Weakly Supervised Role Identification in Teamwork Interactions

Diyi Yang
School of Computer Science
Carnegie Mellon University
Pittsburgh, PA, 15213, USA
diyiy@cs.cmu.edu

Miaomiao Wen
School of Computer Science
Carnegie Mellon University
Pittsburgh, PA, 15213, USA
mwen@cs.cmu.edu

Carolyn Penstein Rosé
School of Computer Science
Carnegie Mellon University
Pittsburgh, PA, 15213, USA
cprose@cs.cmu.edu

Abstract

In this paper, we model conversational roles in terms of distributions of turn level behaviors, including conversation acts and stylistic markers, as they occur over the whole interaction. This work presents a lightly supervised approach to inducing role definitions over sets of contributions within an extended interaction, where the supervision comes in the form of an outcome measure from the interaction. The identified role definitions enable a mapping from behavior profiles of each participant in an interaction to limited sized feature vectors that can be used effectively to predict the teamwork outcome. An empirical evaluation applied to two Massive Open Online Course (MOOCs) datasets demonstrates that this approach yields superior performance in learning representations for predicting the teamwork outcome over several baselines.

1 Introduction

In language technologies research seeking to model conversational interactions, modeling approaches have aimed to identify conversation acts (Paul, 2012; Wallace et al., 2013; Bhatia et al., 2014) on a per turn basis, or to identify stances (Germesin and Wilson, 2009; Mukherjee et al., 2013; Piergallini et al., 2014; Hasan and Ng, 2014) that characterize the nature of a speaker's orientation within an interaction over several turns. What neither of these two perspectives quite offer is a notion of a conversational role. And yet, conversational role is a concept with great utility in current real world applications where language technologies may be applied.

Important teamwork is achieved through collaboration where discussion is an important medium for accomplishing work. For example, distributed work teams are becoming increasingly the norm in the business world where creating innovative products in the networked world is a common practice. This work requires the effective exchange of expertise and ideas. Open source and open collaboration organizations have successfully aggregated the efforts of millions of volunteers to produce complex artifacts such as GNU/Linux and Wikipedia. Discussion towards decision making about how to address problems that arise or how to extend work benefit from effective conversational interactions. With a growing interest in social learning in large online platforms such as Massive Open Online Courses (MOOCs), students form virtual study groups and teams to complete a course project, and thus may need to coordinate and accomplish the work through discussion. In all such environments, discussions serve a useful purpose, and thus the effectiveness of the interaction can be measured in terms of the quality of the resulting product.

We present a modeling approach that leverages the concept of latent conversational roles as an intermediary between observed discussions and a measure of interaction success. While a stance identifies speakers in terms of their positioning with respect to one another, roles associate speakers with rights and responsibilities, associated with common practices exhibited by performers of that role within an interaction, towards some specific interaction outcome. That outcome may be achieved through strategies characterized in terms of conversation acts or language with particular stylistic characteristics. However, individual acts by themselves lack the power to achieve a complex outcome. We argue that roles make up for this decontextualized view of a conversational contribution by identifying distributions of conversation acts and stylistic features as behavior profiles indicative of conversational roles. These

profiles have more explanatory power to identify strategies that lead to successful outcomes.

In the remainder of the paper we first review related work that lays the foundation for our approach. Then we describe a series of role identification models. Experimental results are analyzed quantitatively and qualitatively in Section 4, followed by conclusions and future work.

2 Related Work

The concept of *social role* has long been used in social science fields to describe the intersection of behavioral, symbolic, and structural attributes that emerge regularly in particular contexts. Theory on coordination in groups and organizations emphasizes role differentiation, division of labor and formal and informal management (Kittur and Kraut, 2010). However, identification of roles as such has not had a corresponding strong emphasis in the language technologies community, although there has been work on related notions. For example, there has been much previous work modeling disagreement and debate framed as stance classification (Thomas et al., 2006; Walker et al., 2012). Another similar line of work studies the identification of personas (Bamman et al., 2013; Bamman et al., 2014) in the context of a social network, e.g. celebrity, newbie, lurker, flamer, troll and ranter, etc, which evolve through user interaction (Forestier et al., 2012).

What is similar between stances and personas on the one hand and roles on the other is that the unit of analysis is the person. On the other hand, they are distinct in that stances (e.g., liberal) and personas (e.g., lurker) are not typically defined in terms of what they are meant to accomplish, although they may be associated with kinds of things they do. Teamwork roles are defined in terms of what the role holder is meant to accomplish.

The notion of a natural outcome associated with a role suggests a modeling approach utilizing the outcome as light supervision towards identification of the latent roles. However, representations of other notions such as stances or strategies can similarly be used to predict outcomes. Cadilhac et al. maps strategies based on verbal contributions of participants in a win-lose game into a prediction of exactly which players, if any, trade with each other (Cadilhac et al., 2013). Hu et al. (Hu et al., 2009) predict the outcome of featured article nominations based on user activeness, discussion consensus and user co-review relations. In other work, the authors of (Somasundaran and Wiebe, 2009) adopt manually annotated characters and leaders to predict which participants will achieve success in online debates. The difference is the interpretation of the latent constructs. The latent construct of a role, such as team leader, is defined in terms of a distribution of characteristics that describe how that role should ideally be carried out. However, in the case of stances, the latent constructs are learned in order to distinguish one stance from another or in order to predict who will win. This approach will not necessarily offer insight into what marks the most staunch proponents of a stance, but instead distinguish those proponents of a stance who are persuasive from those who are not.

Roles need not only be identified with the substance of the text uttered by role holders. Previous work discovers roles in social networks based on the network structure (Hu and Liu, 2012; Zhao et al., 2013). Examples include such things as mixed membership stochastic block-models (MMSB) (Airoldi et al., 2008), similar unsupervised matrix factorization methods (Hu and Liu, 2012), or semi-supervised role inference models (Zhao et al., 2013). However, these approaches do not standardly utilize an outcome as supervision to guide the clustering.

Many open questions exist about what team roles and in what balance would make the ideal group composition (Neuman et al., 1999), and how those findings interact with other contextual factors (Senior, 1997; Meredith Belbin, 2011). Thus, a modeling approach that can be applied to new contexts in order to identify roles that are particularly valuable given the context would potentially have high practical value.

3 Role Identification Models

The context of this work is team based MOOCs using the NovoEd platform. In this context, we examine the interaction between team members as they work together to achieve instructional goals in their project work. Our modeling goal is to identify behavior profiles that describe the emergent roles that team members take up in order to work towards a successful group grade for their team project. Identification of effective role based behavior profiles would enable work towards supporting effective team formation in subsequent

work. This approach would be similar to prior work where constraints that describe successful teams were used to group participants into teams in which each member's expertise is modeled so that an appropriate mixture of expertise can be achieved in the assignment (Anagnostopoulos et al., 2010).

In this section, we begin with an introduction of some basic notations. Then we present an iterative model, which involves two stages: teamwork quality prediction and student role matching. Furthermore, we generalize this model to a constrained version which provides more interpretable role assignments. In the end, we describe how to construct student behavior representations from their teamwork collaboration process.

3.1 Notation

Suppose we have C teams where students collaborate to finish a course project together. The number of students in the j-th team is denoted as $N_j, (1 \leq j \leq N_j)$. There are K roles across C teams that we want to identify, where $1 \leq K \leq N_j, \forall j \in [1, C]$. That is, the number of roles is smaller than or equal to the number of students in a team, which means that each role should have one student assigned to it, but not every student needs to be assigned to a role. Each role is associated with a weight vector $W_k \in \mathcal{R}^D$ to be learned, $1 \leq k \leq K$ and D is the number of dimensions. Each student i in a team j is associated with a behavior vector $B_{j,i} \in \mathcal{R}^D$. The measurement of teamwork quality is denoted as Q_j for team j, and $\hat{Q}_j$ is the predicted quality. Here, $\hat{Q}_j$ is determined by the inner product of the behavior vectors of students who are assigned to different roles and the corresponding weight vectors.

Teamwork Role Identification Our goal is to find a proper teamwork role assignment that positively contributes to the collaboration outcome as much as possible.

3.2 Role Identification

Here we describe our role identification model. Our role identification process is iterative and involves two stages. The first stage adjusts the weight vectors to predict the teamwork quality, given a fixed role assignment that assumes students are well matched to roles; the second stage iterates the possible assignments and finds a matching to maximize our objective measure. The
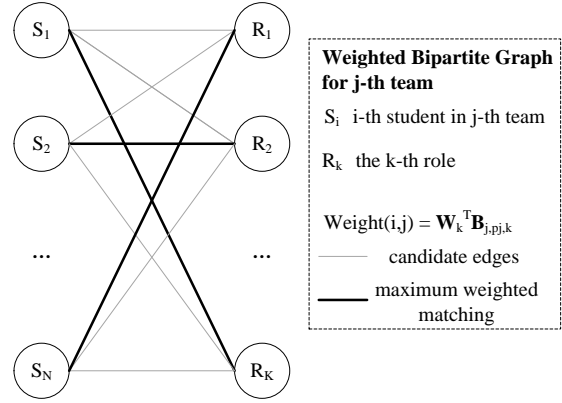

Figure 1: Weighted Bipartite Graph for a Team

two stages run iteratively until both role assignment and teamwork quality prediction converge.

Teamwork Quality Prediction: Given the identified role assignment, i.e. we know who is assigned to which roles in a team, the focus is to accurately predict the teamwork quality under this role assignment. $p_{j,k}$ refers to the student who is assigned to role k in team j. We minimize the following objective function to update the role weight vector W:

$$\min_{W} \frac{1}{2} \sum_{j=1}^{C} (Q_j - \sum_{k=1}^{K} W_k{}^T \cdot B_{j,p_{j,k}}) + \lambda \cdot \|W\|^2 \quad (1)$$

Here, λ is the regularization parameter; large λ leads to higher complexity penalization. To give the optimal solution to Equation 1, which is a classical ridge regression task (Hoerl and Kennard, 2000), we can easily compute the optimal solution by its closed form representation, as shown in the Algorithm 1.

Matching Members to Roles: Once the weight vector W is updated, we iterate over all the possible assignments and find the best role assignment, where the goal is to maximize the predicted teamwork quality since we want our assignment of students and roles to be associated with improvement in the quality of teamwork. The complexity of brute-force enumeration of all possible role assignments is exponential. To avoid such an expensive computational cost, we design a weighted bipartite graph and apply a maximum weighted matching algorithm (Ravindra et al., 1993) to find the best matching under the objective of maximizing $\sum_{j=1}^{C} \hat{Q}_j$. Because this objective is a summation, we can further separate it into C iso-

Algorithm 1: Role Identification

1 Heuristicly initialize the role assignment $p_{j,k}$
2 **while** *assignments have not converged* **do**
  // Teamwork Quality Prediction
3    $X \leftarrow$ a $C \times (K \cdot D)$ matrix
4   **for** $j = 1$ *to* C **do**
5      $X_{j,*} \leftarrow (B_{j,p_{j,1}}, B_{j,p_{j,2}}, \ldots, B_{j,p_{j,K}})$
  // optimal solution to Eq. 1
6    $(W_1, \ldots, W_C) \leftarrow (X^T X + \lambda I)^{-1} X^T Q$
  // Student and Role Matching
  // maximize $\sum_j \hat{Q}_j$
7   **for** $j = 1$ *to* C **do**
8      $(p_{j,*}) \leftarrow$ maximum weighted bipartite matching on Figure 1

lated components for C teams by maximizing each $\hat{Q}_j$. For each team, a weighted bipartite graph is created as specified in Figure 1. By applying the maximum weighted matching algorithm on this graph, we can obtain the best role assignment for each team.

The two stage role identification model is solved in detail in Algorithm 1.

3.3 Role Identification with Constraints

The above role identification model puts no constraints on the roles that we want to identify in teamwork. This might result in more effort to explain how different roles collaborate to produce the teamwork success. Therefore, we introduce a constrained role identification model, which is able to integrate external constraints on roles. For example, we can require our extracted role set to contain a role that makes a positive contribution to the project success and a role that contributes relatively negatively, instead of extracting several generic roles. To address such constraints, in the stage of teamwork quality prediction, we reformulate the Equation 2 as follows:

$$
\begin{aligned}
L = \frac{1}{2} \sum_{j=1}^{C} (Q_j - \sum_{k=1}^{K} W_k^T \cdot B_{j,p_{j,k}}) + \lambda \|W\|^2 \\
- \mu_+ \sum_{k \in S_+} \sum_{d=1}^{D} log(W_{kd}) \\
- \mu_- \sum_{k \in S_-} \sum_{d=1}^{D} log(-W_{kd})
\end{aligned}
$$

(2)

Algorithm 2: Identification with Constraints

1 Heuristicly initialize the role assignment $p_{j,k}$
2 **while** *assignments have not converged* **do**
  // Teamwork Quality Prediction
3    $X \leftarrow$ a $C \times (K \cdot D)$ matrix
4   **for** $j = 1$ *to* C **do**
5      $X_j \leftarrow (B_{j,p_{j,1}}, B_{j,p_{j,2}}, \ldots, B_{j,p_{j,K}})$
  // gradient descent solution to Eq. 2
6    $\mu_+, \mu_- \leftarrow$ large enough values
7   **while** $\mu_+, \mu_- > \epsilon$ **do**
8   **while** *not converge* **do**
9   **for** $k = 1$ *to* K **do**
10          $W_k \leftarrow W_k - \eta \cdot \frac{\partial L}{\partial W_k}$
11      $\mu_+ \leftarrow \theta \cdot \mu_+$
12      $\mu_- \leftarrow \theta \cdot \mu_-$
  // Students and Roles Matching
  // maximize $\sum_j \hat{Q}_j$
13   **for** $j = 1$ *to* C **do**
14      $(p_{j,*}) \leftarrow$ maximum weighted bipartite matching on Figure 1

The external constraints are handled by the log barrier terms (Ahuja et al., 1993). Here, μ_+ and μ_- are positive parameters used to penalize the violation of role constraints. S_+ is the set of roles that we want to assign students who contribute positively to the group outcome (i.e. above average level), and S_- contains the roles that we want to capture students who contribute negatively to the group outcome (i.e. below average level). The solving of Equation 2 cannot directly apply the previous ridge regression algorithm, thus we use the Interior Point Method (Potra and Wright, 2000) to solve it. The detailed procedure is illustrated in Algorithm 2, where the θ is a constant to control the shrinkage and η is the learning rate.

3.4 Behavior Construction

One essential component in our teamwork role identification models is the student behavior representation. To some extent, a proper behavior representation is essential for facilitating the interpretation of identified roles. We construct the representation of student behavior from the following feature types:

Team Member Behaviors: How a team functions can be reflected in their team communication messages. To understand how students collaborate

Type	Behavior Definition	Example Messages
Team Building	Invite or accept users to join the group	*Lauren, We would love to have you.* *Jill and I are both ESL specialists in Boston.*
Task Management	Initiate a task or assign subtask to a team member	*Housekeeping Task 3 is optional but below are the questions I summarize and submit for our team.*
Collaboration	Collaborate with teammates, provide help or feedback	*I figured out how to use the Google Docs.* *Let's use it to share our lesson plans.*

Table 1: Three Different Types of Team Member Behaviors

to contribute to teamwork success, we identified three main team member behaviors based on messages sent between team members as shown in Table 1. These annotations, which came from prior qualitative work analysing discussion contributions in the same dataset (Wen et al., 2015), are used to define component behaviors in this work. We design four variables to characterize the above collaboration behaviors:

1. *Collaboration*: the number of Collaboration messages sent by this team member.

2. *Task Management*: the number of Task Management messages sent by this team member.

3. *Team Building*: the number of Team Building messages sent by this team member.

4. *Other Strategies*: the number of messages that do not belong to the listed behavior categories.

Communication Languages: Teams that work successfully typically exchange more knowledge and establish good social relations. To capture such evidence that is indicated in the language choice and linguistic styles of each team member, we design the following features:

5. *Personal Pronouns*: the proportion of first person and second person pronouns.

6. *Negation*: counts of negation words.

7. *Question Words*: counts of question related words in the posts, e.g. why, what, question, problem, how, answer, etc.

8. *Discrepancy*: number of occurrences of words, such as should, would, could, etc as defined in LIWC (Tausczik and Pennebaker, 2010).

9. *Social Process*: number of words that denote social processes and suggest human interaction, e.g. talking, sharing, etc.

10. *Cognitive Process*: number of occurrences of words that reflect thinking and reasoning, e.g. cause, because, thus, etc.

11-14. *Polarity*: four variables that measure the portion of Positive, Negative, Neutral, Both polarity words (Wilson et al., 2005) in the posts.

15-16. *Subjectivity*: two count variables of occurrences of Strong Subjectivity words and Weak Subjectivity words.

Activities: We also introduce several variables to measure the activeness level of team members.

17-18. *Messages*: two variables that measure the total number of messages sent, and the number of tokens contained in the messages.

19-20. *Videos*: the number of videos a student has watched and total duration of watched videos.

21. *Login Times*: times that a student logins to the course.

4 Experiments

In this section, we begin with the dataset description, and then we compare our models with several competitive baselines by performing 10-fold cross validation on two MOOCs, followed by a series of quantitative and qualitative analyses.

4.1 Dataset

Our datasets come from a MOOC provider NovoEd, and consist of two MOOC courses. Both courses are teacher professional development courses about Constructive Classroom Conversations; one is in elementary education and another is about secondary education. Students in a NovoEd MOOC have to initiate or join a team in the beginning of the course. A NovoEd team homepage consists of blog posts, comments and other content shared within the group. The performance measure we use is the final team project score, which is in the range of 0 to 40. There are 57 teams (163 students) who survived until the end in the Elementary education course, and 77 teams (262 students) who survived for the Secondary course. The surviving teams are the ones in which none of the team members dropped out of the course, and who finished all the course requirements. For the purpose of varying teamwork roles K, we only keep the teams with

at least 3 members. Self-identified team leader are labeled in the dataset.

4.2 Baselines

We propose several baselines to extract possible roles and predict the teamwork quality for comparison with our models. Preprocessing is identical for baselines as for our approach.

Top K Worst/Best: The worst performing student is often the bottleneck in a team, while the success of a team project largely depends on the outstanding students. Therefore, we use the top K worst/best performing students as our identified K roles. Their behavior representation are then used to predict the teamwork quality. The performing scores are only accessible after the course.

K-Means Clustering: Students who are assigned to the same roles tend to have similar activity profiles. To capture the similarities of student behavior, we adopt a clustering method to group students in a team into K clusters, and then assign students to roles based on their distances to the centroid of clusters. Prediction is then performed on the basis of those corresponding behavior vectors. Here, we use K-Means method for clustering. That is, each cluster is a latent representation of a role and each student is assigned to its closest cluster (role).

Leader: Leaders play important roles for the smooth functioning of teams, and thus might have substantial predictive power of team success. We input our role identification model with only the identified leader's behavior representation and conduct our role identification algorithm as illustrated in Algorithm 1. Each team in our courses have a predefined leader.

Average: The average representation of all team members is a good indication of team ability level and thus teamwork success. Here, we average all team members' behavior feature vectors and use that to predict the teamwork quality.

4.3 Teamwork Quality Prediction Results

The purpose of our role identification is to find a role assignment that minimizes the prediction error, thus we measure the performance of our models using RMSE (Rooted Mean Square Error). 10-fold Cross Validation is employed to test the overall performance. Table 2 and Table 3 presents the results of our proposed models and baselines

on our two courses. Our role identification model shown in Algorithm 1, is denoted as **RI**. θ is set as 0.9 and we vary the role number K from 1 to 3 in order to assess the added value of each additional role over the first one.

4.3.1 Who Matters Most In a Team

If we set the number of roles K as 1, what will the role identification pick as the most important person to the teamwork outcome? From Table 2 and 3, we find that, RI performs better than Leader, and either Top K Best gives a good RMSE in one course and Top K Worst gives a good RMSE in the other course. This indicates that, the predefined leader is not always functioning well in facilitating the teamwork, thus we need a more fair mechanism to select the proper leading role. Besides, Top K worst has quite good performance on the Elementary course, which reflects that the success of a teamwork is to some extent dependent on the worst performing student in that team. The best performing student matters for the teamwork outcome on the Secondary course.

4.3.2 Multi-Role Collaboration

From Table 2 and 3, in the setting of K=3, RI achieved better results compared to Top K Best, Top K Worst and K-means methods. One explanation is that our RI model not only considers individual student's behaviors, but also takes into account the collaboration patterns through all teamwork. Besides, RI achieves better performance compared to our baselines as K becomes larger. We also noticed that Top K Best gives a quite good approximation to the teamwork quality on both courses. However, such performing scores that are used to rank students are not accessible until the course ends, and have high correlation with team score. Thus an advantage of our RI model is that it does not make use of that information. Compared with all other results, our RI has a good generalization ability, and achieves both a smallest RMSE of around 10 across both MOOCs.

4.4 Role Assignment Validation

We demonstrate the predicative power of our identified roles to team success above. In this part, we interpret the identified roles guided by different constraints in a team qualitatively, and show how different roles are distributed in a team, how each role contributes to teamwork, and how collaboration happens among the roles.

Table 2: RMSE Comparison of Different Methods on the Elementary Course

	Average	Leader	K-Means	K Worst	K Best	RI	RIC	RIC$_-$	RIC$_+$
K = 1	13.945	16.957	14.212	13.092	20.464	14.982	N/A	N/A	N/A
K = 2	N/A	N/A	13.160	13.428	15.591	11.581	N/A	N/A	N/A
K = 3	N/A	N/A	12.291	15.460	14.251	**9.517**	10.486	27.314	10.251

Table 3: RMSE Comparison of Different Methods on the Secondary Course

	Average	Leader	K-Means	K Worst	K Best	RI	RIC	RIC$_-$	RIC$_+$
K = 1	12.571	15.611	12.583	17.899	10.886	13.297	N/A	N/A	N/A
K = 2	N/A	N/A	12.288	19.268	11.245	10.435	N/A	N/A	N/A
K = 3	N/A	N/A	11.218	22.933	14.079	**10.143**	10.961	24.583	10.427

4.4.1 Constraint Exploration

By incorporating constraints into the role identification process, we expect to guide the model using human intuition such that the results will be more interpretable, although the prediction error might increase because of the limitation of the search space. We present three alternative possible constrained models here. The **RIC** model emphasizes picking one best member, one worst member and another generic member, which is achieved by putting one role to S$_+$ and one to S$_-$ as defined in Equation 2. **RIC$_+$** aims at picking three best team members who collaborate to make the best contribution to the team success, achieved by putting three roles into S$_+$. Similarly, **RIC$_-$** rewards poorly performing students to contribute to teamwork quality, i.e. putting all roles into S$_-$.

Based on results shown in Table 2 and 3, we found that RIC$_+$ and RIC work similar as RI even though RI is slightly better. RIC$_-$ gives quite unsatisfying performance which shows that examining the behavior of a set of poorly performing students is not very helpful in predicting teamwork success. The comparison of RIC$_+$ and RIC$_-$ can be shown clearly in Figure 2, which presents the behavior representation of each role identified by RIC$_+$ and RIC$_-$. Obviously, RIC$_+$ produces positive roles that contribute largely to the teamwork quality across all feature dimensions; such behaviors are what we want to encourage. Those identified roles are diverse and not symmetrical because each role achieves peaks at different feature dimensions. On the contrary, roles identified by RIC$_-$ works negatively towards teamwork quality and they have homogeneous behavior representation curves. Therefore, our constrained models can provide much interpreta-

tion, with a little loss of accuracy compared to RI.

4.4.2 Role Assignment Interpretation

Leading Role Validation: As a validation, we found that one of our identified roles has substantial overlap with team leaders. For instance, in the Elementary course, around 70% of students who are assigned to Role 0 are actual leaders for RIC and RIC$_+$ models. On the Secondary course, around 86% students who are in the position of Role 0 are real team leaders. When it comes to RIC$_-$, such ratio drops to around 2% for all roles. This validates the ability of our models in producing role definitions that make sense.

Information Diffusion: Figure 3 compares the information diffusion among different identified roles of RI, RIC, RIC$_+$ and RIC$_-$. The darker the node, the better grade it achieves. The number associated with each role indicates the average final grades (scale 0-100) of all students who are assigned to this role. The edge represents how many messages sent from one node to another. The thicker the edge, the more information it conveys. From the figure, we found that, RI performs similarly with RIC and roles in RIC$_+$ have much higher grades compared to RIC$_-$. One explanation is that RIC actually does not incorporate many constraints and is less interpretable compared to RIC$_+$ and RIC$_-$. As shown in (c), RIC$_+$ Role 0 contributes more information to Role 1 with an average of 5.5 messages and to Role 2 with weight 6.1. Role 1 and Role 2 also have many messages communicated with others in their team. However, less communication happens in RIC$_-$ roles. This comparison comes much easier when it comes to each role's behaviors on different normalized feature representations as shown in Figure 2 for
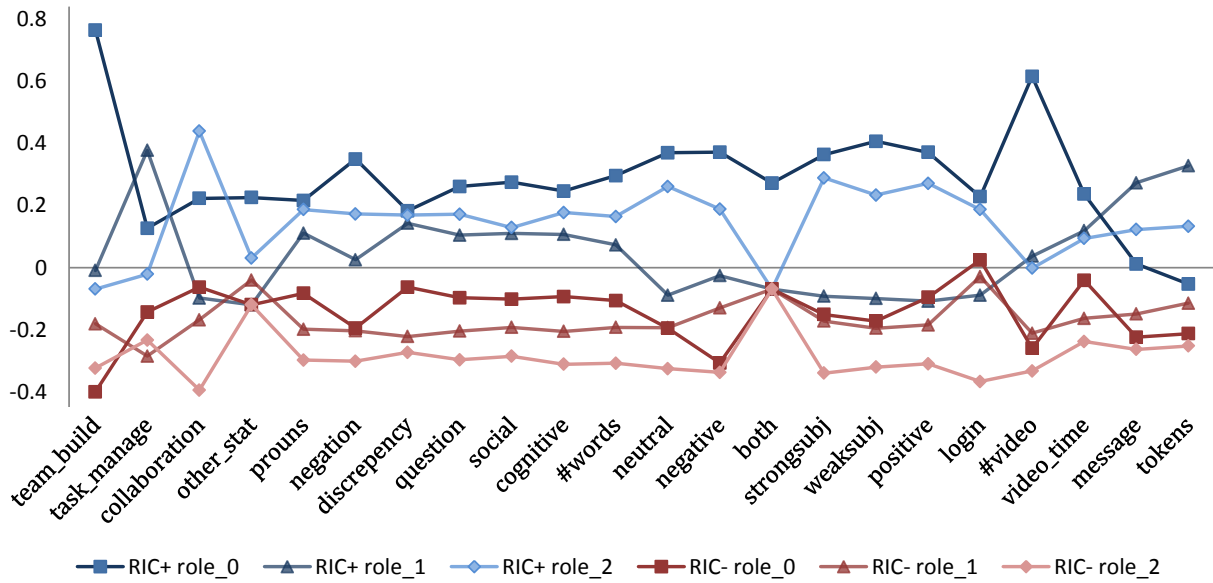
Figure 2: Beahvior Representation of Each Role on the Secondary Course

	Typical Behavior	Representative Post
RIC$_+$	Team Building	I *started a new doc* ... Let me *know your email* if you didn't get the invite.
	Positive	*Great* job team!! Our lesson plan is *amazing* and I learned so much ...
	Collaboration	We plan to meet on Monday to *figure out exactly how to complete* the assignment ...
	Task Management	Here's *what I propose*: *1) to save time, use* ... *2) Tara, do you have plans* ... *3) once* a lesson plan outline is up, *we can* each go in and add modifications..
RIC$_-$	Negative	I'm *confused*. I answered all the questions ... and I didn't see ...
	Strong Subjectivity	I *like* the recycling lesson ... *feeling* so dumb.. really *confused* by Google Docs...
	Negation	I'm *not* able to ... the pictures *don't* show up...I *don't understand* how to create a link..

Table 4: Representative Posts and Corresponding Behavior Feature Comparison on the Secondary Course

RIC$_+$ and RIC$_-$ models. It can be concluded that by incorporating rewarding and penalizing constraints, our model works effectively in picking the behavior profiles we want to encourage and avoid in a teamwork.

Behavior Comparison: Table 4 presents several representative posts and their corresponding behavior features for our identified roles. Most features shown in Table 4 correspond to the peak behaviors associated with roles in Figure 2, which is consistent with our previous interpretation. For example, RIC$_+$ picks the well performing student who adds calmness to the teamwork as indicated by using positive words and adopting collaborative strategies. On the contrary, RIC$_-$ reflects a less cooperative teamwork, such as strong subjectivity, negation and negativity indicated in their posts.

In summary, our role identification models provide quite interpretable identified roles as discussed above, as well as accurate prediction of

teamwork quality. More interpretability can be achieved by incorporating intuitive constraints and sacrificing a bit of accuracy.

5 Conclusion

In this work, we propose a role identification model, which iteratively optimizes a team member role assignment that can predict the teamwork quality to the utmost extent. Furthermore, we extend it to a general constrained version that enables humans to incorporate external constraints to guide the identification of roles. The experimental results on two MOOCs show that both of our proposed role identification models can not only perform accurate predictions of teamwork quality, but also provide interpretable student role assignment results ranging from leading role validation to information diffusion.

Even though we have only explored up to 3 roles in this work that would enable us to use most
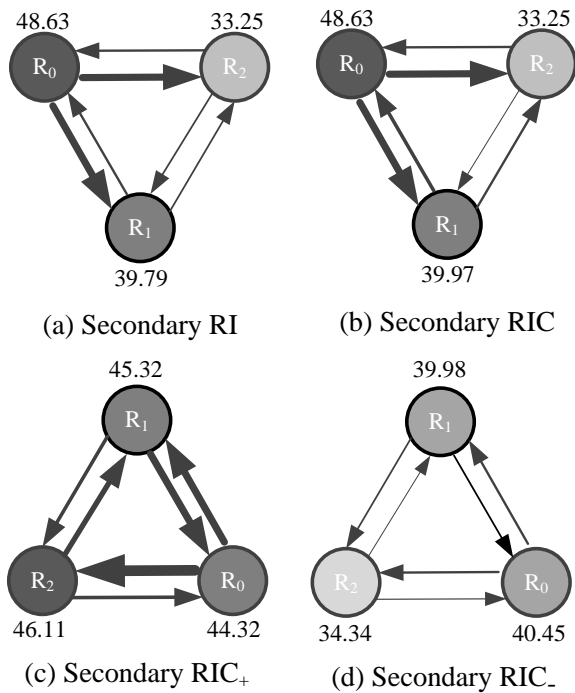
(a) Secondary RI   (b) Secondary RIC

(c) Secondary RIC$_+$ (d) Secondary RIC$_-$

Figure 3: Information Diffusion among Roles

of our data, our role identification method is capable to experiment with a larger range of values of K, such as in the context of Wikipedia (Ferschke et al., 2015). Furthermore, our model can be directly applied to other online collaboration scenarios to help identify the roles that contribute to collaboration, not limited in the context of MOOCs. In the future, we are interested in relaxing the assumptions that people can take only one role and roles are taken up by only one person and incorporating mixed membership role matching strategies into our method. Furthermore, nonlinear relationship between roles and performance as well as the dependencies between roles should be explored. Last but not least, we plan to take advantage of our identified roles to provide guidance and recommendation to those weakly performing teams for better collaboration and engagement in online teamworks.

Acknowledgement

References

Ravindra K. Ahuja, Thomas L. Magnanti, and James B. Orlin. 1993. *Network Flows: Theory, Algorithms, and Applications*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA.

Edoardo M. Airoldi, David M. Blei, Stephen E. Fienberg, and Eric P. Xing. 2008. Mixed membership stochastic blockmodels. volume 9, pages 1981–2014. JMLR.org, June.

Aris Anagnostopoulos, Luca Becchetti, Carlos Castillo, Aristides Gionis, and Stefano Leonardi. 2010. Power in unity: Forming teams in large-scale community systems. In *Proceedings of the 19th ACM International Conference on Information and Knowledge Management*, CIKM '10, pages 599–608, New York, NY, USA. ACM.

David Bamman, Brendan O'Connor, and Noah A. Smith. 2013. Learning latent personas of film characters. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 352–361, Sofia, Bulgaria, August. Association for Computational Linguistics.

David Bamman, Ted Underwood, and Noah A Smith. 2014. A bayesian mixed effects model of literary character. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*, volume 1, pages 370–379.

Sumit Bhatia, Prakhar Biyani, and Prasenjit Mitra. 2014. Summarizing online forum discussions – can dialog acts of individual messages help? In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 2127–2131, Doha, Qatar, October. Association for Computational Linguistics.

Anais Cadilhac, Nicholas Asher, Farah Benamara, and Alex Lascarides. 2013. Grounding strategic conversation: Using negotiation dialogues to predict trades in a win-lose game. In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, pages 357–368, Seattle, Washington, USA, October. Association for Computational Linguistics.

Oliver Ferschke, Diyi Yang, and Carolyn Rosé. 2015. A lightly supervised approach to role identification in wikipedia talk page discussions.

Mathilde Forestier, Anna Stavrianou, Julien Velcin, and Djamel A. Zighed. 2012. Roles in social networks: Methodologies and research issues. *Web Intelli. and Agent Sys.*, 10(1):117–133, January.

Sebastian Germesin and Theresa Wilson. 2009. Agreement detection in multiparty conversation. In *Proceedings of the 2009 International Conference on Multimodal Interfaces*, ICMI-MLMI '09, pages 7–14, New York, NY, USA. ACM.

Kazi Saidul Hasan and Vincent Ng. 2014. Why are you taking this stance? identifying and classifying reasons in ideological debates. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 751–762, Doha, Qatar, October. Association for Computational Linguistics.

Arthur E. Hoerl and Robert W. Kennard. 2000. Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, 42(1):80–86, February.

Xia Hu and Huan Liu. 2012. Social status and role analysis of palin's email network. In *Proceedings of the 21st International Conference Companion on World Wide Web*, WWW '12 Companion, pages 531–532, New York, NY, USA. ACM.

Meiqun Hu, Ee-Peng Lim, and Ramayya Krishnan. 2009. Predicting outcome for collaborative featured article nomination in wikipedia. In *Third International AAAI Conference on Weblogs and Social Media*.

Aniket Kittur and Robert E. Kraut. 2010. Beyond wikipedia: Coordination and conflict in online production groups. In *Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work*, CSCW '10, pages 215–224, New York, NY, USA. ACM.

R Meredith Belbin. 2011. Management teams: Why they succeed or fail. *Human Resource Management International Digest*, 19(3).

Arjun Mukherjee, Vivek Venkataraman, Bing Liu, and Sharon Meraz. 2013. Public dialogue: Analysis of tolerance in online discussions. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1680–1690, Sofia, Bulgaria, August. Association for Computational Linguistics.

George A Neuman, Stephen H Wagner, and Neil D Christiansen. 1999. The relationship between work-team personality composition and the job performance of teams. *Group & Organization Management*, 24(1):28–45.

Michael J. Paul. 2012. Mixed membership markov models for unsupervised conversation modeling. In *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*, EMNLP-CoNLL 12, pages 94–104, Stroudsburg, PA, USA. Association for Computational Linguistics.

Mario Piergallini, A Seza Doğruöz, Phani Gadde, David Adamson, and Carolyn P Rosé. 2014. Modeling the use of graffiti style features to signal social relations within a multi-domain learning paradigm. *EACL 2014*, page 107.

Florian A Potra and Stephen J Wright. 2000. Interior-point methods. *Journal of Computational and Applied Mathematics*, 124(1):281–302.

K Ahuja Ravindra, Thomas L Magnanti, and James B Orlin. 1993. Network flows: theory, algorithms, and applications.

Barbara Senior. 1997. Team roles and team performance: is there reallya link? *Journal of occupational and organizational psychology*, 70(3):241–258.

Swapna Somasundaran and Janyce Wiebe. 2009. Recognizing stances in online debates. In *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP: Volume 1 - Volume 1*, ACL '09, pages 226–234, Stroudsburg, PA, USA. Association for Computational Linguistics.

Yla R Tausczik and James W Pennebaker. 2010. The psychological meaning of words: Liwc and computerized text analysis methods. *Journal of language and social psychology*, 29(1):24–54.

Matt Thomas, Bo Pang, and Lillian Lee. 2006. Get out the vote: Determining support or opposition from congressional floor-debate transcripts. In *Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing*, EMNLP '06, pages 327–335, Stroudsburg, PA, USA. Association for Computational Linguistics.

Marilyn A. Walker, Pranav Anand, Robert Abbott, and Ricky Grant. 2012. Stance classification using dialogic properties of persuasion. In *Proceedings of the 2012 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, NAACL HLT '12, pages 592–596, Stroudsburg, PA, USA. Association for Computational Linguistics.

Byron C Wallace, Thomas A Trikalinos, M Barton Laws, Ira B Wilson, and Eugene Charniak. 2013. A generative joint, additive, sequential model of topics and speech acts in patient-doctor communication. In *EMNLP*, pages 1765–1775.

Miaomiao Wen, Diyi Yang, and Carolyn Penstein Rosé. 2015. Virtual teams in massive open online courses. In *Artificial Intelligence in Education*.

Theresa Wilson, Janyce Wiebe, and Paul Hoffmann. 2005. Recognizing contextual polarity in phrase-level sentiment analysis. In *Proceedings of the conference on human language technology and empirical methods in natural language processing*, pages 347–354. Association for Computational Linguistics.

Yuchen Zhao, Guan Wang, Philip S. Yu, Shaobo Liu, and Simon Zhang. 2013. Inferring social roles and statuses in social networks. In *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '13, pages 695–703, New York, NY, USA. ACM.