Agent-Based Probabilistic Models of Social Interaction

Doctoral Consortium

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

ABSTRACT

Multi-agent simulation is a powerful tool for studying real-world interactions and identifying influential factors that determine the emergent dynamics of social systems. While they are often partially validated using theories from the social sciences or correspondences with real-world data, it is usually difficult to accurately characterize real-world population phenomena with these models. This can be limiting in applications involving large, distributed systems, for example, if we want to know how the opinions of users are evolving over time in a social network. Social media provides an expansive and readily available dataset for analyzing such models, but very rarely are there compatibilities with the textual information available on social media and the latent characteristics that simulated agents exhibit. To develop more effective agent-based approach to modeling interaction and expression patterns on social media, we propose a probabilistic analysis framework in which the internal state of the agent and some contextual situation influences the textual content of their post. We investigate several text-based models that can be validated by and used to analyze social media corpora.

KEYWORDS

Probabilistic Modeling; Multi-agent Simulation; Topic Flow Models

ACM Reference Format:

Chad Crawford. 2018. Agent-Based Probabilistic Models of Social Interaction. In *Proc. of the 17th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2018), Stockholm, Sweden, July 10–15, 2018,* IFAAMAS, 3 pages.

1 INTRODUCTION

With the widespread popularity of social media, a massive corpus of human activity is publicly accessible. Due to the complex nature of language and limited capabilities of textual parsing, analytical tools are necessary to translate social media text into machine-interpretable information. Natural Language Processing tools have been widely applied to a number of different textual analysis problems. These tools usually do not account for the key role conversational context plays in determining communication content and expression.

For example, consider the phrase "I'll kill you." Between strangers, this phrase would be interpreted as a serious threat. Among friends, it could be seen as a harmless joke. Understanding the tone of a conversation, through the words exchanged and the context of the conversation, including the relationship between the participants, can lead to significantly different interpretations of the same phrase.

One of our key interests is in the detection of conversationally influential agents. Most existing work measures influence using some function of the number of retweets/views/likes a post receives, and models diffusion as a process of information passing through many simple nodes [7]. Conversations, on the other hand, have few topological indicators that can exactly specify when someone has been influenced by another.

Many agent-based models simplify interactions between agents, with an emphasis being on their effects – through the change of opinion, spread of ideas or emergence of reputation or trust. With the vast size of social media, there is now a preponderance of interaction information available to validate these models with real-world data. Translating textual information into the basic representations of interactions is often infeasible or under-specified for many models. We are instead interested in constructing models that can *predict* the outcome of interactions. This might be, for example, predicting if the tone of a conversation gradually grows hostile based on the original opinions of the authors. Under such a metric, a model is a better representation of real-world dynamics when it is capable of predicting conversational trends.

Looking at individual words is often insufficient for predicting the actual content of a conversation, which requires more complex analytical tools. We thus desire conversation-aware summarization techniques that can be pre-trained on large amounts of data and then applied to a specific communications to identify trends of interest. Predicting the appearance of topics in a conversation, for example, would be a good indicator of understanding individual user's thought processes. Similarly, tone can be a good indicator of whether a person is receptive to new ideas.

We have developed some initial tools towards this goal: The Markov Topic Flow Model (MTFM) and Latent Community Topic Flow Model (LCTFM) are topic models of conversational text that can identify the primary topic of discussion in a document. We propose several extensions of these basic tools in order to solve real-world problems: A model of conversational influence and of opinion diffusion, both based on real-world textual data.

2 MODELING TOPIC FLOW

Our initial work has been on developing conversation-aware topic models. While there has been a lot of research on topic modeling, few works account for the context of conversation [1, 6].

We hypothesize that conversational relations can be generalized with a Markovian assumption – the topic(s) of a response to a parent document are dependent only on it's parent, and not any earlier predecessors. Such a trend is often observable in social media

*Advisor: Sandip Sen

Proc. of the 17th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2018), M. Dastani, G. Sukthankar, E. André, S. Koenig (eds.), July 10–15, 2018, Stockholm, Sweden.

conversations. Many times, a conversation will quickly diverge from its original topic of discussion and move into other areas.

Our first baseline model, the Markov Topic Flow Model (MTFM), is a basic adaption of the Hidden Markov Model (HMM) for tree-structured networks with multi-word observations. The latent state matrix describes topic transition behaviors on a collection of documents. This model was evaluated on a manually collected dataset from Reddit, a popular social media website for sharing links. However, we discovered that a single topic transition matrix was not enough to adequately explain topic transition trends. When compared to a Mixture of Unigrams model, which assumes that each document is independent [4], the MTFM performed only marginally better on predictive tasks. The challenge of this work was that discussions were often also sensitive to the discussion that they were a part of – for example, a discussion on politics might expand to sports, but it would eventually return to politics. The Latent Community Topic Flow Model (LCTFM) addresses this issue by allowing conversations to have independent transition behaviors – each represented as its own HMM. In order to alleviate the dimensionality issue, in which there would not be enough posts in a conversation to parameterize a topic flow model, conversations are also clustered into latent "communities" with shared transition trends.

We are currently extending these models to scenarios where documents may have a distribution over topics, as is the case for non-conversational state-of-the-art models such as Latent Dirichlet Allocation [2]. The Markov chain formulation used in the MTFM and LCTFM are inadequate to describe transition trends between multi-topic documents. In our extended Bayesian model, we construct our own prior distribution for responses that is dependent on both the topics of the parent and conversation transition trends. The model makes a similar assumption about topics: if a parent spends 20% of its time writing about topic A, then the prior of the response will use 20% of topic A's prior response vector. We have derived a parameter estimation routine for computing the topics, topic transitions and latent topic distributions on a per-topic basis, and are currently experimenting with it on various corpora.

While we have evaluated our model on prediction tasks, we are interested in additional robust evaluation by applying these models to supervised problems. Unfortunately, there are few supervised datasets available with conversational information. We have collected several datasets from the social media websites Reddit and Twitter, with a focus on the tasks of detecting harassment and summarizing political arguments. In the future, we plan to label these conversational datasets using crowdsourcing for both tasks and use it to validate our models.

3 CONVERSATIONAL INFLUENCE

Since language is, by nature, conversational, it should be no surprise that ideas passed through conversation would be more powerful than those spread by other means. For example, often the best type of advertising a company can receive is that through word-of-mouth advertising, spread through direct conversation on a social network. Similarly, conversational influence can determine opinion formation. Persuasive individuals will be able to change the minds of even their opponents with emotive or knowledgeable arguments, while abrasive or confrontational agents may lead to polarization away from their attitudes due to their repugnant or divisive behavior. Unlike regular influence, conversational influence can be personal. The tone of a conversation can be very influential or not at all depending on the participants involved. In some cases, people would never be influenced by a particular person.

We define a conversationally influential individual in terms of their ability to control the topic of discussion. Detecting these types of individuals could be helpful in identifying confrontational individuals that are likely to get involved in heated arguments, or in identifying individuals that can diffuse a tense or emotionally charged situation. Our model of conversational influence is an extension of the LCTFM in which individuals can either follow the flow of a conversation, by repeating the same topic, or introduce a new topic into the conversation. These new topics are not necessarily influential – if the conversation stops after this post, then it fails to influence anybody. Those who respond to this post, however, would "carry on" their influence.

4 OPINION DIFFUSION

We are also interested in the modeling and identifying individuals' opinions on social media. We propose developing a community-based model of opinion diffusion, in which individuals associate or disassociate with social communities based on their inter- and intra-group interactions. Modeling opinions through social group membership resolves two issues: First, social media participants often do not post enough on a single subject for us to effectively learn their attitudes on every subject. Secondly, our salient attitudes are often a determinant of which groups we choose to associate with and how we interact with others [3].

We are interested in both identifying the general attitudes of a group and determining what factors predict the growth or recession of a group on social media using conversational text. This requires, among others, analyzing agreements and debates among users to identify which issues are most important to the social group. We are interested in applying methods from the vast literature on argument mining [5] to our conversational text in order to parse arguments and learn each social group's position on issues.

5 CONCLUSION

We propose a general technique for summarizing documents into topics or other salient, but unobservable variables to aide the construction and evaluation of personalized multi-agent models of social interactions. The MTFM and LCTFM are proposed as models of topic flow in social media conversations under the assumption that responses to some parent document are solely influenced by that parent's topic. These basic models are used as a basis for developing probabilistic multi-agent simulations that can serve a multitude of application-driven purposes.

In our first proposed application, conversationally influential and influenced agents are identified based on the fact that individuals can be influential along some dimensions and influenced along others. The second proposed application domain is in the modeling of opinion diffusion on social networks. These applications can show how agent-based simulations, which were classically supported by qualitative evidence, can be developed such that they can be validated and provide analysis of real-world social media data.

REFERENCES

[1] Hesam Amoualian, Marianne Clausel, and Eric Gaussier. 2016. Streaming-LDA : a copula-based approach to modeling topic dependencies in document streams. *Acm Sigcdd* (2016), 695–704. https://doi.org/10.1145/2939672.2939781

[2] David M Blei, Andrew Y Ng, and Michael I Jordan. 2003. Latent Dirichlet Allocation. *Journal of Machine Learning Research* 3, Jan (2003), 993–1022.

[3] Matthew J Hornsey. 2008. Social identity theory and self-categorization theory: A historical review. *Social and Personality Psychology Compass* 2, 1 (2008), 204–222.

[4] Kamal Nigam, Andrew Kachites McCallum, Sebastian Thrun, and Tom Mitchell. 2000. Text classification from labeled and unlabeled documents using EM. *Machine learning* 39, 2 (2000), 103–134.

[5] Andreas Peldszus and Manfred Stede. 2013. From argument diagrams to argumentation mining in texts: A survey. *International Journal of Cognitive Informatics and Natural Intelligence (IJCINI)* 7, 1 (2013), 1–31.

[6] Alan Ritter, Colin Cherry, and Bill Dolan. 2010. Unsupervised modeling of twitter conversations. *The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics* June (2010), 172–180. http://nparc.cisti-icist.nrc-cnrc.gc.ca/npsi/ctrl?action=rtdoc

[7] Daniel M Romero, Wojciech Galuba, Sitaram Asur, and Bernardo A Huberman. 2011. Influence and passivity in social media. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*. Springer, 18–33.