A Comparative Study of the Properties of Emotional and Non-Emotional Words in the Wordnet: A Complex Network Approach

Plaban Kr. Bhowmick, Animesh Mukherjee, Pabitra Mitra, Anupam Basu
Department of Computer Science and Engineering,
Indian Institute of Technology, Kharagpur, India – 721302
{plaban,animeshm,pabitra,anupam}@cse.iitkgp.ernet.in

Aritra Banik
Department of Computer Science and Engineering,
Indian Statistical Institute, Kolkata-700108
aritrabanik@gmail.com

Abstract

In this paper, we explore certain complex network properties of the Wordnet to devise a suitable technique that successfully distinguishes the emotional words from the non-emotional ones. We observe that the popular centrality based measures of complex networks are not appropriate for differentiating the emotional words from their non-emotional counterparts. Therefore, we propose a sophisticated transformation based on *structural similarity* followed by a clustering of the transformed network and show that this method results in the accurate segregation of the emotional and the non-emotional words. We plan to apply this method to induce emotional word base from Wordnet in future.

1 Introduction

Emotion recognition from text is a new subarea of Natural Language Processing (NLP) and has drawn considerable attention of the NLP researchers in recent times. People express emotion through different channels of communication: facial expressions, speech, bio-signals and language. Among these, facial expression is perhaps the most commonly used evidence for decoding emotion expressed by others. Nevertheless, as perceived by Barrett et al. (2007), emotion words help in quick and easy perceptions of emotion.

Sometimes the words appearing in the text bear emotional cues that help in determining the emotional category of the text segment. This provides the impetus to develop an emotional lexical resource for the task of emotion recognition. Building emotion word base can be either manual (Valitutti et al., 2004) or automatic. Acquisition of emotion words can be carried out by extracting the appropriate parts from resources like dictionaries, thesauruses and other comprehensive resources like the Wordnet.

Worndet is a huge lexicon base where the words are grouped into different lexical categories such as noun, verb, adjective and adverb. Information in Wordnet is represented in terms of *synsets*, where each synset comprises a set of synonymous words. The synsets are related to each other by means of certain semantic relationships: hypernymy (hyponymy), meronymy (holonymy) and antonymy. Essentially, Wordnet can be viewed as a very large graph with nodes representing the words and links representing the semantic relationships.

In order to extract emotion lexical base from Wordnet, the first step one needs to devise a strategy that can distinguish an emotional word from a non-emotional one. One way to achieve this would be to investigate the patterns of interactions among the words and extract certain statistical properties based on them, which in turn can successfully draw a distinction between the emotional and non-emotional words. The theories of *complex networks* presents us with a bunch of techniques to measure such statistical properties that can suitably be applied to bring about such a distinction. In fact, researchers have studied Wordnet in the framework of complex networks reported various interesting properties (see (Mariano and Guillermo, 2002) for instance) although not with the same objective as above.

In this paper, we explore certain complex network properties of the Wordnet to find out an appropriate technique that separates the emotional words from the non-emotional ones. We consider

Proceedings of ICON-2008: 6th International Conference on Natural Language Processing,
Macmillan Publishers, India. Also accessible from http://ltrc.iiit.ac.in/proceedings/ICON-2008

a set of *basic emotions* (Ekman et al., 1982) to construct emotional lexical network from them. For non emotional words, we construct two different networks

- where non-emotional seed words are selected manually and
- where the non-emotional seed words are chosen at random.

In construction of emotional and non-emotional networks, the network construction methodology starts with a set of seed or base words. A BFS traversal of the Wordnet with maximum depth of four hop distance from the seed words is performed to extract the emotional and non-emotional networks. We observe that the popular centrality based measures of complex networks are not suitable in distinguishing the emotional words from the non-emotional words. Therefore, a sophisticated transformation based on *structural similarity* (Hanneman, 2001) has been applied to the networks followed by a clustering of the transformed network which results in the segregation of the emotional and the non-emotional words.

In section 2, we provide a brief overview of some of the complex network measures relevant for this work. We provide the rationale behind selection of the base words in section 3. Lexical network construction methodology has been discussed in details in section 4. In section 5, we provide different complex network studies performed on the constructed networks and the linguistic analysis of the experimental results. In section 6, we discuss some relevant issues regarding the work.

2 Background

In this section, we shall review some of the complex network measures as they provide the basis of the methodology for distinguishing between the emotional and the non-emotional sections of the Wordnet.

2.1 Centrality Measures

The notion of centrality is essential to quantify how central a vertex is in a network. Some of the popular centrality measures are described below.

Degree Centrality (Lee, 2006) is the most basic of all the centrality measures and is simply the degree of the vertex in question.

Betweenness Centrality (Freeman, 1977) counts how many times a particular node comes out to be an intermediate node in the shortest paths between other vertices. Vertices that appear on many shortest paths between other vertices have higher betweenness value.

Closeness Centrality (Lee, 2006) is a measure of how close one vertex is with other vertices in the network. It is defined as the reciprocal of the sum of the geodesic distances between the vertex in question and all other vertices reachable from it.

2.2 Other Topological Properties

Apart from the centrality measures discussed earlier, two other measures are also important in complex network study:

Clustering Coefficient (Newman, 2003) for a vertex is given by the fraction of links between the vertices within its one distance neighborhood divided by the number of links that could possibly exist between these vertices.

Average Nearest Neighbor Degree (ANND) (Barrat et al., 2004) of a node in a network is defined as the average over the degrees of its one-distant neighboring nodes. If a high degree node has a high ANND, then the network is belongs to the *assortative* (Newman, 2002) class; otherwise the network is a *disassortative* one.

2.3 Structural Similarity Measures

Structural similarity (SS) (Bunke and Messmer, 1994; Basak and Niemi, 1988) is a measure of equivalence or two objects in the structural level. For instance, in a social network, two actors are said to be structurally similar if they play similar social roles. Graph-theoretically, two nodes are said to be structurally similar if they have similar neighbor profiles. As perfect structural equivalence is rare, graph theorists are often interested in the extent of structural equivalence.

Assuming that the graph is represented as an adjacency matrix, the row vector corresponding to a node defines its neighborhood. Structural similarity between two nodes is calculated as the extent of overlap between these neighborhoods, i.e., their row vectors. One way to compute the overlap is as follows. As the constructed networks, in this paper, are unweighted, therefore, the adjacency matrix for each of them is essentially a 0-1 matrix and the distance between the neighborhood patterns of two nodes is equivalent to the *Hamming distance* between their respective row-vectors. The lower this distance, the higher is the structural similarity.

3 Selection of Base Lexicons

Emotional words may be divided into two distinct categories (Kovecses, 2000): expressive words and descriptive words. Examples of expressive words include *wow!, shit!, yuk!* etc. *Anger, disgust, happy* etc. are some examples of the descriptive words. In the descriptive word category, some words are more *basic* than others in a sense that these words describe a particular emotion in a more intense manner than the others.

Basic emotions are those for which the respective expressions across culture, ethnicity, age, sex, social structure are invariant (Ortony and Turner, 1990). One of the theories behind the basic emotions is that they are biologically primitive because they possess evolutionary significance related to the basic needs for the survival of the species (Plutchik, 1980).

Following (Ekman et al., 1982), we have selected six basic emotion categories and the corresponding descriptive emotional words are *Anger, Disgust, Fear, Happiness, Sadness, Surprise* as specified by Ekman.

For the selection of the non-emotional words, we have considered two sets of words. First set has been selected manually so that the words belonging to the set are conceptually distant from the emotional words. The second set has been picked up randomly from Wordnet.

4 Lexical Network Construction Methodology

Two types of networks are constructed: one for emotional words and another for non-emotional words.

4.1 Emotional Network

The lexical network construction starts with an initial list consisting of six base emotional words. The considered words for this study belong to the noun category and therefore, the relations considered for network construction hold for nouns only, i.e., hypernymy/hyponyny, synonymy, meronymy/holonymy. With the initial list of nodes, we perform *Breadth First Search* (BFS) on Wordnet upto four levels. We have made no distinctions among the different senses of a word. Thus, different senses of a polysemous word are merged to form a single node for that word. This network will be referred to as N_E in the rest of the paper.

4.2 Non-Emotional Network

The network construction methodology for the non-emotional case is same as that of emotional one. The list of base words at the start of network construction differs from that of emotional network. The choice of base words in this case, has been done in two ways.

- In the first case, the base words are selected manually from the Wordnet. The selected words are *computer, window, debris, water* and *paper*. We call the network constructed from them MN_{NE}.

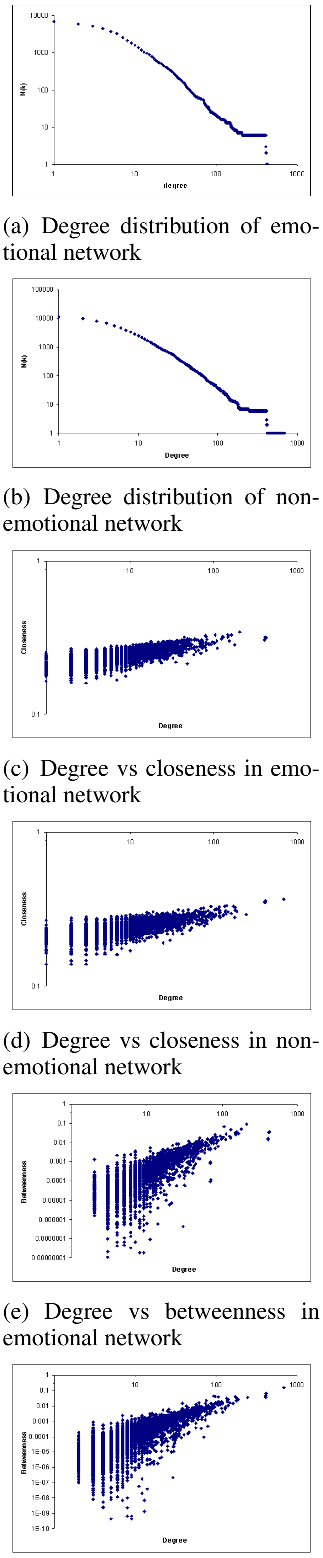
- In the second case, the base words are randomly selected from the Wordnet. The network constructed from these words is termed as RN_{NE}. Note that, the base words in this case are *arc_boutan, caudex, discus, window, water, knight_bac, hammer, microspora, micrococcu, edema, parapraxis, plectron, prosenceph, sanctum, stylus*.

In the next section, we provide the complex network study of the constructed networks.

5 Complex Network Study of Emotional and Non-Emotional Network

Having constructed the networks from the emotional and non-emotional base words, we next compute the various statistical properties outlined in section 2.1, for N_E and MN_{NE}. The objective is to find out if one or more of these properties can qualitatively differentiate N_E from MN_{NE} thereby, pointing to a general methodology for distinguishing emotional words from non-emotional words.

We provide the plots of degree distribution, closeness versus degree and betweenness versus degree in Figure 1. The data points have been plotted in log-log scale. The plot of clustering coefficient (CC) versus degree (log-log scale) has been provided in Figure 2(a) and Figure 2(b). The plots of ANND versus degree are provided in Figure 2(c) and Figure 2(d). The Pajek (Batagelj and Mrvar, 2002) software has been used to compute the different measures.



(a) Degree distribution of emotional network

(b) Degree distribution of non-emotional network

(c) Degree vs closeness in emotional network

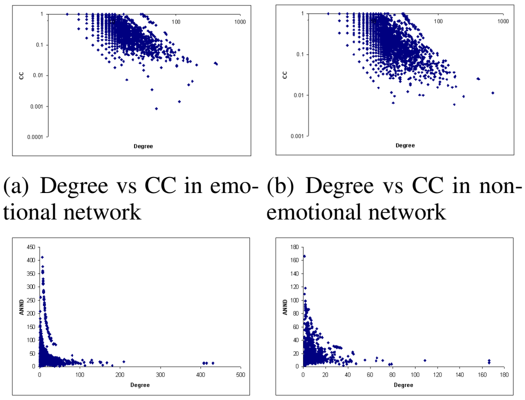
(d) Degree vs closeness in non-emotional network

(e) Degree vs betweenness in emotional network

(f) Degree vs betweenness in non-emotional network

Figure 1: Plots of centrality measures vs degree.

The general properties observed from the plots presented in Figure 1 and Figure 2 are summarized below.



(a) Degree vs CC in emotional network

(b) Degree vs CC in non-emotional network

(c) Degree vs ANND in emotional network

(d) Degree vs ANND in non-emotional network

Figure 2: Plots of CC and ANND versus degree.

- By applying curve-fitting technique on cumulative degree distributions (Figure 1(a) and Figure 1(b)), we observed that both the networks follow a power-law (Clauset et al., 2007) of the form

  $$N(k) \propto k^{-\alpha} \qquad (1)$$

  where $N(k)$ is the number nodes having degree higher than or equal to k. α is 1.55 for emotional network (N_E) and 1.92 for non-emotional network (MN_{NE}). It follows, that there are a very few nodes in the networks that have very high degree whereas the majority of the nodes have a very low degree. Note that, most of the high degree nodes belong to the set of polysemous words (this result is in agreement with those reported in (Mariano and Guillermo, 2002)).

- In both the networks, the variation of closeness centrality with degree is less as evident from Figure 1(c) and Figure 1(d). This indicates that all the nodes are equally close to each other irrespective of their degrees. This may be attributed to the fact that when the polysemous words are merged into a single node, the Wordnet networks exhibits a *small world* effect (Mariano and Guillermo, 2002) where the polysemous words act as short bridges in the networks thereby, bringing all other nodes close to each other.

- The betweenness centralities in both the networks are highly correlated with the degree as depicted in Figure 1(e) and Figure 1(d). The nodes with higher degrees possess higher betweenness values. It can be conjectured that the polysemous words are the most *between* nodes and they form the bridges in the network.

- In CC versus degree plots (Figure 2(a) and Figure 2(b)) of both the networks, it is observed that high degree nodes (i.e., the polysemous nodes) have low CC. Therefore, the neighbors of the polysemous words are possibly non-neighbors of each other.
- From Figure 2(c) and Figure 2(d), it is observed that nodes with high degree have low ANND values (i.e., show strong trends of being disassortative). This leads us to conjecture that two polysemous words are hardly neighbors of each other.

Finally, from Figures 1 and 2, similar qualitative trends are observed in emotional and non-emotional networks and therefore, these well-known techniques seem to be inappropriate in distinguishing between the emotional and non-emotional words. This is perhaps a consequence of the influence of the underlying universal structural properties of the Wordnet. Thus, one needs to suitably transform the above networks and analyze the properties of these transformations in order to bring about the distinction.

5.1 Structural Similarity Measure

The basic hypothesis underlying the experiments presented in this section is that emotional words play similar "social roles" in linguistic networks such as those constructed by us. Therefore, the structural similarity among the emotional words should be stronger than that between the emotional and the non-emotional words. In order to test this hypothesis, we perform two different experiments enumerated below.

5.1.1 Experiment 1: Manual Selection of Non-emotional Lexicon

The steps of this experiment are as follows.

I. Combine the networks N_E and MN_{NE} and call this new network N_{COM}.

II. Compute the structural similarity between every pair of the base emotional words in N_{COM} (i.e., anger-disgust, sadness-fear, and so on).

III. Similarly, compute the structural similarity between each base emotional and base non-emotional word in N_{COM} (i.e., disgust-debris, fear-window and so on).

IV. Construct a weighted network N_{SS} where the set of nodes comprise all the base words (emotional and non-emotional). The edge weight between two nodes (either both emotional or one emotional and the other non-emotional) is their corresponding structural similarity.

V. Perform a hierarchical agglomerative clustering of N_{SS}.

The result of the clustering of N_{SS} in Figure 3 clearly indicates that the above process can successfully distinguish between the emotional and the non-emotional words.

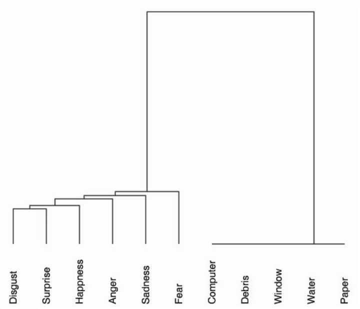


Figure 3: Hierarchical clustering of N_{SS} in manual selection experiment.

5.1.2 Experiment 2: Random Selection of Non-emotional Lexicon

The steps of the experiment are as follows.

I. Combine the networks N_E and RN_{NE} and call this new network N_{COM}.

II. Repeat steps II to V outlined in Experiment 1.

Here again, we observe that the method successfully distinguishes the emotional from the randomly selected words (see Figure 4).

Results of both the experiments indicate that even quite distant emotional words join together before they join the non-emotional words. This transformation based on structural similarity should therefore allow us to detect other emotional words that are close to the seed words in terms of their semantic content. Furthermore, it should also

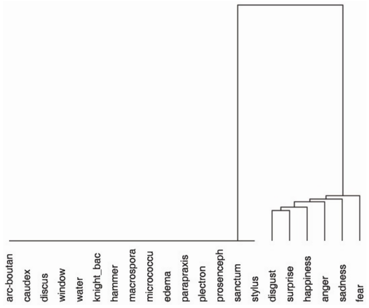


Figure 4: Hierarchical clustering of N_{SS} in random selection experiment.

be possible to develop a ranking of these words depending on how much structurally similar they are to the seed words. Note that this method would allow us to rank all the words qualitatively with respect to each of the seed words. This is useful because it is often the case that words can belong to multiple emotion categories and need to be ranked in each of these categories separately.

6 Conclusions

In this paper, we have presented a methodology that can successfully differentiate the emotional words from their non-emotional counterparts using sophisticated methods of network transformation. More specifically, we have shown that whereas, the popular centrality based measures of complex networks fail to bring forth this difference, structural similarity based techniques neatly performs the same. We therefore, conjecture that such similarity based techniques can be, in general, employed to induce emotional words from large lexical bases like the Wordnet. We plan to conduct the same in future.

Acknowledgement

P. K. Bhowmick is partially supported by Microsoft Corporation, USA and Media Lab Asia, India. A. Mukherjee is thankful to Microsoft Research India for financial support. Authors are thankful to Dr. Monojit Choudhury for his detailed suggestions regarding the work.

References

Alain Barrat, Marc Barthelemy, Romualdo Pastor-Satorras, and Alessandro Vespignani. 2004. The architecture of complex weighted networks. *PNAS, USA*, 101:3747.

Lisa F. Barrett, Kristen A. Lindquist, and Maria Gendron. 2007. Language as context for the perception of emotion. *Trends in Cognitive Sciences*, 11(8):327–332.

Subhash C. Basak and G. J. Niemi. 1988. Determining structural similarity of chemicals using graph-theoretic indices. *Discrete Appl. Math.*, 19(1-3):17–44.

Vladimir Batagelj and Andrej Mrvar, 2002. *Pajek - Analysis and Visualization of Large Networks*, volume 2265. January.

Horst Bunke and Bruno T. Messmer. 1994. Similarity measures for structured representations. In *EWCBR '93: Selected papers from the First European Workshop on Topics in Case-Based Reasoning*, pages 106–118, London, UK. Springer-Verlag.

Aaron Clauset, Cosma Rohilla Shalizi, and M. E. J. Newman. 2007. Power-law distributions in empirical data. *arXiv*, arXiv:0706.1062v1.

Paul Ekman, W. V. Friesen, and P. Ellsworth. 1982. What emotion categories or dimensions can observers judge from facial behavior? In P. Ekman, editor, *Emotion in the human face*, pages 39–55. Cambridge University Press, New York.

Linton C. Freeman. 1977. A set of measures of centrality based on betweenness. *Sociometry*, 40(1):35–41, March.

Robert A. Hanneman. 2001. Introduction to social network methods. *University of California*.

Zoltan Kovecses. 2000. *Metaphor and Emotion: Language, Culture, and Body in Human Feeling*. Cambridge University Press, Cambridge.

Chang-Yong Lee. 2006. Correlations among centrality measures in complex networks.

Sigman Mariano and Cecchi A. Guillermo. 2002. Global organization of the wordnet lexicon. *Proceedings of the National Academy of Sciences*, 99(3):1742–1747.

M. E. J. Newman. 2002. Assortative mixing in networks. *Physical Review Letters*, 89(20).

M. E. J. Newman. 2003. The structure and function of complex networks. *SIAM Review*, 45:167.

Andrew Ortony and Terence J. Turner. 1990. What's basic about basic emotions? *Psychological Review*, 97(3):315–331.

Robert Plutchik. 1980. A general psychoevolutionary theory of emotion. *Emotion: Theory, Research, and Experience*, 1:3–33.

Alessandro Valitutti, Carlo Strapparava, and Oliviero Stock. 2004. Developing affective lexical resources. *PsychNology Journal*, 2(1):61–83.