Learning Topic Flows in Social Conversations

Chad Crawford
University of Tulsa
Tulsa, Oklahoma, USA
chad-crawford@utulsa.edu

Sandip Sen
University of Tulsa
Tulsa, Oklahoma, USA
sandip-sen@utulsa.edu

ABSTRACT

The rise of social media has opened up structurally dynamic means of communication among humans. We are particularly interested in how human agents may influence or be influence by the topic flow of a conversation – and how intelligent agents can use that information to model user behavior. For example, some topics can be more agreeable to some users than others, which would prompt responses containing the same topic. Other topics could ignite fierce debate in which arguments and counter-arguments span many topics. Social media websites such as Reddit and Twitter support a distinct structure in which a single document (a tweet on Twitter or a comment on Reddit) may have multiple distinct responses, so conversations have a tree-like structure opposed to being linear.

We identify and classify the dynamics of topic flow in conversation. Our work explores the roles of topics in branching patterns in conversations that can have tree structures. We evaluate some of the emergent topic patterns that appear when analyzing a real-world dataset from social media, and analyze the effect of accounting for user identity in our model.

1. INTRODUCTION

The rise of social media has lead to an ever-increasing amount of data to analyze and an interest in understanding the implicit meaning of that data. One of the most salient aspects of social media is its conversational nature – comments, posts, or documents are composed with the intention of being addressed to another post. Work on analyzing dialogue in general focuses on decomposing discourse into dialogue acts, which are more structurally-motivated components, which is useful for intelligent agents that need to communicate with a human to receive and perform instructions. Topic flows, on the other hand, deal with the actual content of the post – for example, predicting how people respond when a user makes a controversial comment.

In this work, we define our topics under the "bag of words" assumption. That is, we ignore the ordering of words in each document and focus solely on the presence and frequency of words. This is a popular assumption in the machine learning textual analysis, and can be found in simplistic word weighting mechanisms such as Term Frequency-Inverse Document Frequency (TF-IDF) and also in recent state-of-the-art probabilistic topic models.

Many topic models assume that documents are independent of each other, and do not take advantage of conversational ties between documents. This assumption is particularly problematic for social media where posts are very short and often refer, either directly or indirectly, to the conversation they are involved in. To effectively categorize topics in such systems, the conversational context must also be accounted for in the topic model.

Several social media websites allow users under pseudonymous identities to discuss a number of topics with each other. Platforms such as Twitter and Reddit enable conversations to have a tree-like structure: for example, if a user writes a document (such as a Tweet or a Reddit comment), other users can make multiple direct responses, but a response can only respond to one "parent" document at a time. Hence, conversations grow like trees.

The importance of context is illustrated in the following conversation:

- Sort of serious question here.
  Is he really just making shit up as he goes?
  Are his cohorts also that stunningly deluded or, like the banks with Donny's bankruptcies, are they figuring they'll make a lot more money by playing along with his dense motherfuckery instead of setting him straight?
  - In a way,yes. It is not just 'him' it's all the lobbies that sit behind the GOP. Gun, riot equipment manufacturers, **FOR PROFIT PRISONS** and corporations like Walmart.
  - No. Once again, he is doing everything he campaigned on.

    None of what Trump has done should be a surprise, if you were paying attention.
    - It was stupid when he came up with it last year and it's stupid now. No one is surprised, but exasperated might be a good word to use.

Individual posts, such as the last one in this chain, are hard to understand given their context. However, words from the previous comment, such as "Trump" and "Campaign," are strong indicators of the last post's topic.

Intelligent agents can use topic information to influence conversations that improve social well-being. For example, harassment on social media is a growing issue that needs an effective automated approach to deal with the huge influx of social media posts happening at all times. Using humans to manually moderate conversations on social media is a time-consuming task that is infeasible for larger platforms, and inappropriate for things such as private conversations. Report systems that classify content as harassment based on reports can be easily abused to the point that they can be used as a tool of harassment. Intelligent agents can respect the privacy of users while also being efficient and potentially more cost-effective. Such agents could identify topics that encourage it, and directing harassers away from those topics.

Furthermore, topic flow can also be an important factor in understanding the social behavior of users. We are particularly interested in identifying and understanding certain classes of topic flow – do some topics end a conversation, or are there cyclic patterns in topics that can be observed? In situations where agents debate over topics, topic flow can help identify common discussions and the popular arguments/counterarguments involved.

Our work examines the role of topic flow in conversations by learning significant topics using an unsupervised approach. We construct a simple topic model similar to previous models that is designed for tree-like conversations on large social media outlets, and test it on a sample dataset collected from the social media website Reddit. We then analyze emergent topic flow patterns and discuss associated user patterns.

2. RELATED WORK

Understanding human dialogue has been an important challenge in the multi-agent literature for problems such as negotiation [6] or mediation [1]. Dialogue has also been used as a framework for information sharing via argumentation [3]. However, there has been little work on unsupervised methods of clustering conversations. The capability to effectively identify such topics could help improve detection of implicit preferences in recommender systems, for example.

Topic modeling spans a large number of specialized statistical techniques to efficiently and effectively identify significant word associations. Many of these approaches use the bag-of-words model formulation and ignore the ordering of words in documents. In the domain of probabilistic models, topics are often represented as distributions over words.

One of the simplest and earliest techniques of topic modeling is known as the mixture of unigrams model. This model supposes that each document has a single topic generated from a topic multinomial distribution π, and that each topic for $k = 1, \ldots, K$ carries a multinomial distribution over words, ϕ_k. The popular approach to fitting a model to data involves finding the maximum likelihood estimate (MLE) of the model to the data.

Early models such as Probabilistic Latent Semantic Analysis (PLSA) extend the unigram mixture model by assuming each document is drawn from an independent distribution over topics [7]. PLSA is a mixed-membership model since each 'sample' has its own distribution over topics. However, PLSA has often been criticized for not being a generative model, and as such it is impossible to infer the topic distribution over new documents. Latent Dirichlet Allocation (LDA) [4] extends PLSA with a Bayesian formalization by adding Dirichlet priors to the topic and word distributions. LDA and its variants have been widely used since the Dirichlet prior can express the variability in topic and word distributions that each document contains.

Analysis of human dialogue has been largely a supervised classification problem of classifying content into "dialogue acts" which may be starting a discussion, responding to an argument, commenting on something, asking or answering a question, etc [10]. Recently, Ritter et al. [9] presented an unsupervised model that learns dialog acts using concepts from LDA. One of the initial models in their work models pure topic transitions using a Markov Chain approach, similar to the model we will present in our paper. A similar work by Yano et al. [11] uses topic modeling for comments on blog posts, which has a tree structure to some extent, but only supports one level of replies. Our model combines both of these works in some sense by allowing a tree-like conversation structure where transitions are solely dependent on the topic of the parent document.

3. CONVERSATIONAL TOPIC MODEL

For the sake of simplicity, our topic model assumes that each document is generated from a singular topic. Since comments on social media communities like Twitter and Reddit are short, we believe that there is little loss in the expressibility of the model. Techniques for unigram mixture models with latent topic transition probabilities have been experimented on unsupervised dialogue act modeling [9] and sentence-level document parsing [2]. We adapt the same type of topic model for tree-structured conversations.

We assume that the topic of a reply to a source document is dependent on the source topic. While this does not apply to all types of conversational communities, particular groups – such as those discussing politics – have discussions that are driven by an argument-style nature. In a debate, for example, back-and-forth arguments will often introduce new ideas to support each side's position. While our model does not explicitly take into account the topic of siblings to determine a document's topic, the inference procedure described below will take sibling topics into account, since the shared parent's latent topic is unknown. Replies indirectly influence their sibling's topic.

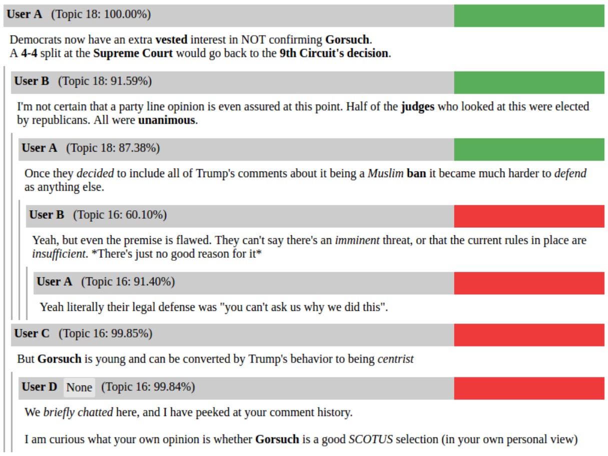
Let $\mathcal{D} = \{d_1, \ldots, d_n\}$ be the set of documents that represent a series of independent conversations, and let B be the $n \times n$ interaction matrix, where $B_{ij} = 1$ iff d_j is a response to d_i. Furthermore, let $\mathcal{W} = \{w_1, \ldots, w_m\}$ be the set of words used among the documents and C be the $n \times m$ conversation matrix, where C_{ip} represents the number of times w_p appears in d_i. Finally, suppose there are K latent topic classes. Documents are generated as follows:

1. For the ith document:
   (a) If it is a root, choose a topic $z_i \sim \text{Multinomial}(\pi)$
   (b) If it is a response to d_j, choose $z_i \sim \text{Multinomial}(A_{z_j})$
2. For each word $w_p, p = 1, \ldots, N_i$, the particular word $w_p \sim \text{Multinomial}(\theta_z)$

Where π is the $K \times 1$ initial topic distribution vector, A is the $N \times N$ topic transition matrix, and θ is the $K \times M$ word distribution matrix, where θ_{kp} represents the likelihood of drawing w_p from the kth topic distribution.

Expectation.

Like the forward-backward algorithm for HMMs, conditional independencies can be exploited to significantly reduce the computational cost of tree inference. We opt to use Pearl's belief propagation algorithm for trees [8], which is an exact inference method that runs in linear time. We refer to our distribution over latent topics as the $N \times K$ matrix Q. Since expectation is performed on a large number of independent trees, the construction of Q can be highly parallelized.



User A (Topic 18: 100.00%) 🟩

Democrats now have an extra **vested** interest in NOT confirming **Gorsuch**.
A **4-4** split at the **Supreme Court** would go back to the **9th Circuit's decision**.

User B (Topic 18: 91.59%) 🟩

I'm not certain that a party line opinion is even assured at this point. Half of the **judges** who looked at this were elected by republicans. All were **unanimous**.

User A (Topic 18: 87.38%) 🟩

Once they *decided* to include all of Trump's comments about it being a *Muslim* **ban** it became much harder to *defend* as anything else.

User B (Topic 16: 60.10%) 🟥

Yeah, but even the premise is flawed. They can't say there's an *imminent* threat, or that the current rules in place are *insufficient*. *There's just no good reason for it*

User A (Topic 16: 91.40%) 🟥

Yeah literally their legal defense was "you can't ask us why we did this".

User C (Topic 16: 99.85%) 🟥

But **Gorsuch** is young and can be converted by Trump's behavior to being *centrist*

User D None (Topic 16: 99.84%) 🟥

We *briefly chatted* here, and I have peeked at your comment history.

I am curious what your own opinion is whether **Gorsuch** is a good *SCOTUS* selection (in your own personal view)

Figure 1: A sample conversation classified using our topic model. Italics represents words with some correlation with the topic, and bold represents a stronger correlation. The percentage on the left of the topic name represents our confidence in the comment having that topic.

Maximization.

Maximization of the model parameters is similar to the EM implementation in [9]. Given the expectation matrix Q, the corresponding maximization step is:

$$\pi_k = \frac{\sum_{i \in \mathcal{S}} Q_{ik}}{||\mathcal{S}||} \tag{1}$$

where $\mathcal{S}$ is the set of source documents, that are not a response to any other document in the corpus. Maximization of A and θ can be expressed as the matrix equations,

$$A \propto Q^{\intercal} B Q, \tag{2}$$

$$\theta \propto C^{\intercal} Q. \tag{3}$$

Since B and C are very sparse, the matrix product equations can be made to be much more efficient.

3.1 User Clustering in a Topic Model

One important aspect of a conversational topic model is its user-oriented nature. The former model assumes that there is a general transition model between topics that each document follows. However, the user that creates a reply will influence what topic they respond with. For example, suppose two users respond to a post about a scene in a film. One user might write a response that focuses on the characters in the scene, while another focuses on the use of camera. We are interested in constructing a topic model that captures the differences between the two users that is equally easy to learn. For this problem, we assume each document in the corpus also has the identity of its author.

It would be ideal to develop a separate transition model for every author of posts in the network. However, the number of parameters needed to specify a transition model is $K(K+1)$, for unique π and A parameters. Therefore, we would much more than $K(K+1)$ posts from each user to reliably develop a topic transition model. Rather, we assume that user behavior can be clustered in several groups, where each group has consistently similar behaviors. Suppose that there are G user groups for U users, and each user is associated with a single group.

We make a simple modification to our model – for each document d_i, the author u_i has an associated group, g_{u_i}. The topic for d_i is chosen using $A^{(g_{u_i})}, \pi^{(g_{u_i})}$ as parameters. While the change to the model itself is minimal, we has to dramatically change the inference procedure.

To perform inference on this problem, we need to estimate the marginal latent variable distributions. Unlike the last problem that could be characterized as a singly connected tree, if a single user has two posts inside a single conversation, the problem will lose that singly connected property. General inference on multiply connected networks is generally NP-complete and must be approximated.

One of the most popular methods of approximate inference is Gibbs sampling [5]. Latent variable probabilities are estimated by assigning random initial values to each latent variable, and then repeatedly updating single latent variables by sampling from their conditional probability distribution, given all latent variable assignments. It has been proven that aggregating the observations from this method converges to the true latent variable marginal probability.

However, Gibbs sampling is slow for our model since it is costly to compute conditional probability distributions, and the number of samples needed to produce reliable estimates is high. Instead, we use an approximation method similar to conditioning [8]. This procedure works by breaking the inference problems into two sub-problems. The individual inference problems on computing latent topics and latent user-group associations are individually singly-connected networks for which inference can be computed exactly. Suppose that $q(z)$ is the latent topic distribution, and $r(y)$ is the latent user-group distribution. Then a natural method for estimating the likelihoods would be the iterative process

$$\begin{aligned} q^{(n+1)}(z) &= \sum_{y \in Y} P(Z, Y) = \sum_{y \in Y} P(Z|Y)P(Y) \qquad (4)\\ &= E_{r^{(n)}}[P(Z|Y)] \qquad (5)\\ r^{(n+1)}(y) &= \cdots = E_{q^{(n+1)}}[P(Y|\Theta, Z)] \qquad (6) \end{aligned}$$

until they converge. In this case, the computation of equation 4 can be done by slightly modifying Pearl's message passing algorithm from earlier, and equation 6 can be easily estimated since it is a 1-level tree.

If q or r are the true marginal distributions, then this equation will compute the exact marginal distribution for the other term.

Empirically, this method converges very quickly when used within the EM algorithm. We have noticed that it converges the slowest for the first step of EM, when parameters are initialized to random values and the latent variable distributions are also random estimates. However, in later steps of the EM algorithm, the error reduces by several magnitudes. The iterative procedure stops when the magnitude of change over both q and r falls below the threshold value 0.001.

Maximization is very similar to the previous problem. Let $T^{(g)}_{ik} = Q_{ik} R_{u_i g}$, where R is the $U \times G$ matrix that is drawn from the latest values from r. Then, the update rules are:

$$\pi^{(g)}_k \propto \sum_{i \in \mathcal{S}} T^{(g)}_{ik} \tag{7}$$

$$A^{(g)} \propto Q^{\intercal} B T^{(g)\intercal} \tag{8}$$

$$\theta \propto Q^{\intercal} C \tag{9}$$

The complexity of the maximization update is similar to the original model – but the iterative inference process takes several times longer, since inference must iterate about 4-5 times before converging, after each EM update.

4. DATASET

We collected comment data from the social media website Reddit[1], which is a link-sharing forum where users can communicate using pseudonymous identities. Reddit hosts a number of communities, called "subreddits", where discussions can focus on particular subjects such as science, politics or hobbies. Our work focuses on the POLITICS subreddit, which is a general discussion forum for US politics.

We chose Reddit because it encapsulates many features of conversation-based topic flow that we were interested in modeling: communities like POLITICS are very issue-driven, and topics can be identified by a set of key words. Furthermore, these communities are strongly discussion-oriented, and conversations can span many replies and go into some depth. Finally, Reddit's social network structure is strongly associated

We scraped approximately 170,000 comments from the most popular posts on Reddit post-inaguration. Conversation trees with 5 or fewer comments in them were removed from the dataset. On Reddit, users are also allowed to disassociate themselves from a comment they authored and can delete a comment completely. These comments, and all subsequent replies, were removed. A conversation on Reddit is a collection of comments that are tied together as replies to each other in some way. We automatically removed any conversations with 3 or fewer comments in them. We filtered out text in the remaining comments by removing common stopwords and stemming all words, which are standard text analysis techniques that can help eliminate noise and avoid having different forms of the same word respectively. Additionally, any extremely common (more than half of the comments) and extremely uncommon (10 or fewer comments) words were filtered out, as a simple measure of reducing noise. On websites like Twitter, the limit on characters often leads to using abbreviations and slang for common terms [9]. Since Reddit has virtually no limit on the length of the comment, we found that such slang and abbreviations were less common and did not do any additional processing. The final Reddit dataset had 130,648 comments and 9,122 words.

In total, there were 4487 conversations in the filtered dataset. Many conversations could span hundreds of comments. However, the length of discussions was fairly short – 'tail' comments that had no subsequent replies composed 65,486 comments which is about half of the dataset.

5. EXPERIMENTAL RESULTS

For the simple topic model, we manually tried out several values of K and chose $K = 30$ empirically. For the user-group topic model, we found that $G = 3$ with the same number of topics was the most convincing result.

A sample conversation that was provided to our model is included in Figure 1. Word strengths are computed using the value $p(z_k|w)$, where z_k is the kth topic and w is the word of interest. This probability is ideal since it represents how unique a word is to the topic; the higher $p(z_k, w)$ is, the more likely that any random document selected containing that word will also be classified with topic z_k. We can easily compute $p(z_k|w)$ using Baye's theorem:

$$p(z_k|w) \propto p(w|z_k)p(z_k) = \theta_{wk} \sum_i Q_{ik}. \tag{10}$$

We will discuss some general statistics about the dataset. There was a large variance in the topic distribution on root and tail topics. Comments under Topic 26 had the highest likelihood of being a root topic at 6%: this topic was mainly about news outlets and Trump. Topic 11 was the least likely to be a root topic at 0.6%, and was mainly about guns.

Topic root likelihoods had little correlation with the number of comments a particular topic had. The topic with the most comments by a far margin had a 2% root topic likelihood, for example. The average "depth" of the conversation also did not have a direct correlation with the root probability. We say that the depth of a conversation is the maximum length of a chain of replies that starts at the root comment. For the guns topic, the average depth of the conversation was about 6.4 comments. The news outlets topic, on the other hand, had an average depth of about 3.7. The depths of other topics varied between 3 and 6.

5.1 Characterizing Topic Flow

We generalize topic flows by aggregating all conversation-level transition statistics into a directed graph, shown in Figure 2b. Edges on the graph from topic node x to y represent that the likelihood of that transition occurring, as given by the topic transition matrix A, was greater than our threshold parameter γ. We chose to use $\gamma = 0.01$ since it captured a good part of the network structure while only representing edges that were highly probable of occurring. Larger topic nodes on the network represent those with a high likelihood of transitioning towards themselves.

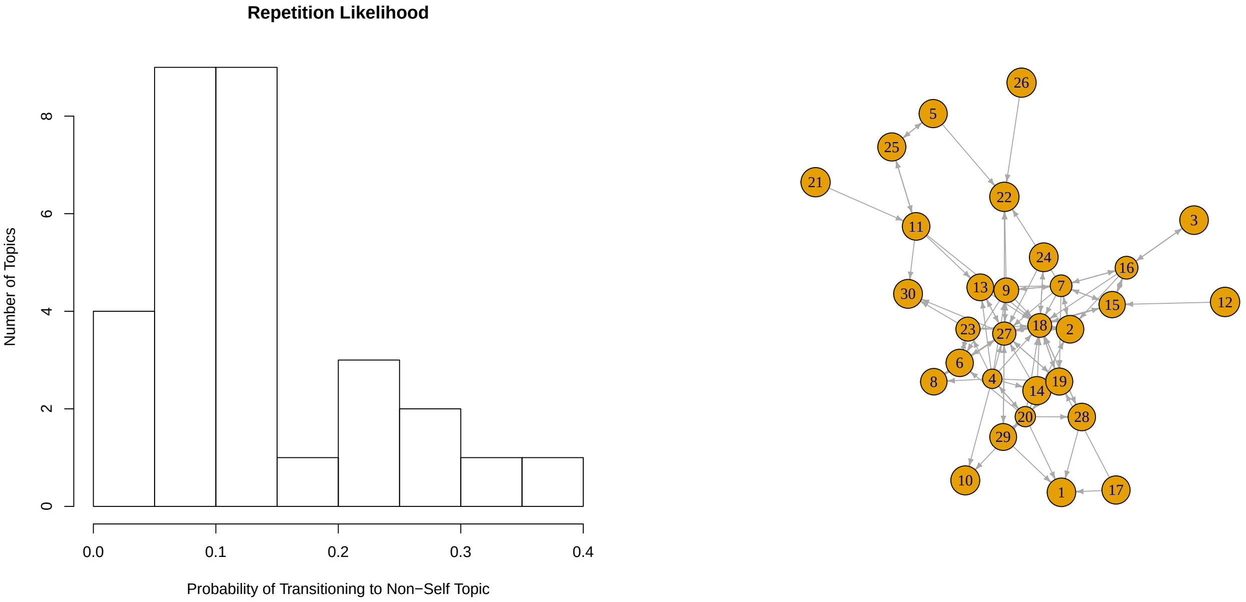
The topic flow network exhibits several scale-free properties. The probability of transitioning to a new topic from each topic is shown in Figure 2a – the density distribution loosely follows a power-law distribution. "Hub" topics such as topic 27 and 18, which had many incoming edges, also had a large number of outgoing edges and a smaller probability of self-transitioning. Topic 27 was on international politics and military conflict. Topic 18 dealt with the constitutionality of Trump's actions and the US immigration ban. Based on analysis of several different conversations, these topics appeared often because one argument subtree would shift towards these topics, such as discussing Trump's competency as a leader or reacting to distinct incidents.

One of our beliefs about topic flow was that topics would emerge to follow an argument-structured pattern such as "Topic A→Topic B→Topic A". However, in our analysis of all such triples on the data, such a pattern appeared in only about 2.3% of posts. This weakness appears because the topic model only accounts for the appearance of words in a conversation, and not their ordering. For example, in conversations on gun control, it was common to see both sides of an argument assigned the same topic because they both mentioned gun-related words.

5.2 User Trends

We initially focus on particular user trends in the original conversational topic model, and then offer some preliminary discussion on results from the explicit user-grouping model. Even though our dataset represented a small snap-

[1] https://www.reddit.com



(a) Histogram of topics by the likelihood of each topic transitioning to a topic other than itself. The higher frequency near 0.1 indicates that most topics self-loop, while fewer have a higher probability of transitioning to a different topic.

(b) Network visualization. The larger a node is, the more likely it is to transition to itself. Directed edges shown represent high likelihoods of transitioning from topic to topic.

Figure 2: Topic flow network

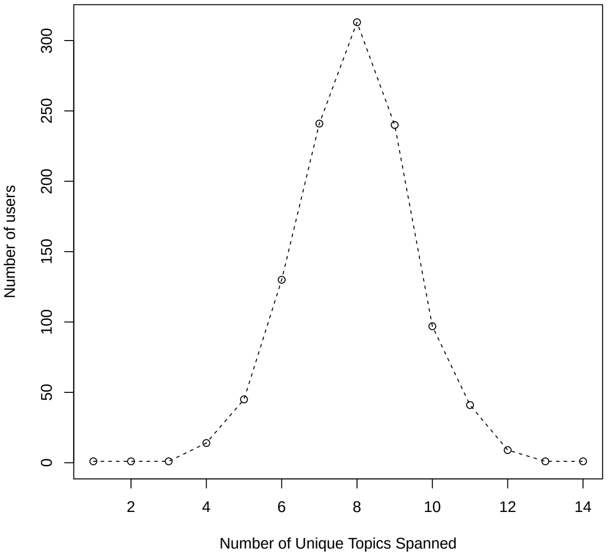


Figure 3: Frequency of users posting across a number of topics.

shot of community discussions, approximately one thousand users had participated in 10 or more different conversations. Unexpectedly, users did not consistently speak in one topic during a discussion. In Figure 1, a sample conversation is shown where there is a shift from Topic 18 to Topic 16. Two of the users, *User A* and *User B* had comments associated with both topics in their discussion. We observed more often that when a conversation shifted towards another topic, subsequent conversation would remain on the same topic for some time. This is because the nature of our topic model was not granular enough to capture the intricacies of such discussions, and rather captured the main subject of discussion.

Furthermore, users were fairly diverse in the topics they covered. Figure 3 shows the distribution of users by the number of topics that their comments have been attributed with. This was collected among users that had participated at least one in 10 or more conversations, and then aggregating unique topics that they had commented with on 10 randomly sample conversations from those they had participated in. The distribution is similar to a binomial distribution – which suggests that users are less prone to discussing only one subject, but will choose from a range of topics that interest them.

To see if the user-grouping model would perform well on a dataset of this type, we tested it and evaluated the groups formed. Unfortunately, only about 10% of users were strongly correlated with one group (that is, 10% of users had a more than 50% likelihood of being associated with a single group). This is likely because of the similar pattern in how users behaved on this dataset – the majority of replies shared the same topic as their parent, so there was not much difference in the user models generated by the groups.

6. CONCLUSIONS

We present an unsupervised topic model that captures the flow of tree-structure conversations, which have become prevalent on many social media analysis, and then analyze the relationship between user behavior and topic flow. Early results give a promising outlook on the significance of understanding topic trends, and there seems to be some common topic trends that run through any conversation. Users also were fairly diverse with the topics of their comments, which suggests that users in the community are interested in a wide range of subjects rather than being focused on single issues.

There are some natural extensions to this work that we have noticed in the development and analysis of our topic modeling framework. First, while the topic model seems effective for shorter comments, some Reddit communities have comments that carry several hundred words with them, and may be more suited towards models similar to PLSA and LDA that assign topics to each word in a document rather than to the entire document. Furthermore, while our model does account for the user part, we were unable to find an appropriate dataset for the user-group topic model. In the future, we would like to test our topic model on data sets that would not have a strong topic self-loop property.

Another potential application of topic flow modeling would be in unsupervised clustering of arguments in a discussion-driven community. For example, for our political dataset, this would involve detecting arguments in favor of/against subjects such as gun control, and then learning how arguments will flow from subject one to another. We had hoped that our current model would be able to detect such patterns, but the bag of words assumption had a significant impact on the topics formed. In many discussions, both “sides” of an argument would often be classified with the same topic since they would mention the same words, and topics would then encapsulate both sides of the argument. To be able to capture argument-aware topics, it is possible that the incorporation of *sentiment* on words in the dataset could improve classification.

REFERENCES

[1] M. Barlier, R. Laroche, and O. Pietquin. A stochastic model for computer-aided human-human dialogue. In *Interspeech 2016*, volume 2016, pages 2051–2055, 2016.

[2] R. Barzilay and L. Lee. Catching the drift: Probabilistic content models, with applications to generation and summarization. *arXiv preprint cs/0405039*, 2004.

[3] E. Black and K. Atkinson. Dialogues that account for different perspectives in collaborative argumentation. In *Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems-Volume 2*, pages 867–874. International Foundation for Autonomous Agents and Multiagent Systems, 2009.

[4] D. M. Blei, A. Y. Ng, and M. I. Jordan. Latent dirichlet allocation. *Journal of Machine Learning Research*, 3(Jan):993–1022, 2003.

[5] G. Casella and E. I. George. Explaining the gibbs sampler. *The American Statistician*, 46(3):167–174, 1992.

[6] G. Gutnik and G. Kaminka. Towards a formal approach to overhearing: Algorithms for conversation identification. In *Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems-Volume 1*, pages 78–85. IEEE Computer Society, 2004.

[7] T. Hofmann. Probabilistic latent semantic indexing. In *Proceedings of the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 50–57. ACM, 1999.

[8] R. E. Neapolitan et al. *Learning bayesian networks*, volume 38. Pearson Prentice Hall Upper Saddle River, NJ, 2004.

[9] A. Ritter, C. Cherry, and B. Dolan. Unsupervised modeling of twitter conversations. In *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, pages 172–180. Association for Computational Linguistics, 2010.

[10] A. Stolcke, K. Ries, N. Coccaro, E. Shriberg, R. Bates, D. Jurafsky, P. Taylor, R. Martin, C. Van Ess-Dykema, and M. Meteer. Dialogue act modeling for automatic tagging and recognition of conversational speech. *Computational Linguistics*, 26(3):339–373, 2000.

[11] T. Yano, W. W. Cohen, and N. A. Smith. Predicting response to political blog posts with topic models. In *Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, pages 477–485. Association for Computational Linguistics, 2009.