Detecting Influencers in Written Online Conversations

Or Biran[1]* **Sara Rosenthal[1]*** **Jacob Andreas[1]****

Kathleen McKeown[1]* **Owen Rambow[2]†**

[1] Department of Computer Science, Columbia University, New York, NY 10027
[2] Center for Computational Learning Systems, Columbia University, New York, NY 10027
* {orb, sara, kathy}@cs.columbia.edu
** jda2129@columbia.edu †rambow@ccls.columbia.edu

Abstract

It has long been established that there is a correlation between the dialog behavior of a participant and how influential he or she is perceived to be by other discourse participants. In this paper we explore the characteristics of communication that make someone an opinion leader and develop a machine learning based approach for the automatic identification of discourse participants that are likely to be influencers in online communication. Our approach relies on identification of three types of conversational behavior: persuasion, agreement/disagreement, and dialog patterns.

1 Introduction

In any communicative setting where beliefs are expressed, some are more influential than others. An influencer can alter the opinions of their audience, resolve disagreements where no one else can, be recognized by others as one who makes important contributions, and often continue to influence a group even when not present. Other conversational participants often adopt their ideas and even the words they use to express their ideas. These forms of *personal influence* (Katz and Lazarsfeld, 1955) are part of what makes someone an opinion leader. In this paper, we explore the characteristics of communication that make someone an opinion leader and develop a machine learning based approach for the automatic identification of discourse participants who are likely to be influencers in online communication.

Detecting influential people in online conversational situations has relevance to online advertising strategies which exploit the power of peer influence on sites such as Facebook. It has relevance to analysis of political postings, in order to determine which candidate has more appeal or which campaign strategy is most successful. It is also relevant for designing automatic discourse participants for online discussions ("chatbots") as it can provide insight into effective communication. Despite potential applications, analysis of influence in online communication is a new field of study in part because of the relatively recent explosion of social media. Thus, there is not an established body of theoretical literature in this area, nor are there established implementations on which to improve. Given this new direction for research, our approach draws on theories that have been developed for identifying influence in spoken dialog and extends them for online, written dialog. We hypothesize that an influencer, or an influencer's conversational partner, is likely to engage in the following conversational behaviors:

Persuasion: An influencer is more likely to express personal opinions with follow-up (e.g., justification, reiteration) in order to convince others.

Agreement/disagreement: A conversational partner is more likely to agree with an influencer, thus implicitly adopting his opinions.

Dialog Patterns: An influencer is more likely to participate in certain patterns of dialog, for example initiating new topics of conversation, contributing more to dialog than others, and engendering longer dialog threads on the same topic.

Our implementation of this approach comprises a system component for each of these conversational behaviors. These components in turn provide the features that are the basis of a machine learning approach for the detection of likely influencers. We test this approach on two different datasets, one drawn from Wikipedia discussion threads and the other drawn from LiveJournal weblogs. Our results show that the system performs better for detection of influencer on LiveJournal and that there are interesting differences across genres for detecting the different forms of conversational behavior.

The paper is structured as follows. After reviewing related work, we define influence, present our data and methods. We present a short overview of the black box components we use for persuasion and detection of agreement/disagreement, but our focus is on the development of the influencer system as a whole and thus we spend most time exploring the results of experimentation with the system on different data sets, analyzing which components have most impact. We first review related work.

2 Related Work

It has long been established that there is a correlation between the conversational behavior of a discourse participant and how influential he or she is perceived to be by the other discourse participants (Bales et al., 1951; Scherer, 1979; Brook and Ng, 1986; Ng et al., 1993; Ng et al., 1995). Specifically, factors such as frequency of contribution, proportion of turns, and number of successful interruptions have been identified as being important indicators of influence. Reid and Ng (2000) explain this correlation by saying that "conversational turns function as a resource for establishing influence": discourse participants can manipulate the dialog structure in order to gain influence. This echoes a starker formulation by Bales (1970): "To take up time speaking in a small group is to exercise power over the other members for at least the duration of the time taken, regardless of the content." Simply claiming the conversational floor is a feat of power. This previous work presents two issues for a study aimed at detecting influence in written online conversations.

First, we expect the basic insight – conversation as a resource for influence – to carry over to written dialog: we expect to be able to detect influence in written dialog as well. However, some of the characteristics of spoken dialog do not carry over straightforwardly to written dialog, most prominently the important issue of interruptions: there is no interruption in written dialog. Our work draws on findings for spoken dialog, but we identify characteristics of written dialog which are relevant to influence.

Second, the insistence of Bales (1970) that power is exercised through turn taking "regardless of content" may be too strong. Reid and Ng (2000) discuss experiments which address not just discourse structure features, but also a content feature which represents how closely a turn is aligned with the overall discourse goal of one of two opposing groups (with opposing opinions on a specific issue) participating in the conversation. They show that interruptions are more successful if aligned with the discourse goal. They propose a model in which such utterances "lead to participation which in turn predicts social influence", so that the correlation between discourse structure and influence is really a secondary phenomenon. However, transferring such results to other types of interactions (for example, in which there are not two well-defined groups) is challenging. In this study, we therefore examine two types of features as they relate to influence: content-related (persuasion and agreement/disagreement), and discourse structure-related.

So far, there has been little work in NLP related to influencers. Quercia et al. (2011) look at influencers' language use in Twitter contrasted to other users' groups and find some significant differences. However, their analysis and definition relies quite heavily on the particular nature of social activity on Twitter. Rienks (2007) discusses detecting influencers in a corpus of conversations. While he focuses entirely on non-linguistic behavior, he does look at (verbal) interruptions and topic initiations which can be seen as corresponding to some of our Dialog Patterns Language Uses.

3 What is an Influencer?

Our definition of an influencer was collectively formulated by a community of researchers involved in the IARPA funded project on Socio Cultural Content in Language (SCIL).

This group defines an influencer to be someone who:

P1 by Arcadian $<pc_1>$There seems to be a much better list at the National Cancer Institute than the one we've got.$</pc_1><pa_1>$It ties much better to the actual publication (the same 11 sections, in the same order).$</pa_1>$ I'd like to replace that section in this article. Any objections?
P2 by JFW $<pc_2><a_1>$Not a problem.$</a_1></pc_2>$Perhaps we can also insert the relative incidence as published in this month's wiki Blood journal
P3 by Arcadian I've made the update. I've included template links to a source that supports looking up information by ICD-O code.
P4 by Emmanuelm Can Arcadian tell me why he/she included the leukemia classification to this lymphoma page? It is not even listed in the Wikipedia leukemia page! $<pc_3>$I vote for dividing the WHO classification into 4 parts in 4 distinct pages: leukemia, lymphoma, histocytic and mastocytic neoplasms.$</pc_3><pa_3>$ Remember, Wikipedia is meant to be readable $</pa_3>$by all. Let me know what you think before I delete the non-lymphoma parts.
P5 by Arcadian Emmanuelm, aren't you the person who added those other categories on 6 July 2005?
P6 by Emmanuelm $<d_1>$Arcadian, I added only the lymphoma portion of the WHO classification. You added the leukemias on Dec 29th.$</d_1>$Would you mind moving the leukemia portion to the leukemia page?
P7 by Emmanuelm $<pc_4>$Oh, and please note that I would be very comfortable with a "cross-coverage" of lymphocytic leukemias in both pages.$</pc_4>$My comment is really about myeloid, histiocytic and mast cell neoplasms who share no real relationship with lymphomas.
P8 by Arcadian $<pa_5><a_2>$To simplify the discussion, I have restored that section to your version. $</a_2></pa_5>$You may make any further edits, and $<pc_6>$I will have no objection.$</pc_6>$
P9 by JFW The full list should be on the hematological malignancy page, and the lymphoma part can be here. $<pc_7>$It would be defendable to list ALL and CLL here.$</pc_7><pa_7>$They fall under the lymphoproliferative disorders.$</pa_7>$

Table 1: Influence Example: A Wikipedia discussion thread displaying Emmanuelm as the influencer. Replies are indicated by indentation (for example, P2 is a response to P1). All Language Uses are visible in this example: Attempt to Persuade ($\{pc_i, pa_i\}$), Claims (pc_i), Argumentation (pa_i), Agreement (a_i), Disagreement (d_i), and the five Dialog Patterns Language Uses (eg. Arcadian has positive Initiative).

1. Has credibility in the group.
2. Persists in attempting to convince others, even if some disagreement occurs
3. Introduces topics/ideas that others pick up on or support.

By *credibility*, we mean someone whose ideas are adopted by others or whose authority is explicitly recognized. We hypothesize that this shows up through agreement by other conversants. By *persists*, we mean someone who is able to eventually convince others and often takes the time to do so, even if it is not quick. This aspect of our definition corresponds to earlier work in spoken dialog which shows that frequency of contributions and proportion of turns is a method people use to gain influence (Reid and Ng, 2000; Bales, 1970). By point 3, we see that the influencer may be influential even in directing where the conversation goes, discussing topics that are of interest to others. This latter feature can be measured through the discourse structure of the interaction. The influencer must be a group participant but need not be active in the discussion(s) where others support/credit him.

The instructions that we provided to annotators included this definition as well as examples of who is *not* an influencer. We told annotators that if someone is in a hierarchical power relation (e.g., a boss), then that person is not an influencer to sub-ordinates (or, that is not the type of influencer we are looking for). We also included someone with situational power (e.g., authority to approve other's actions) or power in directing the communication (e.g., a moderator) as negative examples.

We also gave positive examples of influencers. Influencers include an active participant who argues against a disorganized group and resolves a discussion is an influencer, a person who provides an answer to a posted question and the answer is accepted after discussion, and a person who brings knowledge to a discussion. We also provided positive and negative examples for some of these cases.

Table 1 shows an example of a dialog where there is evidence of influence, drawn from a Wikipedia Talk page. A participant (Arcadian) starts the thread with a proposal and a request for support from other participants. The influencer (Emmanuelm) later joins the conversation arguing against Arcadian's proposal. There is a short discussion, and Arcadian defers to Emmanuelm's position. This is one piece of dialog within this group where Emmanuelm may demonstrate influence. The goal of our system is to find evidence for situations like this, which suggests that a person is more likely to be an influencer.

Since we attempt to find local influence (a person who is influential in a particular thread, as opposed to influential in general), our notion of influencer is consistent with diverse views on social influence. It is consistent with the definition of influencer proposed by Gladwell (2001) and Katz (1957): an exceptionally convincing and influential person, set apart from everyone else by his or her ability to spread opinions. While it superficially seems inconsistent with Duncan Watts' concept of "accidental influentials" (Watts, 2007), that view does not make the assertion that a person cannot be influential in a particular situation (in fact, it predicts that someone will) - only that one cannot in general identify people who are always more likely to be influencers.

4 Data and Annotation

Our data set consists of documents from two different online sources: weblogs from LiveJournal and discussion forums from Wikipedia.

LiveJournal is a virtual community in which people write about their personal experiences in a weblog. A LiveJournal entry is composed of a post (the top-level content written by the author) and a set of comments (written by other users and the author). Every comment structurally descends either from the post or from another comment.

Each article on Wikipedia has a discussion forum (called a Talk page) associated with it that is used to discuss edits for the page. Each forum is composed of a number of threads with explicit topics, and each thread is composed of a set of posts made by contributors. The posts in a Wikipedia discussion thread may or may not structurally descend from other posts: direct replies to a post typically descend from it. Other posts can be seen as descending from the topic of the thread.

For consistency of terms, from here on we refer to each weblog or discussion forum thread as a *thread* and to each post or comment as a *post*.

We have a total of 333 threads: 245 from LiveJournal and 88 from Wikipedia. All were annotated for influencers. The threads were annotated by two undergraduate students of liberal arts. These students had no prior training or linguistic background. The annotators were given the full definition from section 3 and asked to list the participants that they thought were influencers. Each thread may in principle have any number of influencers, but one or zero influencers per thread is the common case and the maximal number of influencers found in our dataset was two. The inter-annotator agreement on whether or not a participant is an influencer (given by Cohen's Kappa) is 0.72.

5 Method

Our approach is based on three conversational behaviors which are identified by separate system components described in the following three sections. Figure 1 shows the pipeline of the Influencer system and Table 1 displays a Wikipedia discussion thread where there is evidence of an influencer and in which we have indicated the conversational behaviors as they occur. Motivated by our definition, each component is concerned with an aspect of the likely influencer's discourse behavior:

Persuasion examines the participant's language to identify attempts to persuade, such as $\{pc_1, pa_1\}$ in Table 1, which consist of claims (e.g. pc_1) made by the participant and supported by argumentations (e.g. pa_1). It also identifies claims and argumentations independently of one another (pc_4 and pa_5).

Agreement/Disagreement examines the other participants' language to find how often they agree or disagree with the participant's statements. Examples are a_1 and d_1 in Table 1.

Dialog Patterns examines how the participant interacts in the discussion structurally, independently of the content and the language used. An example of this is Arcadian being the first poster and contributing the most posts in the thread in Table 1.

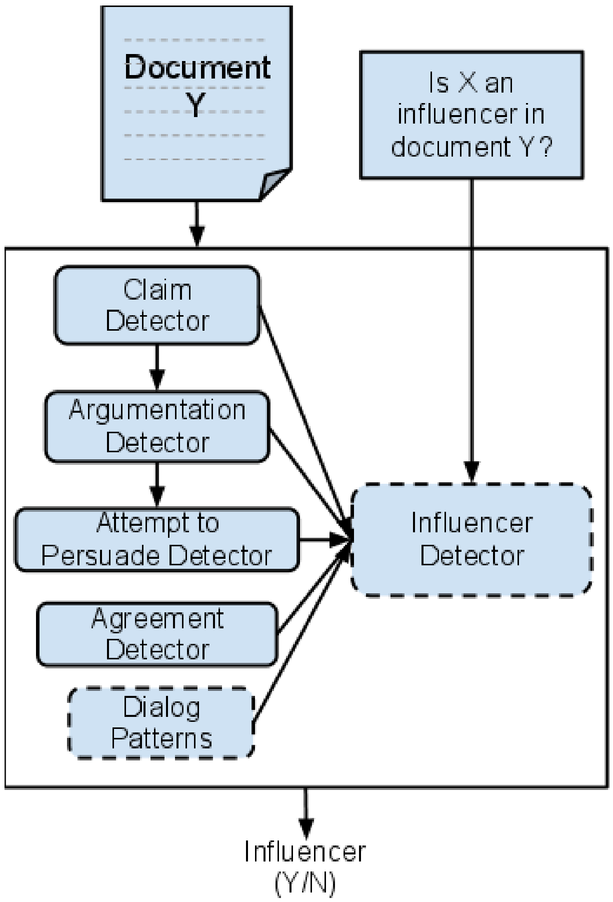

Figure 1: The influencer pipeline. Solid lines indicate black-box components, which we only summarize in this paper. Dashed lines indicate components described here.

Each component contributes a number of *Language Uses* which fall into that category of conversational behavior and these Language Uses are used directly as features in a supervised machine learning model to predict whether or not a participant is an influencer. For example, Dialog Patterns contributes the Language Uses *Initiative*, *Irrelevance*, *Incitation*, *Investment* and *Interjection*.

The Language Uses of the Persuasion and Agreement/Disagreement components are not described in detail in this paper, and instead are treated as black boxes (indicated by solid boxes in Figure 1). We have previously published work on some of these (Biran and Rambow, 2011; Andreas et al., 2012). The remainder of this section describes them briefly and provides the results of evaluations of their performance (in Table 2). The next section describes the features of the Dialog Patterns component.

5.1 Persuasion

This component identifies three Language Uses: Attempt to Persuade, Claims and Argumentation.

We define an attempt to persuade as a set of contributions made by a single participant which may be made anywhere within the thread, and which are all concerned with stating and supporting a single claim. The subject of the claim does not matter: an opinion may seem trivial, but the argument could still have the structure of a persuasion.

Our entire data set was annotated for attempts to persuade. The annotators labeled the text participating in each instance with either *claim*, the stated opinion of which the author is trying to persuade others or *argumentation*, an argument or evidence that supports that claim. An attempt to persuade must contain exactly one claim and at least one instance of argumentation, like the {claim, argumentation} pairs $\{pc_1, pa_1\}$ and $\{pc_3, pj_3\}$ in Table 1.

In addition to the complete *attempt to persuade* Language Use, we also define the less strict Language Uses *claims* and *argumentation*, which use only the subcomponents as stand-alones.

Our work on argumentation, which builds on Rhetorical Structure Theory (Mann and Thompson, 1988), is described in (Biran and Rambow, 2011).

5.2 Agreement/Disagreement

Agreement and disagreement are two Language Uses that model others' acceptance of the participant's statements. Annotation (Andreas et al., 2012) is performed on pairs of phrases, $\{p_1, p_2\}$. A phrase is a substring of a post or comment in a thread. The annotations are directed since each post or comment has a time stamp associated with it. This means that p_1 and p_2 are not interchangeable. p_1 is called the "target phrase", and p_2 is called the "subject phrase". A person cannot agree with him- or herself, so the author of p_1 and p_2 cannot be the same. Each annotation is also labeled with a type: either "agreement" or "disagreement".

6 Dialog Patterns

The Dialog Patterns component extracts features based on the structure of the thread. Blogs and discussion threads have a tree structure, with a blog post or a topic of discussion as the root and a set of

Component	Wikipedia			LiveJournal		
	P	R	F	P	R	F
Attempt to persuade	79.1	69.6	74	57.5	48.2	52.4
Claims	83.6	74.5	78.8	53.7	13.8	22
Argumentation	23.3	91.7	37.1	30.9	48.9	37.8
Agreement	12	31.9	17.4	20	50	28.6
Disagreement	8.7	9.5	9.1	6.3	14.3	8.7

Table 2: Performance of the black-box Language Uses in terms of Precision (P), Recall (R), and F-measure(F).

Conversational Behavior Component	Language Use (Feature)	Users		
		A	J	E
Persuasion	Claims	2/6	2/6	2/6
	Argumentation	Y	Y	Y
	Attempt to Persuade	Y	Y	Y
Agreement/ Disagreement	Agreement	1/1	0/1	0/1
	Disagreement	1/1	0/1	0/1
Dialog Patterns	Initiative	Y	N	N
	Irrelevance	2/4	1/2	1/3
	Incitation	4	1	3
	Interjection	1/9	2/9	4/9
	Investment	4/9	2/9	3/9

Table 3: The feature values for each of the participants, Arcadian (A), JFW (J), and Emmanuelm (E), in the Wikipedia discussion thread shown in Table 1.

comments or posts which are marked as a reply - either to the root or to an earlier post. The hypothesis behind Dialog Patterns is that influencers have typical ways in which they participate in a thread and which are visible from the structure alone.

The Dialog Patterns component contains five simple Language Uses:

Initiative The participant is or is not the first poster of the thread.

Irrelevance The percentage of the participant's posts that are not replied to by anyone.

Incitation The length of the longest branch of posts which follows one of the participant's posts. Intuitively, the longest discussion started directly by the participant.

Investment The participant's percentage of all posts in the thread.

Interjection The point in the thread, represented as percentage of posts already posted, at which the participant enters the discussion.

7 System and Evaluation

The task of the system is to decide for each participant in a thread whether or not he or she is an influencer in that particular thread. It is realized with a supervised learning model: we train an SVM with a small number of features, namely the ten Language Uses. One of our goals in this work is to evaluate which Language Uses allow us to more accurately classify someone as an influencer. Table 3 shows the full feature set and feature values for the sample discussion thread in Table 1. We experimented with a number of different classification methods, including bayesian and rule-based models, and found that SVM produced the best results.

7.1 Evaluation

We evaluated on Wikipedia and LiveJournal separately. The data set for each corpus consists of all participants in all threads for which there was at least one influencer. We exclude threads for which no influencer was found, narrowing our task to finding the influencers where they exist. For each participant X in each thread Y, the system answers the following question: *Is X an influencer in Y?*

We used a stratified 10-fold cross validation of each data set for evaluation, ensuring that the same participant (from two different threads) never appeared in both training and test at each fold, to eliminate potential bias from fitting to a particular participant's style. The system components were identical when evaluating both data sets, except for the claims system which was trained on sentiment-annotated data from the corpus on which it was evaluated.

Table 4 shows the performance of the full system and of systems using only one Language Use feature compared against a baseline which always answers positively (X is always an influencer in Y). It also shows the performance for the best system, which was found for each data set by looking at all possible combinations of the features. The best system for the Wikipedia data set is composed of four features: Claims, Argumentation, Agreement and Investment. The best LiveJournal system is composed of all five Dialog Patterns features, Attempt to Persuade and Argumentation. We found our results to be statis-

System	Wikipedia			LiveJournal		
	P	R	F	P	R	F
Baseline: all-yes	16.2	100	27.9	19.2	19.2	32.2
Full	40.5	80.5	53.9	61.7	82	70.4
Initiative	31.6	31.2	31.4	73.5	72.7	73.1
Irrelevance	21.7	77.9	34	19.2	100	32.2
Incitation	28.3	77.9	41.5	49.5	73.8	59.2
Investment	43	71.4	53.7	50.2	75.4	60.3
Interjection	24.7	88.3	38.6	36.9	91.3	52.5
Agreement	36	46.8	40.7	45.1	82.5	58.3
Disagreement	35.3	70.1	47	19.2	100	32.2
Claims	40	72.7	51.6	54.3	76	63.3
Argumentation	19	98.7	31.8	31.1	85.2	45.6
Attempt to persuade	23.7	79.2	36.5	37.4	48.1	42.1
Best system	**47**	**80.5**	**59.3**	**66.2**	**84.7**	**74.3**

Table 4: Performance in terms of Precision (P), Recall (R), and F-measure (F) using the baseline (everyone is an influencer), all features (full), individual features one at a time, and the best feature combination for each data set.

System	Wikipedia			LiveJournal		
	P	R	F	P	R	F
Baseline	13.9	100	24.5	14.2	100	24.9
Full	36.7	79.2	50.2	46.3	79.8	58.6
Best	40.1	76.6	52.7	48.2	81.4	60.6

Table 5: Performance on the data set of all threads, including those with no influencers. The 'Best System' is the system that performed best on the filtered data set.

Data Set	Our System			Gold Answers		
	P	R	F	P	R	F
Wikipedia	23.6	69.4	35.2	23.8	81.6	36.9
LiveJournal	37.5	48.1	42.1	40.7	61.8	49

Table 6: Performance of the Attempt to Persuade component in directly predicting influencers. A comparison of our system and the component's gold annotation. These experiments were run on the full data set, which is why the system results are not exactly those of Table 4.

tically significant (with the Bonferroni adjustment) in paired permutation tests between the best system, the full system and the baseline of each data set.

When we first performed these experiments, we used all threads in the data set. The performance on this full set was lower, as shown in Table 5 due to the presence of threads with no influencers. Threads in which the annotators could not find a clear influencer tend to be of a different nature: there is either no clear topic of discussion, or no argument (everyone is in agreement). We leave the task of distinguishing these threads from those which are likely to have an influencer to future work.

7.2 Evaluating with Perfect Components

In a hierarchical system such as ours, errors can be attributed to imperfect components or to a bad choice of features, so it is important to look at the potential contribution of the components. As an example, Table 6 shows the difference between our Attempt to Persuade system and a hypothetical perfect Attempt to Persuade component, simulated by using the gold annotations, when predicting influencer directly (i.e., a participant is an influencer iff she makes an attempt to persuade).

Clearly, when predicting influencers, Attempt to Persuade is a stronger indicator in LiveJournal than it is in Wikipedia. However, as shown in Table 2, our Attempt to Persuade system performs better on Wikipedia. This situation is reflected in Table 6, where the lower quality of the system component in LiveJournal corresponds to a significantly lower performance when applied to the influencer task. These results demonstrate that Attempt to Persuade is a good feature: a more precise feature value means higher predictability of influencer. In the future we will perform similar analyses for the other features.

8 Discussion

We evaluated our system on two corpora - LiveJournal and Wikipedia discussions - which differ in structure, context and discussion topics. As our results show, they also differ in the way influencers behave and the way others respond to them. To illustrate the differences, we contrast the sample Wikipedia thread (Table 1) with an example from LiveJournal (Table 7).

It is common in LiveJournal for the blogger to be an influencer, as is the case in our example thread, because the topic of the thread is set by the blogger and comments are typically made by her friends. This fact is reflected in our results: Initiative is a very strong indicator in LiveJournal, but not so in

P1 by poconell $<pc_1>$He really does make good on his promises! $</pc_1><pa_1>$Day three in office, and the Global Gag Rule (A.K.A"The Mexico City Policy") is gone!$</pa_1>$I was holding my breath, hoping it wouldn't be left forgotte. He didn't wait. $<pc_2>$He can see the danger and risk in this policy, and the damage it has caused to women and families.$</pc_2><pc_3>$I love that man!$</pc_3>$
P2 by thalialunacy $<a_1>$I literally shrieked 'HELL YES!' in my car when I heard. :D:D:D$</a_1>$
P3 by poconell $<a_2>$Yeah, me too$</a_2>$
P4 by lunalovepotter $<pc_4><a_3>$He is SO AWESOME!$</a_3></pc_4><pa_4>$Right down to business, no ifs, ands, or buts! :D$</pa_4>$
P5 by poconell $<pc_5>$It's amazing to see him so serious too!$</pc_5><pa_5>$This is one tough, no-nonsense man!$</pa_5>$
P6 by penny_sieve My icon says it all :)
P7 by poconell $<pc_6>$And I'm jealous of you with that President!$</pc_6><pa_6>$We tried to overthrow our Prime Minister, but he went crying to the Governor General. $</pa_6>$

Table 7: Influence Example: A LiveJournal discussion thread displaying poconell as the influencer. All the Language Uses are visible in this example: agreement/disagreement (a_i/d_i), persuasion ($\{pc_i, pa_i\}, pc_i, pa_i$), and dialog patterns (eg. poconell has positive Initiative). This example is very different from the Wikipedia example in Table 1.

Wikipedia, where the discussion is between a group of editors, all of whom are equally interested in the topic. In general, the Dialog Patterns features are stronger in LiveJournal. We believe this is due to the fact that the tree structure in LiveJournal is strictly enforced. In Wikipedia, people do not always reply directly to the relevant post. Investment is the exception: it does not make use of the tree structure, and is therefore an important indicator in Wikipedia.

Attempt to Persuade is useful in LiveJournal (the influencer poconell makes three attempts to persuade in Table 7) but less so in Wikipedia. This is explained by the precision of the gold system in Table 6. Only 23.8% of those who attempt to persuade in Wikipedia are influencers, compared with 40.7% in LiveJournal. Attempts to Persuade are more common in Wikipedia (all participants attempt to persuade in Table 1), since people write there specifically to argue their opinion on how the article should be edited. Conversely, agreement is a stronger predictor of influence in Wikipedia than in LiveJournal; we believe that is because of a similar phenomenon, that people in LiveJournal (who tend to know each other) agree with each other more often. Disagreement is not a strong indicator for either corpus which may say something about influencers in general - they can be disagreed with as often as anyone else.

9 Conclusion and Future Work

We have studied the relevance of content-related conversational behavior (persuasion and agreement/disagreement), and discourse structure-related conversational behavior to detection of influence. Identifying influencers is a hard task, but we are able to show good results on the LiveJournal corpus where we achieve an F-measure of 74.3%. Despite a lower performance on Wikipedia, we are still able to significantly outperform the baseline which yields only 28.2%. Differences in performance between the two seem to be attributable in part to the more straightforward dialog structure in LiveJournal.

There are several areas for future work. In our current work, we train and evaluate separately for our two corpora. Alternatively, we could investigate different training and testing combinations: train on one corpus and evaluate on the other; a mixed corpus for training and testing; genre-independent criteria for developing different systems (e.g. length of thread). We will also evaluate on new genres (such as the Enron emails) in order to gain an appreciation of how different genres of written dialog are.

Acknowledgment

This work has been supported by the Intelligence Advanced Research Projects Activity (IARPA) via Army Research Laboratory (ARL) contract number W911NF-09-C-0141. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright annotation thereon.

References

Jacob Andreas, Sara Rosenthal, and Kathleen McKeown. 2012. Annotating agreement and disagreement in threaded discussion. In *Proceedings of the 8th International Conference on Language Resources and Computation (LREC)*, Istanbul, Turkey, May.

R. F. Bales, Strodtbeck, Mills F. L., T. M., and M. Roseborough. 1951. Channels of communication in small groups. *American Sociological Review*, pages 16(4), 461–468.

R. F. Bales. 1970. Personality and interpersonal behaviour.

Or Biran and Owen Rambow. 2011. Identifying justifications in written dialog. In *Proceedings of the Fifth IEEE International Conference on Semantic Computing*.

M.E. Brook and S. H. Ng. 1986. Language and social influence in small conversational groups. *Journal of Language and Social Psychology*, pages 5(3), 201–210.

Malcolm Gladwell. 2001. *The tipping point: how little things can make a big difference*. Abacus.

Elihu Katz and Paul F. Lazarsfeld. 1955. *Personal influence*. Free Press, Glencoe, IL. by Elihu Katz and Paul F. Lazarsfeld. With a foreword by Elmo Roper. "A report of the Bureau of Applied Social Research, Columbia University." Bibliography: p. 381-393.

E. Katz. 1957. *The Two-Step Flow of Communication: An Up-To-Date Report on an Hypothesis*. Bobbs-Merrill Reprint Series in the Social Sciences, S137. Ardent Media.

William C. Mann and Sandra A. Thompson. 1988. Rhetorical structure theory: Toward a functional theory of text organization. *Text*, 8(3):243–281.

S. H. Ng, D. Bell, and M. Brooke. 1993. Gaining turns and achieving high in influence ranking in small conversational groups. *British Journal of Social Psychology*, pages 32, 265–275.

S. H. Ng, M Brooke, and M. Dunne. 1995. Interruption and in influence in discussion groups. *Journal of Language and Social Psychology*, pages 14(4),369–381.

Daniele Quercia, Jonathan Ellis, Licia Capra, and Jon Crowcroft. 2011. In the mood for being influential on twitter. In *SocialCom/PASSAT*, pages 307–314. IEEE.

Scott A. Reid and Sik Hung Ng. 2000. Conversation as a resource for in influence: evidence for prototypical arguments and social identification processes. *European Journal of Social Psychology*, pages 30, 83–100.

Rutger Joeri Rienks. 2007. *Meetings in smart environments : implications of progressing technology*. Ph.D. thesis, Enschede, the Netherlands, July.

K. R. Scherer. 1979. Voice and speech correlates of perceived social influence in simulated juries. In *H. Giles and R. St Clair (Eds), Language and social psychology*, pages 88–120. Oxford: Blackwell.

Duncan Watts. 2007. The accidental influentials. *Harvard Business Review*.