



Institute of Informatics
Logics and Security Studies

ECO:Effective Communication Online

*Sharon Small, Tomek Strzalkowski and Nick Webb
December, 2010*

ILS Tech Report: 015

Available for download from:
<http://www.ils.albany.edu>

Abstract

The purpose of this research was to develop methods that could assess the effectiveness of communications destined for members of specific diacultural groups, as exemplified by their collective presence on Internet social media, specifically blogs. The research was conducted using data collected from Internet blogs representing opposing cultural groups, and containing (presumably) opposing value systems. Analysis of the collected data led to construction of a simple, stochastic model for representing these value systems. Subsequently, improvements were made using a machine learning approach, which enables us to capture community specific valuation models. These models facilitate comparison between new messages and the target community, reporting mismatches in valuations of concepts between the two. ECO technology represents an advance in methods in cross-cultural communication, where the content of new messages has to be carefully balanced with the value system of the target community, and we expect this technology to be useful for influence operations, advertising, and national security.

Introduction

We performed research to extract and model community specific value systems using Internet social media, specifically blogs, as source data. We define diacultural groups as communities that share certain values and beliefs, where these communities can be ethnic, religious or professional. From the writings of these communities in online forum, we can obtain Valuation System Vectors; a set of salient concepts, and the valuation of those concepts, as represented in those online materials. Once effectively represented, we can perform comparisons between community specific valuation systems and new messages, determining mismatches in valuations between the two, and highlighting these to the user, as a way of suggesting possible fixes to make the message better fit the target community. The example in Figure 1 shows text from a pro-Palestinian blog (on the left) that talks about an “illegal Israeli military blockade”, a negative valuation, and a pro-Israeli message (on the right) that talks about the “legitimacy of the blockade”, a different valuation of the same concepts. What is interesting in this example is that both message and target community talk about “ Hamas terrorists ” negatively, and perhaps this common valuation can be a start point for fruitful dialogue between these two, otherwise disparate parties.

MOTIVATION

Internet social media, e.g., blogs, discussion groups and wikis, are established means of shaping opinions within communities that share certain values and beliefs. These communities, also referred to as *diacultures*, can be defined along many dimensions, including ethnic, religious, or professional. Individuals from the outside wishing to reach into such communities (e.g., to gain support, to ease tensions, to change attitudes, etc.) often have difficulties communicating effectively if their own system of values does not align with that of their audience. A classic example is the perception of equal opportunity in the workplace; another example is the regard for violence or sacrifice held by radical religious or political groups, e.g., in the Middle East.

Our hypothesis is that in order to communicate effectively across diacultural boundaries the communications must not unduly violate the value system established within the target community. This is not to say that all debate is off limits; however, it still must be grounded in the community value system to be effective. An important part of this value system is a set of opinions or valuations applied to salient concepts shared and held by the members of the community. These may include broad notions such as equal opportunity or use of torture, as well as much narrower concepts such as the dress code for business functions. Valuations applied to these concepts could be on a discrete scale, or may be binary (e.g., right-wrong). A community may thus be defined by the set of salient concepts along with the values they assign to them. We may represent these mathematically as vectors in an N-dimensional concept space, which we shall call *Valuation System Vectors (VSV)*.

In order to achieve an effective communication, any message with significant content should take a form that aligns closely with the community valuation system. Whenever references are made (direct or implied) to the salient concepts recognized by this community, they should conform to the valuation system in place. To state this more formally: the message signature vector (MSV) must be similar to this community’s VSV.

The above condition is necessary but of course not sufficient, i.e., valid logical arguments must still be deployed in order for the communication to achieve desired objectives, such as changing someone’s opinion, etc. Nonetheless, even a perfectly logical argument may be quite ineffective when the MSV of the communication carrying this argument is not aligned with the target community’s VSV. Equally, being able to extract the target community’s VSV allows external agencies to rapidly tune their communications for the maximum impact, e.g., in reaching out to new groups, countering radicalization trends, or reversing anti-US sentiments, without undertaking costly and painstaking social fieldwork.

The purpose of our prototype ECO system is to verify the stated research hypotheses, and demonstrate utility of the new technology.

Data Collection

Our first task was to collect a sample collection of Internet blog data. Blogs are excellent representatives of community valuation systems, and we chose blogs that had significant online presence, with an active membership, and clear and distinct valuation systems. We initially chose blogs that dealt with hot button issues between Arab and Israeli perspectives, from blogs that are hosted external to the United States. We chose a small number of blog postings for initial analysis. We estimated that around 1000 valued concepts would give us an appropriate amount of data for constructing a VSV. We performed a density analysis on a small amount of trial data, and determined that around 50,000 words of blog material were required to obtain 1000 valued concepts. We required 1000 valued concepts for each community under consideration, and so we collected 49,000 words (from 84 blog posts) of pro-Israeli data and 52,000 words (from 55 blog posts) of pro-Palestinian data.

For the pro-Israeli side, the blogs used were:

- <http://israel.foreignpolicyblogs.com/>
- <http://samsonblinded.org/blog/>
- <http://www.blogsofzion.com/>

For the pro-Palestinian side, we used these blogs:

- <http://palsolidarity.org/>
- <http://palestineblogs.net/>

Data Annotation

We hired 2 annotators to mark up the key concepts in these texts, along with any valuations indicated in the text. We looked for four types of valuations; Positive and Negative, which are direct valuations made with clear language in the text, and Sympathetic and Unsympathetic, where there was no direct language, but the reader is left with a clear impression that the subject of the text is being valued in some way. For example, in the sentence:

“Israeli Armed Forces are denying food to Palestinians”

It can be judged that the ‘Israeli Armed Forces’ are doing something bad, and so are judged negatively, and by virtue of the fact that the ‘Palestinians’ are having something bad done to them, ‘Palestinians’ in this context are being treated sympathetically. The annotators used our ECO Annotation Tool (EAT) to identify concepts and valuations in the new blog material. All annotators were trained on a sample set of data. The tool is relatively simple, as annotators are only looking for the concepts mentioned, and the valuation assigned

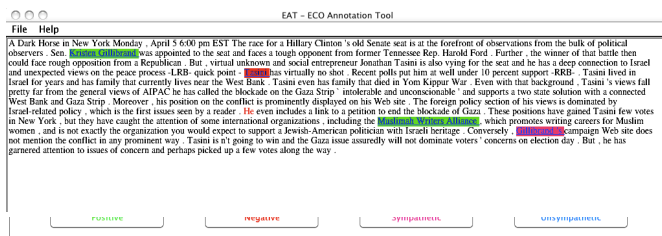


Figure 3. Screenshot of ECO Annotation Tool

In the screenshot above, note that the annotator highlights a concept under consideration, and then chooses from the buttons below to mark the utterance as positive, negative, sympathetic or unsympathetic. The tool also facilitates co-reference, allowing the annotator to select pronouns, and then highlight concepts to which these pronouns refer. However, this process was not fully explored in this round of experiments. At this time, we have not calculated inter-annotator agreement scores between our annotators. For initial purposes, we were interested in obtaining as many judgments, even if they didn't overlap, as the annotators saw in the source text.

ECO Methodology

We mine the blog sources associated with these groups and apply natural language processing algorithms to:

- Determine if we can identify references to salient concepts that occur in the postings: events, personalities, organizations, etc., then perform concept extraction using information extraction tools
- Extract value statements regarding these concepts from the data.
- Derive VSV proxies from representative communications generated by members of the selected community. Further, derive Message Signature Vectors from individual blog postings and commentaries.
- Compare MSVs to VSVs, and display the results to the user in an intuitive way.

We cast the ECO problem as a classification exercise, choosing which concepts should be given a positive or negative valuation, based on the presence of surrounding features. Given the manually annotated judgments from our annotators, we used four other features: each word of the blog; each words part-of-speech category (assigned by the Stanford POS tagger (REF); if the word was some type of named entity, as identified by AeroText™ software; and finally, if the word was present in the MPQA corpus of valuation words (Wiebe et al., 2005).

We decided to use Transformation-Based Learning (TBL) (Brill, 1995) as a mechanism to learn rules for identifying the valuations of our identified concepts, given the features indicated above. TBL is an error-correcting machine learning approach, where misclassifications in the training data are seen as errors, and rules are learnt (based on the presence of surrounding features) that correct these errors, resulting in the correct classification. We chose an implementation of TBL by Lager (1999), which is fast and efficient. The architecture of the overall ECO system is shown in Figure 4. Given our input features (4 automatically generated, and one, the judgments coming from our annotators), we encoded each word of the blog in TBL axiom style (performed by the axiom builder in Figure 4). For initial input data to TBL, all words are assigned the judgment category 'neutral'. At this stage, we collapsed our annotation categories, combining negative and unsympathetic, and positive and sympathetic.

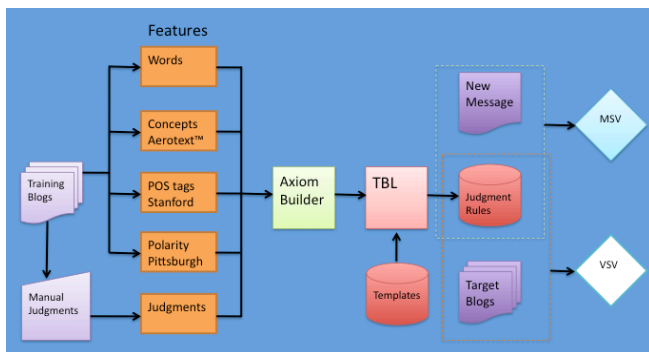


Figure 4. ECO Architecture and Components.

We also have to supply a series of templates to TBL, or proto-rules, which specify combinations of features that can be applied in the learning process. We initially selected a window size of 11 (that is, 5

words each side of the word under consideration) and allow for variation of each feature within that window. The result is a series of instantiated rules for each diacultural perspective that use the features to classify for the judgments. For example, when applied to pro-Palestinian data, the following one of the rules that is generated:

judgment:neutral>negative ← word:'Israel'@[0] & polarity:negative@[4,3,2,1]

This rule states that the judgement of the current word should be transformed from neutral (the default assignment for every word) to negative if the current lexical item is Israel, and there is a negative polarity word in a window of four words *following* the current word. This is a lexical specific rule. In future work, we will look to create more general rules relying further on ‘classes’ of concepts identified by Named Entity extraction mechanisms.

Evaluation

We conducted a small scale evaluation over a subset of our annotated data, splitting the data into 5 training and test data sets and performing a five-fold cross validation. We found that using TBL boosted our overall precision to around 90%, which means that applying the TBL-learned rules nearly always produces correct valuation of the relevant concept, although recall is significantly lower. Larger-scale experiments are needed to evaluate coverage, i.e., when TBL learns enough rules to cover an entire domain. Additionally, using combinations of features in our templates (such as looking for consistent interactions between judged concepts and words occurring as adjectives, as indicated by the POS tagger), we were able to ‘discover’ words acting as valuation words that did not appear in the Pittsburgh subjectivity lexicon. For example, we extracted words such as ‘*fractured*’, ‘*self-defeating*’, ‘*premature*’, and ‘*bereft*’, which all occur multiple times in conjunction with concepts that have been determined to be negatively judged concepts by our annotators, and yet these words do not appear in the subjectivity lexicon.

Create and Compare Vectors

In order to create our vectors, at this stage we simply include counts of the concepts that are automatically labelled as positive and negative by our TBL mechanism. For example, if ‘*Hamas*’ is labelled negatively in 10 instances in pro-Israeli blog material, we give this concept a score of 10 in our vector. At present, we count only those concepts that have a net score of at least 2 (so we do not consider concepts that have one positive and one negative judgment, for example, having a net score of zero).

Having constructed a system that is capable of extracting vectors for textual material, we construct VSVs for both pro-Palestinian and pro-Israeli data, and we compute the similarity scores (by comparing these vectors). Earlier (Figure 1) we showed an excerpt from a pro-Palestinian blog and an individual posting (message) taken from a pro-Israeli blog. As a proof of the ECO concept, in Figure 5 below, we show the VSV and MSV generated by our system for both the pro-Palestinian blog and the pro-Israeli message, respectively. The vectors capture where the message is incompatible with the blog. This is of course as we would expect it; nonetheless, if our intention were to actually post such a message, we would need to first adjust its valuation vector to match that of the target blog.

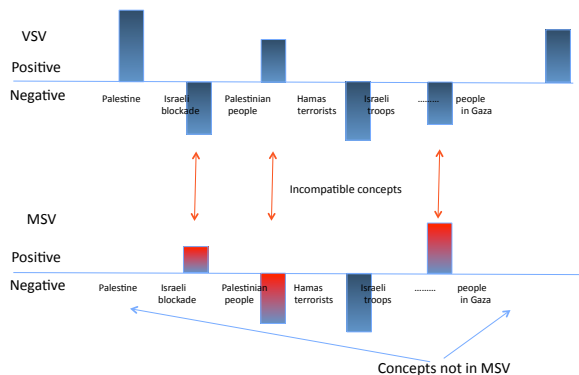


Figure 5. Comparing an MSV (bottom) from a pro-Israeli message to a VSV (top) from a pro-Palestinian blog.

ECO in Action

Figures 6, 7 and 8 represent examples of ECO in action over our collected data. In Figure 6, we have loaded a pro-Palestinian blog into our ECO tool, and we are showing a close up of some of the blog data, and the VSV we have created from this data. In Figure 7, we have loaded a pro-Israeli message and calculated the MSV (again, a section of the message is shown to demonstrate how the vector is created). Finally in Figure 8, we see the graph created when the VSV and the MSV from these examples is compared, giving the user visual feedback on the compatibility of their message to the target community.

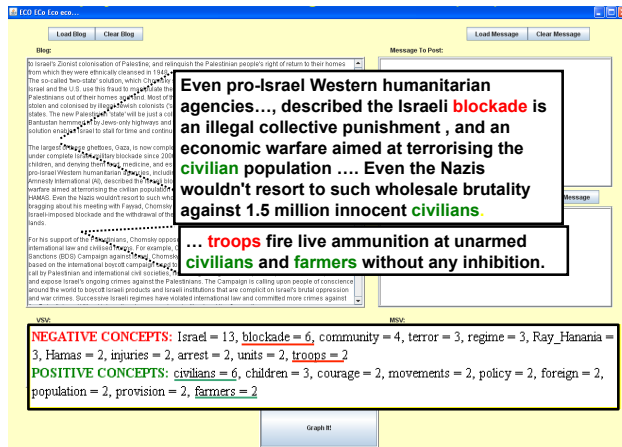


Figure 6. Generating a VSV from pro-Palestinian blog data

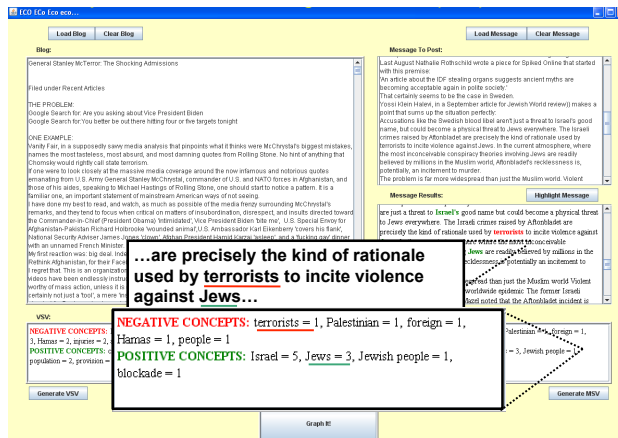


Figure 7. Generating a MSV from a pro-Israeli message

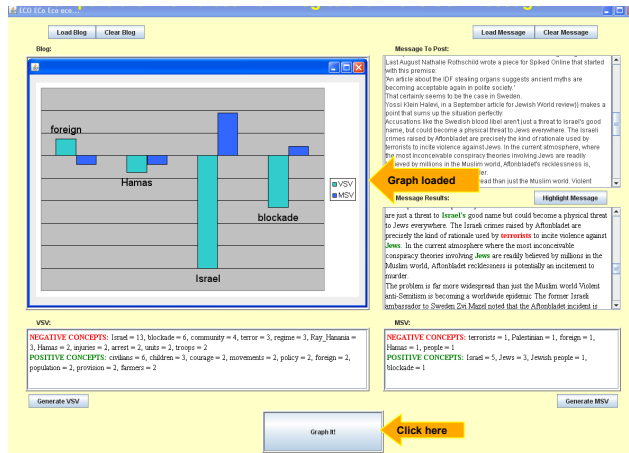


Figure 8. Comparing the MSV with the VSV

Future Work

We have developed a system that is able to compare a potential new message to a target community of blogs, and highlight for the user incompatibilities between the message and the valuation of concepts present in that community. We believe this is an interesting first step in a more complex system, where we can work collaboratively with the user to ‘tune’ their message so as to make it more acceptable to the target community. At present, we only identify those concepts that are viewed as net positive or negative. It may be interesting to look at those concepts that are judged as both negative and positive within a single community, and identify those concepts that are ‘vulnerable’, or are ripe for conversion from one perspective to another. Additionally, if we view concepts as an interrelated network, where the valuation of one concept depends in part on other concepts, we can plan a method to change the perspective on one concept by affecting those directly related to it. Such technology could have significant impact on influence operations, and be used to influence specific beliefs or sow doubt around individual concepts.

References

- Engelmore, R., and Morgan, A. eds. 1986. *Blackboard Systems*. Reading, Mass.: Addison-Wesley.
- Brill, E. (1995) Transformation-Based Error-Driven Learning and Natural Language Processing: A Case Study in Part of Speech Tagging. *Computational Linguistics*, December 1995.
- Lager, T. (1999). The μ -TBL System: Logic Programming Tools for Transformation-Based Learning, Paper presented at the *Third International Workshop on Computational Natural Language Learning (CoNLL'99)*, Bergen, 1999.
- Wiebe, J., T. Wilson and C. Cardie (2005). *Annotating expressions of opinions and emotions in language*. Language Resources and Eva