

An Environment for Collective Perception based on Fuzzy and Semantic Approaches

Giuseppe D’Aniello and Matteo Gaeta

Department of Information and Electrical Engineering and Applied Mathematics,
University of Salerno, Fisciano, Italy

Francesca Loia

Department of Management,
Sapienza, University of Roma, Rome, Italy

Marek Reformat

Department of Electrical and Computer Engineering, University of Alberta,
Edmonton, Alberta, Canada

Daniele Toti

Department of Sciences, Roma Tre University,
Rome, Italy

April 28, 2017

Abstract

This work proposes an environment implementing a methodology for acquiring and exploiting the collective perception of Points of Interests (POIs) in a Smart City, which is meant to support decision makers in urban planning and management. This environment relies upon semantic knowledge discovery techniques and fuzzy computational approaches, including natural language processing, sentiment analysis, POI signatures and Fuzzy Cognitive Maps, turning them into a cohesive architectural blend in order to effectively gather the realistic perception of a user community towards given areas and attractions of a Smart City. The environment has been put to the test via a thorough experimentation against a massive user base of an online community

with respect to a large metropolitan city (the City of Naples). Such an experimentation yielded consistent results, useful for providing decision makers with a clear awareness of the positive as well as critical aspects of urban areas, and thus helping them shape the measures to be taken for an improved city management and development.

1 Introduction

Decision Support Systems (DSS), since their introduction several decades ago, have progressively gained popularity and momentum over the years and have profited from the research products stemming from a wide range of scientific areas, including arti-

ficial intelligence, database systems, operational research and, more recently, knowledge discovery and the Semantic Web. As a matter of fact, plentiful DSS proposals were filed in literature and a few of them even succeeded in becoming commercial applications [1, 2, 3, 4]. The ultimate purpose of these systems is to complement the decision makers' own insight, knowledge and intuitions with structured, rational models and approaches, so that the overall decision-making process may be improved; in addition, thanks to their computational and automated mechanisms, the improvement they produce is not only reflected upon the effectiveness of the process, but is also tangible in terms of its performance, by reducing the time needed for making decisions, increasing the satisfaction of their users and providing better and more accurate motivations behind the decisions themselves. This is also due to the ever-increasing availability of information coming from disparate sources that was put at these systems' disposal in order to further enhance their accuracy. With the advent of the Semantic Web and the explosion of social networks, a deluge of information has flooded the Web and can now be exploited by DSS as well for a variety of purposes and applications, thus potentially strengthening their success and pervasiveness [5]. It is no wonder, then, that decision support systems are becoming increasingly pivotal even for the recent paradigm of Smart City, as defined by IBM [6]. While DSS used within such a context typically tend to rely on the data derived from the range of sensors and logs that capture the "state" of the Smart City from a quantitative perspective, a whole amount of subjective, qualitative but nonetheless useful information can be found scattered across social interactions among people on the web and similar media. The idea that this work wants to stress out is the fact that, by gathering and taking advantage of social data from people and citizens, the resulting discoveries could prove extremely

useful for decision makers to enact policies, implement measures and carry out actions upon an urban area. That is why this work proposes an environment meant to acquire and exploit the collective perception of people and citizens, as coming from social media, towards Points of Interest (POIs) of a Smart City, relying upon an innovative combination of techniques from different research areas, including semantic knowledge discovery, natural language processing, sentiment analysis, as well as fuzzy computational approaches and the concept of POI and Area signatures. This environment aims at providing decision makers with an advanced system that might be able to improve their decision processes and help them shape the present and future of a Smart City.

The paper is structured as follows. In Section 2, the methodological background of the proposed environment is reported. Section 3 describes the architecture of the environment along with its enabling technologies. In Section 4, an extensive experimentation on a large metropolitan city is reported. Finally, in Section 6 conclusions are drawn.

2 Background

This section describes the methodological substrate underlying the proposed environment, which includes the concepts of semantics-based knowledge discovery, POI signatures and Fuzzy Cognitive Maps for Smart Cities, as detailed respectively in the following subsections.

2.1 Semantic Knowledge Discovery

Knowledge discovery based on semantic techniques has been progressively acquiring prominence within the context of intelligent systems as a method to extract, as automatically as possible, meaningful information from unstructured or semi-structured tex-

tual data sources and turn it into a cohesive representation. In this regard, a number of methodologies and tools, either freely-accessible or covered by commercial licenses, are nowadays available for performing a variety of knowledge discovery tasks from textual sources. In this work, CONCEPTUM [7] and a custom-made Sentiment Analysis tool based on SentiWordNet [8] have been used for carrying out the tasks related to the extraction and building of the POI knowledge base and to the evaluation of the sentiment (expressed in terms of positivity/objectivity/negativity) behind the words and expressions used to talk about a POI by the users of a community. Further details on how such activities are performed can be found in Section 3.

2.2 POI signatures

As mentioned in [9], POI signatures are defined as follows. Given the following expressions (1) and (2) related to users of a community:

$$b_i = \frac{\# \text{ observations by the } i^{\text{th}} \text{ user}}{\max \# \text{ different observations by a user}} \quad (1)$$

$$\text{UserActivity}(u) = \left\{ \frac{b_i}{u_i} \right\} \quad (2)$$

with $i = 1, \dots, n$ where n is the number of users, and the following expressions (3) and (4) related to their observations:

$$a_j = \frac{\# \text{ times } o_j \text{ is used}}{\max \# \text{ users making a single observation}} \quad (3)$$

$$\text{ObsPop}(o) = \left\{ \frac{a_j}{o_j} \right\} \quad (4)$$

with $j = 1, \dots, m$ where m is the number of observations, a POI signature is defined as:

$$\text{POISignature}(u, o) = \text{UserActivity}(u) \times \text{ObsPop}(o) \quad (5)$$

As such, the concept of POI signatures allows to characterize a given POI in terms of the observations made by users with respect to it. These observations, for the purposes of this work, are adjectives or expressions used or mentioned by users within the context of their POI-related comments posted on online discussion boards, forums or social networks.

2.3 Fuzzy Cognitive Maps for Smart Cities

Fuzzy Cognitive Maps (FCMs) [10] are graphs meant to represent a variety of relationships among concepts like events, processes or states FCMs allow qualitative reasoning on the states of complex systems [11]. In a FCM, a node of the graph represents a concept C_i , connected via cause/effect relationships to other concepts, and making up a key factor of the modeled system. The strength of the relationship between concepts C_i and C_j is represented by a weight w_{ij} : negative weights represent inverse causality, while positive ones mean direct causality. The activation value A_i of concept C_i is given by:

$$A_i^{k+1} = f(A_i^k + \sum_{j=1, j \neq i}^n A_j^k w_{ji}) \quad (6)$$

where A_i^{k+1} is the activation level of the concept C_i at the $k+1$ iteration, A_j^k is the activation level of the concept C_j at the k iteration, w_{ji} is the weight between concepts C_j and C_i , and f is a threshold function. Considering that a FCM represents cause-effect relationships among different concepts, it has been widely employed for supporting what-if analysis. Indeed, it is possible to consider an initial scenario of simulation given by a activation vector $S^0 = (s_1, \dots, s_n)$ where $s_i \in [0, 1]$ is the activation level of concept C_i . Starting from the activation vector S^0 of the initial scenario, it is possible to compute the activation values of the concepts in the following iteration of the map, resulting in a set of transitions

$A^0 \rightarrow A^1 \rightarrow \dots$ When a limit cycle is found or a fixed state is found, it is possible to know which is the new state of the map or of specific concepts of interest. In such a way, it is clear that, starting from a baseline scenario, it is possible to simulate what will happen when the state of a specific concept changes, thus performing what-if analysis.

In the context of Smart Cities, many papers explored the use of FCMs as an analytical and decision-support tool in different contexts like, for instance, to support urban resilience analysis [12], green issues [13], urban planning [14]. The authors in [9] proposed the use of FCMs as means to model the mental representations that communities of citizens or experts have shaped with respect to urban issues. In such a case, the urban planners and the decision makers have a powerful tool to analyze the impact of the collective perception that the citizens possess about a specific city asset (*e.g.* quality of transportation, level of safety, etc.) and to understand how such a perception may influence the other assets of the city. In order to be able to activate the FCM with the collective perception of the citizens, so as to perform scenario analysis, it is firstly needed to identify high level objectives for the Smart City (*i.e.* why are we performing such scenario analysis?). Starting from such objectives, it is possible to define a set of qualitative indicators related to the different city assets (*e.g.* the level of safety as perceived by the community, the quality of the urban area, etc.). Such indicators, together with the above mentioned objectives allow us to define a Fuzzy Cognitive Map supporting the reasoning on different perspectives that the stakeholders of the city have with regards to the high-level objectives such as “improving quality of life”, “improving safety”, “reducing pollution” and so on. Further details on the definition and use of Fuzzy Cognitive Map activated by the collective perception can be found in [9]. In the evaluation section (Section 4) a complete example of the use of an FCM

for performing what-if analysis in a real case will be provided.

3 An environment for extracting and evaluating the users’ perception of POIs

The environment proposed in this work is a software architecture made up of a number of different components, whose fundamental purpose is to find information related to POIs among online sources and understand how such POIs are perceived by users. As such, the environment includes the following components:

1. a Knowledge Extractor and Builder module (KEB), which encapsulates software components for crawling texts from online sources and processing them via lexical analysis, NLP and wikification techniques, as well as for building taxonomies from the texts themselves;
2. a Sentiment Detector and Analyzer module (SDA), whose purpose is to find textual references to POIs from online messages and evaluate their sentiment values in terms of the positivity, negativity or neutrality of the opinions expressed by online users;
3. a Collective Perception Identifier module (CPI), meant to produce the POI signatures for each POI, or the Area signature for an entire area containing more than one POI, thus evaluating the collective perception (CP) of the community with respect to the POIs or the Area. The CP is then used in order to perform scenario analysis by using Fuzzy Cognitive Maps.

The conceptual architecture of the proposed environment is depicted in Figure 1.

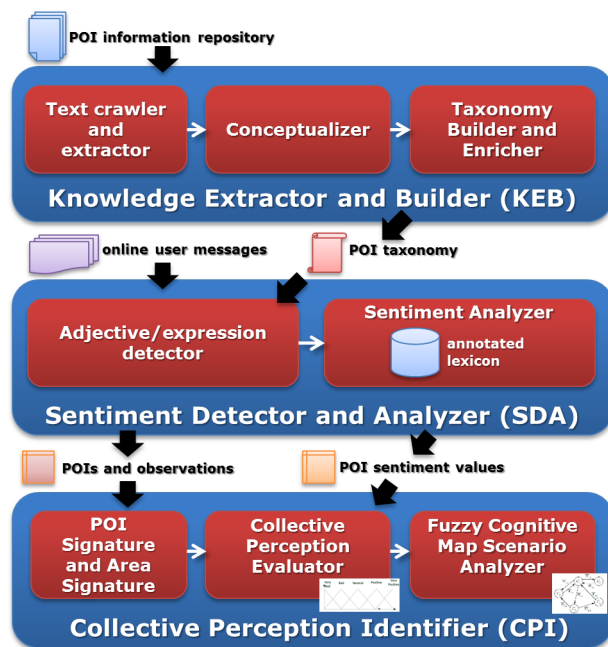


Figure 1: Conceptual architecture of the environment.

3.1 Extraction and building of the POI knowledge base

The Knowledge Extractor and Builder module (KEB) is responsible of acquiring information about the POIs to be considered and turning it into a structured taxonomy of relevant concepts. Given a POI, such a module implements the procedure made up of the following steps.

3.1.1 Text crawling and extraction

Textual repositories containing relevant information about the POI are crawled. Specifically, at this stage of the work the Wikipedia page (article) describing the POI is the one considered. This choice was made to provide the system with an information summary for the selected POI, given the general availability

of Wikipedia articles and the user-driven knowledge usually contained within them. Thus, the automatic crawling of this information overcomes the need to manually identify relevant tags for the specific POI and painstakingly annotate it accordingly. Besides, the user-driven nature of such information makes the subsequent phases of the whole process easier, especially concerning the gathering of users' observations (as described later in Section 3.2) and specifically the detection of POI references from the user community (Section 3.2.1), since users may be in principle more inclined to use similar terms or expressions as those found within the corresponding Wikipedia article.

Since the length of an article may vary and a deluge of potentially irrelevant information may flood the system, in order to minimize the possibility of ending up with an excessive number of irrelevant terms, only the introduction paragraph of the page is currently taken into account. In the event that an introduction paragraph is not present, the first section (with the exception of the "History" section) is considered; it might be also worth mentioning that, when considering a sufficiently small number of POIs, a supervision of this mechanism may be employed so that only the most relevant blocks of text are actually used and passed to the system.

Further developments of this work may come to include specific information pages from additional official and informal sources, including the tourism websites of the city, province or region where the considered POI is located.

3.1.2 Conceptualization

A conceptualization phase is carried out upon the retrieved text, by using the corresponding service provided by the CONCEPTUM system [7]. This phase performs natural language processing upon the given text and returns the semantically most relevant con-

cepts from the text by cross-referencing a common-sense knowledge base (Wikipedia). As such, the identified concepts are not necessarily those with the highest number of occurrences within the text: the number of their occurrences is weighted against their relevance score according to their respective presence and affinity in documents (articles) of the common-sense knowledge base. Specifically, this conceptualization phase processes the retrieved text as follows:

- the text is preliminarily cleaned by removing punctuation, stop-words etc.;
- the cleaned text is tagged with Part-Of-Speech (POS) tags and lemmatized accordingly;
- a shallow syntax parsing is performed upon the text;
- the information from POS tagging and syntax parsing, along with some language-dependent proximity rules, is used to compose terms (*e.g.* adjective + noun, contiguous upper-case nouns etc.)
- while the above mentioned language processing is performed, a Wikification process of the original text, based on Wikipedia Miner [15] and originally applied in [16, 17] is carried out in order to produce a number of topics from the Wikipedia common-sense knowledge base that are semantically relevant to the given text;
- a matching is performed between the composed terms from the language processing step and the common-sense topics from the wikification step, by using syntactical distances (*e.g.* Levenshtein, Jaccard) and synonym expansion. Each match is thus assigned a matching score: those matches scoring higher than a given percentage threshold (currently 70%) are deemed good;

- the resulting good matches (composed term-topic pairs) correspond to the most relevant semantic concepts from the input text; specifically, the topic element from the pair is selected as the concept to be returned for each match.

3.1.3 Taxonomy building and enrichment

A knowledge base is created from the returned concepts. Currently, the generated knowledge base for each POI is a one-level taxonomy including among its terms the concepts obtained from the previous step. Furthermore, a subsequent enrichment step is performed in order to associate with the POI's name a list of potential synonyms or corresponding alternate spellings: this is done via another Wikification-based process provided by CONCEPTUM, which returns a number of "alternate" senses for the POI's name, as featured in the Wikipedia knowledge base. For instance, for the "Subterranean Naples" POI, synonyms and additional terms returned include "Subterranean Caves", "Underground Naples", "Naples City Caves", etc. A refinement that is under research in this regard is related to the creation of a multi-level taxonomy or a full-fledged ontology, by taking advantage of another service provided as a beta version by CONCEPTUM, which detects a subset of relationships (namely hypernymy, relatedness and hierarchy) for the identified concepts of each POI, in order to generate a small semantic graph containing them.

3.2 Gathering and analysis of users' observations from online sources

Once the knowledge bases for the considered POIs are created, the Sentiment Detector and Analyzer module (SDA) takes care of scanning users' comments from online forums, message boards or social networks, in order to (i) find adjectives and expres-

sions related to the considered POIs and (ii) understand and quantify their positivity/negativity values. These two activities are detailed below.

3.2.1 Detection of adjectives and expressions related to POIs

The first activity revolves around the identification of comments, posts and texts uploaded by users where the considered POIs are mentioned, either explicitly or implicitly. The bulk of such texts is therefore crawled and scanned in order to detect the presence of a POI “reference”, *i.e.* either a POI’s explicit name along with its synonyms/alternate spellings, or some of the concepts associated with it via the generated taxonomy. For each sentence containing a POI reference, adjectives directly linked to the reference, as well as expressions as predicative nominals linked by copulas to the reference, are detected and retrieved via syntax parsing. The result of this activity produces a map structure whose keys are the POIs referenced within the scanned texts and whose values are the corresponding adjectives and expressions referring to them.

3.2.2 Analysis of the sentiment associated with the detected adjectives and expressions

The map produced by the previous activity is then passed to the submodule responsible for the sentiment check. Specifically, for each POI, the adjectives and expressions related to it are checked against a lexicon annotated with sentiment values in order to establish their potential positive, negative or objective value. The lexicons currently used by the system are SentiWordNet [8] for the English language, and the one used in [18] as a byproduct of three lexicons (SentiWordNet, MultiWordNet and WordNet itself) for the Italian language. The submodule returns, for each adjective/expression, values in the

[0,1] range that represent the adjective/expression’s positivity, negativity, or neutrality, whose sum total is 1. As such, for the i -eth POI, its corresponding positivity (Sp_i), negativity (Sn_i) or neutrality (Su_i) values are computed as follows:

$$Sp_i = \frac{\sum_{k=1}^K p_k}{K} \quad (7)$$

$$Sn_i = \frac{\sum_{k=1}^K n_k}{K} \quad (8)$$

$$Su_i = \frac{\sum_{k=1}^K u_k}{K} \quad (9)$$

where K is the total number of adjectives/expressions found and evaluated, and p_k , n_k and u_k are the k -th positivity, negativity and neutrality value, respectively, for the k -th adjective/expression.

3.3 An environment for extracting and evaluating the users’ perception of POIs

3.3.1 Evaluation of the collective perception for POIs via POI signatures

The Collective Perception Identifier module (CPI) is responsible for the computation of the collective perception starting from the observations and the sentiment identified by the SDA module described in the previous section. Starting from the map structure containing POIs and observations referred to them, the CPI module computes the POI Signature for each POI mentioned in the analyzed comments. Indeed, by aggregating all the observations that refer to a same POI, it is possible to define the POI Signature, as described in Section 2.

In order to evaluate the collective perception related to a single POI, a Fuzzy Inference System (FIS) is used, whose rule base is reported in Table 1, so

Table 1: If-Then rules to achieve CP (in case of O=Low)

	P	O	N	CP
R1	H	L	L	VP
R2	H	L	M	P
R3	H	L	H	N
R4	M	L	L	P
R5	M	L	M	N
R6	M	L	H	B
R7	L	L	L	N
R8	L	L	M	B
R9	L	L	H	VB

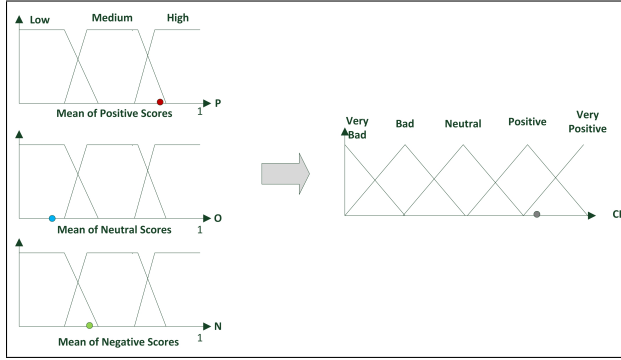


Figure 2: Fuzzy variables and membership of the FIS for computing the collective perception

that it is possible to obtain the value of the CP starting from the sentiment scores Sp_i , Sn_i and Su_i of the POI computed by the SDA module with the formulas in Eq. 7 - 9.

Such scores are intended as the inputs of the FIS and their membership functions are reported in Figure 2. The FIS allows us to obtain a value for the collective perception, obtained by defuzzifying the output of the FIS.

The CP value helps us with the estimation of the perception that a community has with respect to the analyzed POI. Depending on the type of the selected

comments as well as on the characteristics of the considered community, such CP can refer to different, specific assets of the city, allowing us to execute several kinds of what-if analysis as mentioned above. For instance, such collective perception may refer to wider concepts like the quality of the urban space as perceived by the community, or it can refer to more specific assets like the quality of the transportations or the perceived safety. In such cases, it is required that the comments processed by the SDA module are related only to that specific asset. By leveraging the taxonomy created by the KEB module, it is possible to select only those comments that are relevant for the kind what-if analysis chosen and the related city assets.

3.3.2 Area Signature

In most cases, the evaluation of the collective perception of a whole geographical area, instead of a single POI, is of relevance. The collective perception of a whole area is computed by aggregating the POI Signatures of all the POIs that are located in that area. To group the POI Signatures the approach proposed in [19], based on the Ordered Weighted Average (OWA) operator [20], is used:

$$AreaSignature(u_i, o_j) = OWA_{POI \in Area}^{LQ}(POISig_{POI}(u_i, o_j)) \quad (10)$$

Different linguistic quantifiers LQ lead to different characterizations of the selected area. A coarse-grained characterization will be obtained if the quantifier is $LQ = max$, while a narrower characterization will be obtained with the quantifier $LQ = min$ as only the most common pairs (u_i, o_j) will be included in the Area Signature. Roughly speaking, the Area Signature associates a bag of observations to the area so as to reflect the perceptions and the opinions of the community on its POIs. To transform this

bag of observations into the value of Collective Perception, the same approach used for a single POI is employed. The sentiment scores of all the observations included in the Area Signature are computed by the SDA module. The positive, negative and neutral scores of each observation are aggregated in order to obtain three values (P, O, N) which characterize the whole area. Next, the Fuzzy Inference System computes the value of the CP for the area.

Please note that there are two degrees of freedom to modify the characterization of the area:

- by changing the linguistic quantifier LQ in Eq. 10
- by changing the α -cut in Eq. 5. Indeed, by using a lower value of the α -cut, the POI Signature will contain less frequent pairs (u_i, o_j) as well, leading to a wider and less precise characterization of the area, but also revealing meaningful observations although not so popular. With greater values of α -cut a dual behavior is obviously obtained.

4 Experimentation of the system

After a preliminary evaluation carried out earlier in order to assess the feasibility of the proposed approach (as described in [9]), a more comprehensive experimentation has been performed upon the Metropolitan City of Naples, Italy.

4.1 Data

In this regard, an area of Naples, downtown, was selected, upon which decision makers may want to perform their research. The area, depicted in Figure 3, has an extension of 1.57 km², and includes eight tourist attractions: the Royal Palace of Naples, the San Carlo Theater, Piazza Plebiscito, Castel Nuovo,

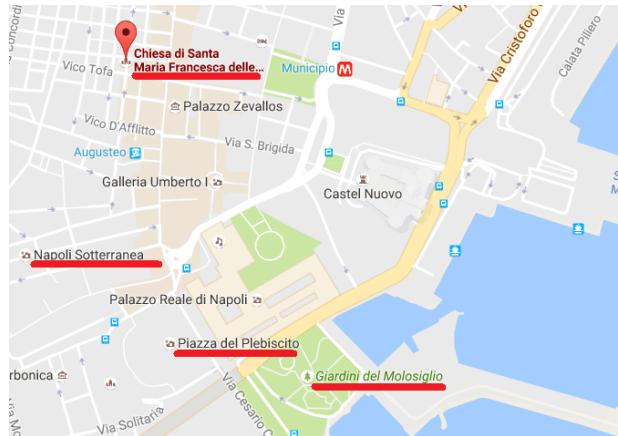


Figure 4: The POIs of the considered area chosen for the experimentation, namely Subterranean Naples (“Napoli Sotterranea”) and Piazza Plebiscito [strongly attractive POIs], and the Gardens of Molosiglio (“Giardini di Molosiglio”) and the Church of Santa Maria Francesca (“Chiesa di Santa Maria Francesca”)[weakly attractive POIs].

The Molosiglio Gardens, Galleria Umberto I, Subterranean Naples, the Basilica of San Francesco di Paola and the Church of Santa Maria Francesca.

From these eight POIs, a representative sample of 4 points was selected:

- 2 strongly attractive POIs (according to the evaluations the POIs received on Google Maps): Subterranean Naples and Piazza Plebiscito;
- 2 weakly attractive POIs: Gardens of Molosiglio and the Church of Santa Maria Francesca;

The position of the four POIs is depicted on the map featured in Figure 4.

The user community considered was “NapoliDaVivere”, an Instagram group with more than 22000 followers and 30 users (who are the administrators



Figure 3: Area of Naples, downtown, chosen for the experimentation, with an extension of 1.57 square meters.

of the group). 30 comments for each POI were analyzed in order to construct the POI signatures and then evaluate the collective perception of the selected area.

4.2 Method

4.2.1 POI Signature

Following the approach described in Section 3, the Knowledge Extractor and Builder (KEB) module generates a taxonomy of the most relevant concepts related to the selected POIs. The Sentiment Detector and Analyzer (SDA) module scans the users' comments for the selected POIs in order to identify the most relevant observations (adjectives and expressions) which refer to important concepts related to the POI. In this process, the SDA leverages the taxonomy of concepts created by the KEB for under-

standing whether an adjective or an expression actually refers to the POI. The identification of all the observations for the POIs allows the system to compute the POI Signatures. Figures 5 thru 8 show the pairs of users and observations (u_i, o_j) for each POI. Users are listed along the rows of the tables, while the observations along their columns. A "1" in the cell (u_i, o_j) indicates that the user u_i expressed the observation o_j in his/her comment regarding that POI. In the last column, the values of the *UserActivity* (Eq. 2) are reported, indicating the degree of activity of the user (*i.e.* how many observations he/she used) with respect to all the other users. In the last row, the values of *ObsPop* (Eq. 4) are listed, representing the popularity of the observation among the users.

The value of the POI Signature is evaluated by using Eq. 5 for the four POIs. In this experimentation scenario, two different values of α -cut are used: α -

Users	Observations	wonderful	interesting	historical	guide	peaceful	centre	uncomfortable	informative	suggestive	worthful	ancient	charming	ObsUser	UserActivity
U1		1							1					2	0,500
U2			1								1	1		3	0,600
U3							1							2	0,400
U4								1						1	0,200
U5		1				1		1	1					4	0,800
U6									1					1	0,200
U7						1				1				3	0,600
U8		1					1							2	0,400
U9		1			1			1	1				1	5	1,000
U10		1									1			2	0,400
U11				1										1	0,200
U12		1						1						2	0,400
U13						1						1		2	0,400
U14							1				1			2	0,400
U15		1						1		1	1			4	0,800
U16			1											1	0,200
U17								1	1					2	0,400
U18					1						1			2	0,400
U19						1								1	0,200
U20				1				1			1			3	0,600
U21		1							1				1	3	0,600
U22					1				1					2	0,400
U23		1									1	1		3	0,600
U24								1					1	2	0,400
U25				1					1					2	0,400
U26						1					1		1	3	0,600
U27		1						1						2	0,400
U28									1	1				2	0,400
U29			1							1				2	0,400
U30		1											1	3	0,600
ObsFreq		11	3	3	3	5	3	9	10	4	10	3	5		
ObsPop		1,000	0,273	0,273	0,273	0,455	0,273	0,818	0,909	0,364	0,909	0,273	0,455		

Figure 5: POI Signature for POI1 (Subterranean Naples)

Users	Observations	amazing	criminal	sacrum	chair	anxious	motorbike	beautiful	noise	experience	dangerous	painting	emotional	ObsUser	UserActivity
U1		1	1											2	0,333
U2			1	1						1			1	4	0,667
U3		1			1								1	3	0,500
U4		1					1		1					3	0,500
U5		1	1					1			1		1	5	0,833
U6								1				1		2	0,333
U7			1					1			1		1	4	0,667
U8					1			1					1	3	0,500
U9				1										1	0,167
U10			1						1	1			1	4	0,667
U11			1				1		1		1			4	0,667
U12				1		1					1			3	0,500
U13			1	1				1					1	4	0,667
U14				1					1					2	0,333
U15													1	1	0,167
U16			1					1			1			3	0,500
U17				1				1						2	0,333
U18					1	1			1				1	4	0,667
U19										1	1	1		3	0,500
U20			1			1								2	0,333
U21			1	1			1		1					4	0,667
U22					1								1	2	0,333
U23			1				1		1					3	0,500
U24			1					1		1			1	4	0,667
U25				1	1									2	0,333
U26			1			1					1			3	0,500
U27							1		1					2	0,333
U28								1	1					2	0,333
U29			1	1						1			1	4	0,667
U30		1										1		2	0,333
ObsFreq		5	14	9	4	4	6	9	9	4	8	3	12		
ObsPop		0,357	1,000	0,643	0,286	0,286	0,429	0,643	0,643	0,286	0,571	0,214	0,857		

Figure 6: Signature for POI2 (Church of Santa Maria Francesca)

Users	Observation													ObsUser	UserActivity
		wide	wonderful	monument	safe	culture	dirty	lights	beautiful	vandals	architecture	charming	elegant		
U1		1	1			1				1				4	1,000
U2		1										1		2	0,500
U3						1				1				2	0,500
U4			1				1							2	0,500
U5		1						1				1		3	0,750
U6			1						1					2	0,500
U7			1								1			2	0,500
U8									1	1				2	0,500
U9		1		1								1		3	0,750
U10								1						1	0,250
U11		1								1				2	0,500
U12		1		1						1				3	0,750
U13							1					1		2	0,500
U14			1		1									2	0,500
U15				1						1				2	0,500
U16		1							1	1				3	0,750
U17			1						1					2	0,500
U18		1		1		1								3	0,750
U19		1												1	0,250
U20							1			1				2	0,500
U21			1						1					2	0,500
U22			1					1		1				3	0,750
U23			1				1			1				3	0,750
U24					1						1			2	0,500
U25		1	1											2	0,500
U26							1			1	1			3	0,750
U27			1											1	0,250
U28								1		1				2	0,500
U29									1	1				2	0,500
U30		1			1		1			1				4	1,000
ObsFreq		11	11	4	3	3	5	3	4	5	14	3	3		
ObsPop		0,786	0,786	0,286	0,214	0,214	0,357	0,214	0,286	0,357	1,000	0,214	0,214		

Figure 7: Signature for POI3 (Piazza Plebiscito)

Users	Observation													ObsUser	UserActivity
		garden	wonderful	dirty	abusive	peaceful	vagabonds	monument	walking	suggestive	worthful	fountain	charming		
U1						1				1				2	0,400
U2					1							1		2	0,400
U3		1	1				1				1			4	0,800
U4				1					1					2	0,400
U5			1								1	1		3	0,600
U6							1	1						2	0,400
U7				1					1		1			3	0,600
U8		1				1				1				3	0,600
U9					1			1						2	0,400
U10			1					1				1		3	0,600
U11					1			1						2	0,400
U12			1		1					1	1	1		5	1,000
U13				1			1							2	0,400
U14			1							1				2	0,400
U15								1	1					2	0,400
U16		1								1	1			3	0,600
U17			1			1	1			1		1		4	0,800
U18						1	1					1	1	4	0,800
U19					1									1	0,200
U20			1								1			2	0,400
U21			1								1			2	0,400
U22					1		1							2	0,400
U23									1		1			2	0,400
U24			1				1			1				3	0,600
U25				1							1			2	0,400
U26					1				1					2	0,400
U27			1								1			2	0,400
U28			1	1				1						3	0,600
U29										1	1			2	0,400
U30			1				1					1		3	0,600
ObsFreq		3	12	5	7	3	9	6	3	6	6	12	4		
ObsPop		0,250	1,000	0,417	0,583	0,250	0,750	0,500	0,250	0,500	0,500	1,000	0,333		

Figure 8: Signature for POI4 (Gardens of Molisiglio)

$cut > 0.5$ to select the most common pairs of users and observations, and $\alpha-cut > 0.3$ to select a larger number of pairs, even if they are less frequent. In this way, it is possible to show the effect of the $\alpha-cut$ on the collective perception value.

The SDA module computes the sentiment scores for each observation. From these, the observations that are too objective (*i.e.* with a value of objectiveness $O > 0.5$) are removed, since they do not contribute to the value of collective perception.

With an $\alpha-cut > 0.5$, the following bags of observation for each POI are obtained:

- POI1 Subterranean Naples: {wonderful, uncomfortable, worthwhile}
- POI2 Church of Santa Maria Francesca: {criminal, beautiful, noise, dangerous, emotional}
- POI3 Piazza Plebiscito: {wonderful}
- POI4 Gardens of Molosiglio: {wonderful, abusive, vagabonds}

By using $\alpha-cut > 0.3$, other observations are added to the four bags:

- POI1 Subterranean Naples: {wonderful, uncomfortable, worthwhile, suggestive, charming}
- POI2 Church of Santa Maria Francesca: {criminal, beautiful, noise, dangerous, emotional, amazing}
- POI3 Piazza Plebiscito: {wonderful, dirty}
- POI4 Gardens of Molosiglio: {wonderful, abusive, vagabonds, dirty, suggestive, worthwhile, charming}

The Collective Perception Identifier module (CPI) computes the arithmetic mean of the sentiment

Table 2: Collective perception of the four POIs with $\alpha-cut > 0.5$ and $\alpha-cut > 0.3$

POI	CP	CP
	$\alpha-cut > 0.5$	$\alpha-cut > 0.3$
Subterranean Naples	0.69	0.72
Church	0.36	0.47
Piazza Plebiscito	0.83	0.63
Gardens Molosiglio	0.33	0.5

scores of all the observations for a same POI, with the two different values of $\alpha-cut$. Such average values of the sentiment scores become the input of the Fuzzy Inference System for computing the collective perception. The values of the collective perception of the four POIs are reported in table 2.

By analyzing the obtained results, it becomes clear that the most attractive POIs (Subterranean Naples and Piazza Plebiscito) are positively perceived by the community, confirming that they are two of the most important touristic attractions in Naples. On the other hand, the perception on the Church is really poor, and this is mainly due to the presence of observations like “criminal” and “dangerous”. This is because the Church lies in a neighborhood of the city that is considered dangerous due to frequent robberies. With respect to Piazza Plebiscito (POI3), it is possible to notice a decrease in the CP when considering $\alpha-cut > 0.3$. In this case, in fact, observations like “dirty” contribute to a bad perception of the area and refer to a situation of dirt and disorder that some users of the community begin to perceive in a negative way. This shows that it is useful to evaluate the collective perception at different values of $\alpha-cut$ in order to identify hidden issues in the analyzed area.

4.2.2 Area Signature and Fuzzy Cognitive Map

The Area Signature is useful for evaluating the collective perception (CP) of the whole area and it can

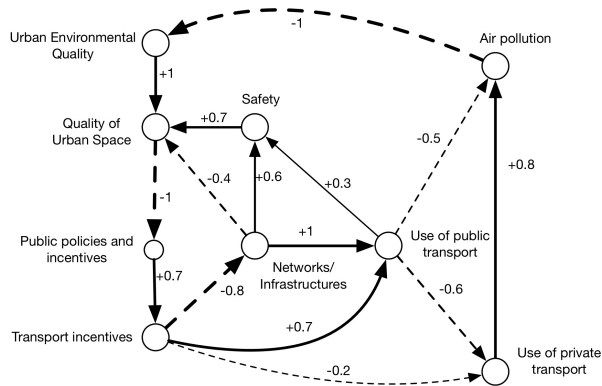


Figure 9: Fuzzy Cognitive Map for the evaluation scenario (from [9])

be used in what-if analysis scenarios based on Fuzzy Cognitive Maps. The Area Signature for the selected area of Naples downtown is evaluated by using the Eq. 10. An arithmetic mean is used as the average function. The collective perception of the whole area with $\alpha\text{-cut} > 0.5$ is $CP = 0.64$, whereas with $\alpha\text{-cut} > 0.3$ is $CP = 0.5$.

Figure 9 shows the FCM used for the what-if analysis. This map is built by integrating the maps proposed in [12] and [21]. It contains the causal relationships between some important concepts and assets related to the management of a city, like environment, transportation, social assets and government. The CP of the area is used as the activation level of the “Quality of Urban Space” concept in this map. This is motivated by the fact that the Quality of Urban Space is a faceted concept which includes aspects such as quality of buildings, cultural and tourist attractions, safety, transportations and so on. For the objective of this evaluation scenario, only one indicator for the Quality of Urban Space is considered, *i.e.* the quality of tourist attractions, specifically of the four considered POIs. It is reasonable to state that very positive perception of urban areas match with higher levels of Quality of Urban Space.

In order to evaluate the impact of the collective perception on the concepts of the FCM, a baseline scenario with which to compare the obtained results is needed. Therefore, the equilibrium states of the map is chosen as a baseline scenario. The activation levels of the concepts of the map at the equilibrium are reported in Figure 10 and 11 in the Scenario_1 column.

Considering the value of $CP = 0.64$, and using it as Quality of Urban Space, the activation levels of the map change as reported in column Scenario_2 in Figure 10. It can be noticed that, in order to face a decrease in the value of the Quality of Urban Space (from 0.657 given by the equilibrium state to 0.64 given by the collective perception), an increase in the public policies of 0.85% is required. Thus, the FCM can help the decision makers understand which asset, among the main assets of the city, they need to intervene on for responding to a slight decrease of the collective perception regarding the quality of the urban space.

When using $\alpha\text{-cut} > 0.3$, the value of $CP = 0.5$ is used as the activation level of the FCM. Figure 11 reports the activation levels in this case with respect to the equilibrium state of the map. Here, the significant decrease of the Quality of Urban Space requires a greater improvement of the public policies (about 10.58%). Besides, also the transportation might need some interventions (1.11%).

5 Discussion

5.1 Theoretical Implications

This work proposes the adoption of a new decisional paradigm which is complementary to the traditional decisional approaches used in urban planning and government. Indeed, even if the traditional approaches use the opinions of experts, they rarely consider the perception that the citizens have of an

	Concept	Scenario_1	Scenario_2	Difference	Percent_change
1	Urban Environmental Quality	0.38021	0.38021	0.00000	0.00082
2	Air Pollution	0.48866	0.48864	-0.00001	-0.00270
3	Quality of Urban Space	0.65701	0.64400	-0.01301	-1.97949
4	Safety	0.60820	0.60819	-0.00001	-0.00161
5	Use of public Transport	0.68602	0.68607	0.00006	0.00809
6	Use of private Transport	0.37203	0.37200	-0.00003	-0.00844
7	Public Policies	0.34141	0.34434	0.00293	0.85828
8	Network and Infrastructure	0.38994	0.38984	-0.00010	-0.02467
9	Transport Incentives	0.55946	0.55997	0.00051	0.09035

Figure 10: Comparison between the two scenarios (equilibrium and $CP = 0.64$ with $\alpha\text{-cut} > 0.5$)

	Concept	Scenario_1	Scenario_2	Difference	Percent_change
1	Urban Environmental Quality	0.38021	0.38025	0.00004	0.01008
2	Air Pollution	0.48866	0.48849	-0.00016	-0.03327
3	Quality of Urban Space	0.65701	0.50000	-0.15701	-23.89712
4	Safety	0.60820	0.60808	-0.00012	-0.01980
5	Use of public Transport	0.68602	0.68670	0.00068	0.09956
6	Use of private Transport	0.37203	0.37165	-0.00039	-0.10387
7	Public Policies	0.34141	0.37754	0.03613	10.58193
8	Network and Infrastructure	0.38994	0.38875	-0.00118	-0.30356
9	Transport Incentives	0.55946	0.56569	0.00622	1.11237

Figure 11: Comparison between the two scenarios (equilibrium and $CP = 0.50$ with $\alpha\text{-cut} > 0.3$)

entire area, and even more rarely the impact that it may have on the different assets of the city, like public policies, safety and transportation. The ultimate purpose of the proposed environment is to increase the awareness of decision makers related to the perception that a community has about a specific area of interest in the city. Such an awareness becomes systematic in the decisional processes for urban planning and development. Moreover, this environment allows to consider the opinions and the points of view of citizens and communities in a way that is less invasive as possible, since it considers observations and comments that the users have published for a completely different purpose. Consequently, it reduces the difficulty in the realization of decision support systems that need to involve a community.

5.2 Managerial Implications

The proposed environment supports different stakeholders in taking decisions with systems that are easily to configure and use. Its approach allows users to set different parameters to fine-tune the process, thus providing a high level of flexibility in the decision analysis. In fact, it is possible to analyze the same data at different levels of granularity, by considering less or more numbers of POIs and observations. Specifically, the *alpha-cut* value can be considered as a tool for the decision maker to understand what happens to the collective perception (or to the other assets of the city) when considering or eliminating some observations. This enables decision makers to easily uncover hidden observations and critical aspects that may help improve their awareness with re-

spect to specific issues of the city, whereas with other approaches based on sentiment analysis techniques a similar result is not so easy to obtain.

Another degree of freedom is given by the linguistic quantifier used in the evaluation of the Area Signature. The decision maker may choose to weigh in different ways the positive or the negative observations, or to give a different importance to the most recent observations with respect to older observations.

6 Conclusion

In this work, an environment for helping decision makers within the context of a Smart City has been introduced. This environment implemented a methodology meant to acquire and exploit the collective perception of Points of Interest in a Smart City, based on an innovative combination of techniques, including semantic knowledge discovery, sentiment analysis, Fuzzy Cognitive Maps and POI signatures, which is unprecedented in this domain. As such, this methodology exploits the concepts of shared perception and awareness, enabling decision makers to take into account the different, often heterogeneous, perspectives of a wide range of stakeholders (citizens included) with respect to relevant city events and phenomena and to the city as a whole. A significant application of this process is testified by the reported experimentation of the proposed environment, carried out against a massive user community discussing the large metropolitan city of Naples, Italy. Results yielded in similar contexts may prove significantly useful for providing decision makers with a clearer awareness of the positive and/or critical aspects of urban areas, thus helping them find or refine the measures to be taken for better managing the city itself and boost its development.

Further refinements of the methodology implemented and of the environment itself include the ex-

tension of the POI knowledge base according to the availability of public touristic repositories of information related to city attractions and points of interest. Also, additional experimentations against other urban contexts and cities, as well as with respect to different concepts aside from the quality of urban space, are expected to be carried out in order to further extend the conclusions of this work and strengthen the effectiveness of the proposed environment.

References

- [1] D. Arnott, "Decision support systems evolution: Framework, case study and research agenda," *European Journal of Information Systems*, vol. 13, no. 4, pp. 247–259, 2004.
- [2] D. Arnott and G. Pervan, "A critical analysis of decision support systems research," *Journal of Information Technology*, vol. 20, no. 2, pp. 67–87, 2005, cited By 210.
- [3] G. Bafoutsou and G. Mentzas, "Review and functional classification of collaborative systems," *International Journal of Information Management*, vol. 22, no. 4, pp. 281–305, 2002, cited By 126.
- [4] J. Fjermestad and S. R. Hiltz, "Assessment of group support systems research: Results," 1999, p. 14, cited By 0.
- [5] F. Antunes and J. Costa, "Integrating decision support and social networks," *Advances in Human-Computer Interaction*, vol. 2012, no. 574276, 2012.
- [6] "Smarter Cities - New cognitive approaches to long-standing challenges," <http://www.ibm.com/smarterplanet/us/en/>

- smarter_cities/overview/, accessed: 2017-02-10.
- [7] D. Toti and M. Rinelli, "On the road to speed-reading and fast learning with CONCEPTUM," in *Proceedings - 2016 International Conference on Intelligent Networking and Collaborative Systems, IEEE INCoS 2016*, 2016, pp. 1–6.
- [8] S. Baccianella, A. Esuli, and F. Sebastiani, "SentiWordNet 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining," in *Proceedings of the Seventh Conference on International Language Resources and Evaluation (LREC'10)*. Valletta, Malta: European Language Resources Association (ELRA), May 2010.
- [9] G. D'Aniello, A. Gaeta, M. Gaeta, V. Loia, and M. Reformat, "Collective awareness in Smart City with Fuzzy Cognitive Maps and Fuzzy sets," in *2016 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, 2016.
- [10] B. Kosko, "Fuzzy cognitive maps," *International journal of man-machine studies*, vol. 24, no. 1, pp. 65–75, 1986.
- [11] G. D'Aniello, V. Loia, and F. Orciuoli, "A multi-agent fuzzy consensus model in a situation awareness framework," *Applied Soft Computing*, vol. 30, pp. 430 – 440, 2015.
- [12] M. Olazabal and U. Pascual, "Use of fuzzy cognitive maps to study urban resilience and transformation," *Environmental Innovation and Societal Transitions*, 2015.
- [13] U. Özesmi and S. L. Özesmi, "Ecological models based on peoples knowledge: a multi-step fuzzy cognitive mapping approach," *Ecological Modelling*, vol. 176, no. 1, pp. 43–64, 2004.
- [14] F. Habib and A. Shokoohi, "Classification and resolving urban problems by means of fuzzy approach," *World Academy of Science, Engineering and Technology*, vol. 36, pp. 894–901, 2009.
- [15] D. Milne and I. H. Witten, "An open-source toolkit for mining wikipedia," *Artif. Intell.*, vol. 194, pp. 222–239, Jan. 2013. [Online]. Available: <http://dx.doi.org/10.1016/j.artint.2012.06.007>
- [16] N. Capuano, C. De Maio, S. Salerno, and D. Toti, "A methodology based on common-sense knowledge and ontologies for the automatic classification of legal cases," in *ACM International Conference Proceeding Series*, 2014.
- [17] N. Capuano, A. Longhi, S. Salerno, and D. Toti, "Ontology-driven generation of training paths in the legal domain," *International Journal of Emerging Technologies in Learning*, vol. 10, no. 7, pp. 14–22, 2015.
- [18] V. Basile and M. Nissim, "Sentiment analysis on Italian tweets," in *Proceedings of the 4th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*. Atlanta, Georgia: Association for Computational Linguistics, June 2013, pp. 100–107. [Online]. Available: <http://www.aclweb.org/anthology/W13-1614>
- [19] R. R. Yager and M. Z. Reformat, "Looking for like-minded individuals in social networks using tagging and e fuzzy sets," *Fuzzy Systems, IEEE Transactions on*, vol. 21, no. 4, pp. 672–687, 2013.
- [20] R. R. Yager, "On ordered weighted averaging aggregation operators in multicriteria decision-

making,” *Systems, Man and Cybernetics, IEEE Transactions on*, vol. 18, no. 1, pp. 183–190, 1988.

- [21] T.-A. Shiau and J.-S. Liu, “Developing an indicator system for local governments to evaluate transport sustainability strategies,” *Ecological Indicators*, vol. 34, pp. 361 – 371, 2013.