A Review of Opinion Mining Methods for Analyzing Citizens' Contributions in Public Policy Debate

Manolis Maragoudakis[1], Euripidis Loukis[1], Yannis Charalabidis[1]

[1]University oftheAegean, Information and Communication Systems Engineering
Department, Gorgyras and Palama Str.
83200 Karlovassi, Samos, Greece
{mmarag, eloukis, yannisx}@aegean.gr

Abstract. Electronic Participation (eParticipation), both in its traditional form and in its emerging Web 2.0 based form, results in the production of large quantities of textual contributions of citizens concerning government policies and decisions under formation, which contain valuable relevant opinions and knowledge of the society, however are exploited to a limited only extent. It is of critical importance to analyze these contributions in order to extract the opinions and knowledge they contain in a cost-efficient way. This paper reviews a wide range of opinion mining methods, which have been developed for analyzing commercial product opinions and reviews posted on the Web, as to the capabilities they can offer for meeting the above challenges. The review has revealed the great potential of these methods for the analysis of textual citizens' contributions in public policy debates, both for assessing contributors' general attitudes-sentiments (positive, negative or neutral) towards the policy/decision under discussion, and also for extracting the main issues they raise (e.g. negative and positive aspects and effects, implementation barriers, improvement suggestions) and the corresponding attitudes-sentiments. Based on the conclusions of this review a basic framework for the use of opinion mining methods in eParticipation has been formulated.

Keywords: Electronic Participation (eParticipation), Electronic Consultation (eConsultation), Web 2.0, Opinion Mining, Sentiment Analysis.

1 Introduction

Electronic Participation (eParticipation) results in the production of large quantities of textual contributions of citizens concerning government policies and decisions under formation. In the 'traditional' Web 1.0 based eParticipation (conducted in official government-initiated eParticipation spaces) are offered to the citizens tools for entering not only simple contributions of yes/no or rating type (e.g. various on-line voting and survey tools), but also more sophisticated contributions of textual type as well (e.g. various on-line discussion forum tools) [1-4]. Similarly, in the emerging Web 2.0 based forms of eParticipation (exploiting highly popular social media for publishing policy messages and collecting citizens' feedback on them) is offered to

the citizens functionality for providing feedback not only of simple 'like/dislike' or rating type, but also more of sophisticated comment type as well (e.g. comments on YouTube videos, or postings in blogs) [5-8]. So hundreds or even thousands of citizens' textual contributions are generated from these eParticipation channels, in numbers much higher than in the off-line public policy debates; these contain valuable opinions on the government policy/decision under discussion, and knowledge on the societal needs or problems this policy/decision attempts to address and the proposed government actions/interventions for this purpose. While the above simple forms of citizens' contribution can be easily analyzed using statistical methods, so that sound conclusions can be drawn from them, this does not hold for the numerous textual contributions collected. The big effort – and therefore the long time and high cost - required for reading hundreds or thousands of citizens' textual contributions on a public policy under formation (in traditional eParticipation spaces, Web 2.0 social media and also in offline public policy debates), summarizing them and extracting the general attitudes of the contributors and main points and issues they raise (e.g. positive and negative aspects and effects of the policies and decisions under discussion, implementation barriers, improvement suggestions) usually leads to a limited exploitation of them. This results in losses of valuable citizens' knowledge and opinions, which would be quite useful to the competent for public organizations for making better and more socially acceptable policies and decisions. Also, this does not allow feedback to be provided to the citizens as to how and to what extent their contributions have been taken into account, resulting finally in 'e-consultation fatigue' and disappointment [3]; providing such a feedback would greatly promote government transparency, accountability and openness.

For these reasons ten years ago in the first report of OECD discussing the potential of eParticipation [1], and also in the subsequent reports on this topic [2-3], is recognized as one of the most important challenges for the practical application of eParticipation ideas the analysis of the vast amounts of unstructured information that citizens' contributions (e.g. in 'threads' of e-conversations) contain using appropriate technologies. These technologies should aim to 'support the summarization and content analysis of contributions', 'help highlight areas of agreement and disagreement' and 'identify the participants main concerns, their level of support for any draft proposals, or their suggestions for action they think necessary to address problems raised'. At the same time it is argued that the lack of such analysis of citizens' textual contributions results in a lack of feedback to the citizens on 'how the results of the engagement have influenced the decision-making process and changed policy outcome' and has negative impact on citizens' trust in government and eParticipation. Subsequently Rose & Sanford [9] from a comprehensive literature review in the area of eParticipation conclude that one of the main research challenges in this area is to use appropriate tools for the analysis of citizens' input to policy-making. Similarly, Macintosh, Coleman & Schneeberger [10] in the introductory paper of the *ePart 2009* Conference Proceedings discussing the research gaps in the eParticipation area state that there is a lack of effective and efficient technologies for the analysis of unstructured eParticipation data and note that this poses a significant challenge to on-going research.

This paper contributes to addressing this critical for the large scale application and institutionalization of eParticipation research and practice gap, by critically reviewing

a wide range of methods and algorithms developed in the area of opinion mining for the analysis of product opinions and reviews posted on the Web. For this purpose we made a systematic search of relevant journal and conference papers using as keywords 'opinion mining' and 'sentiment analysis'. This literature was reviewed focusing on the capabilities that opinion methods provide for meeting the above challenges in the area of eParticipation in a cost-efficient manner (so that we can use them to the largest possible extent), and ii) the main principles of these methods (so that we can select the most suitable ones and use them effectively and efficiently). Based on the results of this review a basic framework has been formulated for the use of opinion mining methods for analyzing citizens' textual contributions in eParticipation. The content of this paper can be very useful for the numerous researchers and practitioners interested in the analysis and exploitation of citizens' textual contributions in various forms of eParticipation.

It should be mentioned that the research presented in this paper has been conducted as part of the project PADGETS ('Policy Gadgets Mashing Underlying Group Knowledge in Web 2.0 Media – www.padgets.eu)[8], which is supported by the 'ICT for Governance and Policy Modelling' research initiative of the European Commission. The research objective of this project is to develop a methodology and a technological platform for the systematic and centrally managed exploitation of the emerging Web 2.0 social media by government organizations in the processes of policy and decision making. This platform will enable publishing content and deploying micro-applications (termed 'Policy Gadgets' - PADGETS) simultaneously to many different Web 2.0 social media, and also retrieving the corresponding users' interactions (e.g. views, likes/dislikes, ratings, comments, etc.) and processing those using advanced methods. Taking into account the high popularity of the targeted Web 2.0 social media it is expected that a large quantity of valuable textual contributions (e.g. blog postings, comments) will be collected from them, so it is of critical importance to make the best possible analysis and exploitation of them in a cost-efficient manner.

The paper consists of five sections. In the following section 2 the objectives and basic concepts of opinion mining are presented. In section 3 is reviewed the main research stream of this area dealing with sentiment analysis at the document and sentence level. Then in section 4 is reviewed the second most important research stream of this area dealing with feature-based sentiment analysis. Section 5 discusses issues related to our initial framework for dealing with sentiment analysis under project PADGETS, which incorporates linguistic peculiarities such as extensive use of idioms and ill-formed sentences (i.e. short sentences which do not necessarily follow syntactic rules). The final section 6 includes the conclusions and also a framework for the use of opinion mining methods for analyzing citizens' textual contributions in eParticipation.

2 Opinion Mining Objectives and Concepts

The Web has dramatically altered the way people express their opinions, offering them the capability to post comments and reviews on commercial products and

express their views on a plethora of issues in forums, discussion groups, chat rooms, social networking groups and blogs. This user-generated content has been recognized as a valuable source of commercial and political information. However, the large amount of this information and its natural language form make it difficult to extract the useful elements, such as the general feeling/sentiment (e.g. positive, negative, or neutral) on the particular topic (e.g. a product/service or a new policy proposal) and the specific issues raised about it by the users/visitors of these websites. For these reasons methods started being developed for supporting the above tasks, and this lead to the development of the sentiment analysis or opinion mining research domain, which according to [11] aims to develop methods for the computational processing of opinions, sentiments and emotions found, expressed and implied in text. Its initial motivation has been to enable firms to analyze online reviews and comments entered by users of their products in various review sites, blogs, forums, etc., in order to draw general conclusions as to whether users liked the product or not (sentiment analysis), and also more specific conclusions concerning features (characteristics) of the product that have been commented positively or negatively (features extraction and analysis).In general, opinions can be expressed on anything, e.g., an item, a product, a service, an individual, an organization, an event, or a topic, so we use the collective term 'object' to denote the target entity that has been commented on; however, comments can be expressed not only on objects, but also on particular 'features' (characteristics) of them [12], e.g. on the "*battery life*" feature of a mobile phone.

This research domain can be subdivided into two main streams. The first research stream deals with the classification of an opinionated text (i.e. a document consisting of several statements, such as a forum or blog post) as expressing a positive, negative or neutral opinion. More recent research in this research stream focuses on the sentence level, dealing with the classification of a sentence as objective (fact) or subjective (opinion), and then on the further classification of the subjective sentences as expressing a positive, negative or neutral opinion [13]. The second research stream of this area deals with the extraction of the most commented features of the commented object, and for each of them, the classification of relevant opinions as positive, negative or neutral is performed. Throughout the following sections 3 and 4 these two research streams are reviewed.

3 Document and Sentence Level Sentiment Analysis

Sentiment analysis is perhaps the most widely studied topic in this domain [14,15,18, 19,20,21,22,23,24]. Sentiment classification is similar in some aspects to, but also different in some other aspects from, standard topic-based text classification, which classifies documents into labeled topic classes, e.g., politics, sciences, sports, etc. In topic-based classification, topic related words are important, while in sentiment-based classification it is opinion words denoting a positive or negative attitude (e.g. great, excellent, amazing, horrible, bad, worst, etc.) that are really important. Initially, sentiment-based classification was focused on the document level.

3.1 Document-Level Sentiment Analysis

Existing supervised learning methods have been readily applied to document-level sentiment classification. Pang et al. [25] followed this approach to classify movie reviews into two classes, positive and negative. It was shown that using unigrams (a bag of individual words) as features in classification performed well with either naïve Bayesian or SVM classifiers. Subsequent papers used a large variety of features and techniques in learning. As most machine learning applications, the main task of sentiment classification is to find a suitable set of classification features. Some of the most widely used features for this purpose are (for a more comprehensive survey see [11]):

- Terms' frequencies and positions: these features are individual words or word n-grams and their frequency counts, while in some cases word positions may also be considered.
- Part-of-speech (POS) tags: It has been observed by a plethora of early research studies that adjectives (e.g. "interesting", "bad", etc.) were important indicators of subjectivities and opinions, so they have been treated as special features.
- Opinion words and phrases: By definition, "opinion words" are words commonly used to express positive or negative sentiments, e.g. "beautiful", "wonderful", "good", and "amazing" denote positive opinion words, whereas "bad", "poor", and "terrible" indicate negative opinion words. Although many opinion words are adjectives (as previously shown), adverbs, nouns (e.g. "rubbish", "junk") as well as verbs (e.g. "hate" and "like") can indicate sentiment as well. Apart from individual words, there are also opinion phrases and idioms, such as in the English language "something cost an arm and a leg" that indicates sentiment.
- Syntactic dependency: Word dependencies generated from syntactic parsing or dependency trees are also used as features by several researchers.
- Negation: Clearly negation words are important because their appearances often change the opinion orientation, e.g. the sentence "I don't like this camera" is negative.

An interesting conclusion in this research stream is that the "domain adaptation"is very often of critical importance. In particular, it has been found that sentiment classification is highly sensitive to the domain from which the training data (e.g. positive and negative opinion words) have been extracted. A classifier trained using opinionated texts from one domain often performs poorly when it is applied or tested on opinionated texts from another domain. The reason is that words and even language concepts that are used in different domains for expressing opinions can be substantially different. It is possible that the same word in one domain may mean positive, but in another domain may mean negative. For example, as remarked in [26], the word 'unpredictable' may have a negative orientation in a car review (e.g., "unpredictable brakes"), but it could have a positive orientation in a movie review (e.g., "unpredictable scenario"). Thus, domain expertise is required. Ontologies are very beneficial in this respect since they encode human domain knowledge in a machine readable format. Nevertheless, since ontology construction is a painstaking process, most articles use labeled opinions originated from one domain and apply general opinion words to unlabeled opinions of another domain [18, 27, 28].

Also, some research has been conducted on the use of unsupervised learning techniques for document-level sentiment analysis. In these techniques, weighted terms

and phrases are the main indicators for sentiment classification. The use of unsupervised algorithms, such as Mutual Information, which utilize a neighborhood of k surrounding terms around a standard syntactic element (e.g. adjuncts or adverbs) that are likely to express opinions, can give good results [26].

3.2   Sentence-Level Sentiment Analysis

In the more recent research in this area the level of granularity increases to the level of sentence [45,64,65], focusing on sentence-level subjectivity and sentiment classification. In particular, the main objective is: given a sentence, one has to perform the following two sub-tasks:
i) subjectivity classification: i.e. to determine whether it is a subjective sentence (opinion) or an objective one (fact),
ii) sentence-level sentiment classification: if the sentence is subjective determine whether it expresses a positive, negative or neutral opinion.

Traditional supervised learning methods are here applicable as well. For instance, one of the early works reported by Wiebe et al. [13] performed subjectivity classification using the naïve Bayesian classifier, while subsequent research has used many other learning algorithms [29,30,31,32].One of the main bottlenecks of using supervised learning for the above purposes is the manual effort involved in annotating the very large number of training examples (i.e. labeling the available examples as either positive or negative). To reduce the manual effort a 'bootstrapping' approach to label training data automatically is reported in [30,33], using a form of re-sampling that aims to estimate the variability of statistical properties of interest.

Yu and Hazivassiloglou [34] propose a technique that identifies subjective (opinion) sentences and also determines their opinion orientations. For the initial identification of subjective sentences it uses supervised learning examining three particular learning methods: sentence similarity, naïve Bayesian classification, and multiple naïve Bayesian classifiers. Then for sentiment classification of each identified subjective sentence it uses a similar method to the one used in [26], but with several words (instead of two used in [26]), and the score function was the well-known log-likelihood ratio. It should be noted that inn [26] the semantic orientation of a phrase is calculated as the Mutual Information between the given phrase and the word "excellent" minus the Mutual Information between the given phrase and the word "poor". The same issue is examined in [35] considering gradable adjectives. In [24] a semi-supervised learning method is applied, and in [36] the decision is made by simply summing up opinion words in a sentence. In [37,38,39] statistical prediction models are built to identify some specific types of opinions in reviews. A significant conclusion of this research is that sentence-level classification is not suitable for compound sentences. Wilson et al. [16] pointed out that a single sentence may contain not only multiple opinions, but also both subjective and factual clauses, making the problem even more difficult. In a more recent work [40] the problem of sentence-level sentiment analysis is studied using machine learning by considering contextual sentiment influencers such as negation (e.g. not and never) and contrary (e.g. nevertheless and however). A list of such influencers are provided in [41].

3.3 Polar Words

For the application of the above sentiment analysis methods it is important to define particular words that bear a positive or negative meaning, which are referred to as 'opinion words' or 'polar words' in the literature and form the so-called opinion lexicon. Examples of positive polar words in the English language include "beautiful", "good", "amazing", "astonishing", etc., while negative polar words include "poor", "bad", and "awful". Apart from individual words, there are also opinion phrases and idioms indicating positive or negative sentiments, e.g., "the service cost me an arm and a leg" (indicating a negative sentiment of having paid too much). For developing a polar word list three main approaches have been proposed:

- The 'manual' approach, which is very time-consuming [20, 42, 43, 44] and thus it is usually used in combination with automated approaches as a final check, since such automated methods usually make mistakes.

- The 'dictionary-based' approach. One of the simpler techniques in this approach is based on 'bootstrapping', using a small set of seed opinion words and an online dictionary such as WordNet [45]. The strategy is to first collect a small set of opinion words with known orientations (positive and negative) manually, and then to grow this set by searching within the WordNet for their synonyms and antonyms. The newly found words are added to the seed list, and this can lead to a next iteration, etc.; this iterative process stops when no more new opinion words are found. This approach is successfully used in [46,56]. After the iteration process is completed manual labeling can be carried out to correct errors. Researchers have also used additional information (e.g. glosses) in WordNet and additional techniques (e.g., machine learning) to generate better lists [46,47,48,49,50]. So far several opinion word lists have been generated [51,52]. The dictionary-based approach and the opinion words collected from it have a major shortcoming: it is unable to find opinion words with domain specific orientations. For example, for a speakerphone if it is 'quiet' this is usually negative; however, for a car if it is 'quiet' this is positive. The corpus-based approach described next can help coping with this problem.

- The 'Corpus-based' approach and sentiment consistency: The methods following this approach rely on syntactic or co-occurrence patterns and also a seed list of opinion words in order to find other opinion words in a large corpus (= set of documents). A representative work of this approach is the one of Hazivassiloglou and McKeown [53]. The technique starts with a list of seed opinion adjective words, and uses them and a set of linguistic constraints or conventions on connective words to identify additional adjective opinion words and their orientations. For instance, one of the constraints concerns the conjunction (AND), and says that conjoined adjectives usually have the same orientation, e.g. in the sentence "This vehicle is beautiful and spacious" if "beautiful" is known to be positive it can be inferred that "spacious" is also positive. This is so because people usually express the same opinion on both sides of a conjunction. Similar rules or constraints also hold for other connectives, such as OR, BUT, EITHER-OR, and NEITHER-NOR. We call this idea sentiment consistency. Also, clustering can be used to produce two sets of words: positive and negative ones. In [54] Kanayama and Nasukawa expanded this approach by introducing the idea of intra-sentential (within a sentence) and inter-sentential (between neighboring sentences) sentiment consistency (called coherency in [54]).

This approach has a major advantage in comparison with the dictionary-based approach: it can identify domain specific opinion words and their orientations if a corpus from only the specific domain is used.

4 Feature-based Sentiment Analysis

Classifying opinionated texts at the document or sentence level is useful, but does not provide a complete view on the commented object: a positive opinionated document does not necessarily mean that the author has positive opinions on all aspects or features of the object; likewise, a negative opinionated document does not mean that the author dislikes everything. In a typical opinionated passage, although the general sentiment on the object may be positive or negative, the author can express opinions on both positive and negative aspects of the object. Document-level and sentence-level classification does not provide such information, so in order to drill down such details we need to go to the object feature level. Therefore feature-based sentiment analysis includes the following two sub-tasks:
- identify object features that have been extensively commented on,
- determine whether the opinions on each of these features are positive, negative or neutral.

4.1 Feature Extraction

Its first task is to specify the topic or the feature of an object commented, which is essential in order to proceed to the second task of classifying its positive or negative meaning. The most widely known approaches that discuss unsupervised learning for identifying explicit features, such as nouns and noun phrases, involve the following two steps ([44],[48]):

I) Finding frequent nouns and noun phrases: Nouns and noun phrases are straightforwardly identified by using a POS tagger, which is a linguistic software tool for labeling each word with its part-of-speech. Their occurrence frequencies are counted, and only the frequent ones are kept (a frequency threshold can be decided). This approach is based on the assumption that when people comment on object features the vocabulary that they use usually converges, and most object features are nouns. Thus, those nouns that are frequently talked about are usually genuine and important features.

II) Finding infrequent features by making use of opinion words: Opinion words are usually adjectives and adverbs that express positive or negative opinions. The basic idea is that the same opinion word can be used to characterize positively or negatively different object features. Opinion words that characterize frequent features can also modify infrequent features, and thus can be used to extract infrequent features.

The precision of first step of the above algorithm was improved by Popescu and Etzioni in [55]. Their algorithm tries to remove those noun phrases that may not be object features. It evaluates each noun phrase by computing a pointwise mutual information (PMI) score between the phrase and 'meronymy' discriminators associated with the object class, e.g., a scanner class. In WordNet, Y is a meronymy

of X if Y is a part of X (e.g. wheel is a meronymy of car). The meronymy discriminators for the scanner class are, "of scanner", "scanner has", "scanner comes with", etc., which are used to find components or parts of scanners by searching on the Web. The algorithm also distinguishes components and parts from attributes and properties using WordNet's is-a hierarchy (which enumerates different kinds of properties) and morphological cues (e.g., "-iness", "-ity" suffixes).

Other related works on feature extraction mainly use the ideas of topic modeling and clustering to capture topics/features in reviews [56, 57, 58]. For example, in [59], Mei et al. proposed a probabilistic model called topic-sentiment mixture to capture the mixture of features and sentiments simultaneously. One topic model and two sentiment models were defined based on language models to capture the probabilistic distribution of words in different topics/features with their associated opinion orientations. Su et al. [58] also proposed a clustering based method with mutual reinforcement to identify implicit features.

After the extraction of object features two additional issues need to be resolved the synonyms issue. It is common that people use different words or phrases to describe the same feature. For example, photo and picture refer to the same feature in digital camera opinions and reviews. Identifying and grouping synonyms is essential for applications. Although WordNet as well as other thesauri and dictionaries help to some extent, they are far from sufficient due to the fact that many synonyms are domain dependent. For example, picture and movie are synonyms in movie reviews, but they are not synonyms in digital camera reviews as picture is more related to photo while movie refers to video. Carenini et al. [60] proposed a method based on several similarity metrics similar to those in information integration [61]. It requires the use of a taxonomy of features to be given for a particular domain. The algorithm merges each discovered feature to a feature node in the taxonomy. The similarity metrics are defined based on string similarity, synonyms and other distances measured using WordNet. Experiments on digital camera and DVD opinions show interesting outcomes.

4.2   Identification of Opinion Orientation

An important issue in feature-based sentiment analysis is how to identify the orientation of opinions expressed on an object feature in a sentence. Clearly, the sentence-level sentiment classification methods discussed previously are also applicable here, i.e. they can be applied to each sentence containing object features. However, there is an additional lexicon-based approach to this problem [51]. This lexicon-based approach basically uses opinion words and phrases in a sentence to determine the orientation of the opinion. Apart from the opinion lexicon, negations and but-clauses in a sentence are also taken into account. In particular, this approach includes the following four steps:

A) Identify opinion words and phrases: Given a sentence that contains an object feature initially are identified all opinion words and phrases. Each positive word is assigned the opinion score of +1, each negative word is assigned the opinion score of -1, and each context dependent word (i.e. word having meaning positive or negative depending on the context) is assigned the opinion score of 0. Suppose for example

that we are given the sentence "The picture quality of this camera is not great, but the battery life is long." After this step, the sentence is turned into "The picture quality of this camera is not great [+1], but the battery life is long [0]" because "great" is a positive opinion word and "long" is context dependent.

B) Handling negations: Negation words and phrases are used to revise the opinion scores obtained in the previous step based on some negation handling rules. After this step, the above sentence is turned into "The picture quality of this camera is not great [-1], but the battery life is long [0]" due to the negation word "not".

C) But-clauses: In English, the word "but" means contrary, so a sentence containing but is handled by applying the following rule: the opinion orientation before but and after but are opposite to each other. After this step, the above sentence is turned into "The picture quality of this camera is not great [-1], but the battery life is long [+1]" due to "but". Apart from but, phrases such as "with the exception of", "except that", and "except for" behave similarly to "but" and are handled in the same way.

D) Aggregating opinions: This final step applies an opinion aggregation function to the resulting opinion scores to determine the final orientation of the opinion on each object feature in the sentence.

4.3   Ontology-based Sentiment Analysis

One of the most recent developments in feature-based sentiment analysis is the use of domain-specific knowledge through ontologies, which constitute a well-known formalism for representing knowledge in a both human and machine comprehensible manner. Using ontologies for web classification and document extraction has been a successful technique that inspired researchers from the sentiment analysis domain to incorporate such knowledge in order to define a taxonomy (or hierarchy) of object features and then build sentiment analysis subsystems that apply rules denoted by the structures of such taxonomies. Most research and commercial systems that incorporate ontologies [62,63] consist mainly of two main modules: an ontology-based extraction module and a sentiment analysis module. The ontology-based extraction module creates a small hierarchical tree using a set of relevant texts from which terminology is extracted. In the analysis module this small tree is compared against the large core ontology tree for analyzing the sentiment labels of its nodes.

6   Sentiment Analysis in project PADGETS

From the aforementioned review on existing techniques on sentiment analysis in every granularity level, ranging from document-based to feature-based opinion mining, we could state that given a language with a plethora of tools and opinionated text in a well-structured format, the performance of sentiment classification could reach very high percentages in the range of 80%-90%. Furthermore, a significant portion of those techniques have been applied and evaluated in domains where users state their opinions in a formal language and using large textual inputs such as electronic commerce, media presentations (i.e. movies or music albums), etc. Nevertheless, as regards to PADGETS, the application languages would be Greek,

Slovenian and Italian, which lack of abundance in linguistic resources. Additionally, the domain of social media is not similar to the aforementioned cases, since users tend to utilize a smaller vocabulary with linguistic idioms or utilize the grammatical phenomenon of ellipsis in the texts. This phenomenon resulted in a novel approach that incorporates Social Network Analysis theory in order to deal with connecting opinions. To our knowledge, this is the first time such an approach is being exploited for sentiment classification. The idea is borrowed from the social network analysis (SNA) domain, which considers humans participating in social networking activities as *actors* and studies their *relationships* (see the PageRank algorithm of Google® for further details). SNA views the above process as a graph, where actors are *nodes* and *edges* state a kind of relations between two actors (e.g. friends in Facebook, Followers in Twitter, etc.). Opinions in social media may be suffering of short texts, encompassing linguistic idioms that are **more difficult** to be analyzed using traditional tools. However, if we consider opinions as actors and their inter-reference as links, we could form a common basis between opinion mining and SNA. The achievement of such a correspondence lies to the utilization of the **prestige** attribute, commonly found in SNA. Prestige is a refined measure of prominence of an actor. An actor is considered prestigious if he/she is the object of many references by other actors. Similarly, an opinion is significant (either positive or negative) if it is being addressed by other opinions. Furthermore, an opinion which is referenced by prestigious opinion is more significant than ones which are not referenced by opinions of great prestige. An analogue to real life is the case where a person who is recommended by a CEO of a company is more prestigious than a person who is recommended by the clerk of the company. By connecting opinions (where applicable) and grading their **rank prestige** accordingly, we could identify which sub-graph contains nodes of positive of negative sentiment and proceed with the classification of all other nodes. A reference can of course be positive (agreement) or negative (disagreement).

5   Conclusions

Both the traditional and the emerging eParticipation forms result in the production of large quantities of citizens' textual contributions concerning policies and decisions under formation, in numbers much higher than in the off-line public policy debates. It is of critical importance to use appropriate technologies for analyzing them in a cost-efficient manner, in order to extract the valuable opinions and knowledge they contain, and then integrate them in the policy/decision making processes and provide feedback to the citizens. In this paper we review methods developed in the area of opinion mining as to the capabilities they provide for meeting the above challenges. From this review it has been concluded that a useful body of knowledge has been developed in this area consisting of methods for addressing mainly the following three problems:
- classification of an opinionated text as expressing as a whole a positive, negative or neutral opinion (document-level sentiment analysis),
- classification of each sentence of such a text as objective (a fact) or subjective

(opinion), and then focus on the latter and classification of each of them as expressing a positive, negative or neutral opinion (sentence-level sentiment analysis),

- extraction of the particular features/subtopics commented by the authors of these texts, and for each of them identification of the orientation of the opinions expressed about it as positive, negative or neutral (feature-level sentiment analysis).

Based on the conclusions of this review a basic framework for the use of opinion mining methods in eParticipation can be formulated consisting of five stages:

I. Classify each particular posting on the policy/decision under discussion as positive, neutral or negative, using methods of document-level sentiment analysis (section 3.1), and then calculate relative frequencies of positive, neutral and negative postings.

II. For each posting identify its subjective sentences (expressing opinions) and classify each of them as positive, neutral or negative using methods of sentence-level sentiment analysis (section 3.2), and then calculate relative frequencies of positive, neutral and negative subjective sentences.

III. Compare and integrate findings of the above steps I and II, and also findings from the analysis of other types of citizens' non-textual feedback (e.g. numbers of users who voted in favor or against the policy/decision under discussion in an e-vote tool, or rated it positively or negatively in an e-survey tool, or even liked or disliked a relevant content we published in social media). This will allow conclusions to be drawn as to the citizens' general sentiments/feelings (positive or negative) on this government policy/decision.

IV. By further processing all postings on this policy/decision using feature extraction methods (section 4.1) identify the main issues raised and commented by citizens.

V. Finally for each issue classify each of the postings' sentences containing it as positive, neutral or negative using methods of sentence-level sentiment analysis - opinion orientation (section 4.2), and then calculate relative frequencies of positive, neutral and negative subjective sentences. This will allow the identification of the main issues raised by the citizens and the particular sentiments/feelings on them (e.g. positive and negative aspects and effects of the policies/decisions under discussion, implementation barriers, improvement suggestions, etc.).

It should be noted that for the practical application of the above opinion mining methods it is of critical importance to have sufficient language resources, such as lexicons, POS taggers, name entity recognizers, and tokenizers. The availability of these resources varies among languages. Further research is in progress by the authors for the application of the above framework for analyzing citizens' comments on policy messages posted by government organizations in several social media, as part of the PADGETS project mentioned in the introductory section.

References

1. Organization for Economic Co-operation & Development (OECD) (2001). OECD Handbook on Information, Consultation and Public Participation in Policy-Making, OECD Publications Service, Paris.
2. Organization for Economic Co-operation & Development (OECD) (2003). Engaging Citizens Online for Better Policy-making - Policy Brief, OECD Publications Service, Paris.
3. Organization for Economic Co-operation & Development (OECD) (2004). Promise and

Problems of e-Democracy: Challenges of Online Citizen Engagement, OECD Publications Service,Paris.

4. Xenakis, A., Loukis, E. (2010). An Investigation of the Use of Structured e-Forum for Enhancing e-Participation in Parliaments', International Journal of Electronic Governance, Volume 3, Issue 2, pp. 134 – 147.

5. Osimo D. (2008). Web 2.0 in Government: Why and How? JRC Scientific and Technical Reports. European Commission, Joint Research Centre, Institute for Prospective Technological Studies.

6. Chadwick, A. (2009). Web 2.0: New Challenges for the Study ofE-Democracy in an Era of InformationalExuberance, I/S: A Journal of Law and Policy for the Information Society, 5 (1), pp. 9-41.

7. Mergel, I. A., Schweik, C. M. and Fountain, J. E. (2009). The Transformational Effect of Web 2.0 Technologies on Government. Available at SSRN: http://ssrn.com/abstract=1412796

8. Charalabidis, Y., Gionis, G., Ferro, E., Loukis, E. (2010). Towards a Systematic Exploitation of Web 2.0 and Simulation Modeling Tools in Public Policy Process. IFIP Second International Conference on e-Participation - ePart 2010, August 29 - September 2, 2010, Lausanne, Switzerland.

9. Rose,J., Sanford, C. (2007). MappingeParticipation Research: Four Central Challenges,Communications of the Association for Information Systems, 20, pp. 909- 943.

10. Macintosh, A., Coleman, S.and Schneeberger,A. (2009). eParticipation: The Research Gaps, ePart 2009, LNCS 5694, Springer-Verlag Berlin Heidelberg,pp. 1–11.

11. Pang, B., Lee, L. (2008).Opinion mining and sentiment analysis. Foundations and Trends in Information Retrieval 2(1-2), pp. 1–135.

12. Liu, B. (2006). Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data. Springer.

13.Wiebe, J., Wilson, T. Cardie, C. (2005). Annotating expressions of opinions and emotions in language.Language Resources and Evaluation, 1(2).

14. Godbole, N.,Srinivasaiah, M. and Skiena, S. (2007). Large-scale sentiment analysis for news and blogs, Proceedings of the International Conference on Weblogs and Social Media (ICWSM).

15. Liu, B., Hu, M., Cheng, J. (2005). Opinion observer: Analyzing and comparing opinions on the web, Proceedings of WWW.

16. Wilson, T.,Wiebe, J.,Hwa, R. (2004). Just how mad are you? Finding strong and weak opinion clauses, Proceedings of AAAI, pp. 761–769.

17. Parrott,W. (2001). Emotions in Social Psychology, Psychology Press, Philadelphia.

18. Aue,A.,Gamon, M. (2005).Customizing sentiment classifiers to new domains: A case study, Proceedings of Recent Advances in Natural Language Processing (RANLP).

19. Choi, Y.Breck, E. and Cardie, C. (2006). Joint extraction of entities and relations for opinion recognition, Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP).

20. Das,S. R., Chen, M. Y. (2007). Yahoo! for Amazon: Sentiment extraction from small talk on the Web, Management Science, vol. 53, pp. 1375–1388.

21. Dave, K.,Lawrence, S., Pennock, D. M. (2003). Mining the peanut gallery: Opinion extraction and semantic classification of product reviews, Proceedings of WWW, pp. 519–528.

22. Devitt,A., Ahmad, K. (2007). Sentiment analysis in financial news: A cohesion based approach, Proceedings of the Association for Computational Linguistics (ACL), pp. 984–991.

23. Gamon, M. (2004), Sentiment classification on customer feedback data: Noisy data, large feature vectors, and the role of linguistic analysis, Proceedings of the International Conference on Computational Linguistics (COLING).

24. Gamon, M.,Aue, A.,Corston-Oliver, S.,RinggerE. (2005). Pulse: Mining customer opinions from free text, Proceedings of the International Symposium on Intelligent Data Analysis (IDA), pp. 121–132.

25. Pang, B.,Lee, L.,Vaithyanathan, S. (2002).Thumbs up? Sentiment classification using machine learning techniques, Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP), pp. 79–86.

26. Turney, P. (2002). Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews," Proceedings of the Association for Computational Linguistics (ACL), pp. 417–424.

27. Blitzer, J.,Dredze, M., Pereira, F. (2007). Biographies, Bollywood, boom-boxes and blenders: Domain adaptation for sentiment classification, Proceedings of the Association for Computational Linguistics (ACL).

28. Yang, H. Si, L.,Callan, J. (2006). Knowledge transfer and opinion detection in the TREC2006 blog track, Proceedings of TREC.

29. Hatzivassiloglou,V., Wiebe, J. (2000). Effects of adjective orientation and gradability on sentence subjectivity," Proceedings of the International Conference on Computational Linguistics (COLING).

30. Riloff,E. ,Wiebe, J. (2003). Learning extraction patterns for subjective expressions, Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP).

31. Nasukawa,T., Yi, J. (2003).Sentiment analysis: Capturing favorability using natural language processing, Proceedings of the Conference on Knowledge Capture (K-CAP).

32. Esuli,A., Sebastiani, F. (2006). SentiWordNet: A publicly available lexical resource for opinion mining, Proceedings of Language Resources and Evaluation (LREC).

33. Riloff, E.,Wiebe, J., Wilson, T. (2003). Learning subjective nouns using extraction pattern bootstrapping, Proceedings of the Conference on Natural Language Learning (CoNLL), pp. 25–

34. Yu,H., Hatzivassiloglou, V. (2003). Towards answering opinion questions: Separating facts from opinions and identifying the polarity of opinion sentences, Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP).

35. Kennedy,A.,Inkpen, D. (2006). Sentiment classification of movie reviews using contextual valence shifters," Computational Intelligence, vol. 22, pp. 110–125.

36. Kim,S. M., Hovy, E. (2004). Determining the sentiment of opinions," Proceedings of the International Conference on Computational Linguistics (COLING).

37.Kim,S. M., Hovy, E. (2006). Automatic identification of pro and con reasons in online reviews, Proceedings of the COLING/ACL Main Conference Poster Sessions, pp. 483–490.

38. Kim,S. M., Hovy, E. (2007). Crystal: Analyzing predictive opinions on the web, Proceedings of the Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP/CoNLL).

39. Kim,S. M., Pantel, P.,Chklovski, T., Pennacchiotti, M. (2006). Automatically assessing review helpfulness," Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP), pp. 423–430, Sydney, Australia, July.

40. Wilson, T.,Wiebe, J., Hoffmann, P. (2005). Recognizing contextual polarity in phrase-level sentiment analysis. Proceedings of the Human Language Technology Conference and the Conference on Empirical Methods in Natural Language Processing (HLT/EMNLP), pp. 347–354.

41. Polanyi,L.,Zaenen, A. (2004). Contextual lexical valence shifters, Proceedings of the AAAI Spring Symposium on Exploring Attitude and Affect in Text.

42. Morinaga, S. Yamanishi, K.,Tateishi,K., Fukushima, T. (2002). Mining product reputations on the Web, Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), pp. 341–349.(Industry Track).

43.Tong, R. M. (2001). An operational system for detecting and tracking opinions in on-line discussion.Proceedings of the Workshop on Operational Text Classification (OTC).

44.Yi, J.,Nasukawa, T.,Bunescu, R.,Niblack, W. (2003). Sentiment analyzer: Extracting sentiments about a given topic using natural language processing techniques, Proceedings of the IEEE International Conference on Data Mining (ICDM).

45. Fellbaum, C. (ed.)(1998), Wordnet: An Electronic Lexical Database. MIT Press.

46. Andreevskaia,A., Bergler, S. (2006). Mining WordNet for a fuzzy sentiment: Sentiment tag extraction from WordNet glosses. Proceedings of the European Chapter of the Association for Computational Linguistics (EACL).

47. Esuli,A.,Sebastiani, F. (2005). Determining the semantic orientation of terms through gloss analysis," Proceedings of the ACM Conference on Information and Knowledge Management (CIKM).

48. Esuli,A.,Sebastiani, F. (2006). Determining term subjectivity and term orientation for opinion mining," Proceedings of the European Chapter of the Association for Computational Linguistics (EACL).

49. Esuli,A.,Sebastiani, F. (2007). PageRankingWordNetsynsets: An application to opinion mining, Proceedings of the Association for Computational Linguistics (ACL).

50. Kamps, J.,Marx, M.,Mokken,R. J., De Rijke,M. (2004). Using WordNet to measure semantic orientation of adjectives. In Proc. of LREC'04, pp. 1115-1118.

51. Ding, X., Liu, B., Yu, P. S. (2008). A holistic lexicon-based approach to opinion mining, Proceedings of the Conference on Web Search and Web Data Mining (WSDM).

52. Qiu, G.,Liu, B., Bu,J., Chen,C. (2009).Expanding Domain Sentiment Lexicon through Double Propagation, International Joint Conference on Artificial Intelligence (IJCAI-09).

53. Hatzivassiloglou,V., McKeown, K. (1997). Predicting the semantic orientation of adjectives, Proceedings of the Joint ACL/EACL Conference, pp. 174–181.

54. Kanayama,H., Nasukawa, T. (2006). Fully automatic lexicon expansion for domain-oriented sentiment analysis, Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP), pp. 355–363.

55. Popescu,A. M.,Etzioni, O. (2005). Extracting product features and opinions from reviews, Proceedings of the Human Language Technology Conference and the Conference on Empirical Methods in Natural Language Processing (HLT/EMNLP).

56.Liu, Y., Huang, J.,An, A., Yu,X. (2007). ARSA: A sentiment-aware model for predicting sales performance using blogs," Proceedings of the ACM Special Interest Group on Information Retrieval (SIGIR).

57.McDonald, R.,Hannan, K.,Neylon, T.,Wells, M., Reynar, J. (2007).Structured models for fine-to-coarse sentiment analysis, Proceedings of the Association for Computational Linguistics (ACL), pp. 432–439, Prague, Czech Republic.

58. Su, Q.,Xu, X.,Guo, H.,Wu, X., Zhang, X.,Swen, B., SuZ.,(2008). Hidden Sentiment Association in Chinese Web Opinion Mining.Proceedings of WWW'08, pp. 959-968.

59. Mei, Q. Ling, X. Wondra, M. Su, H., Zhai, C. X. (2007). Topic sentiment mixture: Modeling facets and opinions in weblogs," Proceedings of WWW, pp. 171–180.

60. Carenini, G.,Ng, R. T., Zwart, E. (2005).Extracting knowledge from evaluative text, Proceedings of International Conference on Knowledge Capture (K-CAP), pp. 11-18.

61. Qiu, G., Liu, B., Bu J., Chen,C. (2009). Expanding Domain Sentiment Lexicon through Double Propagation, International Joint Conference on Artificial Intelligence (IJCAI-09).

62. I-sieve Technologies, Sentiment Analysis Beyond Impressions, retrieved from http://i-sieve.com/

63. Wei,W., Gulla, J.A. (2010).Sentiment Learning on Product Reviews via Sentiment Ontology Tree", Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics, pages 404–413, Uppsala, Swede, Association for Computational Linguistics.

64. Riloff, E.,Patwardhan, S.,Wiebe, J. (2006).Feature subsumption for opinion analysis, Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP).

65. Riloff,E., Wiebe, J. (2003). Learning extraction patterns for subjective expressions, Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP).