Novel Relationship Discovery Using Opinions Mined from the Web

Lun-Wei Ku, Hsiu-Wei Ho and Hsin-Hsi Chen

Department of Computer Science and Information Engineering
National Taiwan University
Taipei, Taiwan
{lwku, xwhe}@nlg.csie.ntu.edu.tw; hhchen@csie.ntu.edu.tw

Abstract

This paper proposes relationship discovery models using opinions mined from the Web instead of only conventional collocations. Web opinion mining extracts subjective information from the Web for specific targets, summarizes the polarity and the degree of the information, and tracks the development over time. Targets which gain similar opinionated tendencies within a period of time may be correlated. This paper detects event bursts from the tracking plots of opinions, and decides the strength of the relationship using the coverage of the plots. Companies are selected as the experimental targets. A total of 1,282,050 economics-related documents are collected from 93 Web sources between August 2003 and May 2005 for experiments. Models that discover relations are then proposed and compared on the basis of their performance. There are three types of models, collocation-based, opinion-based, and integration models, and respectively, four, two and two variants of each type. For evaluation, company pairs which demonstrate similar oscillation of stock prices are considered correlated and are selected as the gold standard. The results show that collocation-based models and opinion-based models are complementary, and the integration models perform the best. The top 25, 50 and 100 answers discovered by the best integration model achieve precision rates of 1, 0.92 and 0.79, respectively.

Introduction

Opinion extraction identifying subjective information from designated sources is fundamental for summarization, tracking, etc. (Ku, Li, Wu and Chen, 2005). Much work is done on this topic. Pang et al. (2002) recognized opinionated documents. Riloff and Wiebe (2003) distinguished subjective sentences from objective ones. Kim and Hovy (2004) proposed a sentiment classifier for English words and sentences. These works used closed sets of documents. Compared to these, the Web provides huge heterogeneous information for opinion extraction. Dave's (2003), Hu's (2004) and Morinaga's (2002) research focused on extracting opinions of product reviews. Bai, Padman and Airoldi (2005) categorized movie reviews by opinion polarities. While both utilized information from the Web, the extracted opinions were only on a single target. The comparison of opinions towards multiple targets is not mentioned.

Summarization is a good way to provide an overview of public opinions. Hu and Liu (2004) proposed an opinion summarization about consumer products. Liu, Hu and Cheng (2005) then illustrated an opinion summarization using bar graphs. Wiebe *et al*. (2002) proposed a method for opinion summarization by analyzing the relationships among basic opinionated units within a document. In fact, there are many events embedded within opinions, thus event burst detection is indispensable. However, the above research only retrieves relevant documents for summarization, and event-based summarization is not touched on.

Relationship discovery aims to explore relations among multiple targets. Collocation (Manning and Schutze, 1999) has been employed to discover relationships among terms based on their co-occurrences in a physical context, such as documents, sentences, etc. The results of opinion tracking on multiple targets provide another kind of information for relationship discovery. If the targets involved in the same sequence of events gain similar opinionated tendencies, these targets may be correlated. Such ideas have not been previously explored.

In this paper, an event-based opinion summarization is proposed. Companies are selected as targets. Collocations and opinions are used for relationship discovery. For opinion-based models, original curves, digitized curves and smoothed curves of tracking plots are generated to test their effects on relationship discovery. For collocation-based models, collocations at word, sentence and document levels are extracted to discuss the impact of how closely they are collocated. A total of eight models, including collocation-based, opinion-based and integration models, are proposed and compared. Finally, the possibility of predicting with opinions the short-term behavior of a target is examined.

Experiment Materials

Total 1,078 companies in Taiwan are the target candidates for discovering relationship. There are two kinds of experimental materials related to companies: one is a corpus of documents for mining opinions, and the other is the stock statistics for extracting answer keys.

Corpus Description

Total 1,282,050 economics-related documents are collected automatically from 93 Web sources of between August 2003 and May 2005 for model training. For opinion-based models, documents relevant to the listed companies first are retrieved by an Okapi IR system. Because the analyses of opinions need sufficient information, we count the number of relevant documents for each company and select, as the targets for relationship discovery, the top 250 from the listed companies. Here a relevant document for a target means a document mentioning the target. In reality, mention does not always indicate relevance. Therefore, five relevant document sets for each company, are prepared for experiments, i.e., top 2,000, 5,000, 8,000, 10,000, and All. On average, there are 10,441 relevant documents for each company. In contrast, for collocation-based models, all documents are used to count the co-occurrences.

Gold Standard Acquisition

Generally speaking, investors want to know the relationship among designated companies in the stock market. That makes relationship discovery a practical application. To match the phenomena in the real world, the gold standards are mined from the stock prices from August 2003 to May 2005, i.e., the same as the period within which the corpus is collected. We postulate if two stocks' prices are correlated, the companies have other relationships. Chi-square method is adopted to find such pairs from a total of 31,125 (C_2^{250}) company pairs. The change of one stock price is defined in Formula (1). The Taiwan large cap stock index is used as the basis for comparison because it reflects stock market trends of large companies; hence, the change of large-cap price is compared to a company's stock to decide the stock's ups (↑) and downs (↓).

$$d_i = \frac{p_i - p_{i-1}}{p_{i-1}} - \frac{q_i - q_{i-1}}{q_{i-1}} \quad (1)$$

where p_i (p_{i-1}) is the price of the stock and q_i (q_{i-1}) is the price of the large-cap at day i (i-1), and d_i is the percent of the difference of the stock prices. If d_i is positive, an "up (↑)" appears in day i; if d_i is negative, an "down (↓)" appears in day i.

$$\chi^2 = \sum_i \sum_j \frac{(f_{ij}^o - f_{ij}^e)^2}{f_{ij}^e} \quad (2)$$

The chi-square formula is defined in (2), where f_{ij}^o is the exact observed value of f_{ij}, and f_{ij}^e is the expect value of f_{ij}. Considering the ups (↑) and downs (↓) of prices of stocks of two companies A and B, a chi-square contingency table with one degree of freedom is generated as in Table 1.

		Stock Price of Company A	
		↑	↓
Stock Price of Company B	↑	f_{11}	f_{12}
	↓	f_{21}	f_{22}

Table 1. Chi-square contingency table, one degree of freedom

In addition, when no change in price is considered, a chi-square table with four degrees of freedom is generated as in Table 2.

		Stock Price of Company A		
		↑	-	↓
Stock Price of Company B	↑	f_{11}	f_{12}	f_{13}
	-	f_{21}	f_{22}	f_{23}
	↓	f_{31}	f_{32}	f_{33}

Table 2. Chi-square contingency table of freedom 4

In Tables 1 and 2, all the f_{ij} indicate the number of days the stock changes in a particular manner. With different degree of freedom and significance level, correlated company pairs are extracted as the gold standard. Table 3 shows the number of pairs in the gold standard in different conditions.

Degree of freedom	$\chi^2_{.950}$	$\chi^2_{.990}$	$\chi^2_{.995}$
1	7,815	4,239	2,008
4	2,489	1,366	703

Table 3. Numbers of company pairs

In this case, specific company pairs with strong relationship are more informative than many company pairs with weak relationship. Because precision is more important than recall here, the strictest condition is adopted to generate the gold standard. A total of 703 pairs are selected as the gold standard under $\chi^2_{.995}$ (significance level of 0.005) and four degrees of freedom.

Opinion Tracking, Burst Detection, and Event-based Opinion Summarization

Opinion tracking tells how people change their opinions over time. Tracking opinions about a single target is fundamental to know the variability of the reputation of the target. Calculating the overall opinion scores for a specific target every day generates a tracking plot (Ku, Liang and Chen, 2006). We call a day *positive* or *negative* for a target according to the opinion tendency. Here positive and negative days are tracked separately to detect positive and negative events. The grey curve of Figure 1 illustrates the tracking plot of a company (TSMC) in positive days.

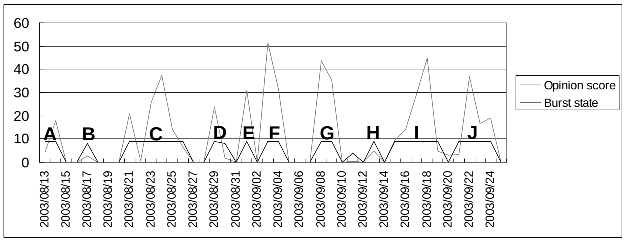

Figure 1. Company Plot (TSMC, Positive)

To detect the proper period of events, an approach of burst detection approach (Kleinberg, 2002) is adopted. It models a stream of data with a state automaton, and bursts appear as state transitions. The black curve of Figure 1 shows the resulting plot of burst detection. In this way, opinion summaries can be generated based on the documents within the same detected period of a burst event, and the events embedded in opinions are identified. In Figure 1, symbols A-J denote the events detected. The tracking plot in Figure 1, the simple opinion score, is not used because the durations of positive periods can be vague, see I and J; hence the event burst detection approach is preferable for event identification. Take event A as an example. The duration of A is from 13^{th} to 15^{th} of August 2003. The brief summary generated is "The expect income of TSMC will increase in the fourth season."

Next, the concepts of events and opinions are similar to those of causes and consequences. If targets (companies) have relationships, an event (cause) will have similar consequences for them and can be identified from opinions. This effect is illustrated by Figure 2.

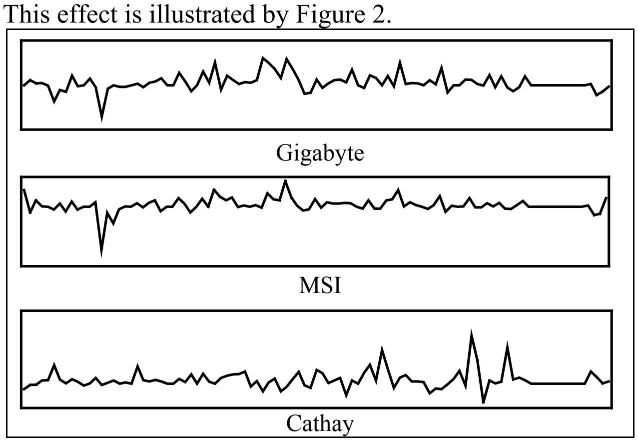

Figure 2. Opinion Tracking

The plots in Figure 2 track three companies, and the plots of Gigabyte and MSI are more similar than those of MSI/Cathay or Gigabyte/Cathay. In reality, Gigabyte and MSI are companies producing motherboards, while Cathay is a financial holdings company. Therefore, if tracking plots of opinions for two companies are similar, we can postulate that they are more closely related than those with different plots.

Relationship Discovery

Relationship discovery tells whether there is a certain relationship between targets. Targets can be any kind of objects, e.g., persons, companies, products, etc. With economics-related documents extracted from the Web, companies are selected as targets for relationship discovery in this paper. Two collocation-based models and four opinion-based models are proposed.

Collocation-based Models

Collocation-based models discover the relationship of two objects based on their co-occurrences in a context. Many statistical methods are proposed. Mutual information (MI) and t-test are selected as the collocation-based models in this study. Below, formula (3) defines mutual information and (4) defines t-test.

$$I(A,B) = \log_2 \frac{P(A,B)}{P(A)P(B)} \tag{3}$$

where $P(A,B)$ is co-occurrence probability of two companies A and B, and $P(A)$ and $P(B)$ are occurrence probabilities of A and B. The higher the score $I(A,B)$ is, the greater the relationship.

$$t = \frac{\bar{x} - \mu}{\sqrt{\frac{s^2}{N}}} \tag{4}$$

where $\bar{x}$ is the sample mean; s^2 is the sample variance; N is the sample size; μ is the mean of distribution. The confidence level of t-test is 0.005 and the value of t is 2.576. Relationships exist when the t-test is passed.

The degree of collocation is separated into three levels in both models: document level, sentence level and word level. Collocation at document level counts the number of documents in which two companies co-occur. Similarly, collocation at sentence level counts the number of sentences in which two companies co-occur. How frequent two companies are neighbor to each other in documents defines collocation at word level.

Opinion-based Models

Opinion-based models discover the relationship of two objects based on the similarity of their tracking plots. The strength of relationship is in terms of an overlap ratio of two plots. Curve overlap (CO), digitalized curve overlap (DCO), and curve overlap with burst detection (BDCO) are proposed.

Curve Overlap (CO):

$$CO(A,B) = \frac{\sum_{i=1}^{n} \left(\operatorname{sgn}(R_i \cdot S_i) \cdot \frac{\min(|R_i|,|S_i|)}{\max(|R_i|,|S_i|)} \right)}{n} \tag{5}$$

where R_i and S_i are opinion scores of companies A and B in a specific day i, respectively, and n is the number of days in the tracking period. This method emphasizes opinion scores.

Digitalized Curve Overlap (DCO):

$$DCO(A,B)=\frac{\sum_{i=1}^{n}\left(\operatorname{sgn}(R_i\cdot S_i)\cdot\frac{\min(\operatorname{sgn}(|R_i|),\operatorname{sgn}(|S_i|))}{\max(\operatorname{sgn}(|R_i|),\operatorname{sgn}(|S_i|))}\right)}{n} \quad (6)$$

Only the sign of the opinion score is used to calculate the overlap of two curves in DCO. That is, only polarities of opinions are considered. The degree of opinions has no effect in this model.

Curve Overlap with Burst Detection (BDCO):

First, $BD(X,t,i)$ is defined as the state of burst detection at day *i* considering the tracking plot of target *X*. Since positive opinions and negative opinions are processed separately in burst detection, variable *t* identifies the tendency of the analyzed plot. If *t* equals to 1, function *BD* returns states from the positive tracking plot; if *t* equals to -1, function *BD* returns states from the negative tracking plot. States returned are utilized to calculate the curve overlap in BDCO. This method flattens out extremes.

$$\begin{aligned}
R_{BDi} &= k\cdot\max(BD(A,1,i),BD(A,-1,i))\\
S_{BDi} &= k\cdot\max(BD(B,1,i),BD(B,-1,i))\\
k &= \begin{cases}1 & \text{for } \max(BD(X,1,i),BD(X,-1,i))=BD(X,1,i)\\ -1 & \text{for } \max(BD(X,1,i),BD(X,-1,i))=BD(X,-1,i)\end{cases}
\end{aligned} \quad (7)$$

$$BDCO(A,B)=\frac{\sum_{i=1}^{n}\left(\operatorname{sgn}(R_{BDi}\cdot S_{BDi})\cdot\frac{\min(|R_{BDi}|,|S_{BDi}|)}{\max(|R_{BDi}|,|S_{BDi}|)}\right)}{n} \quad (8)$$

The plot of burst detection states is a smoothed curve of the tracking plot (see Figure 1). In this model, relationship is discovered from a macro view of plots.

Chi-square:

Chi-square is adopted as the fourth opinion-based mining mehtod. Daily opinion scores are extracted from the results of opinion tracking. Compared to Table 1, signs of scores instead of ups and downs of stock prices are used for exploring relationship of targets. Opinion score 0 means there is no relevant documents on that day, and a chi-square contingency table of one degree of freedom is used.

Experiments and Evaluation

Relevance is important in relationship discovery. Using relevant documents better reflects the actual comparative performance of both collocation-based and opinion-based models. However, many companies adopt good terms, such as "happy", "lucky", "peace", etc, in their names. Mining with these kinds of names may retrieve irrelevant documents. To examine the influence of relevance on relationship discovery, the experiments are conducted with and without companies whose names contain general terms.

MI and t-test

Tables 4 and 5 show the results of MI and t-test. Experimental results show that MI performs better than t-test, and in Table 4 at sentence level MI achieves precision rates of 0.96, 0.94, and 0.74 when proposing the top 25, 50, and 100 company pairs, respectively. Tables 6 and 7 show the results of MI and t-test when the companies whose names are general terms are filtered out. There is a slight improvement, and in Table 6 the best precision rates of MI rise up to 0.96, 0.96 and 0.75.

MI	Document Level			Sentence Level			Word Level		
N	P	R	F	P	R	F	P	R	F
25	0.240	0.009	0.017	0.960	0.034	0.066	0.480	0.017	0.033
50	0.260	0.019	0.035	0.940	0.067	0.125	0.460	0.033	0.061
100	0.250	0.036	0.062	0.740	0.105	0.184	0.440	0.063	0.110
200	0.235	0.067	0.104	0.540	0.154	0.239	0.400	0.114	0.177
500	0.184	0.131	0.153	0.322	0.229	0.268	0.298	0.212	0.248

Table 4. Performance of MI with general names

t-test	Document Level			Sentence Level			Word Level		
N	P	R	F	P	R	F	P	R	F
25	0.440	0.016	0.030	0.480	0.017	0.033	0.600	0.021	0.041
50	0.260	0.019	0.035	0.460	0.033	0.061	0.460	0.033	0.061
100	0.240	0.034	0.060	0.440	0.063	0.110	0.410	0.058	0.102
200	0.210	0.060	0.093	0.335	0.095	0.148	0.373	0.080	0.131
500	0.148	0.105	0.123	0.254	0.181	0.211	0.340	0.098	0.151

Table 5. Performance of t-test with general names

MI	Document Level			Sentence Level			Word Level		
N	P	R	F	P	R	F	P	R	F
25	0.280	0.012	0.022	0.960	0.040	0.076	0.520	0.021	0.041
50	0.280	0.023	0.043	0.960	0.079	0.146	0.560	0.046	0.085
100	0.250	0.041	0.071	0.750	0.124	0.212	0.490	0.081	0.139
200	0.240	0.079	0.119	0.535	0.176	0.265	0.435	0.143	0.216
500	0.184	0.152	0.166	0.320	0.264	0.289	0.316	0.260	0.286

Table 6. Performance of MI without general names

t-test	Document Level			Sentence Level			Word Level		
N	P	R	F	P	R	F	P	R	F
25	0.440	0.018	0.035	0.480	0.020	0.038	0.640	0.026	0.051
50	0.260	0.021	0.040	0.440	0.036	0.067	0.540	0.045	0.082
100	0.260	0.043	0.074	0.430	0.071	0.122	0.470	0.077	0.133
200	0.230	0.076	0.114	0.335	0.110	0.166	0.413	0.102	0.164
500	0.156	0.129	0.141	0.258	0.213	0.233	0.420	0.138	0.208

Table 7. Performance of t-test without general names

	Top 2000			Top 5000			Top 8000			Top 10000			All		
	P	R	F	P	R	F	P	R	F	P	R	F	P	R	F
CO	0.4724	0.0782	0.1172	0.5522	0.0919	0.1374	0.6026	0.1018	0.1524	0.6064	0.1041	0.1551	0.5096	0.0913	0.1351
DCO	0.2233	0.0341	0.0551	0.3340	0.0505	0.0818	0.3953	0.0551	0.0899	0.3630	0.0533	0.0864	0.3037	0.0491	0.0789
BDCO	0.0900	0.0124	0.0203	0.0493	0.0057	0.0094	0.0970	0.0142	0.0231	0.0923	0.0142	0.0229	0.1073	0.0178	0.0286
χ^2	0.5503	0.0690	0.1135	0.5843	0.0718	0.1183	0.6213	0.0775	0.1275	0.6253	0.0789	0.1296	0.5997	0.0772	0.1264

Table 8. Average performance of opinion-based models using different quantities of relevant documents

Top 8000	CO			DCO			BDCO			χ^2		
N	P	R	F	P	R	F	P	R	F	P	R	F
25	0.8400	0.0299	0.0577	0.4400	0.0156	0.0302	0.1200	0.0042	0.0082	0.9600	0.0341	0.0659
50	0.7800	0.0555	0.1036	0.5400	0.0384	0.0717	0.1000	0.0071	0.0132	0.7000	0.0498	0.0930
100	0.6400	0.0910	0.1594	0.4200	0.0597	0.1046	0.1100	0.0156	0.0273	0.5300	0.0754	0.1320
200	0.4750	0.1351	0.2104	0.2950	0.0839	0.1307	0.0850	0.0241	0.0376	0.4000	0.1138	0.1772
500	0.2780	0.1977	0.2311	0.2100	0.1494	0.1746	0.0500	0.0355	0.0415	0.2320	0.1650	0.1929

Table 9. Performance of opinion-based models in relationship discovery with general names

Top 8000	CO			DCO			BDCO			χ^2		
N	P	R	F	P	R	F	P	R	F	P	R	F
25	0.9200	0.0379	0.0728	0.6400	0.02635	0.05063	0.6800	0.0280	0.0538	0.9600	0.0395	0.0759
50	0.8600	0.0708	0.1309	0.6600	0.05436	0.10045	0.6400	0.0527	0.0974	0.7000	0.0577	0.1065
100	0.6800	0.1120	0.1924	0.4600	0.07578	0.13012	0.4600	0.0758	0.1301	0.5500	0.0906	0.1556
200	0.4800	0.1582	0.2379	0.3300	0.10873	0.16356	0.3450	0.1137	0.1710	0.4000	0.1318	0.1983
500	0.2760	0.2273	0.2493	0.2160	0.17792	0.19512	0.2240	0.1845	0.2023	0.2360	0.1944	0.2132

Table 10. Performance of opinion-based models in relationship discovery without general names

CO, DCO, BDCO, and χ^2

Table 8 shows the average performance of opinion-based models using different quantities of relevant documents. Precision, recall, and f-measure P, R, F are evaluation measures. Top 2000, 5000, 8000, 10000, and All relevant documents of two companies are retrieved for relationship discovery. Retrieving top 8000 relevant documents is the best strategy for CO, DCO, BDCO, and χ^2. From Table 8, insufficient relevant documents (Top 2000, 5000) or noises (Top 10000, All) worsen the performance. Table 9 shows the comparison of the four opinion-based models using top 8000 relevant documents, proposing a different number of company pairs. Table 10 shows the same comparison with general names filtered out. In opinion-based models, the effect of filtering out general terms improves performance more than in collocation-based models. Both selecting the proper number of documents and filtering out general terms expel many non-relevant documents. These results tell that except for chi-square, the opinion-based models are more sensitive to the degree of relevance than the collocation-based models. The chi-square model focuses more on the distribution of ups and downs than the curve shape. Therefore, the chi-square model is less sensitive to the difference of curves than the other three models.

From Table 10, CO is the best model. It achieves, respectively, the precision rate of 0.92, 0.86 and 0.68 when top 25, 50, and 100 company pairs are proposed. CO performs better than DCO, which utilizes the digitized tracking plot, and than BDCO, which discovers relationship from the smoothened tracking plot. We can conclude that the weights of opinions (CO vs. DCO) and the changes in a short period (CO vs. BDCO) are both important clues for relationship discovery. χ^2 achieves precision comparable with CO. However, the precision rate drops fast when more company pairs are proposed.

Proposed company pairs of two types, MI in collocation models and CO in opinion-based models, are examed. Table 11 shows the intersection and difference of the two answer sets. Only about half of the proposed company pairs of CO and MI are in their intersection, and this quantity decreases when more company pairs are proposed. This result tells that MI and CO proposes different company pairs. Table 12 further shows the ranks of company pairs in the set difference.

Top N	MI∩CO	MI-CO	CO-MI
25	16	8	7
50	27	21	16
100	43	32	25
200	67	40	29
500	103	57	35

Table 11. Intersection and difference

Top N	Ranks of CO-MI in MI	Ranks of MI-CO in CO
25	614.43	180.25
50	806.75	439.29
100	1305.12	722.03
200	1487.86	1085.88
500	2487.69	3663.12

Table 12. Average rank of CO-MI and MI-CO

All company pairs of CO-MI and MI-CO are checked, respectively, to see which ranks they are in the company pairs proposed by MI and CO. If the ranks of company pairs found by one model are low in the other model, then the other model may not find the answers found by this model. Table 12 shows that the average ranks of CO-MI in MI tend to be lower than those of MI-CO in CO. In other words, CO can find company pairs that do not co-occur so often as those found by MI.

Two integration models are proposed to test whether considering both opinions and collocations help in relationship discovery. Model CO+MI considers scores of both CO and MI. Model CO ∩ MI considers those company pairs in both answer sets proposed by CO and MI. The formula of model CO+MI is defined as follows.

CO+MI:

$$CO + MI(A,B) = \alpha \frac{MI(A,B)}{\rho} + \beta \frac{CO(A,B)}{\upsilon} \quad (9)$$

where α=0.5, β=0.5, ρ and υ are normalization constants.

CO∩MI:

CO∩MI checks MI and CO answers in a round robin way to select common proposed candidates. In summary, CO+MI integrates two types of information by scores, while CO∩MI integrates by rankings.

Compared with the collocation-only and the opinion-only models, both integration models perform better. The overall performances of the eight models in Figure 3 show that CO+MI is the best. Top 25, 50 and 100 answers achieve precision rates of 1, 0.92 and 0.79, respectively.

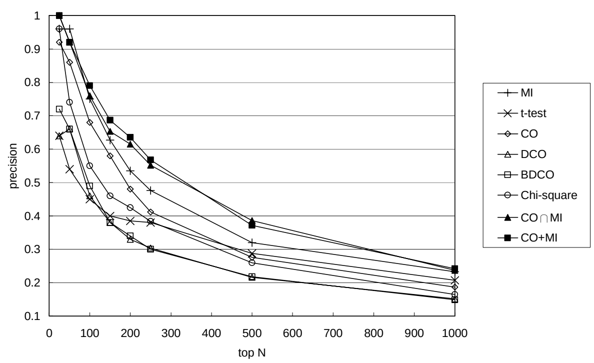

Figure 3. Performances of all models

Conclusion and Future Work

This paper proposes algorithms for event-based opinion summarization utilizing the detected event bursts. Opinion tracking plots are used for burst detection. Summaries of events are generated according to the length of opinion tendencies. Events are correlated with opinions, and targets related to each other (in this case companies) may react to the same events in a similar way. In this paper, tracking plots of opinions are used further in relationship discovery. A total of eight models are proposed. CO+MI, considering both opinions and collocations, performs the best. Top 25, 50 and 100 company pairs discovered by CO+MI achieve precision rates of 1, 0.92 and 0.79, respectively.

This paper shows that the tendencies and weights of opinions are both useful in relationship discovery. This improvement of opinion extraction may help in relationship discovery. In addition, the opinion-based models are sensitive to the relevance of documents. Improving the performance of relevance retrieval for documents and sentences may be the key to improve the performance of opinion-based models.

Models of relationship discovery can be applied to targets other than companies. Almost anything can serve as targets. Finding suitable methods for building up the gold standard and developing evaluation criteria for different types of targets are the future work.

Acknowledgements

Research of this paper was partially supported by National Science Council, Taiwan, under the contract NSC94-2752-E001-001-PAE and NSC95-2752-E001-001-PAE.

References

Bai, X., Padman, R. and Airoldi, E. 2005. On Learning Parsimonious Models for Extracting Consumer Opinions. In *Proceedings of the 38th Annual Hawaii International Conference on System Sciences*, Track 3, Volume 03, 75.2.

Dave, K., Lawrence, S., and Pennock, D.M. 2003. Mining the Peanut Gallery: Opinion Extraction and Semantic Classification of Product Reviews. In *Proceedings of the 12th International World Wide Web Conference*, 519-528.

Hu, M. and Liu, B. Mining Opinion Features in Customer Reviews. 2004. In *Proceedings of the 19th National Conference on Artificial Intelligence*, 755-760.

Ku, L.-W., Li, L.-Y., Wu, T.-H. and Chen., H.-H. 2005. Major Topic Detection and Its Application to Opinion Summarization. In *Proceedings of 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 627-628.

Ku, L.-W., Liang, Y.-T. and Chen, H.-H. 2006. Opinion extraction, summarization and tracking in news and blog Corpora." In *Proceedings of AAAI-2006 Spring Symposium on Computational Approaches to Analyzing Weblogs*, AAAI Technical Report. 100-107

Liu, B., Hu, M. and Cheng, J. 2005. Opinion Observer: Analyzing and Comparing Opinions on the Web. In *Proceedings of the 14th International World Wide Web Conference*, 342-351.

Kim, Soo-Min and Hovy, Eduard. 2004. Determining the Sentiment of Opinions. In *Proceedings of the 20th International Conference on Computational Linguistics*, 1367-1373.

Kleinberg, J. 2002. Bursty and Hierarchical Structure in Streams. In *Proceedings of the 8th ACM SIGKDD International Conference on Knowledge Discover and Data Mining*, 91-101. ACM Press.

Manning, C.D. and Schutze, H. 1999. Foundations of Statistical Natural Language Processing. MIT Press.

Morinaga, S., Yamanishi, K., Tateishi, K. and Fukushima, T. 2002. Mining Product Reputations on the Web. *Proceedings of the 8th ACM SIGKDD International Conference on Knowledge Discover and Data Mining,* 341-349. ACM Press.

Pang, B., Lee, L., and Vaithyanathan, S. 2002. Thumbs up? Sentiment Classification Using Machine Learning Techniques. In *Proceedings of the 2002 Conference on EMNLP*, 79-86.

Riloff, E. and Wiebe, J. 2003. Learning Extraction Patterns for Subjective Expressions. In *Proceedings of the 2003 Conference on EMNLP*, 105-112.

Wiebe, J., Breck, E., Buckly, C., Cardie, C., Davis, P., Fraser, B., Litman, D., Pierce, D., Riloff, E., and Wilson, T. 2002. NRRC Summer Workshop on Multi-perspective Question Answering, Final Report. In *ARDA NRRC Summer 2002 Workshop*.