A Motif Approach for Identifying Pursuits of Power in Social Discourse

David B. Bracewell, Marc Tomlinson, and Hui Wang
Language Computer Corporation
Richardson, Texas 75080
Email:{david,marc,hui}@languagecomputer.com

Abstract—In this paper, we investigate whether the social goals of an individual's utterances can be recognized through analysis of a discourse's intentional structure. Specifically we focus on identifying individuals pursuing power within a group. Individuals pursue power in order to increase their control of the actions and goals of the group. Following work in discourse processing we decompose the problem into identifying the social intention of the discourse segments and the intentional structure of the overall discourse. The set of social intentions come from eight psychologically-motivated social acts. We then build a motif-based representation of the discourse's social intentional structure that captures the interactions of the intentions between discourse participants. Finally we show how these structures can be used to identify the social goal of pursuit of power. Our best results achieve an accuracy of 84.2% for predicting pursuit of power in discussions communicated in English and 80.6% for discussions communicated in Chinese.

I. INTRODUCTION

Understanding discourse in the modern world of social media presents a plethora of new challenges. One central challenge for the community is advancing our understanding of the semantics and pragmatics of social media. Communication over social media often has more to do with the social implications of the utterance than its content. However, current computational approaches to understanding the semantics and pragmatics of discourse focus only on identifying the actions of individuals towards the dialogue through dialogue acts [1], [2], and examination of the information content of the dialogue through topical analysis [3], [4]. Instead, in this paper we focus on identifying the social pragmatics of a discourse and present a computational model of the social goals of discourse participants. In particular, we focus on identifying individuals that are pursuing power within an online community.

As in the many discourse theories that converged on the importance of discourse cues for parsing the discourse structure of dialogues [5], [6], we claim that knowledge about social cues and their expression in language is central for recognizing social acts in communication and intentions in discourse. Recent advances in syntactic, semantic, and discourse processing allow us to contemplate the automatic processing of socio-linguistic-driven properties by identifying the language uses that capture the manifestations of social acts within discourse segments. As in other textual inference tasks, e.g. textual entailment [7] and conversation entailment [8], there are a variety of syntactic and semantic constructs that are helpful in recognizing the social implications of a dialogue.

We base our system for identifying pursuits of power on the seminal work of Grosz and Sidner [6], who introduce three structural components: *linguistic structure*, *intentional structure*, and *attentional state*. The linguistic structure encompasses how utterances combine together to form larger segments of discourse. The intentional structure is defined as having a single overarching motivation for the discourse and smaller sub-components which define the purpose of each of the discourse segments. Finally, the attentional structure is a property of the discourse itself, not of its participants, and it acts to keep track of the current focus of the participants in the discourse.

For discourse taking place over social media, we must focus on the social intentions of the discourse participants. We identify linguistic structure based on topic shifts [9]. Each topic in the discourse becomes a discourse segment. We posit that the central focus of any discourse segment, i.e. the attentional state, is the most salient individual pursuing power. For the intentional structure, we turn our focus from the intentions of the discourse to the social intentions of the participants in order to understand their social goals. Therefore, we introduce the concept of social intentions and social intentional structure.

The social intentional structure of discourse is defined over the interactions between the initiating conversational participant (ICP), the person whom we are ascertaining is pursuing or not pursuing power, and the other conversational participants (OCP). The social intentions of each discourse segment are represented using social acts. Social acts capture the socio-cognitive processes that act on individuals during communication. We define a set of eight social acts, which are drawn from literature in the fields of psychology and organizational behavior. Social acts are captured using a combination of lexical, grammatical, and semantic cues as well as n-gram models.

Using social acts, we define the intentional structure of discourse as motifs, which are patterns of interactions (social acts) between the initiating conversational participant (ICP) and the other conversational participants (OCP). In order to capture motifs, discourses are first transformed into sequences of social acts. Motifs are then learned over the sequences of social act using a generative model.

In this contribution we show that individuals with similar social goals (pursuit of power) exhibit similar motifs over the intentional structure of their discourse. The similarity of the

motifs allows for predicting whether or not an individual is or is not pursuing power. We show the validity of this approach for discourses communicated in English and Chinese.

II. RELATED WORK

Related work can be found in a variety of fields. Work in the area of social relationship extraction can be divided into several areas. The field of socio-linguistics boasts well-established studies of interpersonal relationships. For example, [10] present a thorough linguistic analysis of casual conversations that covers topics such as humor, attitude, friendliness, and gossip. This is accomplished through a comprehensive expert analysis of the dialogue at multiple levels. However, research using Natural Language Processing to automatically identify social relationships in text is still in its infancy.

Strzalkowski et al (2010) break down social phenomena into mid-level social language uses [11]. They focus on the use of discourse features (e.g. topic control) to identify language uses that might be indicative of some social constructs. In contrast, we focus on identifying multiple types of social acts in combination with a dynamic model of discourse for predicting pursuits of power.

Similar work is presented in Bracewell et al (2011). They identify a set of social actions for identifying collegiality between individuals [12]. Additionally, Tomlinson et al. (2012) introduce a set of 10 social acts pertaining to pursuit of power in Arabic [13]. In this work we extend and refine the definitions for the social acts and demonstrate their effectiveness for predicting the social role of pursuing power in English and Chinese.

Alternatively, network-based approaches have also been gaining popularity in the areas of social roles, events, and groups. Roles, events, and groups can influence and constrain the interpersonal relationships between individuals. Agarwal and Rambow (2010) look at extracting two types of social events in text: interaction and observation. These events are then used in extracting social network models from text [14]. They use Support Vector Machines with tree kernels based on phrase structure and dependency trees.

Social roles may influence interpersonal relationships; thus, Brendel and Krawczyk use the relations of people in their social network to determine their roles [15]. Similarly, Fisher et al. (2006) examine social networks from Usenet newsgroups [16]. They find that the type of newsgroup determines the type of relationships and roles.

Non-verbal features have been used for identifying relationships and social roles in face-to-face interactions. Zancanaro et al. (2006) examine the identification of social roles in face-to-face interactions [17]. They use visual and auditory information from scenes to train a model to detect and classify the social roles that are present. Gatica-Perez (2009) offers non-verbal analysis of dominance in group situations [18].

Work has also been conducted to identify the relative status of individuals through automated analysis of their language. Bramsen et al. (2011) look for the presence of *upspeak* (speech directed towards individuals of higher status) and *downspeak* (speech directed towards individuals of lower status) within the Enron email corpus using an n-gram based approach combined with human-engineered features [19]. They achieve an accuracy of 78.1% for detecting the relative status difference between individuals. In contrast to identifying a static social relationship between individuals, we look at detecting an individual's intentions to manipulate an existing social relationship.

Related work in social and organizational psychology has a long history of examining why people acquire power and what it means to be powerful. Generally it is thought that individuals have a basic *need* for power [20] and that this is derived from a desire for agency, allowing individuals control over their own actions [21]. Realizations of that power come in two main types, formal power (power given to an individual by an authority) and informal power (power based on an individual's characteristics, e.g. expertise and skills [22].

There is a long history of work in discourse understanding that focuses on understanding the pragmatics of the discourse. More recent work has focused on inferring information, such as conversational intent, about the discourse participants. Zhang and Chai (2010) introduce conversation entailment, which is designed to answer a variety of hypotheses about dialogue participants [8]. The hypotheses can be about factual information, beliefs and opinions, desires, or communicative intentions.

Boella et al. [23] examine how social goals, namely individuals' motivation to take reputation into account, explain how dialogue progresses in non-cooperative settings. However, in the world of social media where anonymity is easily provided, taking the reputation of others into account is often not a goal of individuals. That is why we instead focus on the social intentions of the discourse participants. We posit that it is the social intentions of the participants that keep the dialogue moving.

Other work has focused on the coherence of discourse [24], [25], [5], [26]. Mann and Thompson introduce Rhetorical Structure Theory (RST), which was originally developed during the study of automatic text generation [26]. They posit that the coherence of a text is attributed to the rhetorical relations between non-overlapping texts called the nucleus and satellite. The definition of the relations are not morphological or syntactic, but instead are focused on function and semantics.

III. A SOCIAL DISCOURSE MODEL FOR AN INDIVIDUAL PURSUING POWER

Power is exhibited in many forms, through physical intimidation, wealth (money, physical resources, or knowledge), or position within a hierarchy. There are variety of methods to pursue power, which are often determined by cultural norms. Moreover, because of the shear variety of methods to pursue power, it is difficult to develop a robust cross-domain text-based recognition approach to identify those whom are in pursuit. Instead, we focus on detecting differences in the way people use language when they are attempting to pursue power.

Agreement	Statements that a group member makes to indicate that they shares the same view about something another member has said or done.
Challenge Credibility	Attempts to discredit or raise doubt about another group members qualifications or abilities.
Disagreement	Statements a group member makes to indicate that they does not share the same view about something another member has said or done.
Disrespect	Statements that a group member makes to insult another member of the group.
Establish Credibility	Statements that a speaker makes to demonstrate their knowledge or personal experience in order to make them look better in the eyes of the group.
Leadership Avoidance	Attempts to avoid being in a position of control over the group.
Managerial Influence	Statements that a speaker makes to control the discussion with the goal of increasing sway over the group.
Solidarity	Statements that a group member makes to strengthen the groups sense of community and unity.

Fig. 1: The set of eight social acts that capture social moves by individuals in power or seeking power.

We look to mimic human understanding of power and follow the non-conscious cues provided within a dialogue.

A. Social Acts and Intentions

Social acts signal the dialogue participants' social intentions for a discourse segment. The set of eight acts, presented in Table 1, have been derived from work in psychology on power, status, and leadership [27], [22], [28], [29], [30].

1) Agreement & Disagreement: Agreement can act as an affordance to an individual or as a means to establish solidarity between individuals. Likewise disagreement can act as a way of undermining or challenging credibility. However, because of the special status of agreement and disagreement we consider them as two separate social acts.

Agreement can be manifested through simple phrases, such as "I agree", through negations of disagreement, such as "I am not disagreeing with you", and through more complex phrases, such as "What Adam says is in principle correct." Similarly, disagreement is manifested through simple "I disagree" phrases as well as negations of agreement, such as "I definitely do not agree with what you said."

2) Challenge Credibility: Challenging credibility can be used by an individual to lower the status of other group members [29]. These challenges can be in demands to prove credibility, such as "prove your lies" and aggressive accusing questions, such as "what does that have to do with what we are talking about?". Challenging credibility can also occur through gossip, such as "X doesn't know what he is talking about". This tactic can be used by group members to moderate the power of a leader who has overstepped their boundaries [28].

3) Disrespect: Disrespected individuals often feel they have been unjustly treated due to the disrespectful action, causing a social imbalance between them and the perpetrator (Miller, 2001). This social imbalance causes a power differential between the two individuals, thus giving the perpetrator power over the individual. Examples of disrespect include "You are a gigantic hypocrite you know that?" and "Do you speak English well?"

4) Establish Credibility: Establishing credibility reflects an attempt by an individual to demonstrate their credibility and fitness for leadership [28]. Evidence for establishment of credibility manifests itself in many different ways. The most common in our data set is an explicit mention of the individuals credentials, such as "I am a physicist". Alternatively a person can demonstrate their credibility by providing the group with cited information, such as "Article 10.5 paragraph 3 says...". Finally an individual can justify their opinion through the use of logic or citation of personally relevant anecdotes.

5) Leadership avoidance: Individuals may make attempts to avoid leadership (power) that others wish to bestow upon them, for example turning down a promotion. The act of avoidance by an individual is considered a negative sign that the individual is pursing power. We identify explicit comments by an individual to avoid making decisions, such as "I think I will have to decline". However, avoidance can also signal defiance and can also be a positive indicator of pursuit of power.

6) Managerial Influence: Managerial influence is used by individuals to signal that they are a leader. Examples of managerial influence include "Can we focus the discussion" and "Are we still trying to find out where the scholarly consensus is on the matter of Lukan authorship?"

Figure 4 has a number of examples of managerial influence, such as *A* proposing to move the page and *B* giving factual reasons why "Military of East Timor" is an incorrect name for the page.

7) Solidarity: Further, language indicative of a desire for group solidarity encapsulates the establishment and maintenance of shared group membership. Group membership can be expressed at either the relational level (e.g. Father, co-worker, etc.) or the collective level (e.g. single mothers) [31]. Language indicative of a desire for group solidarity demonstrates that an individual identifies with the group, an important characteristic of leaders [28] and cooperators [32]. This solidarity can be expressed explicitly (e.g. "We're all in this together"), covertly (e.g. as through the use of inclusive first-person pronouns), or through unconscious actions and linguistic cues, such as the use of in-group jargon, certain syntactic constructions, and mimicry.

8) Recognition of Social Acts: The social acts are recognized using a combination of lexical, grammatical, and semantic cues associated with each social act. The cues are hand-coded using a template system that works on top of semantic parsers built using PropBank, NomBank, and FrameNet. In addition, a set of rules defined using a Packrat grammar [33] are used for extracting social acts in Chinese. Finally, we employ a gappy n-gram model, similar to the Motif model presented in section IV in order to improve recall.

S) 1．使用的博客经过了实认证，特别是名人博客．
[*Establish Credibility*]
"1. The blogs we used, especially famous people's blogs implemented real-name authentication. [*Establish Credibility*]"
2．　　百科内的某些观点正是缘于博客，却禁止对该博客的引用．不加入 Blog 地址的话会使得来源更难于验证．[*Establish Credibility*]
"2. Some comments in wiki came from blogs, but it was forbidden to cite the blogs. It is more difficult to confirm the sources without adding the blog address.[*Establish Credibility*]"
3．这些内容根本不需要大众媒体验证即可确信是代表某个知名人士的言论，反之也不会有媒体整天围着博客转来报道这些博文．[*Establish Credibility*]
"3. It can be known that these contents represent certain famous person's comment without the authentication of mass media. On the contrary there is no media monitoring these blogs all day long and reporting these posts.[*Establish Credibility*]"

T）基本上认为博客不是第二手来源。[*Disagreement*]
如果真要加入博客中的论点，最好找别的媒体报道这篇博客中的看法的文章作为来源。[*Managerial Influence*]
"Normally it is considered that blogs are not second hand sources.[*Disagreement*] If you really want to use the comments in the blog, it is the best to use the articles written by other media that reported the comments in these blogs.[*Managerial Influence*]"

Fig. 2: Example discourse in Chinese.

In order to capture the social acts, we first need a corpus with an ample amount of annotations. Our corpus for English consists of 21,067 English sentences with 8,149 (38.7%) of them having one or more social acts annotated. For Chinese, our corpus consists of 24,339 sentences with 11,537 (47.4%) of them having one or more social acts annotated. On average each English sentence is assigned 1.02 social acts and each Chinese sentence is assigned 1.77 social acts. Kappa values for English annotations range from 0.18 for Managerial Influence to 0.52 for Establish Solidarity. For Chinese, the Kappa values range from 0.07 for Disagreement to 0.52 for Establish Credibility.

The annotated corpus of social acts is split into 80% for training and 20% for testing. The combination of rules and gappy n-grams gives the best overall F-measure for classification of social acts. Our best models obtain a 50.4% F-Measure for English and a 52.1% F-Measure for Chinese. Figures 2 and 4 are example excerpts from Wikipedia Talk discussions tagged with social acts in Chinese and English respectively.

B. The Intentional Structure of Discourse for an Individual Pursuing Power

The intentional structure of a discourse reflects the goals of the discourse participants. When the initiating conversational participant is trying to increase their power their discourse segments should reflect their intentions, likewise the other conversational participants will either assist or hinder the ICP's efforts. An individual is pursuing power when they attempt to gain control over the goals, actions, or rewards accorded to themselves or others [22]. From our list of social acts we create a model of the social actions individuals pursuing power should utilize. As shown in Figure 3, we have identified social acts that change an individual's level of power within a group. This list was designed to be culturally neutral and define the basic building-blocks for pursuit of power. This model is based on the reciprocal-influence model of power developed by Keltner et al. [28].
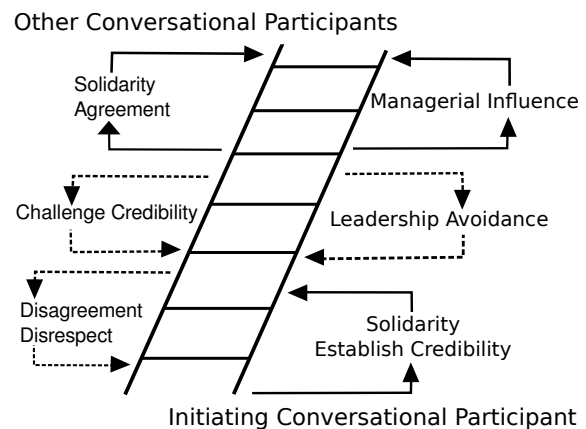
Fig. 3: Ladder model of the path to power. Social acts on the left are directed towards the initial conversational participant (ICP), while those on the right are made by the ICP towards other participants

IV. A MOTIF MODEL FOR INFERRING PURSUIT OF POWER FROM DISCOURSE

The previous section presented an idealized psychological model of how status can shift due to the social implicatures of discourse segments. However, the model does not account for the interactions that can occur between the social acts. For instance, *Leadership Avoidance* shown at the right time can actually establish someone's status through that individual's defiance of another individual in power. Thus it is critical to model the pattern, or structure, of social acts that occur between discourse participants if one is to understand their intentions. Here we present a motif based model for automatically learning to map social intentional structure onto an individual's social goals.

Motifs are widely used the fields of Bioinformatics and Social Network Analysis. In Bioinformatics, a motif is a short segment of DNA that acts as a binding site allowing other
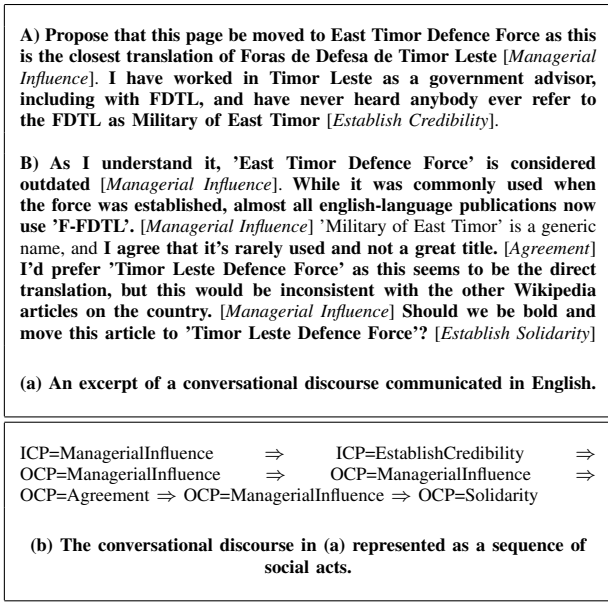
Fig. 4: Example of conversational discourse and its representation as a sequence of social acts.

molecules to bind to the DNA. In Social Network Analysis, motifs are sub-graphs that appear more than normal in a network. Motifs in social networks have been used to identify social efforts, such as leadership and team bonding [34].

The building blocks of the motifs that define the social intentional structure of a discourse are social acts. In particular, motifs generate discourses which are made up of sequences of social acts. A social act sequence is focused around a single ICP and OCP who are interacting with the ICP. The ICP is the individual whom we are determining is or is not pursuing power. A social act sequence consists of the social acts performed in the turns where the ICP is the speaker and turns where OCPs are responding to the ICP. An example of a social act sequence taken from a conversational discourse is shown in figure 4.

The example excerpt, in Figure 4(a), contains two turns in which the individuals are discussing a possible name change for a Wikipedia article. The ICP, person A, begins the conversation by employing Managerial Influence through the proposal of moving the page and follows by Establish Credibility through justifying the move. The group member, or OCP, in the second turn further elaborates on the reasons for the name change and shows Agreement and Solidarity toward the ICP. The sequence of the social acts employed by these two individuals is shown in Figure 4(b).

Motifs are represented as gappy patterns over social act sequences, such as the one in Figure 4(b). They consist of one or more social act between which a gap may exist. Gaps match zero or more social acts regardless of type. Each gap has an associated width, which is the maximum number of social acts it can match. Gaps serve to generalize the motif, by matching any social act, and facilitate the capture of long range dependencies, by allowing zero or more social acts to
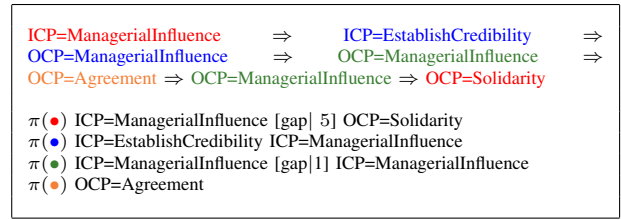
ICP=ManagerialInfluence ⇒ ICP=EstablishCredibility ⇒
OCP=ManagerialInfluence ⇒ OCP=ManagerialInfluence ⇒
OCP=Agreement ⇒ OCP=ManagerialInfluence ⇒ OCP=Solidarity

$\pi(\bullet)$ ICP=ManagerialInfluence [gap| 5] OCP=Solidarity
$\pi(\bullet)$ ICP=EstablishCredibility ICP=ManagerialInfluence
$\pi(\bullet)$ ICP=ManagerialInfluence [gap|1] ICP=ManagerialInfluence
$\pi(\bullet)$ OCP=Agreement

Fig. 5: Example of conversational discourse and its representation as a sequence of social acts.

be captured. An example motif is:

ICP=ManagerialInfluence [gap|5] OCP=Solidarity

The example motif consists of two social acts: ICP=ManagerialInfluence and OCP=EstablishSolidarity. The gap between the two social acts will match any combination of zero to five social acts. The example motif captures the social act sequence listed in Figure 4(b).

Identification of motifs is done through a model that generates social act sequences as a set of gappy patterns. The model is based on the work done by Gimpel and Smith [35] for machine translation. The present work differs from Gimpel and Smith in that we treat the discovery of Motifs as a supervised learning process by defining the probability of a motif as a conditional over a multinomial. The gappy patterns are built by assigning each social act in a sequence a color. Social acts in the sequence assigned the same color become a pattern with gaps added between social acts that are not adjacent in the sequence. More formally, assuming we have a social act sequence with n social acts, $s_{1:n}$, we assign a color to each of the social acts using a vector of color assignment variables $c_{1:n}$. A color C_j is represented as $C_J = \{ i : c_i = j \}$ and the pattern for a color is represented as $\pi(C_j)$. An example of color assignments, $c_{1:n}$, and patterns $\pi(C_j)$ for the example social act sequence in Figure 4(b) is shown in figure 5.

The generative process works by sampling the following:
1) The number of social acts, n, in a social act sequence as a $Poisson$ distribution with parameter β.
2) The number of unique colors in a social act sequence given a $Uniform$ distribution.
3) The color c_i for each social act s_i in the social act sequence as a $Uniform$ distribution.
4) The probability of the pattern associated with each color C_j for sentences with label l as a conditional over a $Multinominal$ with parameter μ.

Thus, to generate patterns we must calculate the probability of generating a social act sequence of length n, with m colors, label l, and color assignments $c_{1:n}$ as

$$p(s_{1:n}, c_{1:n}, m | \beta, \mu) = \frac{1}{Z}(\frac{\beta^n}{n!}e^{-\beta})(\frac{1}{n})(\frac{1}{m})^n \prod_{j=1}^{m} p_\mu(\pi(C_j) \mid l)$$

where Z is a normalization factor.

The multinomial distribution, p_μ, is modeled using a Dirichlet process. A Dirichlet process (DP) can be treated as a

TABLE I: The breakdown of annotations between in-house and Mechanical Turk annotations.

	English			Chinese		
	No. Discourses	No. Participants	Pct. Agreement	No. Discourses	No. Participants	Pct. Agreement
In-House	48	339	76.4%	275	2,806	85.6%
Mechanical Turk	139	1,000	67.5%	121	695	80.0%
Combined	149	1,096	69.5%	375	3,287	82.8%

probability distribution over random distributions. It facilitates an unbounded set of parameters $\mu \sim DP(\alpha, P_0)$, where α is the concentration parameter and P_0 is the base distribution. The base distribution is made up of: (1) a $Poisson$ distribution with parameter υ over the number of social acts in the sequence; (2) a uniform distribution for each social act; (3) a uniform distribution over the number of gaps given the number of social acts; and (4) a uniform distribution over the arrangement of gaps given the number of gaps and social acts.

Collapsed Gibbs sampling is used to sample the posterior distribution $p(\{c^{(i)}, m^{(i)}\}_{i=1}^{K} \mid \{s^{(i)}\}_{i=1}^{K}, \upsilon, \alpha)$, where K is the total number of social act sequences. The Gibbs sampler makes repeated iterations. During each of the iterations it samples a new color for each of the color assignments (c_i). A new color is assigned to c_i by first removing the current color and then choosing from either one of the other m colors in the sentence or a creating a new color. The probability of choosing a new color is proportional to:

$$\frac{N_{\pi(\{i\})} + \alpha P_0(\pi(\{i\}) \mid l)}{N + \alpha}$$

where $N_{\pi(\{i\})}$ is the count of pattern π over all social act sequences with label l and N is the total count of all the patterns. The probability of assigning an existing color j to c_i is proportional to:

$$\frac{N_{\pi(C_j \cup \{i\})} + \alpha P_0(\pi(C_j \cup \{i\}) \mid l)}{N_{\pi(C_j)} + \alpha P_0(\pi(C_j) \mid l)}$$

where $C_j \cup \{i\}$ states that c_i is being added to C_j.

Identification of pursuit of power using motifs is done using logistic regression. Logistic regression is used to learn weights for each motif which denote the importance of the motif in identifying if an individual is in pursuit of power. Thus, an individual is deemed as pursuing power in a discourse when $H(z) = 1$, where $H(z)$ is calculated as:

$$H(z) = \begin{cases} 1, & \frac{1}{1+e^{-z}} > 0.5 \\ 0, & \frac{1}{1+e^{-z}} \leq 0.5 \end{cases}$$

where $z = \sum_{i=1}^{m} w_i * \phi(\pi_i, w_{i:n})$ and ϕ returns 1 if pattern π_i is present in the social act sequence made up of social acts $s_{1:n}$.

V. DATA COLLECTION

Wikipedia is a social media success story. The site allows anyone to edit its encyclopedic content that is used the world over for resolving debates and informing science. Associated with each Wikipedia content page is a WikiTalk page that allows contributors to discuss modifications to the page. Because of the influence of Wikipedia many contributors desire a high-power role on the WikiTalk pages as this allows them control over the site's content and the goals and actions of other contributors. In this paper we focus on identifying individuals that are pursuing these roles inside of the Wikipedia talk forums.

Annotation was performed by both in-house and Mechanical Turk annotators. The in-house annotators went through extensive training for identifying pursuit of power, which included recognizing social acts in the discussions. In contrast, the Mechanical Turk annotators were given a simple definition of pursuit of power with no other training. The group of English annotators consisted of 2 in-house and 93 Mechanical Turkers. The group of Chinese annotators consisted of 2 in-house and 37 Mechanical Turkers. Mechanical Turk annotators had to pass a written qualifying test that judged their ability to understand language and the definition of pursuit of power.

In total 149 discourses containing 1,069 discourse participants in English and 375 discourses containing 3,287 discourse participants in Chinese were multiply annotated. Each discourse participant was annotated by 2 to 7 annotators. We say there is agreement when a super majority of the annotators agree, where a super majority is defined as one minus the total number of annotators for the annotation or all annotators when there are only two annotators. The inter-annotator agreement rates (calculated as percentage agreement) were 69.5% and 82.8% for English and Chinese respectively. The gold standard corpus is created from the portion of the annotations with agreement. The English portion of the corpus consists of 429 participants pursuing power and 314 participants not pursuing power. The Chinese portion of the corpus consists of 1,386 participants pursing power and 1,334 not pursuing power.

VI. EXPERIMENTAL RESULTS

This section presents the experiments testing the effectiveness of motifs for identifying pursuits of power in discourse. Results are obtained by splitting the pursuit of power corpus, described in section V, into 80% for training and 20% testing per language. The motif model is compared to a baseline as well as two other machine learning methods, which use the social acts as features. The accuracy results are presented in Table II.

The baseline method always answers the majority class, which is "Pursuing Power" for both the English and Chinese data. The baseline method has an accuracy of 58% for English and 64.1% for Chinese as seen in Table II. In addition to the simple baseline, we examine using two machine learning

TABLE II: Accuracy results for inferring if an individual is or is not pursuing power.

Model	English	Chinese
Baseline	58.0%	64.1%
Logistic Regression	67.1%	79.9%
SVM	71.5%	**80.8**%
Motif	**84.2**%	80.6%

ICP=Disagreement [gap|1] ICP=EstablishCredibility
ICP=ManagerialInfluence [gap|2] OCP=Disagreement
ICP=ChallengeCredibility [gap|4] ICP=ChallengeCredibility

(a) Motifs for discourses communicated in English.

ICP=ChallengeCredibility [gap|1] ICP=ManagerialInfluence
ICP=ManagerialInfluence [gap|5] OCP=Solidarity
ICP=ManagerialInfluence [gap|2] OCP=Disrespect

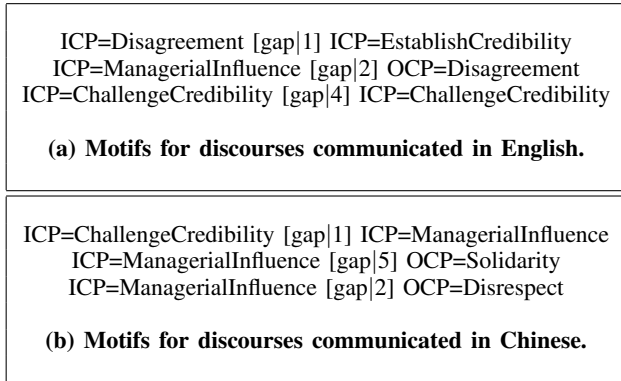**(b) Motifs for discourses communicated in Chinese.**

Fig. 6: Examples of motifs discovered for individuals pursuing power in discourses communicated in English and Chinese.

algorithms, logistic regression and support vector machines (SVM). The features of these models consist of number of times the ICP employed each of the social acts as well as the number of times the OCP employed each of the social acts. These models determine if an ICP is pursuing power by looking at an aggregate of the social actions performed by the ICP and by those responding to the OCP. Both methods, logistic regression and support vector machines, perform better than the baseline. SVMs slightly outperform logistic regression by 4.4% and 0.9% respectively for English and Chinese. These models do not capture the interaction between the ICP and OCP, but are capable of still predicting an ICP's pursuit of power.

The results for the motif model shown in Table II have a maximum gap width of 3, which was determine through experimentation. The motif model outperforms the baseline for identifying pursuits of power by 30.6% and 15.1% respectively for English and Chinese. For discourse communicated in English, the motif model has an increase in accuracy of 12.7% over the SVM. This indicates that the interactions between the ICP and OCP provides valuable information for determining pursuits of power in English communicated discourses. However, in discourse communicated in Chinese the motif model had an accuracy 0.2% below the SVM. This suggests that pursuits of power in Chinese communicated discourses are less about the interactions and more about the sum of the actions. Examples of motifs that highly correlate to pursuits of power for English and Chinese speaking discourse are shown in figure 6.

VII. CONCLUSION

We have introduced a system for the identification of pursuits of power in social discourse. The system focuses on the social intentional structure of discourse. The social intentions of individuals are captured using social acts, which capture the socio-cognitive processes that individuals undergo during communication. These social acts are chained together to form the intentional structure using a motif model. The motifs are built as patterns over the interactions between the ICP for whom we are trying to ascertain is or is not pursuing power and the OCP. We have shown that the motif model of social intentional structure is able to accurately identify pursuits of power with an accuracy of 84.2% and 80.6% in English and Chinese respectively.

ACKNOWLEDGMENT

REFERENCES

[1] J. Allen and M. Core, "Draft of DAMSL: Dialog Act Markup in Several Layers," pp. 1–32, 1997.

[2] S. Jekat, R. Klein, E. Maier, I. Maleck, M. Mast, T. Berlin, J. J. Quantz, and J. J. Quantz, "Dialogue acts in verbmobil," Tech. Rep., 1995.

[3] D. M. Blei, A. Ng, and M. Jordan, "Latent dirichlet allocation," *JMLR*, vol. 3, pp. 993–1022, 2003.

[4] G. Ambwani and A. R. Davis, "Contextually-mediated semantic similarity graphs for topic segmentation," in *Proceedings of the 2010 Workshop on Graph-based Methods for Natural Language Processing*, ser. TextGraphs-5. Stroudsburg, PA, USA: Association for Computational Linguistics, 2010, pp. 60–68. [Online]. Available: http://dl.acm.org/citation.cfm?id=1870490.1870500

[5] J. R. Hobbs, "Coherence and coreference *," *Cognitive Science*, vol. 3, no. 1, pp. 90–67, 1979.

[6] B. J. Grosz and C. L. Sidner, "Attention, Intention, and the Structure of Discourse," *Computational Linguistics*, vol. 12, no. 3, pp. 175–204, 1986. [Online]. Available: http://www1.cs.columbia.edu/~julia/cs4705/grosz-sidner86.pdf

[7] S. Mirkin, J. Berant, I. Dagan, and E. Shnarch, "Recognising Entailment within Discourse," *Computational Linguistics*, no. August, pp. 770–778, 2010. [Online]. Available: http://eprints.pascal-network.org/archive/00007009/

[8] C. Zhang and J. Y. Chai, "Towards Conversation Entailment : An Empirical Investigation," in *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, no. October. Association for Computational Linguistics, 2010, pp. 756–766. [Online]. Available: http://portal.acm.org/citation.cfm?id=1870732

[9] J. Cassell, Y. I. Nakano, T. W. Bickmore, C. L. Sidner, and C. Rich, "Non-verbal cues for discourse structure," in *Proceedings of the 39th Annual Meeting on Association for Computational Linguistics*, ser. ACL '01. Stroudsburg, PA, USA: Association for Computational Linguistics, 2001, pp. 114–123. [Online]. Available: http://dx.doi.org/10.3115/1073012.1073028

[10] S. Eggins and D. Slade, *Analysing casual conversation*. Cassell, 1997.

[11] T. Strzalkowski, G. A. Broadwell, J. Stromer-galley, S. Shaikh, S. Taylor, and N. Webb, "Modeling Socio-Cultural Phenomena in Discourse," in *Internation Conference on Computational Linguistics (Coling)*, no. August, 2010, pp. 1038–1046.

[12] D. B. Bracewell, M. Tomlinson, Y. Shi, J. Bensley, and M. Draper, "Who's playing well with others: Determining collegiality in text," in *Proceedings of the 5th IEEE International Conference on Semantic Computing (ICSC 2011).* Palo Alto: IEEE Computer Society, 2011.

[13] M. Tomlinson, D. Bracewell, M. Draper, Z. Almissour, Y. Shi, and J. Bensley, "Pursing power in arabic on-line discussion forums," *Proceedings of LREC2012*, pp. 1359–1364, 2012.

[14] A. Agarwal and O. Rambow, "Automatic Detection and Classification of Social Events," *Computational Linguistics*, no. October, pp. 1024–1034, 2010.

[15] R. Brendel and H. Krawczyk, "Detection of Roles of Actors in Social Networks Using the Properties of Actors' Neighborhood Structure," in *2008 Third International Conference on Dependability of Computer Systems DepCoS-RELCOMEX.* IEEE, 2008, pp. 163–170. [Online]. Available: http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=4573053

[16] D. Fisher, M. Smith, and H. T. Welser, "You Are Who You Talk To: Detecting Roles in Usenet Newsgroups," *Proceedings of the 39th Annual Hawaii International Conference on System Sciences HICSS06*, vol. 00, no. C, pp. 59b–59b, 2006. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1579410

[17] M. Zancanaro, B. Lepri, and F. Pianesi, *Automatic detection of group functional roles in face to face interactions.* New York, New York, USA: ACM Press, Nov. 2006. [Online]. Available: http://portal.acm.org/citation.cfm?id=1180995.1181003

[18] D. Gatica-Perez, "Automatic nonverbal analysis of social interaction in small groups: A review," *Image and Vision Computing*, vol. 27, no. 12, pp. 1775–1787, Nov. 2009. [Online]. Available: http://dx.doi.org/10.1016/j.imavis.2009.01.004

[19] P. Bramsen, M. Escobar-Molano, A. Patel, and R. Alonso, "Extracting social power relationships from natural language," in *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies - Volume 1*, ser. HLT '11. Stroudsburg, PA, USA: Association for Computational Linguistics, 2011, pp. 773–782. [Online]. Available: http://dl.acm.org/citation.cfm?id=2002472.2002570

[20] D. C. McClelland, *Human motivation.* Cambridge University Press, 1987, vol. 23. [Online]. Available: http://books.google.co.uk/books?id=vic4AAAAIAAJ

[21] M. Van Dijke and M. Poppe, "Motivations Underlying Power Dynamics in Hierarchically Structured Groups," *Small Group Research*, vol. 38, no. 6, pp. 643–669, 2007. [Online]. Available: http://sgr.sagepub.com/cgi/doi/10.1177/1046496407304339

[22] J. R. P. French and B. Raven, "The Bases of Social Power," in *Studies in Social Power*, ser. Studies in social power, D. Cartwright, Ed. Institute for social research, 1959, vol. 35, ch. 9, pp. 150–167. [Online]. Available: http://books.google.com/books?hl=en\&lr=\&id=DL2AsvuzJHUC\&oi=fnd\&pg=PA61\&dq=The+Bases+of+Social+Power\&ots=hNBD1vsfeJ\&sig=XFbHBUpGpaHoBPv9uHgo4vI02Cs

[23] G. Boella, R. Damiano, and L. Lesmo, "Social goals in conversational cooperation," in *Proceedings of the 1st SIGdial workshop on Discourse and dialogue - Volume 10*, ser. SIGDIAL '00. Stroudsburg, PA, USA: Association for Computational Linguistics, 2000, pp. 84–93. [Online]. Available: http://dx.doi.org/10.3115/1117736.1117746

[24] R. Barzilay and M. Lapata, "Modeling local coherence: an entity-based approach," in *Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics*, ser. ACL '05. Stroudsburg, PA, USA: Association for Computational Linguistics, 2005, pp. 141–148. [Online]. Available: http://dx.doi.org/10.3115/1219840.1219858

[25] D. Byron and A. Stent, "A preliminary model of centering in dialog," in *Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics - Volume 2*, ser. ACL '98. Stroudsburg, PA, USA: Association for Computational Linguistics, 1998, pp. 1475–1477. [Online]. Available: http://dx.doi.org/10.3115/980691.980811

[26] W. C. Mann and S. A. Thompson, "Rhetorical Structure Theory: Toward a Functional Theory of Text Organization," *Text*, vol. 8, no. 3, pp. 243–281, 1988. [Online]. Available: http://www.cis.upenn.edu/~nenkova/Courses/cis700-2/rst.pdf

[27] C. Anderson, O. P. John, D. Keltner, and A. M. Kring, "Who attains social status? Effects of personality and physical attractiveness in social groups." *Journal of Personality and Social Psychology*, vol. 81, no. 1, pp. 116–132, 2001. [Online]. Available: http://doi.apa.org/getdoi.cfm?doi=10.1037/0022-3514.81.1.116

[28] D. Keltner, G. A. Van Kleef, S. Chen, and M. W. Kraus, "A reciprocal influence model of social power: Emerging principles and lines of inquiry," *Advances in experimental social psychology*, vol. 40, pp. 151–192, 2008. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S0065260107000032

[29] D. Owens and R. Sutton, "Status contests in meetings: Negotiating the informal order," *Groups at work: Theory and research*, vol. 14, pp. 299–316, 2001. [Online]. Available: http://books.google.com/books?hl=en\&lr=\&id=usasMjZdpugC\&oi=fnd\&pg=PA299\&dq=Status+Contests+in+Meetings:+Negotiating+the+Informal+Order\&ots=S40IxKI5PD\&sig=tC3jBaNnjAd-_1P7WDH-vdypRa8

[30] P. Smith and A. Galinsky, "The nonconscious nature of power: Cues and consequences," *Social and Personality Psychology Compass*, vol. 4, no. 10, pp. 918–938, 2010. [Online]. Available: http://onlinelibrary.wiley.com/doi/10.1111/j.1751-9004.2010.00300.x/full

[31] M. B. Brewer and W. Gardner, "Who is this "We"? Levels of collective identity and self representations." *Journal of Personality and Social Psychology*, vol. 71, no. 1, pp. 83–93, 1996. [Online]. Available: http://doi.apa.org/getdoi.cfm?doi=10.1037/0022-3514.71.1.83

[32] M. Deutsch, "Cooperation and competition," *Conflict, Interdependence, and Justice*, pp. 23–40, 2011. [Online]. Available: http://www.springerlink.com/index/G413J25074750R1L.pdf

[33] B. Ford, "Parsing expression grammars," *ACM SIGPLAN Notices*, vol. 39, no. 1, pp. 111–122, Jan. 2004. [Online]. Available: http://dl.acm.org/citation.cfm?id=982962.964011

[34] W. Richards and O. Macindoe, "Characteristics of small social networks," MIT-CSAIL, Tech. Rep., 2010.

[35] K. Gimpel and N. A. Smith, "Generative Models of Monolingual and Bilingual Gappy Patterns," *Proceedings of the Sixth Workshop on Statistical Machine Translation*, no. 2009, pp. 512–522, 2011. [Online]. Available: http://www.aclweb.org/anthology/W11-2165