the recommendation of services (Lee and Lee 2007) in which the similarity function used to find in past cases similar to the current case (context) is based on the algorithm of the k-nearest neighbors:

$$sim(N,C)\frac{\sum_{i=1}^n f(N_i,C_i)\times W_i}{\sum_{i=1}^n W_i} \qquad (12)$$

Where N_i is the value of characteristic i of the new case, C_i is the value of characteristic i of the old case, n is the number of characteristics, $f(N_i, C_i)$ is the distance function between N_i and C_i, and W_i is the weight of characteristic i.

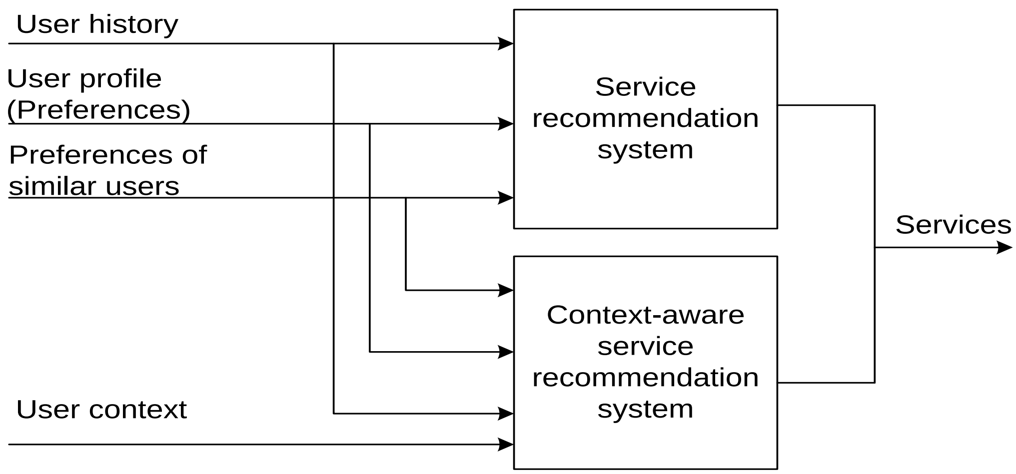

Figure 3. Service recommendation systems and context-aware service recommendation systems

b.i Context-aware services

The use of contextual information in the service recommendation process in a PCS is achieved in two ways according to Adomavicius et al. (2011): recommendations through context-driven querying and searches (including the current user context) and through contextual preference elicitation and estimation (techniques that model and learn user preferences using collaborative filtering, content filtering, or various intelligent data analysis techniques). In recommendations through the incorporation of contextual information, the context for the selection of a service in the past is the central element on which the present recommendation in a current context is based. With the context-aware collaborative filtering technique, several studies employ the Pearson correlation coefficient (Table 2, Section a) to introduce the contextual information relevant to the selection of services by multiple users in different contexts. Chen (2005) used this coefficient to measure the similarity between two sets of contextual information (Table 2, Section b) based on the assumption that if user preferences for a product do not differ in different contexts, then the ratings given in one particular context should apply to another context. Thus, if the ratings of a product are similar in two different contexts, then these two values are relevant to one another.

A similar approach is adopted by Chang and Song (2012), where the spatio-temporal similarity of the user's service ratings is evaluated by the Pearson correlation coefficient (Table 2, Section c). The assumption is that two users are more similar when they choose the same co-located services at the same time. Furthermore, Li et al. (2008) assumed that the more two users have a common location history, the more they share common interests and preferences. The proposed similarity is thus a hierarchical-graph-based similarity measurement (HGSM).

The above approaches are a set of assumptions based on the spatio-temporal context in a user's history. In the best cases, this choice can be used as the final step to evaluate and choose between two or more selected services.

The ontological representation of services as well as contextual information is largely used for the measurement of semantic similarities in the recommendation of services. In García-Crespo et al. (2009), the services described by an ontology are recommended based on the user's preferences and history, with a defined threshold that decides the relevance of the recommended service; the context is represented by the actual location of the user. The semantic similarity algorithms used are thus based on characteristics (e.g., Paolucci et al. 2002). These ontologies are also found in McGovern (2013) to describe the occupation, interests, and so forth of user m, where the semantic similarity measures are calculated between attributes of the same type (e.g., food, occupation) and each attribute is described by a more specific taxonomy that facilitates the calculation of similarity. As a result, a developer can design a richer application for a number of similar users.

Table 2: Semantic similarity measures applied to service recommendations

Semantic similarity in recommendation systems	Semantic similarity type	Studies	
Similarity between contexts	- Concept abduction (Liu et al.) - Feature-based semantic similarity measures (García-Crespo et al.) - Semantic similarity and scalar distance (according to the context definition)	- Liu et al. (2010) - García-Crespo et al. (2009)	
Case-based reasoning (Similarity between cases/contexts)	k-nearest neighbors $sim(N,C) = \frac{\sum_{i=1}^{n} f(N_i,C_i) \times W_i}{\sum_{i=1}^{n} W_i}$	-Lee and Lee (2007)	
Collaborative filtering (similarity between users)	- Pearson coefficient of correlation: $sim(x,y) = \frac{\sum_{s \in S_{xy}}(r_{x,s}-\bar{r}_x)(r_{y,s}-\bar{r}_y)}{\sqrt{\sum_{s \in S_{xy}}(r_{x,s}-\bar{r}_x)^2 . \sum_{s \in S_{xy}}(r_{y,s}-\bar{r}_y)^2}}$ - Cosine method: $sim(x,y) = \cos(\vec{x},\vec{y}) = \frac{\sum_{s \in S_{xy}} r_{x,s} . r_{y,s}}{\sqrt{\sum_{s \in S_{xy}} {r^2}_{x,s}} . \sqrt{\sum_{s \in S_{xy}} {r^2}_{y,s}}}$ - Euclidean distance between users who have rated the same product: $sim(u,u') = \frac{1}{1 + \sqrt{\sum_{j=0}^{n}(r_j - {r'}_j)^2}}$	- Adomavicius and Tuzhilin (2005) - Liu et al. (2010)	Section a
Collaborative filtering (context-aware)	- Pearson coefficient of correlation :(contextual relevance): $Rel_t(x,y,i)^* = \frac{\sum_{u=1}^{n}(r_{u,i,x_t} - \bar{r}_i).(r_{u,i,y_t} - \bar{r}_i)}{\sigma_{x_t}.\sigma_{y_t}}$	- Chen (2005)	Section b
	- Pearson coefficient of correlation (spatio-temporal similarity): $sim(x,y)^{**} = \frac{\sum_{s \in S_{xy}}(T_{x,s} - \bar{T}_x).(T_{y,s} - \bar{T}_y)}{\sqrt{\sum_{s \in S_{xy}}(T_{x,s} - \bar{T}_x)^2 . \sum_{s \in S_{xy}}(T_{y,s} - \bar{T}_y)^2}}$	- Chang and Song (2012)	Section c
Content-based measures	Cosine between keyword vectors : $u(c,s) = \cos(\vec{w}_c, \vec{w}_s) = \frac{\vec{w}_c . \vec{w}_s}{\lVert\vec{w}_c\rVert_2 \times \lVert\vec{w}_s\rVert_2})$	Adomavicius and Tuzhilin (2005)	

$\mathrm{Rel}_t(x,y,i)^*$ is the relevancy of ratings for a product i between two contextual variables x and y (of the same type) and is equivalent to their semantic similarity. r_{u,i,x_t} is the rating of a user u for a product i in a context x.

$sim(x,y)^{**}$, S_{xy} are the co-located services accessed by users x and y, $T_{x,s}$ and $T_{y,s}$ respectively denote users x and y accessing service s, and $\bar{\mathrm{T}}_x$ and $\bar{\mathrm{T}}_y$ respectively denote the mean value of time wehn users x and y access service s.

c. Semantic similarity measures and applications

In a PCS, applications must be sensitive to their execution context, which can be any element that influences the behavior of the application (Capra et al. 2001). To provide the functionalities expected by the user along with the desired quality, applications must therefore be able to reason about changes in context and reconfigure their behavior to meet well-defined objectives (Kakousis et al. 2010). Dalmau et al. (2009) categorize these adaptation objectives as the adaptation of data, services, and presentation. The first type of adaptation relates to the provision of complete and formatted information based on raw data. The adaptation of services concerns the architecture of the application (Figure 4), while that of presentation relates to the interfacing between the user and the equipment.

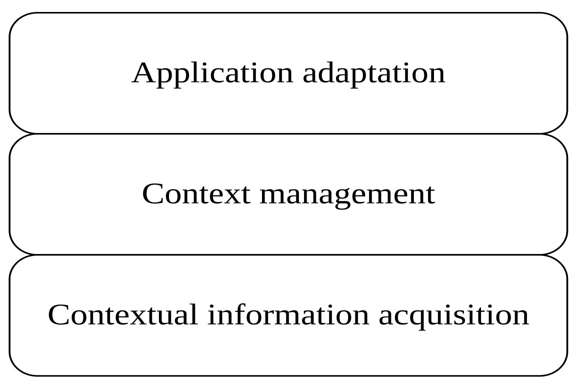

Figure 4. Context-aware applications architecture (Dalmau et al. 2009)

The semantic similarity between components of an application is defined as follows: two components are similar if the substitution of one by the other allows the user to do the same task. For example, a PowerPoint slide is similar to an Acrobat slide because both allow the presentation of texts and pictures. Ranganathan et al. (2005) applied this definition to measure the semantic similarity between components of an application by changing its architecture.

Ontologies are used to describe the semantic properties of an application's components (its function, applicable hardware, readable data formats, etc.). The semantic similarity measure uses the DL algorithm of Gonzalez-Castillo et al. (2001), and it is defined by the relative location of the components in the domain ontology, in which the two concepts C1 and C2 are similar to a certain level:

- C1 is equivalent to C2, with a similarity level of 0;
- C1 is a sub-concept of C2, with a similarity level of 1;
- C1 is a super-concept of C2 with a satisfiable intersection with C2, or C1 is a sub-concept of a super-concept of C2 with a satisfiable intersection with C2; the similarity level is 2+i, where i is the number of nodes on the path in the ontology hierarchy from C2 to the relevant super-concept of C2.

Preuveneers et al. (2009) adopted the same approach based on the modular architecture of a context-aware application for measuring the semantic similarity between components. The authors used the semantic similarity measure of Kirsch-Pinheiro et al. (2006) for content adaptation by having defined user profiles that contain information characterizing the user's context and a set of filtering rules when matching with user's current context to provide the proper content to the user.

d. Service discovery

Service discovery in a PCS is the process of locating the appropriate services to meet the needs of the entity making the request (person or device) (Huaglory Tianfield 2011; Yau et al. 2006). This process is characterized by the following phases: service query, matching, and delivery of the most appropriate service (Broens et al. 2004; Thompson 2006). The context in service discovery in a PCS is defined in Doulkeridis et al. (2006) by "the implicit information related both to the requesting user and service provider that can affect the usefulness of the returned results" (p. 4). This information is used by Yau et al. (2006) in order to:

(1) Expand the service requests to provide more relevant information that is not explicitly specified by users;
(2) Describe users' preferences to different services;
(3) Further categorize services to retrieve better results;
(4) Define the policies to provide services among service providers;
(5) Infer the service semantics based on service descriptions in the matchmaking phase in service discovery.

Finally, the information can improve the two evaluation factors “precision” and “recall” of the semantic similarity measures as well as the relevancy of services provided in a PCS (Klein and Bernstein 2004; Yau 2006). The two factors are defined as follows:

Recall = Number of relevant services retrieved in a service discovery / Total number of relevant services available

Precision = Number of relevant services retrieved in a service discovery /Total number of services identified

The semantic similarity measures involving the context used in the service discovery are applied in the phases shown in Figure 5 below:

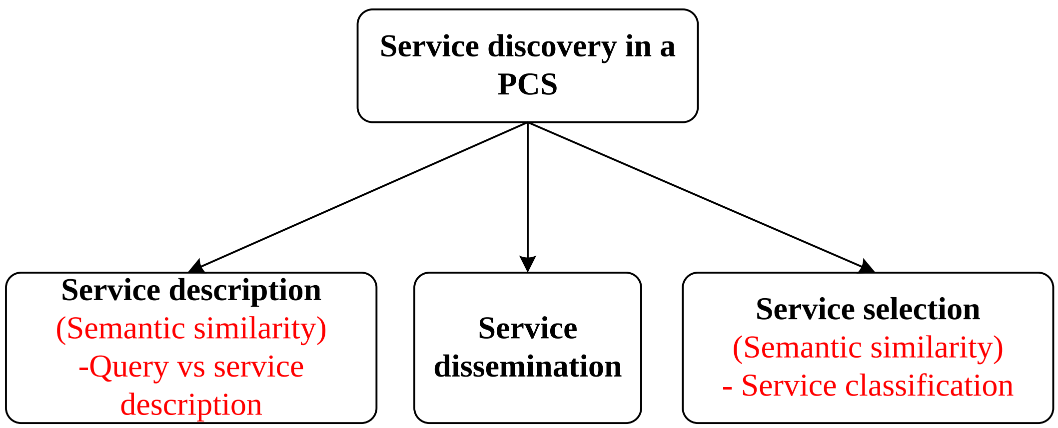

Figure 5. Service discovery in a PCS and semantic similarity measures

The semantic similarity between the query and service depends on their representation, which is either an ontological representation or expressed in DL language. For the ontological representation of queries and services, a common ontology to describe both the queries and services is required in order to implement measures such as edge counting (Aydoğan and Yolum 2007; Ge and Qui 2008; Rada et al. 1989). Moon et al. (2008) used

WordNet as the ontology to find the synonym of a DTD expressed in XML. The subsumption relation in a common ontology (Bandara et al. 2007) is the tool used for measuring the semantic similarity between the symbolic attributes of the query and the available services. The semantics of each attribute of the query and services, as described by an ontology with “is-a” and “part-of” relations, is shared by all nodes of the PCS (Kang et al. 2007). The semantic similarity measure between attributes is thus given as follows:

$$Sim(c_1, c_2) = \begin{cases} e^{-al.\frac{e^{\beta h} - e^{-\beta h}}{e^{\beta h} + e^{-\beta h}}} & si\ c_1 \neq c_2 \\ 1 & Else \end{cases} \quad (13)$$

Where l is the shortest path between the concepts c_1 and c_2, h is the level of LCS in the ontology, and $\alpha \geq 0$ and $\beta \geq 0$ are two scaling parameters for the contribution of l and h.

Finally, a graphical approach based on Tversky's semantic similarity measure is introduced in Ganter and Stumme (2002), where a service is more relevant if it has more contextual attributes (user preferences) in common with the query R. These ontology-based measures always depend on the structure of the ontology, which may change from one designer to another, and as a result, they are not always consistent.

For the description of the query R and available services O as expressed in DL, the matchmaking in this case is categorized according to the five categories listed below (Gonzalez-Castillo et al. 2001; Ruta et al. 2012):

(1) *Exact:* all features requested in R are exactly the same as those provided by O and vice versa.

(2) *Full-subsumption*: All features requested in R are contained in O.

(3) *Plug-in*: All features offered in O are contained in R.

(4) *Potential-intersection*: An intersection exists between the features offered in O and those requested in R.

(5) *Partial-disjoint*: Some features requested in R are in conflict with some of those offered in O.

For the classification of the identified services, semantic similarity measures are used to limit the number of services identified in accordance with their degree of relevance. The context is an element used in this classification. Ruta et al. (2012) represent context through the geographic proximity of the query to the service provider. This aims to classify services in terms of their functional and non-functional properties (e.g., context, quality of service) and according to the four levels of matching as defined by Paolucci et al. (2002). The current contextual information and services enriched with contextual information (e.g., age, location) both being described by ontologies, are compared node by node (Kirsch-Pinheiro et al. 2008). The semantic similarity measure thus depends on the shared proportion of nodes and arcs between the two graphs:

$$Sim_l(E_i, E_j) = \frac{Sim_l(l_i, l_j) + \sum_1^p Sim_l(C_{E_i}, C_{E_j})}{(p+1)} \tag{14}$$

Where l_i, l_j are the edge labels and C_{E_i}, C_{E_j} are the edge extremities.

The contextual information (attribute-value) described by Broens et al. (2004) is the final phase in the matching process between a query R and service description S in order to classify the results of the previous phases. The process of matching is achieved by step-by-step filtering. During each step, a property of the service (service type, input, output, contextual attribute, etc.) that is present in the query but not present in service is used to eliminate the services not relevant to the query.

Table 3: Table summarizing the studies on semantic similarities in pervasive computing

	Studies	Measure support	Type of similarity	Specificities	Contextual elements
Semantic similarity between contexts	Gicquel (2012)	Ontology	Pirró and Euzenat (2010)	User interactions contextualized according to the user profile and physical location	- User Profile - Physical location
	Wen'an Zhou (2012)	Spatio-temporal data	Pearson coefficient	Increases the ratio of user satisfaction	- Space-time
	Benazzouz (2012)	Ontology	Jiwei Zhong (2002) Rada et al. (1989)	Method is able to detect recurring patterns and improve the efficiency of context-aware services	- Current context
	Hartmann et al. (2008)	Ontology	Wordnet, Wikipedia, Wiktionary, c-vector	String-based measures have higher performance than semantic ones	- Current context
	Dietze et al. (2008)	Vectorized data	Euclidean distance	Context-adaptation across distinct mobile situations	-Technical environment - User objectives -Current location
	Kirsch-Pinheiro et al. (2006)	Graph	Common elements	Analyzes the user's current context and selects from among the user's predefined profiles	-User profile
	Meissen et al. (2005)	Ontology/taxonomy	Subsumes path within the dimensions	Delivers relevant information at the right time to mobile users	-Space-time -Current context
Semantic similarity applied to the recommendation of services	McGovern (2013)	Ontology/taxonomy	Comparison between the taxonomies of same-type attributes	Determines if a given group of users have a quantitative similarity determinant	- User proximity - Occupation - Food -Interests
	Chang and Song (2012)	Spatio-temporal data	Pearson coefficient	Adaptation based on user-to-object, space-time interaction patterns	- Space-time
	Liu et al. (2010)	Ontology/DL	Concept abduction, scalar measure	Multi-context and multi-criteria service recommendations based on collaborative filtering	- Current context - user preferences
	García-Crespo et al. (2009)	Ontology	Features of Paolucci et al.(2002)	Fusion of context-aware pervasive systems, GIS systems, social networks, and semantics	- GIS - Social networks
	Li et al. (2008)	User location	Hierarchical-graph-based similarity measurement (HGSM).	Geographically mines the similarity between users based on their location histories	- Location
	Lee and Lee (2007)	Features	k-nearest neighbors	Context-aware music recommendation system using case-based reasoning	- User profile (Listening history) - Current context
	Chen (2005)	Hierarchical structure within each context type	Pearson coefficient	System to predict a user's preference based on past experiences of like-minded users	- Current context
	Van Setten et al. (2004)	Context ontology and domain-specific rules	CBR Similarity functions	Recommendation system (COMPASS)	- Current context - Space-time - User interests

Semantic similarity applied to applications	Ranganathan et al.(2005)	Ontology/DL	Gonzalez-Castillo et al. (2001)	Allows mobile, ubiquitous applications to be adaptive, self-configuring, and self-repairing (built on top of GAIA)	- User location - Device location - Whether device already has applications running - Neighborhood - Current activity
	Preuveneers et al. (2009)	Graph	Kirsch-Pinheiro et al. (2006)	Addresses context in large-scale networks and context-aware redeployment of running applications in a distributed setting	- Current context - Location - Identifying attribute of the device
Semantic similarity applied to service discovery	Aydoğan and Yolum (2007)	Ontology	RP similarity (modified Rada et al. 1989)	Incremental learning architecture in which both consumers and producers use a shared ontology to negotiate a service	- User preferences
	Bandara et al. (2007)	Ontology	Subsumption relation, features (Tversky), scalar measure	Ranking mechanism to order available services according to their suitability	- User preferences and interests
	Kang et al. (2007)	Ontology	Li et al. (2003)	Service clustering supports scalable semantic queries with low communication overheads and balanced load distribution among resolvers	- User preferences
	Ruta et al. (2012)	Ontology/DL	Logic based	Ranks identified resources based on a combination of their semantic similarity with respect to the user request and their geographical distance from the user itself (example of tourism)	- Query and service provider geographical proximity
	Mokhtar et al. (2007)	Ontology/ EASY-L	Paolucci	Supports efficient, semantic, context-aware service and quality-of-service aware identification in addition to the existing SDPs (Ariadne) Framework: EASY	- Current context - User profile
	Kirsch-Pinheiro et al.(2008)	Ontology/graph	Local similarity measures between concepts and global measures between graphs Music PROJECT-	A graph-based algorithm for matching contextual service descriptions using similarity measures and allowing inexact matches	- Current context - Space-time
	Broens et al. (2004)	Ontology	Li and Horrocks (2004), clustering	Uses ontologies to capture the semantics of the user's query, services, and the contextual information	- Current context - Space-time - User preferences

IV. CONCLUSION

In this article, a survey of the semantic similarity measures applied in the field of pervasive computing was presented. The works related to the application of semantic similarity measures between contexts/situations, service recommendations, applications, and service discovery. Semantic similarity measures in the field of pervasive computing mainly relate to the notion of context and its representation. The most common representations of context are through ontologies given the qualities that they provide (possibility of reasoning, sharing, and reusing through digital media, etc.) despite their high costs. This representation allows the application of various measures of semantic similarity based on the structure of the ontology and the characteristics of the concepts. In most applications, context is represented by the spatio-temporal information of the user as well as his preferences and interests (recommendation and discovery of services), which is used as a service classification factor according to relevance. Pearson's correlation coefficient is the most frequent semantic similarity measure in the area of the service recommendations using the technique of collaborative filtering, which can be modified to include the contextual information.

REFERENCES

[1] Adomavicius, G., & Tuzhilin, A. (2011). Context-aware recommender systems. In Recommender Systems Handbook (pp. 217-253). Springer US.

[2] Adomavicius, G., & Tuzhilin, A. (2005). Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. Knowledge and Data Engineering, IEEE Transactions on, 17(6), 734-749.

[3] Al-Mubaid, H., & Nguyen, H.A. (2006). Using MEDLINE as standard corpus for measuring semantic similarity of concepts in the biomedical domain, In Proc. of the IEEE 6th Symposium on Bioinformatics and Bioengineering, (pp. 315-318).

[4] Aydoğan, R., & Yolum, P. (2007, May). Learning consumer preferences using semantic similarity. In Proceedings of the 6th International Joint Conference on Autonomous Agents and Multiagent Systems (p. 229). ACM.

[5] Bandara, A., Payne, T., De Roure, D., & Lewis, T. (2007). A semantic approach for service matching in pervasive environments. Technical Report Number: ECSTR-IAM07-006, University of Southampton.

[6] Benazzouz, Y. (2012). Découverte de contexte pour une adaptation automatique de services en intelligence ambiante. Doctoral dissertation, Ecole Nationale superieure de mines, Saint-Etienne, France.

[7] Bisson, G. (2000), La similarite: Une notion symbolique/numerique. IMAG-CNRS, Projet SHERPA, Unité de recherche INRIA Rhone-Alpes, (p. 3).

[8] Broens, T., Pokraev, S., Van Sinderen, M., Koolwaaij, J., & Costa, P. D. (2004). Context-aware, ontology-based service discovery. In Ambient Intelligence (pp. 72-83). Springer Berlin Heidelberg.

[9] Bulskov H., Knappe R., & Andreasen T. (2002). On measuring similarity for conceptual querying. In the Proc. of the 5th Int'l Conf. on Flexible Query Answering Systems (pp. 100-111). Springer-Verlag.

[10] Capra, L., Emmerich, W., & Mascolo, C. (2001). Reflective middleware solutions for context-aware applications. In Metalevel Architectures and Separation of Crosscutting Concerns (pp. 126-133). Springer Berlin Heidelberg.

[11] Chang, J., & Song, J. (2012, May). Research on context-awareness service adaptation mechanism in IMS under ubiquitous network. In Vehicular Technology Conference (VTC Spring), 2012 IEEE 75th (pp. 1-5). IEEE.

[12] Chen, A. (2005). Context-aware collaborative filtering system: Predicting the user's preference in the ubiquitous computing environment. In Location-and Context-Awareness (pp. 244-253). Springer Berlin Heidelberg.

[13] Dalmau, M., Roose, P., & Laplace, S. (2009). Context-aware adaptable applications: A global approach. International Journal of Computer Science Issues, Vol.1, (pp. 13-25).

[14] d'Amato, Claudia (2007). Similarity-based learning methods for the semantic web, PhD thesis, Universita Degli Studi di Bari Faculta di Scienze Dipartimento di Informatica, pp 97 .

[15] d'Amato, C., Fanizzi, N., & Esposito, F. (2009). A semantic similarity measure for expressive description logics. Universita Degli Studi di Bari Faculta di Scienze Dipartimento di Informatica arXiv preprint arXiv:0911.5043.

[16] Dey, A.K. (2001). Understanding and using context. College of Computing & GVU Center, Georgia Institute of Technology, Atlanta,Personal and Ubiquitous Computing, Vol.5, (pp. 4-7).

[17] Dietze, S., Gugliotta, A., & Domingue, J. (2008). Bridging the gap between mobile application contexts and semantic web resources: Context-aware mobile and ubiquitous computing for enhanced usability: adaptive technologies and applications. Information Science Publishing (IGI Global).

[18] Doulkeridis, C., Loutas, N., & Vazirgiannis, M. (2006). A system architecture for context-aware service discovery. Electr. Notes Theor. Comput. Sci., 146(1), 101-116.

[19] Efstratiou, C. (2004). Coordinated adaptation for adaptive context-aware applications. Doctoral dissertation, Computing Department, Lancaster University, UK, (pp. 173).

[20] Ehrig, M., Haase, P., Hefke, M., & Stojanovic, N. (2005). Similarity for ontologies-a comprehensive framework. ECIS 2005 Proceedings, 127.

[21] El Sayed, A., Hacid, H., & Zighed, D. (2007). A new context-aware measure for semantic distance using a taxonomy and a text corpus. In Information Reuse and Integration. IEEE International Conference on (pp. 279-284). IEEE.

[22] Ganter, B., & Stumme, G., 2002, Formal concept analysis: Methods and applications in computer science. TU Dresden, http://www.aifb.uni-karlsruhe.de/WBS/gst/FBA03.shtml.

[23] García-Crespo, A., Chamizo, J., Rivera, I., Mencke, M., Colomo-Palacios, R., & Gómez-Berbís, J.M. (2009). SPETA: Social pervasive e-tourism advisor. Telematics and Informatics, 26(3), 306-315.

[24] Ge, J., & Qiu, Y. (2008). Concept similarity matching based on semantic distance. In Semantics, Knowledge and Grid, 2008. SKG'08. Fourth International Conference on (pp. 380-383). IEEE.

[25] Germán, S. (2010). Adaptation d'architectures logicielles collaboratives dans les environnements ubiquitaires. Contribution à l'interopérabilité par la sémantique. Doctoral dissertation, Systèmes (EDSYS), France.

[26] Gicquel, P.Y (2012). Similarités sémantiques et contextuelles pour l'apprentissage informel en mobilité. RJC EIAH'2012, 45.

[27] Gomaa, W.H., & Fahmy, A.A. (2013). A survey of text similarity approaches. International Journal of Computer Applications, 68(13), 13-18.

[28] Gonzalez-Castillo, J., Trastour, D., & Bartolini, C. (2001). Description logics for matchmaking of services. HP Laboratories technical report, 265.

[29] Harispe, S., Ranwez, S., Janaqi, S., & Montmain, J.(2013). Semantic measures for the comparison of units of language, concepts or instances from text and knowledge representation analysis, A Comprehensive Survey and a Technical Introduction to Knowledge-based Measures Using Semantic Graph Analysis, LGI2P/EMA Research Center, Parc scientifique, France.

[30] Hartmann, M., Zesch, T., Muhlhauser, M., & Gurevych, I. (2008). Using similarity measures for context-aware user interfaces. In Semantic Computing, 2008 IEEE International Conference on (pp. 190-197). IEEE.

[31] Henricksen, K., Indulska, J., & Rakotonirainy, A. (2006). Using context and preferences to implement self-adapting pervasive computing applications. Software: Practice and Experience, 36(11-12), 1307-1330.

[32] Hirst, G., & St Onge, D. (1998). Lexical chains as representations of context for the detection and correction of malapropisms. In C. Fellbaum (ed.), WordNet: An Electronic Lexical Database, Cambridge, MA: The MIT Press.

[33] Janowicz, K. (2008). Kinds of contexts and their impact on semantic similarity measurement. In Pervasive Computing and Communications. Sixth Annual IEEE International Conference on (pp. 441-446). IEEE.

[34] Jiang J.J., & Conrath D.W. (1997). Semantic similarity based on corpus statistics and lexical taxonomy. Proceedings of International Conference on Research in Computational Linguistics, August 22-24; Taipei, Taiwan.

[35] Kakousis, K., Paspallis, N., & Papadopoulos, G.A. (2010). A survey of software adaptation in mobile and ubiquitous computing. Enterprise Information Systems, 4(4), 355-389.

[36] Kang, S., Kim, D., Lee, Y., Hyun, S.J., Lee, D., & Lee, B. (2007). A semantic service discovery network for large-scale ubiquitous computing environments. ETRI journal, 29(5), 545-558.

[37] Keßler, C. (2007). Similarity measurement in context. In Modeling and Using Context (pp. 277-290). Springer Berlin Heidelberg.

[38] Keßler, C., Raubal, M., & Janowicz, K. (2007). The effect of context on semantic similarity measurement. In On the Move to Meaningful Internet Systems: OTM 2007 Workshops (pp. 1274-1284). Springer Berlin Heidelberg.

[39] Kirsch-Pinheiro, M., Vanrompay, Y., & Berbers, Y. (2008). Context-aware service selection using graph matching. In 2nd Non Functional Properties and Service Level Agreements in Service Oriented Computing Workshop (NFPSLA-SOC'08), ECOWS. CEUR Workshop proceedings (Vol. 411).

[40] Kirsch-Pinheiro, M., Villanova-Oliver, M., Gensel, J., & Martin, H. (2006). A personalized and context-aware adaptation process for web-based groupware systems. In 4th International Workshop on Ubiquitous Mobile Information and Collaboration Systems, CAiSE'06 Workshop (pp. 884-898).

[41] Klein, M., & Bernstein, A. (2004). Towards high-precision service retrieval. IEEE Internet Computing, January, 30-36.

[42] Lavirotte, S., Lingrand, D., & Tigli, J.Y. (2005). Définition du contexte: fonctions de coût et méthodes de sélection. In Proceedings of the 2nd French-speaking Conference on Mobility and Ubiquity Computing (pp. 9-12). ACM.

[43] Leacock, C., & Chodorow, M. (1998). Combining local context and WordNet similarity for word sense identification. In WordNet: An Electronic Lexical Database, C. Fellbaum, MIT Press.

[44] Lee, J.S., & Lee, J.C. (2007). Context awareness by case-based reasoning in a music recommendation system. In Ubiquitous Computing Systems (pp. 45-58). Springer Berlin Heidelberg.

[45] Li, Q., Zheng, Y., Xie, X., Chen, Y., Liu, W., & Ma, W.Y. (2008). Mining user similarity based on location history. In Proceedings of the 16th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (p. 34). ACM.

[46] Li, Y., Bandar, Z.A., & McLean, D. (2003). An approach for measuring semantic similarity between words using multiple information sources. IEEE Transactions on Knowledge and Data Engineering, 15(4), 871-882.

[47] Lin, D (1998). An information-theoretic definition of similarity. In Proceedings of the 15th International Conference on Machine Learning, July 24-27 1998; Madison, Wisconsin, USA.

[48] Li, L., & Horrocks, I. (2004). A software framework for matchmaking based on semantic web technology. International Journal of Electronic Commerce, 8(4), 39-60.

[49] Liu, L., Lecue, F., Mehandjiev, N., & Xu, L. (2010). Using context similarity for service recommendation. In Semantic Computing (ICSC), 2010 IEEE Fourth International Conference on (pp. 277-284). IEEE.

[50] Liu, Q., Ma, H., Chen, E., & Xiong, H. (2013). A survey of context-aware mobile recommendations. International Journal of Information Technology & Decision Making, 12(1), 139-172.

[51] Maedche, A., & Staab, S. (2002). Measuring similarity between ontologies. In Knowledge Engineering and Knowledge Management: Ontologies and the Semantic Web (pp. 251-263). Springer Berlin Heidelberg.

[52] McGovern, J. (2013). Context similarity evaluation: Inferring how users can collectively collaborate together in a pervasive environment. In Cloud and Green Computing (CGC), 2013 Third International Conference on (pp. 553-557). IEEE.

[53] Meissen, U., Pfennigschmidt, S., Voisard, A., & Wahnfried, T. (200,). Context-and situation-awareness in information logistics. In Current Trends in Database Technology-EDBT 2004 Workshops (pp. 335-344). Springer Berlin Heidelberg.

[54] Meng, L., Huang, R., & Gu, J. (2013). A review of semantic similarity measures in wordnet. International Journal of Hybrid Information Technology, 6(1), 1-12.

[55] Michel, M.D., & Deza, E. (2007). Dictionnaire des distances. In Encyclopedia of Distances.

[56] Mihalcea, R., Corley, C., & Strapparava, C. (2006). Corpus-based and knowledge-based measures of text semantic similarity. In AAAI (Vol. 6, pp. 775-780).

[57] Mokhtar, S.B., Preuveneers, D., Georgantas, N., Issarny, V., & Berbers, Y. (2008). EASY: Efficient semAntic Service discoverY in pervasive computing environments with QoS and context support. Journal of Systems and Software, 81(5), 785-808.

[58] Moon, H. J., Kim, S., Moon, J., & Lee, E. S. (2008). An Effective data processing method for fast clustering. In Computational Science and its Applications–ICCSA 2008 (pp. 335-347). Springer Berlin Heidelberg.

[59] Nicklas, D., & Henricksen, K. (2008). Context modeling and reasoning: Key concepts for Pervasive computing. 5th IEEE Workshop on Context Modeling and Reasoning (CoMoRea'08) @PerCom Hong Kong.

[60] Paolucci, M., Kawamura, T., Payne, T.R., & Sycara, K. (2002). Semantic matching of web services capabilities. Lecture Notes in Computer Science, 2342, 333–347.

[61] Petit, M., (2005). L’informatique contextuelle. Technical Report, South Britany University (UBS), France.

[62] Pirró, G., & Euzenat, J. (2010). A feature and information theoretic framework for semantic similarity and relatedness. In The Semantic Web–ISWC 2010 (pp. 615-630). Springer Berlin Heidelberg.

[63] Preuveneers, D., Victor, K., Vanrompay, Y., Rigole, P., Pinheiro, M.K., & Berbers, Y. (2009). Context-aware adaptation in an ecology of applications. Context-Aware Mobile and Ubiquitous Computing for Enhanced Usability: Adaptive Technologies and Applications, 1-25.

[64] Ramparany, F., Benazzouz, Y., Gadeyne, J., & Beaune, P. (2011). Automated context learning in ubiquitous computing environments. In SSN (pp. 9-21).

[65] Ranganathan, A., Shankar, C., & Campbell, R. (2005). Application polymorphism for autonomic ubiquitous computing. Multiagent and Grid Systems, 1(2), 109-129.

[66] Rada, R., Bicknell, H., Mili, E., & Blettner, M (1989). Development and application of a metric on semantic nets. IEEE Transaction on Systems, Man, and Cybernetics, 1(19), 17-30.

[67] Resnik, P (1995). Using information content to evaluate semantic similarity. In Proceedings of the 14th International Joint Conference on Artificial Intelligence, August 20-25; Montréal Québec, Canada.

[68] Rodriguez M.A., & Egenhofer, M.J. (2003). Determining semantic similarity among entity classes from different ontologies. IEEE Transactions on Knowledge and Data Engineering, 15, 442-456.

[69] Rubinstein, H. & Goodenough, J.B. (1965). Contextual correlates of synonymy. Communications of the ACM, 8(10).

[70] Ruta, M., Scioscia, F., Di Sciascio, E., & Piscitelli, G. (2012). Semantic matchmaking for location-aware ubiquitous resource discovery. International Journal on Advances in Intelligent Systems, 4(3/4), 113-127.

[71] Sánchez, D., Batet, M., Isern, D., & Valls, A. (2012). Ontology-based semantic similarity: A new feature-based approach. Expert Systems with Applications, 39(9), 7718-7728.

[72] Saruladha, K. (2011). Semantic similarity measures for information retrieval systems using ontology. Doctoral dissertation, Department of Computer Science, School of Engineering and Technology, Pondicherry University, chapter 2.

[73] Saruladha, K., Aghila, G., & Raj, S. (2010). A survey of semantic similarity methods for ontology based information retrieval. In Machine Learning and Computing (ICMLC), Second International Conference on (pp. 297-301). IEEE.

[74] Schilit, B., Adams, N., & Want, R. (1994). Context-aware computing applications. In IEEE Workshop on Mobile Computing Systems and Applications . Santa Cruz, CA, US.

[75] Sharma, L., & Gera, A. (2013). A survey of recommendation system: Research challenges. International Journal of Engineering Trends and Technology (IJETT), 4(5), 1989-1992.

[76] Simonin, J., & Carbonell, N. (2007). Interfaces adaptatives, Adaptation dynamique à l'utilisateur courant. arXiv preprint arXiv:0708.3742.

[77] Sussna M. (1993). Word sense disambiguation for free-text indexing using a massive semantic network. In Proc. of Second Int'l Conf. Information Knowledge Management (CIKM '93).

[78] Thompson, M.S. (2006). Service discovery in pervasive computing environments. Doctoral dissertation, Virginia Polytechnic Institute and State University.

[79] Tversky, A (1977). Features of similarity. Psycological Review, 84(4).

[80] Van Setten, M., Pokraev, S., & Koolwaaij, J. (2004). Context-aware recommendations in the mobile tourist application COMPASS. In Adaptive Hypermedia and Adaptive Web-based Systems (pp. 235-244). Springer Berlin Heidelberg.

[81] Viterbo, J., Mazuel, L., Charif, Y., Endler, M., Sabouret, N., Breitman, K., & Briot, J. (2008). Ambient intelligence: Management of distributed and heterogeneous context knowledge. In CRC Studies in Informatics Series (pp. 1-44). Chapman & Hall.

[82] Wu, Z., & Palmer, M. (1994). Verb semantics and lexical selection. In Proceedings of the 32nd Annual Meeting of the Associations for Computational Linguistics (pp. 133-138).

[83] Yau, S.S., & Huang D. (2006). Mobile middleware for situation-aware service discovery and coordination. In P. Bellavista and A. Corradi (eds.), Handbook of Mobile Middleware, (pp. 1059-1088).

[84] Zhang, F., Liu, W., & Bi, Y. Review on Wordnet-based ontology construction in China, International Journal on Smart Sensing and Intelligent Systems, vol. 6, No. 2, April 2013.

[85] Zhong, J., Zhu, H., Li, J., & Yu, Y. (2002). Conceptual graph matching for semantic search. In Proceedings of the 10th International Conference on Conceptual Structures (ICCS) (pp. 92-196). Springer-Verlag, London.

[86] Zouari, M. (2011). Architecture logicielle pour l'adaptation distribuée: Application à la réplication de données. Doctoral dissertation, Université Rennes 1, France.