Using Network Structure to Identify Groups in Virtual Worlds

Fahad Shah and **Gita Sukthankar**
Department of EECS
University of Central Florida
4000 Central Florida Blvd, Orlando, FL
sfahad@cs.ucf.edu and gitars@eecs.ucf.edu

Abstract

Humans are adept social animals capable of identifying friendship groups from a combination of linguistic cues and social network patterns. But what is more important, the content of what people say or their history of social interactions? Moreover, is it possible to identify whether people are part of a group with changing membership merely from general network properties, such as measures of centrality and latent communities? In this paper, we address the problem of identifying social groups from conversation data and present results of an empirical study on identifying groups in a virtual world. Virtual worlds are interesting because group membership is more shaped by common interests and less influenced by cultural and socio-economic factors. Our finding is that a combination of network measures is more predictive of group membership than language cues, and that both types of features can be combined to improve prediction.

Introduction

fMRI and fossil record studies have revealed that humans possess highly-specialized neuronal machinery to identify social interactions such as organizational hierarchies, cheating, and altruism from subtle social signals (Powell 2004). Language is clearly an important part of this process, and much of our linguistic apparatus is actually dedicated to expressing social roles and relationships (Clippinger 2010). In fact, it has been hypothesized that language evolved directly to support social coordination (Pinker 1994). In this paper, we examine the problem of extracting social structure from unstructured conversational exchanges—is it possible to determine an actor's social group from a combination of network structure and conversation content in public chat data from a virtual world? Previous work (Shah, Usher, and Sukthankar 2010) indicates that there are measurable differences between the social networks extracted from different group, the question is whether they can be utilized to create sufficiently predictive features for identifying the user's social group.

Text-only chat exchanges lack many informative verbal cues such as prosody and information about socio-economic factors that are highly predictive of real-world social groupings. In virtual worlds, groups are often drawn together by common interests such as shopping, gaming, or scripting. For this paper, we conducted collected chat data from Second Life, a massively multi-player online environment that allows users to construct and inhabit their own 3D world. Users are able to broadcast chat messages to all avatars within a given radius of their avatar using a public chat channel. The physical environment in Second Life is organized into a 2D arrangement, known as the SLGrid. The SLGrid contains many regions, with each region boasting an array of specialized attractions such as shopping markets, gaming grounds, libraries, information kiosks, and scenic views. Users usually frequent a small set of SL regions that offer activities of interest.

Also, Second Life has two specialized types of regions to enable users to better explore game functionality:

orientation areas: specially-designed areas to teach new users how to use the SL navigational controls.

sandboxes: construction areas where more experienced users can construct SL buildings without needing to own SL land.

This paper describes an approach to identifying a user's Second Life regional group on the basis of social networks mined from user communications; we evaluate the relative contribution of conversation features vs. network-based ones when learning supervised classifiers to predict a user's region. Based on our experiments, supervised classifiers trained with network-based features (a combination of community and centrality features) outperform the unigram word features that measure the content of the user utterances. This finding holds true even though the actors in the regions differ substantially across different days. Experiments reveal that there is enough similarity between networks emerging in the same regions on different days to classify user groups with an accuracy of 44%. We hypothesize that that the form of the network follows the functionality required by the users to pursue their activities. For instance, orientation areas have large groups of transient users speaking briefly to the small set of expert helper users. This characteristic "fingerprint" appears to be sufficiently predictive to enable classification. Combining all measures (network, community, and content) yields the best overall accuracy at correctly predicting the regional origin of user dialog.

Table 1: Second Life region descriptions

Region	Region Description
Help Island Public	Orientation area
Help People Island	Orientation area
Mauve	Sandbox
Kuula	Sandbox
Moose Beach	Scenic area
Pondi Beach	Scenic area

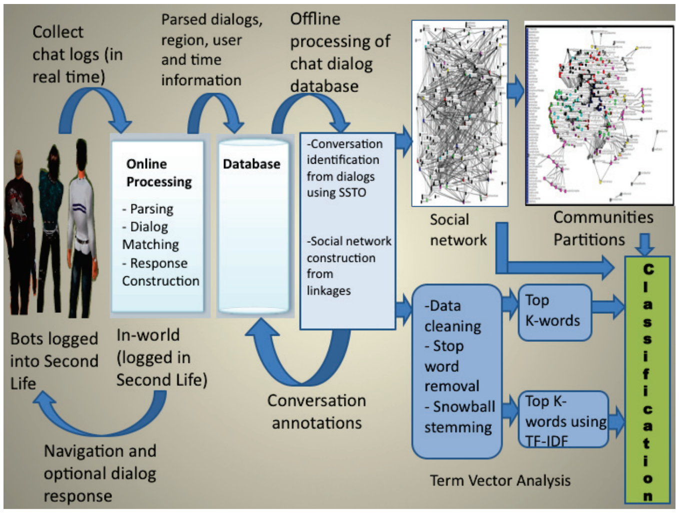
Collect chat logs (in real time)
Parsed dialogs, region, user and time information
Offline processing of chat dialog database
Online Processing
- Parsing
- Dialog Matching
- Response Construction
Database
-Conversation identification from dialogs using SSTO
-Social network construction from linkages
Social network
Communities Partitions
Bots logged into Second Life
In-world (logged in Second Life)
Navigation and optional dialog response
Conversation annotations
-Data cleaning
- Stop word removal
- Snowball stemming
Top K-words
Top K-words using TF-IDF
Term Vector Analysis
Classification

Figure 1: Multi-agent architecture for Second Life data collection

Approach

In this section, we describe our method for identifying user's regional groups. Table 1 contains a description of the Second Life regions that we used for our study. Our goal is to identify each user's region-based group based on a combination of network, community, and conversational content features. Our study includes a three basic types of regions: 1) orientation areas, 2) sandboxes for scripting and building, and 3) scenic areas. Many regions include multiple attraction types, but we categorized each region by the dominant attraction near the bot's position.

Data Collection

Figure 1 shows the overall data collection architecture. Multiple bots, stationed in different SL regions, listen to all the messages within their hearing range on the public chat channel. The bots forward chat messages to the server, which parses and conditions messages for storage in the dialog database. Occasionally the server sends the bots navigational commands and an optional dialog response if the bot is addressed by name. Linkages between SL actors are inferred offline by partitioning the unstructured data into separate conversations; these linkages are used to construct the graphs used in the social network analysis and community detection.

To collect data on social interactions in Second Life, we launched six agent bots (each with a different Second Life account and avatar) in different regions listed in Table 1, over multiple consecutive days and randomly selected 4 days of data for analysis.

Because of an inability to use statistical machine learning approaches due to the lack of sufficiently labeled data and absence of a tagger/parser that can interpret chat dialog data, we opted to use the Shallow Semantics and Temporal Overlap algorithm (SSTO), a rule-based algorithm that relies on shallow semantic analysis of commonly occurring linguistic cues that frequently occur in chat data (Shah and Sukthankar 2011). SSTO employs the following cues:

salutations: Salutations are frequent and can be identified using keywords such as "hi", "hello", "hey". The initial speaker is marked as the *from* user and users that respond within a designated temporal window are labeled as *to* users.

questions: Question words (e.g., "who", "what", "how") are treated in the same way as salutations. We apply the same logic to requests for help (which are often marked by words such as "can", "would").

usernames: When a dialog begins or ends with all or part of a username (observed during the analysis period), the username is marked as *to*, and the speaker marked as *from*.

second person pronouns: If the dialog begins with a second person pronoun (i.e., "you", "your"), then the previous speaker is considered as the *from* user and the current speaker the *to* user; explicit mentions of a username override this.

temporal co-occurrences: SSTO includes rules for linking users based on temporal co-occurrence of utterances. These rules are triggered by a running conversation of 8–12 utterances.

The output of SSTO is a to/from labeling for the chat dialogs with directed links between users. Unfortunately SSTO is not extremely accurate so the resulting social networks have both extraneous and missing links, but is more robust to poorly conditioned input than other network text analysis techniques for this task. The resulting networks were then used for extracting feature sets for the classification task.

Network Features

Using the linkages extracted from the raw chat logs, we construct social networks of the users in each Second Life region monitored by the bots. These social networks were used as the basis for both the network and community features used to train a set of supervised classifiers. For the network features, we calculate measures of centrality for each node in the network using the UCINET software (Borgatti, Everett, and Freeman 2002). The centrality measures of nodes in the network correspond to how connected nodes are to other nodes. We use three measures of centrality, degree (measure of centrality indicates the strength of relationships for a given actor), closeness (the average distance of the actor to all other actors in the network) and betweenness (a view of the actor based on its presence on the geodesic paths between pairs of other actors in the network), as the set of network features (Linton 1979).

Community Features

Based on previous work (Tang and Liu 2009), we hypothesize that community features could be valuable for our user classification problem. As described in (Newman 2006), modularity (denoted by Q below) measures the chance of seeing a node in the network versus its occurrence being completely random. It can be defined as the sum of the random chance $A_{ij} - \frac{k_i k_j}{2m}$ (where A_{ij} is the entry from adjacency matrix and $m = \frac{1}{2}\sum_i k_i$ the total edges in the network) summed over all pairs of vertices i, j that fall in the same group, where s_i equals 1 if the two vertices fall in the same group and -1 otherwise:

$$Q = \frac{1}{4m}\sum (A_{ij} - \frac{k_i k_j}{2m}) s_i s_j. \quad (1)$$

If B is defined as the modularity matrix given by $A_{ij} - \frac{k_i k_j}{2m}$, which is a real symmetric matrix with column vectors whose elements are s_i then Equation 1 can be written as $Q = \frac{1}{4m}\sum_{i=1}^{n} (u_I^T s)^2 \beta_i$, where β_i is the eigenvalue of B corresponding to the eigenvector u. u_i are the normalized eigenvectors of B such that $s = \sum_i^n a_i u_i$ and $a_i = u_i^T s$.

We use the leading eigenvector approach to spectral optimization of modularity to perform a loose community partitioning (s being a real number). For the maximum positive contribution to the modularity we use all the eigenvectors that make positive contribution to the modularity (obtained through cross-validation) and take the coefficients for them as the membership information. We use a membership threshold of 0.1 to determine whether the contribution is sufficient for community membership.

Content Features

In addition to the network and community features mined from the link structure of the networks, we believe that the language content present in the dialogs offers additional clues about the user's regional identity. We opt for a shallow approach using n-gram analysis.

To extract word features for our dataset, we followed the procedure outlined below:

1. First we partition the data into chunks consisting of all spoken dialogs for the time period under consideration. Documents in our dataset simply correspond to all the utterances from each unique user within the time period.
2. Next, we tokenize the dialogs to extract tokens from this chunk for each user.
3. A stop words list is used to eliminate commonly occurring articles, pronouns, helping verbs and salutations.
4. The Snowball Stemmer (Porter 2001) is used to stem the words to their root in order to create unique terms for each region.
5. The top k terms from each region are selected based on classification cross-validation; terms are removed until the classification performance declines significantly. Adding bigrams and trigrams drops the classification performance, possibly because of the large number of typos, acronyms, and emoticons present in our dataset.

Table 2: Hourly token counts

Hour	Number of Tokens
1	3433
2	3175
3	3215
4	2261
Total	11,994
Reduced	603

Table 3: Network user counts

	Help Island Public	Help People Island	Mauve	Kuula	Pondi Beach	Moose Beach	Total
Hour 1	34	19	7	25	27	25	137
Hour 2	30	18	5	16	15	23	107
Hour 3	32	15	6	19	18	24	120
Hour 4	28	22	5	19	16	21	109

6. We evaluated two feature variants: 1) binary encoding of the presence/absence of terms and 2) weighting the terms according to a TF-IDF measure. Both feature variants resulted in approximately equivalent classification accuracy so we used the binary encoding for our feature set.

The dataset collected by our Second Life bots consists of over 500 hours of data for a period of 4 days. We randomly selected one hour of data from the four different days for the term vector analysis. Table 2 summarizes the token counts over the dataset. There were 11,994 tokens overall, which after processing reduced to 603 terms that were then used for classifier training. For our classification task we evaluated the performance of four supervised learning algorithms using the Weka machine learning workbench (Waikato 2009): 1) decision-trees 2) Bayesian belief nets 3) k-nearest neighbor and 4) Naive Bayes and report the best result per feature group.

Results

We obtained conversation data from six different regions in Second Life over four days of data collection (80,000 total utterances). A general description of the activities that the users tend to perform in each region is shown in Table 1. The regions fell into three different general categories: 1) orientation areas for new users to learn how to interact with Second Life, 2) sandbox areas that permit users to experiment with building construction, and 3) general entertainment areas (e.g., beaches). The number of actors in each region for the four hours under consideration is shown in Table 3.

The classification task is to identify the region given the feature set. We evaluated the following feature sets:

Table 4: Classification accuracy by feature set

Feature Set	Classification Accuracy
All Centrality	25.2%
Closeness Only	25.8%
Community #	27.5%
Words (unigrams)	34%
All Centrality and Community	21.4%
All Centrality and Words	33.9%
Closeness and Community	**44.4%**
Closeness and Words	38.1%
Community and Words	34.9%
All Features	**54.3%**

- all the centrality measures (betweenness, closeness and degree),
- the most predictive centrality measure, closeness, only,
- community membership, expressed as the number of communities each user belongs to,
- the top-k words from the term vector model.

The statistics shown in the Table 4 were obtained by combining the results from classification for leave-one-out validation for the best classifier. For the leave-one-out validation we made four splits, where the data from three hours was used for training the classifier and the data from the fourth was used as the test set. We selected this scheme to evaluate whether the learned patterns persist over time and generalize to completely different sets of actors and social networks from the same region. We performed the classification for each of these four splits using all the four algorithms and selected the best performing algorithm for each feature set.

We can make the following conclusions based on the results shown in Table 4:

- Unsurprisingly, using a combination of all the features provides the best classification performance, 54.3%
- Using any of the features individually, or any combination, provides classification performance better than random chance levels (about 17% for six class problem), showing all the features carry important information.
- For individual features, the best performance is obtained from using the words (34%)
- For the combined features, the best performance is obtained by combining closeness with community membership information.

Conclusion

Our empirical study of region-based user groups in the Second Life virtual world reveals that the combination of network and community features can be more predictive of user groups that the actual content of the conversations. This is especially the case in real-world open ended conversations as described in this paper. One explanation for this phenomena is that more meaningful dialogs might be exchanged on private chats; we notice from our dataset that the conversations occurring in the public chat forums are very similar even in regions where the users are participating in different activities. Due to the volume of user traffic and the variable duration of user stays, the network structure and the number of communities differ substantially across regions, apparently resulting in more predictive classification features. This holds true even across data sampled from the same region on different days. Even though many of the actors are different from day-to-day, the network and community features retain some similarity. Our hypothesis that the form of the network follows the function that the users need to pursue their activities is not supported by the observed confusion matrices since the classifier mispredictions do not follow simple activity based trends. In future work, we plan to analyze more task-oriented groups in Second Life, gamers and shoppers.

Acknowledgments

This research was funded by AFOSR YIP award FA9550-09-1-0525.

References

Borgatti, S. P.; Everett, M. G.; and Freeman, L. C. 2002. UCINET 6 For Windows: Software for Social Network Analysis.

Clippinger, J. 2010. Human nature and social networks.

Linton, F. C. 1979. Centrality in social networks: Conceptual clarification. *Social Networks* 1(3):215–239.

Newman, M. 2006. Modularity and community structure in networks. In *Proceedings of the National Academy of Sciences*, volume 103, 8577–8582.

Pinker, S. 1994. *The Language Instinct*. Harper Collins.

Porter, M. 2001. Snowball: A language for stemming algorithms. Retrieved Feb 2011 `http://snowball.tartarus.org/texts/introduction.html`.

Powell, K. 2004. Brains sniff out scam artists: Evolution may have programmed us to compute fairness. Retrieved Feb 2011 `http://www.nature.com/nsu/020812/020812-1.html`.

Shah, F., and Sukthankar, G. 2011. Constructing social networks from unstructured group dialog in virtual worlds. In *International Conference on Social Computing, Behavioral-Cultural Modeling, and Prediction (SBP11)*.

Shah, F.; Usher, C.; and Sukthankar, G. 2010. Modeling group dynamics in virtual worlds. In *Proceedings of the Fourth International Conference on Weblogs and Social Media*.

Tang, L., and Liu, H. 2009. Relational learning via latent social dimensions. In *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 817–826. ACM.

Waikato, U. 2009. Weka. Retrieved July 2009 `http://www.cs.waikato.ac.nz/ml/weka/`.