A Survey Paper on Extraction of Web Services for Effective User Review Sentiment Analysis

Jyotsna A Pawar¹, Prof. N. R. Wankhade²

¹Department of Computer Engineering, KCT’s Late G.N. Sapkal College of Engineering, Anjaneri, Nashik, India
²Department of Computer Engineering, KCT’s Late G.N. Sapkal College of Engineering, Anjaneri, Nashik, India

Abstract: This paper presents a survey on features Extraction of Web Services for effective User Review Sentiment Analysis. Recognition of opinion features from online user reviews is a task to find on which feature user is going to put his opinion. For opinion feature identification there are number of existing techniques but, features are extracted from single corpus. Quality of Service has become a standard way of evaluating web services and selecting the one that suites user interests the best. Traditional methods adopt a fixed set of QoS parameters and typical ones include response time, fee, and availability. There currently lacks an effective way of identifying quality features that users are actually interested in when choosing a service. Meanwhile, the traditional way of collecting QoS values relies on either public information released by service providers or test results from repeatedly invoking a service. The last section of paper will discuss the disadvantages of current systems and will state the future scope.

Keywords: Opinion mining, Opinion feature, Intrinsic Domain Relevance, Extrinsic Domain Relevance, QoS, web services, sentiment analysis.

1. Introduction

The sudden growth in social media such as reviews, blogs, comments are postings on the Web, are increasingly using the content in these media for decision making. There are many web portals that solicit feedback from end users on the usage of web services and other online services. The input from users forms a large user review repository. Each review is a free text description, which not only reveals user opinion on the service but also implies the sentiment orientation: positive or negative. Quality of Service (QoS) has become a way of evaluating web services and selecting the one that suites user interests the best. Typical QoS features include reliability, response time, security, and invocation fee. QoS plays an essential role in various web service management tasks, such as selecting a service that fulfills both the functional and non-functional requirement specified by a user. Opinion mining (also known as sentiment analysis) aims to analyze people’s opinions, sentiments, and attitudes toward entities such as products, services, and their attributes [1]. Sentiments or opinions expressed in textual reviews are typically analyzed at various resolutions. For example, document-level opinion mining identifies the overall subjectivity or sentiment expressed on an entity (e.g., cellphone or hotel) in a review document, but it does not associate opinions with specific aspects (e.g., display, battery) of the entity.

To extract opinion features in opinion mining many techniques have been proposed. Supervised learning model[2],[3] is expected to work well in given domain, but the model must be developed if it is applied to various domains. Unsupervised natural language processing techniques [4], [5], [6] describe opinion features by defining domain-independent syntactic rules that tranche the dependence roles and local context of the feature terms. Topic modeling approach [7] [8] is to mine generic topics.

Current QoS approaches use the QoS features (or parameters), which were primarily determined by domain experts. Those parameters can be domain-independent, such as availability, security, cost, and reliability, as well as domain specific, such as latency for weather services and accuracy for traffic services. The evaluation of QoS features mainly rely on two resources. First, some web service providers make related information, such as the security level and invocation fee, available to users. Second, some users or third party agents may run tests and collect QoS values, such as for availability and reliability. Although these approaches have been increasingly adopted, they still have their limitations. First, the predefined QoS features may not always reflect what users are interested in. For example, users may care about if a service is always compatible to the previous version when updated, so that no versioning issue will occur once the service is included in a software package. However, this concern may not be foreseen by domain experts when determining the QoS list. Second, it is limited to rely on the QoS information published by service providers, which may be misleading or unauthentic, or rely on the testing result in a particular time period or in a particular geographic area, as the QoS values can vary under different temporal and/or spatial settings.

2. Literature Survey

An Opinion in reviews is examined at document, sentence or phrase level for classification of overall subjectivity in a single review. The use of a global structured model that learns to predict sentiments ondifferent levels of granularity for a textual review is invested by Mcdonald et al. [9]. The primary advantage of the proposed model is that it allows classification decisions from one level in the text to influence decisions at another. A regression method based on the bag of opinions model was proposed for review rating prediction from sparse text patterns [10]. Review rating estimation is a much more complicated problem compared to binary
sentiment classification. Generally, sentiments are expressed differently in different domains. The sentiment classification methods discussed above can be tuned to work very well on a given domain; however, they may fail in classifying sentiments in a different domain. A cross-domain sentiment classifier using an automatically extracted sentiment thesaurus was proposed by Bollegala et al. [11].

The effects of dynamic adjectives, semantically oriented adjectives, and gradable adjectives on predicting subjectivity are studied by Hatzivassiloglou and Wiebe [12]; they proposed a supervised classification method to predict sentence subjectivity. Pang et al. [13] proposed three machine learning methods, naive Bayes, maximum entropy, and support vector machines, to classify whole movie reviews into positive or negative sentiments. They found that standard machine learning techniques produced good results in comparison to human-generated baselines. Moreover, machine learning methods did not perform as well on sentiment classification as on traditional topic based categorization. To prevent a sentiment classifier from considering irrelevant or even potentially misleading text, Pang and Lee [14] proposed to first employ a sentence-level subjectivity detector to identify the sentences in a document as either subjective or objective, and subsequently discarding the objective ones. They then applied the sentiment classifier to the resulting subjectivity extract, with improved results.

Differently, sentiment analysis at the phrase (word) level mainly focuses on classifying sentiment polarities of opinion phrases (words). Generally, the sentiment polarity of an opinion word is usually context-dependent as well as domain-specific. Wilson et al. [15] presented an approach to predicting contextual sentiments at the phrase level by applying machine learning techniques on a variety of feature factors. Yessenalina and Cardie [16] presented a compositional matrix-space model for phrase-level sentiment analysis. One of the benefits of the proposed approach is that by learning matrices for words, the model can handle unseen word compositions (e.g., unseen bigrams) as long as the component unigrams have been learned. A two-level affective reasoning method was proposed to mimic the integration of conscious and unconscious reasoning to address word-level sentiment analysis tasks [17].

To classify review documents as thumbs up (positive) or thumbs down (negative) in [18] an unsupervised learning method was proposed. The sentiment of each review document is predicted by the average sentiment orientations of phrases in the review. Domain-dependent contextual information is also considered for better estimation of the phrase sentiments. One limitation of this work is its reliance on an external search engine. A rule-based semantic analysis approach to classify sentiments for text reviews is proposed by Zhang et al. [19]. Word dependence structures to classify the sentiment of a sentence is use in this approach, and predicted document-level sentiments via aggregating the sentence sentiments. Due to the lack of comprehensiveness in their rules rule-based approaches typically suffer from poor coverage. Maas et al. [20] presented an approach to document-level and sentence-level sentiment classification tasks, which uses a mix of unsupervised and supervised techniques to learn word vectors by capturing semantic term-document information as well as rich sentiment content.

Supervised learning models are including hidden Markov models. Supervised models perform well on a given domain, but it need to training again when it applied to a different domain. For a wide variety of products and services in different domains, supervised methods are not efficient because it is very expensive to construct labeled data for each product or service. In addition, this model requires a decent-sized set of labeled data for model learning on every domain [11]. Supervised learning models that require labeled data have been successfully used to build sentiment classifiers for a given domain. In supervised learning model there is problem in domain adaption [10].

Unsupervised technique can be applied to any domain because it does not require any fix sentiment lexicons. Syntactic relationships between features and opinions can be used to locate opinion features in a sentence by using carefully generated syntactic rules [23] in an unsupervised approach. Syntactic relations which are identified by this methods help to locate features associated with opinion words. Unsupervised NLP approaches extract opinion features by mining syntactic patterns of features present in review sentences [24] An unsupervised approach to capturing sentiment-oriented aspects for online reviews is challenging, since the reviews are short and usually each aspect is mentioned only once in a review. Due to the colloquial nature of online reviews this approach could extract large number of invalid features. Unsupervised corpus statistic approach resists the colloquial nature of son line review in order give a large review corpus to extract valid feature. This approach uses the results of statistical analysis to understand the distributional characteristic of the opinion feature.

An association rule mining (ARM) approach is proposed by Hu and Liu [25] to mine frequent item sets and generate opinion features. The opinion features are nouns & noun phrases with high sentence-level frequency. ARM has limitation for the task of feature identification, as it depends on frequency of item sets i.e., frequent but invalid features are not extracted correctly and rare but valid features maybe neglected. To address the problem of feature-based opinion mining, A mutual reinforcement clustering (MRC) approach is introduced by Suet et al. [26] to mine the associations between product feature categories and opinion word groups. MRC technique use multisource knowledge including semantic and textual structure. Based on a co-occurrence weight matrix generated from the given review corpus. The common support clustering approach fully feat the relationship between product features and opinion words. MRC approach is able to extract infrequent features than many other corpus statistics methods. The occasional features are extracted only when, there are the mutual relationships between feature and opinion groups are found during the clustering phase is accurate. MRC’s precision is low due to the difficulty in obtaining good clusters on real-life reviews.

Yu et al. [27] proposed an aspect ranking algorithm which is based on the probabilistic regression model. The use of this ranking algorithm is to automatically identify important
product aspects from online consumer reviews. Aspect ranking algorithm finds the important aspects by parallel taking the aspect frequency and the influence of consumers opinions given to each aspect on their overall opinions. Moreover, this algorithm focus on ranking product aspects that are actually coarse-grained clusters of specific features, not on extracting feature terms that are commented on explicitly in reviews.

Another techniques of unsupervised topic modeling, such as latent Dirichlet allocation (LDA) [28], approach does not count the relationships among sentences, thus dismissing that the same aspect may have quite various word usages in different sentences. This approach is a procreative three-way (term-topic-document) model that have been used to solve aspect-based opinion mining tasks. The models which are formulated primarily for mining latent topics or aspects, are in reality use to identify coarse-grained topics or aspects that represent to distinguishing properties of the commented entities. This method may not utter opinion feature explicitly in the reviews, but rather user-defined clusters of specific opinion features [28], [29], [30], [22]. The techniques are good in exposing latent structures of review data but they may be less successful in dealing with identifying specific feature terms commented on explicitly in reviews. Most of these existing techniques to feature extraction typically only use the knowledge or patterns mined from a given single review corpus, while it completely ignoring the potential variations present in a different domain independent corpus.

A. QoS Management

In [31], an approach was proposed to identify those component services that do not deliver the expected QoS values in a business process. It integrates dependency matrix based and Bayesian network based diagnosis to reduce diagnose cost and improve the accuracy. In [32], an approach was proposed to realize a soft probabilistic contract for QoS management in a composite web service. Instead of enforcing a hard constraint on QoS values in a contract, this work acknowledged the uncertain nature of QoS values and provided flexibility for reaching agreement between users and services. The work in [31] and [32] both focused on manage QoS values of web services and their conformance to service level agreement, which is different from our focus.

B. QoS Collection and Evaluation

In [33], the importance of QoS in web service selection, discovery, recommendation, and composition is acknowledged. A comprehensive QoS evaluation on existing web services was conducted to generate a QoS dataset to benefit the research community of service computing. Over 20 thousands real world web services were tested and the evolution focused on respond time and throughput. In [34], a language was proposed to express QoS monitoring requirement from user perspective. A monitoring request is processed as data steams to cater for its continuous nature. It is evaluated continuously as new testing result is generated. A sliding window is used and the request is responded with the statistical observation during the sliding window. In [35], an approach was proposed to automatically learn and monitor QoS values of web services through Aspect Oriented Programming (AOP). Under AOP, a service stub is generated for each evaluated service and an invocation was performed for the evaluation. This approach focuses on a subset of common QoS features that can be measured through invocation, such as response time, execution time, availability, and accuracy. In [36], a forecasting approach was proposed to predict dynamic QoS values. It leverages both Auto Regressive Integrated Moving Average (ARIMA) and Generalized Auto Regressive Conditional Heteroscedastic (GARCH) to capture the volatility of QoS data and predict the future values. In sum, these methods focus on evaluating the values of predefined QoS parameters.

C. QoS Evaluation for Composite Services

In [37], an approach was proposed to compute QoS values of a composite service based on the workflow of the service and QoS values of each component service. Instead of using fixed values, this approach models each QoS parameter as a probabilistic distribution to better capture the real world scenarios. In [38], an approach was proposed to integrate QoS values and analyze QoS requirement in reconfigurable web services choreographies. It reconfigures a composite service given a QoS goal, with a major focus on latency, throughput, accuracy, and data quality. These approaches focus on measuring and analyzing QoS for composite services, based on the QoS values of individual services.

D. Quality feature extraction

In [39], Quality of Experience (QoE) parameters were extracted from analyzing user reviews. It uses POS tagging to identify frequent nouns in reviews as potential QoE features. Similar nouns are aggregated and grouped into clusters using semantic lexicon, such as SentiWordNet. Representative nouns in each cluster are selected as QoE parameters. This work is most close to our work as it also exploits user reviews as the input for quality feature extraction from services. The difference lies in the way of extracting the features and the extraction result. Instead of choosing frequent nouns, our approach extracts features that are associated with user sentiment orientation towards a service. By seamlessly integrating feature extraction with sentiment analysis, our approach is able to extract better features that are more relevant to the quality aspects of services while being more indicative of users’ positive or negative opinions.

Our experimental results clearly justify the effectiveness of our approach. Feature extraction from user reviews has also been investigated in natural language processing and machine learning [40], [41], [42]. Most of these approaches rely on POS tagging [40], association rule mining [40], unsupervised learning [41], or semi-supervised learning over unlabeled data [42] for feature extraction. In contrast, our approach adopts a supervised learning strategy that extracts quality attributes, performs sentiment analysis, and assigns sentiment orientation to the quality attributes using a single integrated model.

3. Discussions

The literature study reveals that all the system is having few disadvantages in classifying the Quality features of web services through user review sentiment analysis. Most of
these subsisting techniques to feature extraction typically only use the knowledge or patterns mined from a given single review corpus, while it altogether ignoring the likely variations present in a different domain independent corpus. Few systems sentiment classifier considered irrelevant text which degrades the performance of the system. System based on ARM extracts large number of invalid features which also affects the performance. Hence a strong system must be developed which will process the huge text fast and with effective classifications of the reviews, generating the best review on the basis on user sentiment analysis. IEDR (Intrinsic and Extrinsic domain relevance) is use to identify the opinion features and this will make the system generate more accurate results.

4. Conclusion

All the techniques for extracting and qualifying the reviews based on user sentiment analysis are discussed at length in this paper. The past and current system has been reviewed in terms of the techniques used, their results and performance. In literature survey many journal paper have been researched to explain in deep the problems related to user sentiment analysis. In future, more emphasis can be given on IEDR which may prove to be the best approach as it evaluates the domain relevance of an opinion feature and extract the effective opinion features along with feature clustering approach.

References


