2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining

AttitudeBuzz: Using Social Media Data to Localize Complex Attitudes

Jason Cohn*, Alex Kuntz† and Larry Birnbaum‡
Northwestern University
Intelligent Information Lab
Ford EDC, 2133 Sheridan Rd., Evanston, IL USA
Email: *jasoncohn@u.northwestern.edu, †alexanderkuntz2014@u.northwestern.edu, ‡l-birnbaum@northwestern.edu

***Abstract*—AttitudeBuzz is a system that analyzes and presents complex social attitudes based on geolocated social media data. The system uses a machine learning model to apply highly domain-specific sentiment analysis to such data, specifically Twitter, by learning modulators around a configurable lexicon central to the domain of inquiry. Training data are acquired from geographical areas where a specific attitude or opinion is known to dominate. We apply AttitudeBuzz to the domain of homophobic attitudes expressed on Twitter. The resulting user interface is presented and the machine learning model described and analyzed.**

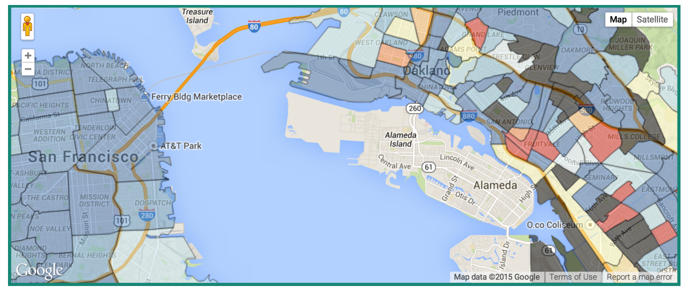

Fig. 1. AttitudeBuzz Interface: Heatmap portraying classifier results for neighborhoods in the San Francisco Bay area.

I. Introduction

AttitudeBuzz is an interactive platform for geographically quantifying and comparing highly subjective sentiments expressed through social media. With nominal modifications, the system is potentially extensible to a wide range of domains. In this paper, we showcase AttitudeBuzz operating within a single domain: the classification of hate speech targeting sexual preference minorities on Twitter.

We collect a corpus of tweets containing terms (e.g. gay, queer, dyke, etc.) used to both denote sexual identity and, in many cases, to disparage people who express those same sexual identities. The goal of AttitudeBuzz is to provide an intuitive user interface informed by a machine learning model sophisticated enough to capture such highly nuanced domains.

II. Method

Under the hood, AttitudeBuzz employs a variety of techniques to analyze sentiment in a complex domain of discourse. The use of geolocated social media data is central to AttitudeBuzz. Though access to geocoded data is limited (e.g., users must opt-in for their tweets to include geotags), previous intelligent systems in areas including crisis mapping have used similar data successfully [5].

A. Data Collection and Linguistic Reclamation

Selection of a minimally biased training set is of great concern to the attitudeBuzz system, which aims to classify highly subjective attitudes.

While instances of hate speech in our domain are often obvious, the task of identifying non-pejorative tweets involving terms such as queer presents a larger degree of difficulty and ambiguity. Often, such expressions are likely to be considered non-pejorative only when uttered by members of the expressly indicated sexual minority themselves. This phenomenon of linguistic reclamation (also called reappropriation) has been much studied by linguists and sociologists [1]. Reappropriation is a potential source of noise in our corpus, which appears more subjective in its non-pejorative examples as a result. Any se- rious model of this domain must account for the phenomenon.

Detecting the occurrence of reappropriation is itself a difficult undertaking. AttitudeBuzz approaches this problem through a combination of geolocated data and demographic inference. To construct a set of non-pejorative training examples, we search for tweets originating in any of a dozen urban neighborhoods considered to be historically gay friendly [8]. While we cannot ensure that all such tweets are truly non-pejorative, this technique gathers large amounts of training data in an automated, scalable manner.

The preponderance of highly insular LGBT neighborhoods within urban centers has been the subject of much sociological research [7]. Our training data labels any tweet both containing a sexual identity term and originating from a gay friendly neighborhood as non-pejorative. In this way, linguistic reclamation within a geographically localized community is used as a proxy for gay-friendly attitudes in social media more generally [7].

B. Process

We start by constructing a two-category training corpus. Using the Twitter API [6],tweets containing sexual identity

ASONAM '15, August 25-28, 2015, Paris, France
ACM ISBN 978-1-4503-3854-7/15/08
DOI: http://dx.doi.org/10.1145/2808797.280933S

terms and originating from historically gay-friendly neighborhoods are harvested and labeled as non-pejorative. A second set of tweets containing such terms, but of arbitrary geographic origin (within the United States) are also collected. Tweets deemed to be universally and objectively homophobic with little subjectivity are labeled as pejorative. Next, the training corpus is randomly permuted and split into training and test sets. A support vector machine (SVM) classifier is trained using the open-source LibLinear package [2]. The SVM is used to classify a corpus of geolocated tweets (about 600K in number) containing the sexual identity terms. Finally, the user interface is populated with labeled tweets as well as depicting heat maps representing .aggregate attitudes in urban areas for which shape files are available.

III. Analysis

The AttitudeBuzz homophobia model presented above contains a number of complexities not usually encountered in basic machine learning applications. Crucially, these include a high degree of subjectivity in both the training data and final classifications as well as our reliance on linguistic reclamation. These complexities may be unavoidable for a system such as AttitudeBuzz, which strives to represent the inherently nuanced character of subjective sentiment around a fraught topic, informed by user-generated social media data. Despite these challenges, AttitudeBuzz is shown to outper- form a standard lexicon-based sentiment model in correctly classifying gay-friendly neighborhoods.

A. Baseline Model

We employed a standard lexical model as a baseline to better understand AttitudeBuzz. This model consists of a lexicon of words with each item given a binary score designating whether it is commonly used to express positive or negative sentiment. The same tweets examined by AttitudeBuzz are tokenized and scored by the baseline.

B. Performance

Both models were assessed on a set of 26 gay-friendly neighborhoods, each containing at least 20 unique tweets. Gay-friendly neighborhoods used to train the AttitudeBuzz model were not included in this set. Additionally, a large number of urban neighborhoods were chosen randomly as a basis for comparison. The models score each neighborhood by the percentage of positive tweets as determined by their respective metrics.

In both models, we see approximately normal distributions in the rankings of the random neighborhoods. The AttitudeBuzz model, however, consistently assigns positive ratings to the gay-friendly neighborhoods. Comparatively, the baseline models ratings of gay-friendly neighborhoods follow those of randomly chosen neighborhoods. The difference in average rating assigned to gay-friendly neighborhoods by the two models, normalized by subtracting the mean of the random neighborhood distribution, was found to be statistically significant ($p = .0003, df = 40.523$). This difference in mean was observed to be in excess of 16%.

Apart from difference in average ratios of gay-friendly neighborhoods, we note the wider spread of the data in the baseline model, versus the relative consistency in the AttitudeBuzz Model.

IV. Conclusion

The AttitudeBuzz platform is a promising case study in the application of machine learning to real-world domains of social interest. We have seen how a complicated and domain-specific social phenomenon, linguistic reappropriation, can be accounted for and leveraged to assess and present the nuanced nature of attitudes towards complex social issues. These adaptations represent a step towards the application of of intelligent systems to better understanding and engaging with critical and complex social issues.

Acknowledgments

The authors would like to thank Miriam Boon for important discussions about geography and linguistic reclamation. We also acknowledge the efforts of *nohomophobes.com* [**?**] as one inspiration for our research. This work was supported in part by the John S. and James L. Knight Foundation, the National Science Foundation (under grant IIS-0917261), and Google.

References

[1] R. Brontsema, *A queer revolution: Reconceptualizing the debate over linguistic reclamation.*, Colorado Research in Linguistics 17, 1 (2004).

[2] R. Fan, E. Change, K.W. Hsieh, C.J. Wang, and C.J. Lin, *LIBLINEAR: A library for large linear classification.*, Journal of Machine Learning Research 9 (2008), 18711874.

[3] *Google Maps Javascript API*, https://developers.google.com/maps/documentation/javascript.

[4] *No Homophobes Project.*, http://www.nohomophobes.com/.

[5] S.E. Middleton, L. Middleton, and S. Modafferi, *Real-time crisis mapping of natural disasters using social media.*, IEEE Intelligent Systems 29, 2 (2014), 917.

[6] *Twitter Streaming API*, https://dev.twitter.com/streaming/.

[7] B.A.. Weightman, *Commentary: Towards a Geography of the Gay Community.*, Taylor & Francis Group, 2009.

[8] *Wikipedia. List of gay villages*, Wikipedia, the free encyclopedia, 2015. [Online; accessed 22-Nov-2015].