Originally published in: Griesbaum, Joachim/Heuwing, Ben/Ruppenhofer, Josef/Werner, Katrin (Eds.): Proceedings des 8. Hildesheimer Evaluierungs- und Retrievalworkshop (HiER 2013), Hildesheim, 25.-26. April 2013. - Hildesheim: Institut für Informationswissenschaft und Sprachtechnologie, Universität Hildesheim, 2013. Pp. 41-50.
https://nbn-resolving.org/urn:nbn:de:gbv:hil2-opus-1929

Predicate Acquisition for Opinion Holder Extraction

A Data-Intensive Approach

Michael Wiegand

Spoken Language Systems, Saarland University
Gebäude C7 1
D-66123 Saarbrücken
michael.wiegand@lsv.uni-saarland.de

Abstract

Opinion holder extraction is one of the most important tasks in sentiment analysis. We will briefly outline the importance of predicates for this task and categorize them according to part of speech and according to which semantic role they select for the opinion holder. For many languages there do not exist semantic resources from which such predicates can be easily extracted. Therefore, we present alternative corpus-based methods to gain such predicates automatically, including the usage of prototypical opinion holders, i.e. common nouns, denoting for example experts or analysts, which describe particular groups of people whose profession or occupation is to form and express opinions towards specific items.

Introduction

Opinion holder extraction is one of the most important subtasks in sentiment analysis. It deals with the automatic extraction of entities uttering an opinion. For example, the opinion holders in Sentences (1) and (2) are *the vet* and *Russia*, respectively.

(1) The owner put down the animal although the vet had **forbidden** him to do so.
(2) Russia **favors** creation of "international instruments" to regulate emissions.

The extraction of sources of opinions is an essential component for complex real-life applications, such as opinion question answering systems or opinion summarization systems.

Most of recent research focused on supervised learning methods (Choi et al., 2005; Choi et al., 2006; Wiegand & Klakow, 2010). While such approaches typically yield robust classifiers, they rely heavily on a large amount of manually annotated sentences. Moreover, these methods are little telling with regard to the linguistic aspects underlying the classification task.

The Role of Predicates

By far the strongest clue for the task of opinion holder extraction are the predicates that syntactically relate to the candidate opinion holders (Wiegand & Klakow, 2011). The relevant types of predicates may be verbs (3), (deverbal) nouns (4) or adjectives (5).

(1) I **believe**$_{\mathbf{verb}}$ that this is more than that.
(2) This includes a growing **reluctance**$_{\mathbf{noun}}$ by foreign companies to invest in the region.
(3) Ordinary Venezuelans are even less **happy**$_{\mathbf{adj}}$ with the local oligarchic elite.

Moreover, the argument position in which the opinion holder is realized varies among the different lexical units. For verbs, the opinion holder may either be in agent position (6) or patient position (7). The former argument position, however, has been identified as – by far – the most frequently occurring construction (Bethard et al. (2004) report that 90% of the opinion holders on their datasets are agents.)

(4) China$_{agent}$ had always firmly **opposed** the US Taiwan Affairs Act.
(5) Washington **angered** Bejing$_{patient}$ last year.

Table 1 compares on the MPQA-corpus (Wiebe et al., 2005) a state-of-the-art supervised classifier that incorporates a large set of different features (the best classifier from Wiegand & Klakow (2010)) with a classifier that only contains knowledge of which predicates typically have an opinion holder as an argument. The predicate-based classifier should be considered an upper bound, since it already knows which predicates co-occur in principle with opinion holders and also in which argument position they can be found. (This information has been extracted from the labeled MPQA-corpus.) The table supports the usefulness of predicates for opinion holder extraction.

Table 1 Comparison of supervised and predicate-based classification.

Method	**Precision**	**Recall**	**F-Score**
Supervised Classifier	59.10	66.57	62.61
Predicate-based Classifier (Oracle)	47.04	68.62	55.82

Automatic Acquisition of Predicates

In a situation in which no labeled training data for supervised learning are available, a predicate-based classifier seems to be a promising solution. For the acquisition of the predicates, one can either employ manually compiled lexical resources, such as FrameNet (a er et al., 1998) or Levin's verb classes (Levin, 1993) (this has been explored in Kim & Hovy (2006) and Wiegand & Klakow (2011)), or develop a data-driven induction method. In the following we will present an approach for the latter type of method. It is particularly attractive since it is only dependent on large amounts of unlabeled textual data rather than manually-compiled resources. As an unlabeled text corpus for this kind of induction, we make use of the North American News Text Corpus (LDC95T21).

Prototypical Opinion Holders

Even though an unlabeled corpus does not contain any explicit information about which entities are actual opinion holders, there exists a set of common nouns denoting particular groups of people who, since their profession or occupation is to form and express some opinion towards specific items, disproportionately often represent actual opinion holders. Therefore, we use the mentions of these nouns as a proxy for actual opinion holders. Typical examples of such prototypical opinion holders (protoOHs) are *advocates*, *analysts*, *critics*, *experts*, *reviewers* or *supporters*. Sentences (8) and (9) illustrate two mentions of such nouns.

(6) Experts$_{protoOH}$ **agree** it generally is a good idea to follow the manufacturer's age recommendation.
(7) Shares of Lotus Development Corp. dropped sharply after analysts$_{protoOH}$ **expressed** concern about their business.

Table 2 illustrates the verbs most highly correlating with protoOHs (we use *Pointwise Mutual Information*). All of these predicates have the opinion holder in agent position[1]. Intuitively, the verbs that are thus gained are plausible predictors for opinion holders.

Table 2 List of verbs most highly correlating with protoOHs.

accuse acknowledge anticipate argue agree attribute believe call caution charge change cite claim concede criticize conclude contend consider complain describe disagree doubt fear expect estimate find forecast hope look note praise predict point question recommend say see speculate suggest tell testify think try view warn write wonder worry

One of the advantages of the usage of prototypical opinion holders is that predicates with parts of speech other than verbs can also be extracted.

1 Semantic roles were recognized by a mapping from grammatical functions (e.g. nsubj(complain-v, proponents-n) typically represents a noun in agent position while dobj(anger-v, fans-n) a noun in patient position).

Predicates that Select Opinion Holders in Patient Position

The downside of protoOHs as a means to extract predictive predicates for opinion holder extraction is that it only reliably detects predicates having the opinion holder in agent position. If a protoOH occurs in patient position, it often does not function as an opinion holder (10).

(8) They **criticized**$_{\mathbf{verb}}$ their opponents$_{\text{protoOH}}$.

For that reason, we need to devise a complementary induction method to extract predicates selecting opinion holders in patient position.

Our induction method for these verbal predicates rests on the observation that the past participle of those verbs, such as *shocked* in Sentence (11), is very often identical to some predicative adjective (12) having a similar if not identical meaning. (For the predicative adjective, the opinion holder is, however, its agent and not its patient.)

(9) He had **shocked**$_{\mathbf{verb}}$ me. (opinion holder: patient)
(10) I was **shocked**$_{\mathbf{adj}}$. (opinion holder: agent)

Instead of extracting those verbs directly (11), we take the detour via their corresponding predicative adjectives (12). This means that we collect all those verbs from our unlabeled corpus for which there is a predicative adjective that coincides with the past participle of the verb. Table 3 shows some verbs that could thus be extracted.

Table 3 Examples of induced predicates that take opinion holders in patient position.

anguish astonish astound concern convince daze delight disenchant disappoint displease disgust disillusion dissatisfy distress embitter enamor engross enrage entangle excite fatigue flatter fluster flummox frazzle hook humiliate incapacitate incense interest irritate obsess outrage perturb petrify sadden sedate shock stun tether trouble

Evaluation

For the evaluation, we consider two datasets belonging to two different domains. Cross-domain evaluation is important in order to obtain a realistic view of the effectiveness of different classification methods. The first domain that we consider, ETHICS (5700 sentences), is the largest domain subset of the MPQA-corpus (Wiebe et al., 2005). It contains news items that discuss human rights issues (mostly in the context of combating global terrorism). The second domain, FICTION (614 sentences), is a set of summaries of fictional work (two Shakespeare plays and one novel by Jane Austen). We will compare both predicate-based and supervised classifiers. As the second domain corpus is much smaller than the first, it can only be used as a test corpus. We use ETHICS for in-domain evaluation for supervised classifiers (we carry out 5-fold cross validation) while FICTION is used as an out-of-domain dataset (i.e. we train on ETHICS and test on FICTION).

Another difference between the two domain corpora lies in the distribution of opinion holders in patient position as stated in Table 4. The proportion is notably higher on the out-of-domain dataset than on the in-domain dataset. This raises the question whether a supervised classifier trained on ETHICS with its very few opinion holders in patient position is able to identify all types of opinion holders on FICTION.

Table 4 Proportion of opinion holder as patients.

ETHICS	FICTION
1.47	11.59

Table 5 compares predicate-based classification with supervised classification. For supervised classification, we use the best classifier from Wiegand & Klakow (2010) that incorporates no external lexicons. For predicate-based classification we consider an automatic induction method (as presented in the previous section) and a classifier using manually compiled resources. The resources comprise *strong subjective expressions* of the Subjectivity Lexicon (Wilson et al., 2005) with a subset of Levin's verb classes (Levin, 1993). The latter lexicon is specifically designed to identify opinion holders in patient position. Our choice of manually-compiled lexicons was shown to be most effective on a large-scale analysis of different resources (Wiegand & Klakow, 2011). Note that for each predicate-based classification method (i.e. *Induction* and *Lexicon*), we consider two configurations: AG just detects opinion holders in agent position while AG+PT detects opinion holders in both agent and patient position.

Table 5 shows that the supervised classifier only outperforms the predicate-based classifier on the in-domain evaluation. On the out-of-domain dataset *Lexicon* is notably better while *Induction* is only slightly worse than the supervised classifier. The supervised classifier is much worse on the out-of-domain dataset since supervised classifiers are very susceptible to domain mismatches. That classifier fails to correctly identify many opinion holders in patient position since only few of them have been included in the training set (i.e. ETHICS corpus). *Induction* is the weakest performing classifier but considering the little manual effort that is required to build it, its performance on ETHICS in comparison to *Lexicon* is already fairly remarkable.

The configurations of the predicate-based classifiers that also detect opinion holders in patient position are only effective on FICTION which can be explained by the fact that this dataset contains significantly more of such opinion holders than ETHICS (see Table 4). Since we observe an improvement on FICTION with *Induction* from AG to AG+PT, we can also conclude that our induction method to extract predicates that take opinion holders in patient positions works.

The gap between *Induction* and *Lexicon* is much larger on FICTION. This can be explained by the fact that *Induction* has a bias towards news texts (hence the good performance on ETHICS) since it was generated with the help of a large news corpus.

Table 5 Comparison of different classifiers (evaluation measure: F-Score).

	Predicate-based classifier				**Supervised classifier**
	Induction		**(Manually-compiled) Lexicon**		
Domain	**AG**	**AG+PT**	**AG**	**AG+PT**	**(No lexicons included)**
ETHICS (in-domain)	50.77	50.99	52.22	52.27	59.52
FICTION (out-of domain)	46.59	49.97	54.84	59.35	51.21

Finally, Table 6 examines whether the supervised classifier can be improved if information from either of the predicate-based classifiers is added. The table clearly shows that on both domains the addition of either of the predicate-based classifiers results in some improvement. However, the gain in improvement is greater on the out-of-domain dataset from which we conclude that the predicate-based classifier is an important means to bridge a domain mismatch between source and target corpus. Although *Induction* itself has a domain bias, this method is less restrictive than supervised classification. This can partly be explained by the fact that the amount of unlabeled data that *Induction* uses is much larger than the amount of labeled training data that is available for supervised learning.

Table 6 Combining the supervised classifier with predicate-based classifiers (evaluation measure: F-Score).

	Supervised classifier	**Supervised classifier +Induction**	**Supervised classifier+Lexicon**
ETHICS (in-domain)	59.52	60.42	61.50
FICTION (out-of domain)	51.21	59.58	65.67

Conclusion

We presented a data-driven approach for opinion holder extraction that focuses on the extraction of predictive predicates. Two induction approaches were examined, one relying on prototypical opinion holders and another that exploits the similarity between predicative adjectives and verbs. The former is designed to extract opinion holders in agent position (which is also the most frequently occurring argument position in general), while the latter is specifically designed for the extraction of predicates that take opinion holders in patient position. The induction methods were compared with a predicate-based approach using a manually compiled lexicon and supervised learning. The results are promising and suggest particular significance of predicate-based classification for cross-domain opinion holder extraction.

References

Baker, C., Fillmore, C., and Lowe, J. (1998). The Berkeley FrameNet Project. In *COLING*.

Bethard, S., Yu, H., Thornton, A., Hatzivassiloglou, V., and Jurafsky, D. 2004. Extracting Opinion Propositions and Opinion Holders using Lexical Clues. In Shanahan, J., Qu, Y., and Wiebe, J. (Eds.) *Computing Attitude and Affect in Text: Theory and Applications*.

Choi, Y., Cardie, C., Riloff, E., and Patwardhan, S. (2005). Identifying Sources of Opinions with Conditional Random Fields and Extraction Patterns. In *HLT/EMNLP*.

Choi, Y., Breck, E., and Cardie, C. (2006). Joint Extraction of Entities and Relations for Opinion Recognition. In *EMNLP*.

Kim, S., and Hovy, E. (2006). Extracting Opinions, Opinion Holders, and Topics Expressed in Online News Media Text. In *ACL-Workshop on Sentiment and Subjectivity in Text*.

Levin, B. (1993). *English Verb Classes and Alternations: A Preliminary Investigation.* University of Chicago Press.

Wiebe, J., Wilson, T., and Cardie, C. (2005). Annotating Expressions and Opinions and Emotions in Language. *Language Resources and Evaluation*, 39(2/3): 164-210.

Wiegand, M., and Klakow, D. (2010). Convolution Kernels for Opinion Holder Extraction. In *HLT/NAACL*.

Wiegand, M., and Klakow, D. (2011). The Role of Predicates in Opinion Holder Extraction. In *RANLP-Workshop on Information Extraction and Knowledge Acquisition*.

Wilson, T., Wiebe, J., and Hoffmann, P. (2005). Recognizing Contextual Polarity in Phrase-level Sentiment Analysis. In *HLT/EMNLP*.