Quantitative Modeling and Analysis of Argumentation Polarization in Cyber Argumentation

Joseph W Sirrianni, *Member, IEEE,* Xiaoqing (Frank) Liu, and Douglas Adams,

***Abstract*—Cyber argumentation platforms offer specially designed environments for users to discuss and debate their stances and viewpoints on important issues. However, argumentation polarization often occurs in discussions and debates on these cyber argumentation platforms. Several researchers investigated argumentation polarization qualitatively in the past, but none have developed a quantitative model for measuring the degree of argumentation polarization. We addressed this important and challenging issue by developing an innovative argumentation polarization model to measure argumentation polarization by incorporating four important attributes of argumentation polarization: 1). The total number of argumentation poles, 2) The population size of the argumentation poles, 3) similarity within argumentation poles, and 4) the dissimilarity between argumentation poles. Its baseline model was derived from an economic polarization model proposed by Esteban and Ray, which measures polarization using three features: 1) Homogeneity within each group, 2) Heterogeneity across groups, and 3) a small number of significant groups. We adapted their model by incorporating population sizes for poles, normalizing for population size, and normalizing for parameter selection, to fit our formulation of argumentation polarization. This model was evaluated using an empirical study conducted using our cyber argumentation platform, the Intelligent Cyber Argumentation System (ICAS). This model was evaluated with two other distribution-based opinion polarization models that were applied in online discussion contexts, both analytically and empirically, since there are no existing argumentation polarization models. The analytical and empirical evaluations indicate that our model performs more effectively in terms of the definition of argumentation polarization and the four attributes.**

***Index Terms*—Polarization, Cyber Argumentation, Online Deliberation, Polarization Modeling, Agreement Analysis**

I. Introduction

CYBER argumentation platforms are specialized online discussion and debate platforms that encourage productive deliberation through the platform's design and provide analysis on the deliberation process and outcomes [1]–[7]. These platforms are alternatives to popular online platforms, such as social media and networking platforms, which serve as the de facto public forums for online deliberation and debate on important topics [8]. Those platforms, while popular, often result in undesirable deliberation outcomes, such as echo chambers [9]–[13], where only like-minded users discuss ideas with one another which isolates them from diverse opinions, viewpoints, and ideas, often leading to extreme opinions [11], animosity toward out-group members [14], and a more polarized environment [12]. Likewise, discussions in online social media and networking platforms also tend to have very low deliberation quality [15]–[19].

Thus, many researchers in online deliberation and cyber argumentation have investigated various argumentation structures and features to promote higher quality online deliberation [19]–[21]. Many research groups have proposed specialized cyber argumentation platforms that use a variety of approaches to promote more productive deliberation and provide analysis on the deliberation process and outcomes [1]–[7]. However, providing structure and features alone is not enough to ensure a productive discourse environment. One major factor that can detrimentally affect online discussions and debate is polarization.

Polarization has been shown to have many detrimental effects on discussion in both offline [22] and online [13], [23] settings. In online settings, polarization has shown to lower the quality of the crowd-wisdom from discussions [13], increased tension between ideological groups [14], contributed to decreased civility in public discourse [24], and polarize attitudes at the individual level [25]. Much attention has been paid to opinion polarization in online settings, such as social media [9], [23], [26]–[30], among political blogs [31], [32], and in comment sections [15]. However, polarization in online argumentation has not received much research attention.

Polarization in discussions in cyber argumentation, which we refer to as argumentation polarization, differs in some key ways from opinion polarization in other contexts, such as in social media and network platforms. We define argumentation polarization as the degree to which the discussion participants in cyber argumentation form distinct, internally consistent argumentation poles or groups of significant size that conflict to some degree with one another based on the degree of their agreement or disagreement with the discussion topic. Discussions in cyber argumentation are typically centralized, meaning all of the posts in a discussion are located in the same place and are available to all participants. This centralization of the discussions means that participants are exposed to a wide variety of viewpoints and opinions from their peers and are encouraged to engage with them. In cyber argumentation, the participants must defend their stances and attack opposing users' stances by making arguments. Thus, opinions formed through argumentation are more informed through consideration of alternative views, ideas, and experiences [33]–[37], compared to unjustified opinions collected through polling

J. Sirrianni is with the Department of Computer Science and Computer Engineering, University of Arkansas, Fayetteville, AR, 72701, USA (email: jwsirria@uark.edu).

X. F. Liu is with the College of Engineering, Southern Illinois University, Carbondale, IL, 62901, USA (email: xiaoqing.liu@siu.edu).

D. Adams is with the Department of Sociology and Criminology, University of Arkansas, Fayetteville, AR, 72701, USA (email: djadams@uark.edu).

Manuscript received May 26, 2020; revised August 10, 2020.

or surveying, which do not reflect these considerations [38]. Furthermore, because discussions in cyber argumentation are centralized, users cannot easily isolate their views within an echo chamber, as often happens in social media and networking platforms. Social media and networking platforms focus on user interaction through social connections, while cyber argumentation focuses on interaction through discussion and debate in a shared space. Thus, the network approaches taken to model polarization in social media, which rely on social connections as the main mechanism for opinion influence [23], [39], [40], do not apply to argumentation polarization, where social connections are de-emphasized.

Developing a polarization model specifically for argumentation polarization is important because the opinion data produced for cyber argumentation has specific qualities that need to be considered. Furthermore, because the opinions in cyber argumentation are justified through the argumentation, the opinions reported through cyber argumentation are more specific and nuanced than opinions collected without justification. As such, the polarization model ought to consider even small differences in opinions between users as somewhat significant and adjust its measurement of polarization accordingly. An argumentation polarization model must place more emphasis on the differences between user opinion, even when those differences are small, when measuring the polarization in cyber argumentation.

Some prior research has been done to investigate argumentation polarization. Avrapalley et al. [41] proposed a method of assessing polarization in cyber argumentation for collaborative decisions using Fuzzy c-means clustering. They clustered discussion participants based on their opinions in several debates under a shared issue. They asserted that each cluster represented a polarized group and each participant's fuzzy membership with the clusters represented the extent to which they as individuals are polarized. Klein et al. [42] proposed a method of assessing cyber balkanization (which is related to polarization [28]) in Reddit forums by encoding the discussions into signed interaction graphs and analyzing the different graph connections between the users. However, these investigations only focus on the presence of argumentation polarization and do not quantitatively model argumentation polarization to quantify the degree of argumentation polarization.

In this work, we address this important issue by developing an innovative argumentation polarization model to measure the degree of argumentation polarization among the participants in cyber argumentation. To develop this model, we identify four important attributes of argumentation polarization that characterize its presence in cyber argumentation, based on the distribution of the participants' opinion stance (agreement or disagreement) toward the discussion topic. Given the distribution of the participants' opinion stances, our four identified attributes that characterize argumentation polarization are: 1) The total number of argumentation poles, 2) the population size of the argumentation poles, 3) the similarity within the argumentation poles, and 4) the dissimilarity between argumentation poles.

Given these four attributes, we develop a novel argumentation polarization model by adapting a multi-modal economic polarization model proposed by Esteban and Ray [43]. Their original model used three features to measure polarization among different income groups: 1) homogeneity within each group, 2) heterogeneity across groups, and 3) a small number of significant groups, which are similar to our attributes 1, 3, and 4. However, we still need to incorporate attribute 2 into their model to make it suitable for argumentation polarization. Their original model measures polarization as a result of conflict between any two individuals, regardless of whether they are in a pole or not. Argumentation polarization, on the other hand, requires that polarization may only occur as the result of conflicts between two significant groups/poles of users. Thus, we adapted their model to consider the population sizes of each user group (attribute 2) by introducing a minimum pole strength threshold requirement to the model. This threshold ensures that only users in a sufficiently strong pole will produce polarization, and users who are not in a pole will not produce polarization. Additionally, we also normalized the model by total discussion population size to ensure that discussions containing more participants are not always more polarized than discussions containing fewer participants (as is the case in their original model) and we normalize the index based on selected parameters in the model to ensure the polarization index output by the model is between 0 and 1. These adaptations ensure that the resulting modified argumentation polarization model (MAP) considers all four important attributes of argumentation polarization and is normalized by the population size.

We evaluate our model on an empirical dataset of online discussions collected using our cyber argumentation platform, the Intelligent Cyber Argumentation System (ICAS). ICAS is equipped with a powerful fuzzy logic argument reduction system, which was used to approximate each user's overall agreement stance toward the discussion topic, based on their discussion posts. The resulting distribution of user agreement (i.e. their opinion stance toward the discussion topic) in a discussion was fed into the model as input to evaluate the degree of argumentation polarization in the agreement distribution.

To our knowledge, we are the first to present a model of argumentation polarization, so we are not able to directly compare our model to other existing argumentation polarization models. Instead, we compare our model to two other distribution-based polarization models that have been applied to online discussions, Flache and Macy's model (FM) [44] and Morales et al.'s model (MBLB) [27]. These models have been applied in contexts similar to cyber argumentation, and demonstrate two popular approaches to modeling polarization: a variance-based approach and a bi-modal based approach respectively. We compare these models to our model, both analytically and on the empirical dataset, in terms of their ability to capture the four attributes of argumentation polarization. Our analytical and empirical results indicate that our model performs more effectively in terms of the definition of argumentation polarization and our four attributes.

We further justify our model's multi-modal approach to argumentation polarization by examining the topic selection made by the participants in their arguments in the discussion. The results of this analysis show that a bi-modal assumption does not sufficiently capture the important differences between participants and a multi-modal approach is more effective.

This article makes the following contributions:

1) We present a novel model of argumentation polarization that measures the polarization within an agreement distribution in cyber argumentation. We identify four key attributes of polarization within the agreement distribution that must be considered when modeling polarization, which our model captures.
2) We compare our model to two other polarization models used in online discussion and online deliberation literature, both theoretically and on empirical data. Our results indicate that the other models do not sufficiently capture our four attributes of argumentation polarization, which results in some unorthodox polarization values in the empirical data.
3) We justify our model's multi-modal approach based on our analysis of frame and topic selection in discussions from our empirical data.

II. Related Work

A. Cyber Argumentation Platforms

Cyber argumentation and online deliberation systems are specialized platforms that focus on facilitating and analyzing high-quality online discussion. gIBIS [1] was one of the first online deliberation platforms and used the IBIS method [45] to structure and visualized discussions around issues. Convince Me [46] both aimed to increase the reasoning of its participants, as well as facilitate discussions, by providing a knowledge-eliciting interface and simulation-driven feedback. The HERMES system [2] improved discussions for online collaborative decision making by integrating database querying capabilities and providing analytics around the amount of evidence related to an argument, among other analytics. The Araucaria system [3] provided argumentation diagraming to explicitly represent various relationships between different elements in an online discussion. The CoPe_it! tool [4] provided different visualizations and levels of formality (sets of rules in the system) to facilitate many different types of online discussions. The Synergy platform [47] analyzes an argument's probabilities of being accepted by the participants, in addition to providing structure to the discussions. The Deliberatorium [6], [7], [42], [48], [49] platform uses moderators along with a formalized argument map discussion structure to support large scale discussions. Other debate centric platforms originating from for-profit and non-profit enterprises such as Debatepedia (Debatepedia.org), Kialo (kialo.com), debate.org, procon.org, and idebate.org, provide structured discussions and are notable for attracting large audiences or for generating data used in online argumentation research (see [50], [51] for examples).

Many of these platforms identified the unstructured nature of online forums, social media, and social networking platforms as a major issue in facilitating high-quality online deliberation, and addressed it by enforcing explicit argumentation structures centered around discussions of issues, such as Toulmin's model of argumentation [52], IBIS [45], and Zeno [53]. Many platforms also improved understanding and readability of online argumentation by visualizing key elements of argumentation [1], [3], [4] and sectioning off unrelated content from one another [2], [48].

Our platform used in this research, the Intelligent Cyber Argumentation System, is the newest iteration of the platform originally developed by Liu et al. [5]. ICAS is a platform designed for large scale deliberation, like the Deliberatorium, and provides a fuzzy-logic based argument reduction system for approximating user opinion on the discussion topic, which will be discussed in the next section.

B. Polarization Modeling

In the broader polarization literature, there are typically three approaches to modeling and measuring opinion polarization: 1) A survey-based approach, 2) A network-based approach, and 3) A distribution-based approach.

The survey-based approach is the most common and uses the difference between user responses on surveys to measure the total amount of polarization between groups [26], [54]–[59]. Polarization itself is measured by combining the survey responses using some relevant measurement, spanning from simple measurements [54], [56], such as averages, variance, and differences, correlation analysis between responses [55], regressions [26], [57], and prediction models [29]. Survey approaches, however, have the weakness that the opinion expressed by the users may not be well considered or well-informed [38].

The network-based approach models polarization as the degree to which a social network is clustered into distinct, conflicting groups [23], [31], [32], [39], [40], [60]. This approach to polarization is popular among those studying polarization in social media [23], [39], [40] or other online content like blogs [31] and those modeling polarization using agent-based network simulations [60]–[62]. Typically these models approach opinion formation as a result of social interactions through network connections [23], [39], [40] and measure polarization as the degree of group cohesion in the network [23], [31], [32], [63].

The distribution-based approaches models polarization as a function of a single or multiple variables distributed across a population [27], [43], [44]. This approach conceptualizes polarization as a division of a population into distinct, conflicting, internally-cohesive groups along some axis. While some studies used basic metrics [64], such as the number of peaks in the distribution [65], to measure polarization, often studies present their own unique models of measuring polarization on a given distribution [27], [43], [44], depending on the type of polarization being modeled.

Of these approaches, only the distribution-based approach applies to argumentation polarization. Survey-based approaches gather opinion information from surveys, which do not require users to justify their opinions, as argumentation polarization does. Network-based approaches require the subjects to be arranged in a social network, where users formulate their opinions based in part on their network connections. However, cyber argumentation does not require users to form a social network and instead encourages opinion formulation through deliberation. Thus, a distribution-based approach, using the user's agreement or disagreement toward the discussion topic as the distribution variable, is most applicable to argumentation polarization.

Work by Bramson et al. breaks down distribution based polarization measurements into nine different, pairwise-independent senses of polarization [66], [67]. These nine senses of polarization, spread, dispersion, coverage, regionalization, community fragmentation, distinctness, and group divergence, are designed to help examine polarization models to determine which senses of polarization are and are not captured by the model. The authors note that different types and contexts of polarization can focus on different subsets of their proposed senses. We discuss our model's coverage and the coverage of other relevant models of these nine senses in later sections.

III. Background

A. ICAS

In this section, we will briefly outline our argumentation platform ICAS that was used to collect our empirical dataset and derive the user opinion distribution used as input into the opinion polarization models evaluated in this work.

ICAS is a large-scale issue-based online argumentation platform. ICAS is the newest iteration of the argumentation system first developed by Liu et al. [5]. It uses the IBIS model [45] to structure its discussions in a tree-like structure. At the root of each discussion tree is the topic issue. Under each root, at the first level of the tree, there are one or more positions, which serve as the solutions or policies to address the root issue. Discussions in ICAS take place under each position node in the tree. Under each position, there is a sub-tree of arguments that argue for or against their parent position or another argument to a certain degree.

ICAS is unique to other platforms in that it allows users to explicitly express partial agreement (or disagreement) toward another argument or position. When creating an argument in ICAS, the user must provide two components, the argument text and an agreement value. The agreement value is an explicit value associated with the argument that indicates the argument's stance (supporting, neutral, or attacking) and intensity toward the argument/position to which it is replying. Each agreement value is encoded on a scale from -1.0 to +1.0, where the value's sign corresponds to the stance polarity (positive is agree, negative is disagree, and zero is neutral), and the value's magnitude corresponds to the intensity of the opinion (e.g. 1.0 is strong, 0.2 is weak, etc.). The agreement value is selected using a sliding input bar that allows the user to select a value from -1.0 to +1.0 at 0.2 intervals, where each interval is associated with a text description (e.g. -1.0 is complete disagreement, -0.8 is strong disagreement, etc.). Fig. 1 gives an example tree structure for an issue in ICAS.

Discussions in ICAS are asynchronous, unmoderated, and users as assigned anonymous screen names during discussion.

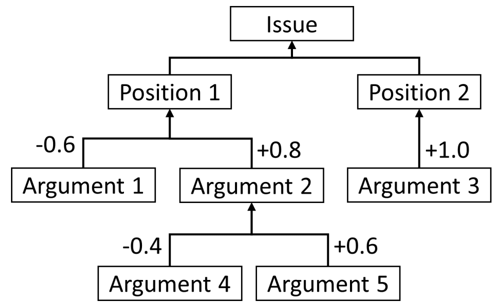


Fig. 1. Example discussion structure used in ICAS. The agreement levels of each argument towards their parent nodes are displayed above the argument.

B. ICAS Argument Reduction

This section outlines the argument reduction process in ICAS that is used to approximate each user's agreement toward a position. Each discussion in ICAS happens at the position level. Thus, every position will have an agreement distribution which describes the extent to which each user agrees or disagrees with the position. This agreement distribution is the input to our argumentation polarization model described in the next section.

To derive the user agreement distribution, we first need to determine each user's overall agreement toward the position. This is done by averaging the agreement values of each argument a user made in the discussion of the position. This process is simple for arguments that are made to directly address the root position, such as arguments 1, 2, and 3 in Fig. 1, however, we must also consider the agreement values from arguments further down the tree, such as arguments 4 and 5. Those arguments do not directly relate to the position, meaning their agreement values are about other posts. Therefore, to capture the impact these arguments had on the user's agreement, we first need to reduce their values such that they related directly to the root position, instead of their current parent, so they can be considered when calculating a user's overall agreement. This argument reduction process is done using ICAS integrated fuzzy logic engine to approximate the agreement values of an argument towards its root position.

ICAS's fuzzy logic engine uses trapezoidal fuzzy logic rules and 25 inference rules to reduce an argument's agreement value one level up the tree at a time until the argument's agreement value relates directly to the parent position. The reduction updates the agreement value an argument has toward its parent to relate instead to its grandparent by identifying the implicit relationship between the child and the grandparent (such as implicit attack/support) and updating the agreement value using the fuzzy logic update rules. This reduction step repeats until the child argument is reduced to relating directly toward its root position.

An example of argument reduction is shown in Fig. 2. For a more in-depth explanation of the reduction process, please refer to Liu et al. [5].

Once each argument has been reduced to directly relate to the root position, ICAS calculates each user's overall agreement on the position by averaging all of their argument's reduced agreement values together. The averaged value is

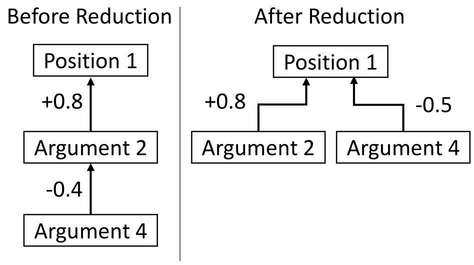


Fig. 2. An example of an argument reduction. The example shows Argument 4's relationship and agreement value before (left) and after (right) the reduction step.

used as the user's overall agreement toward the position. This process is performed for each user, whose agreement values are collected into one user agreement distribution on a position, which is used as input into the argumentation polarization model.

IV. Argumentation Polarization Formulation

To model argumentation polarization as a function of the agreement distribution of the users in a discussion, we must first identify the key attributes of the agreement distribution that characterize polarization.

A. Polarized vs Non-Polarized Argumentation Distributions

The first consideration to make in investigating argumentation polarization is to identify distributions that are at the extremes of polarization. The most polarized scenario for a distribution is widely considered to occur when the entire population is split evenly among the most two most extremes [27], [43], [44]. In this scenario, both poles have a maximum distance between one another, collectively contain the entire population (no outliers), are entirely internally consistent, and have equal strength (in terms of population and similarity). This is the only scenario for maximum polarization.

For a distribution that contains no polarization, there are two scenarios. The first scenario is when only one pole forms in the distribution. If all of the users in the distribution agree with one another, there is only one pole, and there are no rival poles to cause polarizing tension. Likewise, even if some individual users dissent from the pole, they are treated as outliers and have no impact on the polarization, unless they coalesce into a like-minded pole with a substantial enough population size to rival the original pole.

The second scenario is when all of the users disagree with one another, to the extent to which that is possible, resulting in the uniform distribution. If each user is uniformly distributed across all agreement values, then there are no groups of similar users of significant size, and thus there are no poles. Every section of the distribution has the same population and similarity as every other section of the same size. Since no conflicting groups of significant size form in the uniform distribution, according to our definition of argumentation polarization, no polarization can occur in the distribution. Thus, a uniform distribution yields no polarization.

B. Attributes of Argumentation Polarization

From our examination of the maximally and minimally polarized distributions, we can observe four main attributes of the distribution that determine the degree of argumentation polarization. These attributes can be seen as the key determinants of polarization in the agreement distribution in argumentation.

Attribute 1: The total number of argumentation poles.

The biggest indicator of argumentation polarization is the total number of poles in the distribution. As discussed in the previous subsection, if a distribution does not contain at least two poles, then no polarization can exist. So, the distribution needs at least two argumentation poles for argumentation polarization to occur.

However, in the maximal polarization distribution, there are only two poles. The addition of more than two poles would lower the argumentation polarization because the population would be split between all of them, which would weaken each individual pole. From this observation, we can say that the population of an argumentation pole is an important factor in the pole's overall strength, which leads to attribute 2.

Attribute 2: The population size of the argumentation poles.

The population size within an argumentation pole is the main component of its strength. Poles with more users within them are stronger than poles with fewer users. If more users are grouped into conflicting poles, more polarization should occur. Likewise, the ratio between two poles' population size is important when determining the amount of conflict between them. If one pole has significantly more population in it than the other, then the stronger pole overwhelms the weaker pole, resulting in lower polarization. If the poles have a similar population size then they are more equal in strength and thus produce more polarization. However, population size is not the only component of a pole's strength, which leads to attribute 3.

Attribute 3: The similarity within the argumentation poles.

In addition to the population, the strength of the agreement pole is also related to the agreement similarity of the users within it. The more unified the users in the pole are, the stronger the pole. In our distribution with maximum polarization, each pole was entirely internally consistent. If the poles were not entirely consistent and had some internal variance, they would be weaker poles and the overall argumentation polarization would be lower. Thus, the internal distances between the users in the pole impact the strength of the pole and the overall argumentation polarization.

Attribute 4: The dissimilarity between the agreement poles.

The polarizing tension created between argumentation poles is measured by their difference in agreement. The larger the difference between two poles the more polarization is occurring as a result. In our distribution with maximum polarization, the poles were at a maximum distance from one another. If they were to move closer to each other, then the overall polarization would lower.

These four attributes characterize the presence and intensity of argumentation polarization in the agreement distribution of a discussion in cyber argumentation. In general, polarization is greater when there are fewer poles (but more than two) (attribute 1), when more of the population is a member of a pole (attribute 2), when argumentation poles are internally similar (attribute 3), and when distinct argumentation poles are externally dissimilar (attribute 4).

From Bramson et al's examination of distribution-based polarization measurements, they identify nine pair-wise independent senses that describe various aspects of polarization [66], [67]. These nine senses were designed to be broad enough to cover several types of polarization in a variety of contexts, so all nine senses may not apply to every context. While our attributes were developed independently from these senses (we were unaware of Bramson et al.'s work at the time), they map neatly to four of the nine senses that are most indicative of argumentation polarization: community fracturing, group divergence, group consensus, and size parity. The other five senses do not directly reflect the absence or presence of argumentation polarization in the context of cyber argumentation and so are not covered by attributes. For example, the spread of the distribution may not be indicative of argumentation polarization if extreme views are held by outlier users in the distribution (which was common in our empirical data). Thus our model does not explicitly cover these senses, though these senses are captured implicitly in the model due to its formulation, which we outline in the next section.

V. Argumentation Polarization Model

The argumentation polarization model will take in a distribution of user agreement toward a position as input and calculate the total amount of argumentation polarization according to our four identified attributes. Our model is adapted from a distribution-based polarization model by Esteban and Ray [43] for measuring economic polarization. Their model is designed using an axiomatic approach based on three basic features they observed about economic polarization: 1) homogeneity within each group, 2) heterogeneity across groups, and 3) a small number of significantly sized groups. While designed for economic polarization, this model has been applied to other domains to measure polarization within a distribution [68]–[71]. As we can see, their features are very similar to our attributes 1, 3, and 4, which makes this model a good fit for argumentation polarization. But, we still need to adapt the model to consider attribute 2 (population sizes of poles). Additionally, we want to ensure that the polarization index produced from the model is normalized by the population size so that discussions with more population do not always result in increased polarization, so we also normalized the model by population size.

Their original extended model is shown in (1) and (2). The model takes in a distribution (π, y), which is a set of user-value pairs (π_i, y_i), where π_i is the total number of users with value y_i in the distribution. Their model measures polarization as the linear representation of the distances between each user along the distribution, weighted by the mass of similarly clustered users. The max function $(\max(|y_i - y_j| - D, 0))$ calculates the distance between users. The parameter D, is a threshold value that determines if two users are similar $(|y_i - y_j| \leq D)$ (do not produce polarization) or are dissimilar $(|y_i - y_j| > D)$ (produces polarization) with each other. The identity function (I_i) for a user is the total amount of similarity a user has with their near-by neighbors. Here, identity is equivalent to the internal strength of the user's pole (i.e. the more similar others are, the more the user identifies them as in their pole). The function $(w(y_j))$ is a user-defined function that determines how the degree of similarity is weighted in the identity function. The variable α acts as an identity intensifier and determines how important identity (i.e. a pole's internal strength) is in the model. Esteban and Ray prove that α must be bounded by $\alpha \in (0, \alpha^*]$ where $\alpha^* \cong 1.6$, to satisfy their axioms.

$$P(\pi, y) = \sum_{i=1}^{N} \sum_{j=1}^{N} \pi_i \pi_j I_i^{\alpha} \max\{|y_i - y_j| - D, 0\} \quad (1)$$

$$I_i \equiv \sum_{j:|y_i - y_j| \leq D} \pi_j w(y_i) \quad (2)$$

Their model captures their three basic features: 1) homogeneity within each group, 2) heterogeneity across groups, and 3) a small number of significantly sized groups, which are analogous to our attributes 1, 3, and 4. The first feature is captured through the identity function. If a user's neighbors are more similar to them (homogeneous), in terms of distance along the distribution, then their identity value will be larger, which will result in more polarization. The second feature is captured through the max distance function. If two groups of users are further away from one another (heterogeneous), their total distance along the distribution will increase, which will result in a greater amount of polarization. The third feature is captured through the product of the distance function and identity function. Since the population size is finite, if more of the population is distributed among a smaller number of poles, each pole's identity value will be larger, resulting in greater polarization value. For brevity, we will not outline how the model performs on specific theoretical cases, and instead will refer to their original paper [43] for these cases and a more in-depth discussion of the base model. Our modifications do not change the fundamental operation of the model relating to these features, so their examples and discussion are still is applicable to our modified model.

Their original model does not explicitly consider the population sizes of poles (attribute 2). Instead, the model implicitly considers population size by weighing the polarization produced from a user by their identity function (I_i) which is the sum the similarity of their neighbors. So, users with more neighbors (i.e. are in a bigger pole) will produce more polarization than users with fewer neighbors. This approach is sufficient for economic polarization because their distribution domain is unbounded (from 0 to infinity). So, for example, the polarization for a uniform distribution in an unbounded domain would have each user in their own pole $(I_i = 1)$ and would yield a small polarization value (assuming we normalize by population size). However, our agreement distribution domain is bounded (from -1 to +1). So a uniform distribution in our distribution domain would not yield small identity values for the users and instead would measure a substantial amount of polarization. From our formulation of argumentation polarization, a uniform distribution does not contain polarization, so the model should output a polarization index value of zero when given a uniform distribution, which their model, in its original form, does not.

Thus, we adapt the model by defining a threshold value for a user's internal pole strength (i.e. their identity value (I_i)). If a user's identity value is greater than the threshold, T, they are considered a member of a significant pole and can produce polarization. This threshold checks if a user is in a valid pole or not, as shown in (3). This threshold ensures that users are only considered a member of an agreement pole if their pole is sufficiently strong enough to produce polarization.

$$pole(\pi_i, y_i) = \begin{cases} \text{True} & \text{if } I_i > T \\ \text{False} & \text{if } I_i \leq T \end{cases} \tag{3}$$

Since we are using the uniform distribution as the baseline for a non-polarized distribution, we can set the value of T to be the expected identity value from a uniform distribution. If a user does not have an identity value greater than T, then they are not identifying more than they would in a uniform distribution, and thus are not considered part of a pole. Using this definition, we set T as a function of y_i (the user's agreement value) and the parameter D. The function $T(y_i)$ shown in (4) calculates the expected identity value of the user at y_i if they were in a uniform distribution, bounded by $[-1, +1]$. The second term ($|y_i| + D > 1$) accounts for the distribution boundary.

$$T(y_i) = \begin{cases} D/2 & \text{if } |y_i| + D \leq 1 \\ \frac{2D^2 - (|y_i| + D - 1)^2}{4D} & \text{if } |y_i| + D > 1 \end{cases} \tag{4}$$

In addition to the threshold T, we also want to normalize the polarization model to the total population size. Esteban and Ray's original model does not weight by population, so adding more users in the distribution will always increase polarization, regardless of where they are in the distribution. Therefore we normalize the model by the total population size N. The resulting adapted model is shown in (5) and (6).

$$P(\pi, y, N) = \sum_i^N \sum_j^N \left(\frac{\pi_i \pi_j}{N} I_i^\alpha \max\{|y_i - y_j| - D, 0\} \right) \\ \text{where } i : pole(\pi_i, y_i),\ j : pole(\pi_j, y_j) \tag{5}$$

$$I_i = 1 + \sum_{j:|y_i - y_j| \leq D} \left(\frac{\pi_j}{N} * \left(1 - \frac{y_i - y_j}{D} \right) \right) \tag{6}$$

For the identity function (6), we implemented the $w(y_i)$ function as the linear distance away from the target user. Notice that the identity function is normalized by population size. We add one to the summation to ensure the identity value is always greater than one, which is important to ensure that parameter α operates as an identity intensifier and behaves as intended.

The resulting adapted model shown in (5) now explicitly considers the population size of an argumentation pole (attribute 2). If a user does not have an identity value greater than T, then they are not within a pole with sufficient population and similarity to be considered a valid argumentation pole. In this model, a uniform distribution would not produce any polarization.

However, depending on the parameter values selected for variables D and α, the maximum value of (5) may be greater than or less than one. To fix our model's output range to be between zero and one, we normalize the model using min-max normalization. The maximum value of P, given the parameters, can be calculated by providing the maximally polarized distribution: $[((N/2), -1), ((N/2), +1)]$. Since the model normalizes the distribution by population size, we can derive $P_{\max}$ as a function of D and α as shown in (7).

$$P_{\max} = \left(1 - \frac{D}{2}\right)(1.5^\alpha) \tag{7}$$

Using this maximum (and since we know the minimum value is zero), P can be normalized with respect to the input parameters is shown in (8). We define the value of P_{norm} as the polarization index value in our final model.

$$P_{\text{norm}}(\pi, y, N) = \frac{P(\pi, y, N)}{P_{\max}} \tag{8}$$

Our adapted model still satisfies Esteban and Ray's original axioms used to develop their model, if we assume that all of the poles in the axioms are above threshold T (except pole p in Axiom 4, which they describe as insignificant) [43].

In terms of our four attributes of opinion polarization in cyber argumentation, our model also accounts for each of these attributes as well. The similarity within each pole (attribute 3) is modeled by the identity function I_i, which increases the polarization produced by a pole as a factor of its total internal similarity. The number of poles (attribute 1) is also modeled by the identity function by increasing the amount of polarization produced by a pole based on the proportion of the population contained within it. More poles mean each pole has a smaller portion of the population and thus a lower identification value, which in aggregate lowers the total polarization. The dissimilarity across different poles (attribute 4) is modeled by the max function. The difference in population size (attribute 2) is modeled by the threshold T and by weighing the polarization produced by each pole by their proportion of the population; maximum polarization between two poles occurs when they contain equal population proportionality.

The model has a runtime complexity of $O(n^2)$ (identity values can be computed in advance) where n is the number of unique (π_i, y_i) pairs. The total number of pairs depends on the total number of unique agreement values, and since the distribution is bounded from -1.0 to +1.0, the total possible number of pairs is bounded by the precision of the agreement values. Small differences in agreement (< 0.01) have very little impact on the polarization value, so the number of unique pairs can be limited by rounding the values to a lower precision. For example, rounding the values to two decimal places will ensure that the number of unique pairs will be less than or equal to 201.

Referring back to Bramson et al.'s nine senses of polarization [66], [67], this model covers four senses explicitly (from the attributes) and five senses implicitly. Community fracturing (attribute 1), group divergence (attribute 4), group consensus (attribute 3), and size parity (attribute 2) are all covered explicitly by the four attributes. Spread, dispersion, coverage, regionalization, and distinctness are captured implicitly, but only under some conditions. The biggest condition being that our model tries to ignore outliers using the threshold parameter T. So, revisiting our previous example, the spread of the distribution could be ignored in some cases where the extreme users in the distribution are treated as outliers (i.e. non-poles) by the model. Due to our formation of a single combined polarization model, each sense cannot be examined independently of one another. However, in most common scenarios these five senses are implicitly captured.

VI. Argumentation Polarization Model Parameters

In the previous section, we introduced the polarization model. In this section, we will discuss the two user-defined parameters, D, and α, their role in modeling opinion polarization, and recommended value to use for measuring polarization.

A. Parameter α

The parameter α is designed as an identity intensifier and determines how important internal pole strength is in the model. An α value of zero means internal pole strength does not impact the polarization at all, making the model more similar to inequality measures [43], while an α value of 1.59 (near maximum) would indicate that the model will strongly weigh internal pole strength when measuring polarization.

The α parameter chiefly affects the importance of the internal similarity within the agreement poles in the model (attribute 3). Fig. 3 shows how various α values affect the polarization index value for a simulated bi-modal distribution with different standard deviations within the poles (higher standard deviation creates less internal similarity). The greater the value of α, the more sensitive the model is to changes in the internal similarity of the poles.

In practice, we recommend larger values of α. Argumentation polarization describes the degree to which the users have formed internally similar groups based on their agreement with the topic, making internal identification very important in the modeling of argumentation polarization. An α value that is too low will not take internal similarity of the poles into account, thus we recommend an α value between 0.8 and 1.59.

B. Parameter D

The parameter D acts as the maximum agreement distance threshold that determines if two users are similar to one another (i.e. are in the same pole) or not. For example, if user A has an overall agreement value of $+0.5$, and user B has an agreement value of $+0.68$, then these two users would identify with one another if D was greater than or equal to

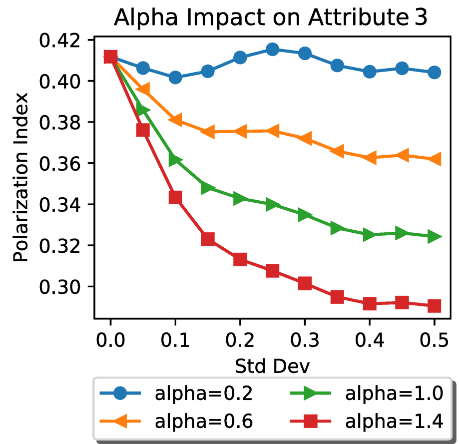


Fig. 3. Comparison of different α values on the polarization index for a simulated bimodal distribution with different standard deviations within the poles. Other parameters: $D = 0.3$, $T = 0$.

0.18. This parameter acts as the model's consideration of the similarity between two user agreement values. If the D value is set to a small value, then the model will consider smaller differences between agreement more important. Likewise, for larger values for D, the model will only consider larger differences important.

Given the bounds of the agreement distribution, we recommend D values between 0.2 and 0.5. If the value of D is too large, then the model will group users who have large differences in agreement, which will affect the total number of poles the model will consider. For example, if $D = 1.0$, it is not possible to have three non-overlapping poles in the distribution. However, D values that are too small will cause the model to over exaggerate the minor differences in user agreement in its measurement of argumentation polarization.

VII. Experiments and Comparison with Polarization Models

In this section, we compare our proposed modified argumentation polarization model (MAP) with two other distribution-based polarization models that have been applied to online deliberation research. We compare the models both theoretically on their design in terms of both our four identified attributes of argumentation polarization and using Bramson et al.'s nine senses of polarization [66], [67].

We also compare the models' performance on an empirical dataset of cyber argumentation discussions collected using our cyber argumentation platform ICAS. In that comparison, we demonstrate how the different approaches taken by each model affect how polarization is reported for three illustrative discussions in our ICAS empirical dataset.

A. Flache and Macy's Model (FM)

In 2014, Gabbriellini and Torroni presented an agent-based dialogue simulation model for argumentative reasoning [72]. As part of their analysis, they measure the polarization of the simulated dialogs using a polarization measurement presented by Flache and Macy [44].

Flache and Macy's model (FM) is relatively straightforward. Given a distribution of user opinions on a bounded scale of -1 to +1, the distance between two users' opinions is d_ij. The level of polarization of the population N is the variance of the distribution of all distances between every user as shown in (9).

$$P = \frac{1}{N(N-1)} \sum_{i=1}^{N} \sum_{\substack{j=1, \\ i \neq j}}^{N} \left(d_{ij} - \overline{d}\right)^2 \tag{9}$$

Where $\overline{d}$ is the average opinion distance across all pairs of opinions (excluding self-distances).

B. Morales, Borondo, Lasada, and Benito's Model (MBLB)

Morales, Borondo, Losada, and Benito's polarization model (MBLB) [27] was originally used to measure polarization in a Twitter discussion about the late Venezuelan present, Hugo Cháves, in 2015.

Their model assumes a bi-modal distribution. Given a distribution of opinion values of the range -1 to +1, polarization is measured by calculating the center of gravity (average) of each side of the distribution, both positive (range: (0,+1]) and negative (range: [-1,0)), and subtracting each center of gravity from one another. Then the value is scaled by the maximum possible distance between the centers of gravity and by the difference in both sides' population sizes. The model is shown in (10).

$$P = \left(1 - \frac{|N_{pos} - N_{neg}|}{N}\right) \frac{|gc^{+} - gc^{-}|}{2} \tag{10}$$

Where N is the total population size, N_{pos} is the population size with opinion greater than 0, N_{neg} is the population size with opinion less than 0, and gc^{+} and gc^{-} are the averages of the opinion values in populations, N_{pos} and N_{neg} respectively.

C. Theoretical Comparisons of Models

We compare the assumptions and approaches made the two models introduced in the previous section in terms of our definition and attributes of argumentation polarization, and the nine senses of polarization proposed by Bramson et al. [66], [67].

The MBLB model is most distinct from the other two in that it takes a bi-modal approach to model polarization. Bi-modal approaches are very common in analysis of polarization [29], [39], [40], [58]. This approach assumes that polarization is the result of a population dividing into two groups that are in opposition to one another. The direct analog to argumentation polarization would be to assign the participants to agree and disagree groups based on their agreement polarity. However, a consequence of grouping participants in this way is that it strips out the strength of their agreement or disagreement in the model.

Research in political psychology has suggested that careful consideration of opposing arguments surrounding an issue can cause ambivalence among the participants [73]–[75], resulting in a weak commitment to attitudes about an issue [73], [74] Likewise, group polarization research suggests that some users become more extreme in their opinion as a result of group discussion, especially when users mostly engage with their like-minded peers [76]–[79] and ignore the arguments and viewpoints with which they disagree. Since participation in online argumentation may produce different outcomes for different types of users, it is somewhat misleading to group the participants based only on whether they agree or disagree with a position, regardless of their agreement strength. In cyber argumentation, more ambivalent users likely have more in common with one another, even if they are on the other "side" of the agreement spectrum than with more extreme users.

This approach that assumes the group boundary exists at the origin may not be conducive to cyber argumentation, since it does not consider the ambivalence of the participants. Furthermore, this model does not consider intra-group similarity (our attribute 3) or Bramson et al.'s sense of group consensus [66], [67]. As previously discussed, this sense/attribute is important to argumentation polarization, as the strength of the user agreement/disagreement is very consequential in group discussions. In addition to group consensus, MBLB does not consider the spread, regionalization, or coverage senses, which indicates that this model is unable to detect the presence (or absence) of distinct groups in various regions of the distribution.

The FM model characterizes polarization as the variance of the distribution of distances between the users' agreement. Using the variance is a common polarization modeling technique [54], [80], [81]. This approach does not explicitly consider the formation of poles and instead only focuses on the distances between the user agreement.

In terms of Bramson et al.'s senses of polarization [66], [67], this model only explicitly covers dispersion and does not cover the other eight senses, including community fracturing, distinctness, group divergence, and size parity which we explicitly identify in our four attributes as important for capturing argumentation polarization. As a result, the FM model may register polarization being present in a distribution that contains zero poles. Consider our previous example of a uniform distribution. As previously discussed, the uniform distribution does not contain any poles, since there are no significant clusters of users, and as such there can be no polarization. However, the FM model does not consider the formation of clusters or poles, so it will measure a non-zero polarization value.

As discussed in Section IV, the presence or absence of poles is important to argumentation polarization. Argumentation polarization is not merely a lack of consensus, instead it is explicitly characterized by the formation of internally similar groups in the agreement distribution. In terms of our attributes, FM ignores the number of poles (attribute 1).

D. Empirical Comparison of Models

In the previous subsection, we outlined the major theoretical differences between our model and two other distribution-based models used in polarization analysis of discussions in terms of our definition and attributes of argumentation polarization. In this section, we build on that analysis by comparing how the models measure polarization on empirical discussion data collected using our cyber argumentation platform ICAS.

We compare the polarization results for each model on three discussions from two of the four issues in the empirical dataset. Since each of the polarization models only considers the distributions of the user agreement, not discussion topics themselves, and all of the participants were the same across all issues, we can compare the discussions of the two issues together without issue.

These three discussions were selected because they best demonstrate how the approaches of the models affect how polarization is calculated from the distribution. By comparing the polarization results from each model to the underlying user agreement distribution, we illustrate how the FM and MBLB models can produce polarization values that run counter to our conceptualization of argumentation polarization outlined in our four attributes of argumentation polarization, while our MAP model produces results that are consistent with the attributes.

1) Empirical Dataset: In the spring of 2018, we conducted an empirical study on an undergraduate class of general sociology. We offered extra credit for students if they used our argumentation platform ICAS to discuss and debate four issues relating to their coursework. The study began on April 10th and ended on May 5th, 2018. The students were asked to discuss four issues; each issue had four positions under them. The positions were constructed such that for each issue, there was a strong conservative position, a moderately conservative position, a moderately liberal position, and a strong liberal position. The students were asked to post 10 arguments under any of the positions for each issue. The issues in the study pertained to the content they were covering in class. The issues where:

- Healthcare: Should individuals be required by the government to have health insurance?
- Same Sex Adoption: Should same sex married couples be allowed to adopt children?
- Guns on Campus: Should students with a concealed carry permit be allowed to carry guns on campus?
- Religion and Medicine: Should parents who believe in healing through prayer be allowed to deny medical treatment for their child?

The study contained 308 users (N=308, Gender: 40% Male, 60% Female, Race: 79% White, 21% Non-White) who posted a total of 10,573 arguments under the four issues. On average, users posted 2.6 arguments per discussion. The study was completed with IRB approval (Protocol # 1710077940).

Using the agreement reduction method described in Section III-B, each user was assigned an overall agreement value for each position discussion. If the user did not participate in the discussion they were not included in the distribution. These overall agreement values were used as the input agreement distribution for the polarization models.

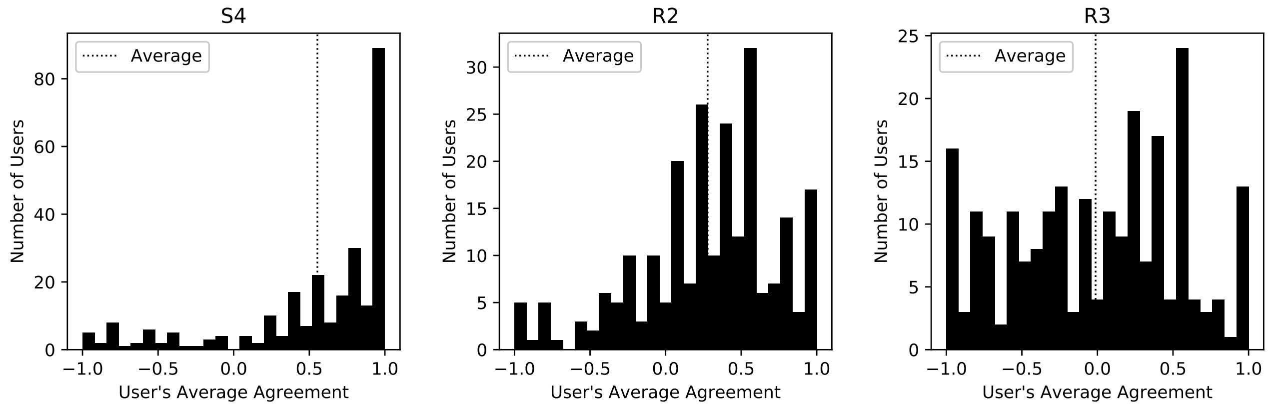
2) Comparison of Models on Empirical Data: Table I shows the polarization index value for each of the models on each of the discussions of the positions. For brevity, we will examine only the Same Sex Adoption and Religion and Medicine issues. The labels for the positions are assigned as such: the first letter reflects the issue the position is under (S = Same Sex Adoption, R = Religion and Medicine), and the number represents the ideological tilt of the position (1 = Strong Conservative, 2 = Moderately Conservative, 3 = Moderately Liberal, 4 = Strong Liberal). Fig. 4 shows the agreement distributions for each of the positions that are assigned the highest polarization value by each model. Our model, MAP calculated that R2 was the most polarized position (0.0297), FM calculated S4 (0.2914), and MBLB calculated R3 (0.4662). MAP was run with parameters $D = 0.5$, $\alpha = 1$.

TABLE I
THE RANK AND POLARIZATION VALUE FOR ALL OF THE POSITIONS IN THE RELIGION AND MEDICINE AND SAME SEX ADOPTION ISSUES FOR EACH POLARIZATION MODEL.

Rank	MAP (Score)	FM (Score)	MBLB (Score)
1	R2 (0.0297)	S4 (0.2914)	R3 (0.4662)
2	R1 (0.0264)	S1 (0.2437)	S3 (0.4633)
3	S3 (0.0059)	R3 (0.2183)	R1 (0.3000)
4	R4 (0.0048)	S3 (0.2133)	S2 (0.2674)
5	S2 (0.0029)	S2 (0.2102)	S1 (0.2269)
6	R3 (0.0002)	R1 (0.1781)	S4 (0.2032)
7	S4 (0)	R4 (0.1741)	R2 (0.1994)
8	S1 (0)	R2 (0.1652)	R4 (0.1496)

FM's most polarized discussion is S4 with a polarization value of 0.2914. From our attributes of argumentation polarization, S4 is not very polarized since almost all of its population concentrated between agreement +0.5 and +1.0, creating only one major pole (attribute 1). Intuitively, positions R2 and R3 are more polarized than S4, as the users in these distributions are more spread out. However, FM's approach to modeling polarization does not consider the formation of poles, and instead only considers the variance between the distances between users. The variance of the distance distribution in S4 is greater than R2 or R3 because the concentration of users in the agreement distribution at +1.0 creates greater distance with every other user. In R2 and R3, the users are more evenly distributed from one another, resulting in distances that are closer to the average distance. Since FM only considers the variance of the distances and does not consider the formation of poles in the agreement distribution (attributes 1 and 2), its measurements on the empirical study produce results that are inconsistent with our definition of argumentation polarization. Similar scenarios can be observed in positions S1, R4, and G4 (see Appendix).

MBLB's most polarized discussion is R3 with a polarization value of 0.4662. MBLB's definition of polarization focuses on the concentration of user agreement at the extremes, as does the model's approach. MBLB primarily gives higher polarization values to distributions that are evenly split across the origin and are closer to the extremes. This is true of position R3's distribution, however, MBLB does not explicitly look for poles, it instead assumes they are there. Thus, R3's distribution is assumed to be in two poles (one on the positive side and one on the negative), even when the distribution, as shown in Fig. 4c, has its population relatively evenly spread. The population in R3 is evenly distributed (48% negative,



(a) Agreement Distribution for Position S4 (b) Agreement Distribution for Position R2 (c) Agreement Distribution for Position R3

Fig. 4. Histograms of users by their overall average agreement for each position.

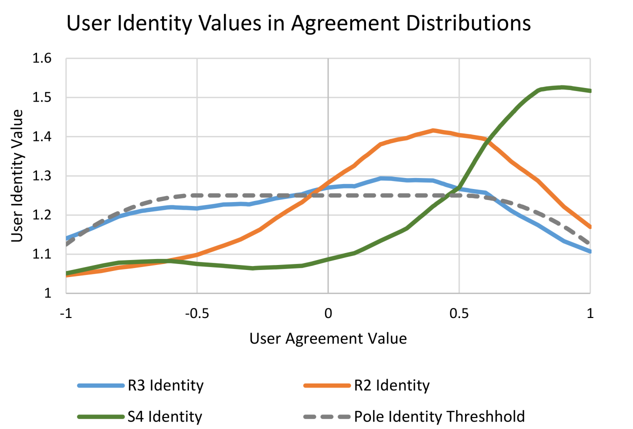


Fig. 5. The population pole sizes for poles centered at each agreement value in positions S4, R2, and R3. The dashed line is the uniform distribution threshold.

52% positive) with either side's center of gravity (i.e. average agreement value) roughly in the middle of the distribution (negative center: -0.51, positive center: +0.45). In the MBLB model, this distribution is interpreted as decently polarized, however, when looking at the actual distribution in Fig. 4c, the distribution more closely resembles a uniform distribution than a bi-modal distribution. A similar scenario can be observed in position H4 (see Appendix).

Our model MAP, for comparison, measured a polarization index value of 0.0002 for position R3, one hundred times lower than for position R2. This low polarization value was due mostly to only very few of the users in R3 being above the pole strength threshold T. Fig. 5 shows the identity values of the users in the distribution compared to the threshold T. The majority of R3's population was distributed too evenly to have an identity value greater than the threshold. MAP's most polarized discussion was position R2. Position R2 had the most users with an identity value that was greater than T that were far enough away from one another (a distance greater than D=0.5) to generate polarization. Position S4, by contrast, only had users in poles that were between +0.5 and +1.0, which

TABLE II
THE TOPIC DESCRIPTIONS FOR EACH OF THE TOPICS IN THE RELIGION AND MEDICINE ISSUE.

Topic Number	Topic Description
0	Focuses on arguing that the child's health is more important the parent's religious freedom.
1	Focuses on autonomy of children and who should be allowed to make decisions for whom.
2	Religion's coexistence with Medicine; How religion should play into life and death decisions.

were not far enough away from one another to conflict. Even though position R2 had the most users in conflicting poles, the total distance between the conflicting pole users was fairly low, resulting in a low polarization index value of 0.0297 for R2.

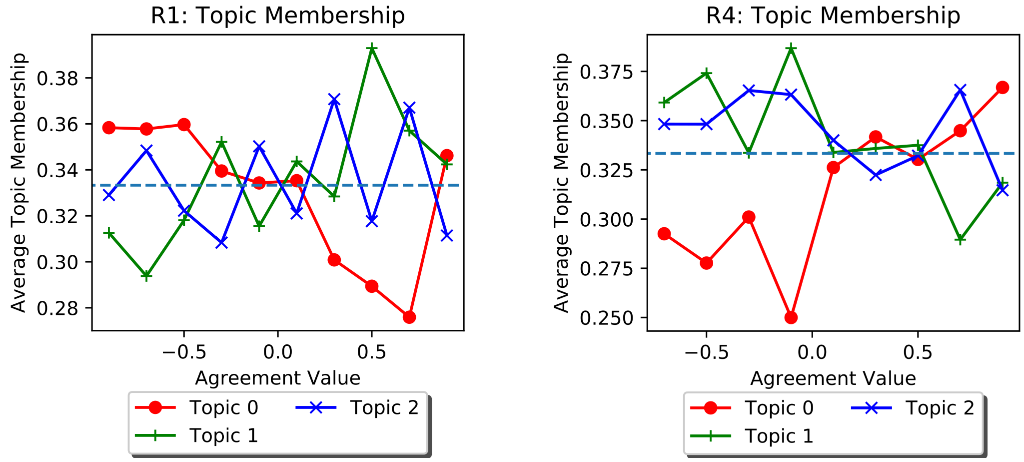
The polarization results and accompanying histograms for the other discussions can be found in the Appendix.

VIII. Justifying a Multi-modal Approach using Topic Modeling

Previously, we discussed the limitations of a bi-modal approach to polarization because it excludes ambivalence and the reservations that users have concerning their agreement toward a position in cyber argumentation. Instead, our model uses a multi-modal approach. While this assumption of a distinction between ambivalent users and users near the extremes is supported theoretically, we want to confirm this difference empirically from our dataset to present a stronger argument.

Thus we examined the agreement groups in terms of topic and framing that users selected during argumentation. Framing and argument topic selection indicate what an author selects as important to discuss [82]. Discussions containing various topics and framings indicate which aspects associated with the issues that the participants are focusing on, and thus indicate their underlying values and thoughts. Topic modeling techniques have shown to be an effective approach to capturing framing in online discussions [30].

We used an LDA model in MALLET [83] to perform topic modeling over all of the arguments under the issues in the



(a) The topic membership for position R1.

(b) The topic membership for position R4.

Fig. 6. The average topic membership for each user grouped by the user's overall agreement value toward the discussion position. The topic membership is determined by all of the arguments posted by the user by issue.

empirical data. The number of topics per issue was selected based on the highest coherence score; we tested two to five topics per issue. For brevity, we will only examine the Religion issue.

The Religion issue had three topics, summarized in Table II. Fig. 6 shows the topic distributions for R4 (the most liberal position) and R1 (the most conservative position). Fig. 6 shows that Topic 0, which focused on prioritizing health over religion (a traditionally liberal argument), tended to be more popular among more liberal participants (i.e. those who disagreed with R1 and agreed with R4). On the other hand, Topic 1, which focused on the autonomy of children versus parental authority (a more conservative argument), tended to be more popular with more conservative participants (i.e. those who agreed with R1 and disagreed with R4). Topic 2, which discussed the coexistence between religion and medical treatment, was not preferred by either liberal or conservative users.

The tendencies of users to discuss these topics were not uniformly consistent across all user agreement groups. Users near the extremes tended to favor one topic over another. Those who disagreed with R4 favored Topics 1 and 2 over Topic 0, those who more strongly agreed with R4 favored Topics 0 and 2 over Topic 1, those who more strongly agreed with R1 favored Topics 1 and 2 over Topic 0, and those who strong disagreed R1 favored Topics 0 and 2 over Topic 1. On the other hand, more moderate or ambivalent users were more likely to discuss all three topics instead of only one or two.

Taken as a whole, the topic distribution by agreement value indicates that users at various levels of agreement, even within the same ”side” of the distribution, have different concerns surrounding the issue. Furthermore, it suggests that more moderate users are discussing a wider variety of topics than those at the extremes, which reflects the literature suggesting that ambivalence is caused in part by considering many arguments. Thus, the approach of assuming a bi-model distribution, based only on whether the user agrees or disagrees, inadvertently combines groups of users who behave differently, discuss different topics, and have different underlying values. Thus, we assert that a multi-modal approach is more suitable for detecting this divide between users than a bi-modal approach.

IX. Discussion

Polarization is a broad concept that can be examined from many different perspectives. For example, polarization can be examined as the quality of respect and attitude from one group toward another (affect polarization) [56], or it can be examined as phenomena that push people toward extreme ideologies (group polarization) [84]. It is unlikely that any one model or definition can cover every facet of polarization.

In this work, we focus on covering polarization from the argumentation perspective. Argumentation polarization focuses on the polarization of the users' agreement toward a discussion topic in cyber argumentation. Unlike other types of polarization, argumentation polarization assumes that the users' attitudes have been developed through informed, thoughtful deliberation. Argumentation requires the participants to carefully consider, defend, and formulate their stance on an issue, which leads to opinions and stances that are better informed through consideration of alternative views, ideas, and experiences [33]–[37] than unjustified opinions [38].

The argumentation process may have different outcome effects on different users. As previously discussed, some users may moderate their opinions and become more ambivalent as they carefully consider arguments from the opposing sides of the debate [73]–[75], while other users may ignore or dismiss the arguments and posts made by those they disagree with and engage mostly with their like-minded peers, becoming more extreme in their stance [76], [77]. Our analysis of topic selection seems to support both of these theories; users on the extremes of the agreement distribution tended to focus on one or two topics while ignoring the topic favored by their opposition, while more moderate users tended to consider all of the topics more equally.

Our model's multi-modal approach is able to make a distinction between users on the extreme and users who are more ambivalent, and pays more attention to the disagreement between users, even when they are on the same side of the distribution. The model's parameters can be adjusted to make the model more sensitive to smaller differences in the user agreement or increase the importance of the formation of argumentation poles.

Polarization as a concept differentiates itself from other phenomena, such as inequality or a lack of consensus, in that it requires the formation of distinct, strong poles that conflict with one another. Our comparison with the FM and MBLB models highlights the importance of how each model handles the presence or absence of poles. FM did not consider poles at all, which resulted in odd modeling behavior in our empirical results. MBLB, on the other hand, always assumes there are only two poles on either side of the distribution, even if the underlying distribution does not match that assumption, which is reflected in the empirical results. Our model does not make prior assumptions about the number and locations of poles in the distribution and instead tries to identify users in poles by proposing a threshold value, T, based on the user's identity with their neighbors. This approach handles different types of distributions better than the bi-modal assumption of MBLB, while still considering the importance of poles, which is key to our definition of argumentation polarization.

Our empirical results showed that argumentation polarization in the discussions in ICAS was lower than one might have expected due to the controversial nature of the issues discussed. Many discussions resulted in either near consensus (such is the case in S4) or near-uniform distribution (such as for R3). For the distributions for each position, please refer to Fig. 7. While these results are only among undergraduate students in a controlled setting, it does suggest that for these issues, users are not polarized as a result of argumentation discussion.

X. Conclusion

Online deliberation through cyber argumentation platforms offers an environment for users to discuss and debate their opinions, viewpoints, and stances on important issues. These well-structured debates help users develop well-informed opinions and stances. However, argumentation polarization often arises in discussions of controversial topics. Thus, we present an argumentation polarization model, adapted from an economic polarization model, to measure the polarization among users in terms of their agreement with the debate topic in cyber argumentation. We discussed how argumentation polarization manifests itself in the distribution of user agreement toward the discussion topic and identified four key attributes of argumentation polarization that a model must capture to effectively measure it. We presented a model that measures the argumentation polarization of an agreement distribution derived from discussions in our cyber argumentation platform. We justified our model's design and adaptations, in terms of our four attributes of argumentation polarization, and compared it to two other distribution-based polarization models. Those models' approaches: treating polarization as the variance of distance between users, and treating polarization as bi-modal phenomena. Both produced results on our empirical data that did not align with our definition and key attributes of argumentation polarization. Our model's multi-modal approach more closely aligns with our observations of argumentation polarization in the empirical data, better adapts to the behaviors of the participants, as shown in our topic modeling analysis, and allows different parameter values to increase or decrease the granularity of the measurement. Thus, our proposed model is effective at measuring argumentation polarization in terms of our definition and attributes of argumentation polarization.

Appendix A
Polarization Results on all Positions

The full list of the polarization values for each of the positions in the dataset for each of the models is shown in Table III. The labels for the positions are assigned as such: the first letter reflects the issue the position is under (G = Guns on Campus, H = Healthcare, S = Same Sex Adoption, R = Religion and Medicine), and the number represents the ideological tilt of the position (1 = Strong Conservative, 2 = Moderately Conservative, 3 = Moderately Liberal, 4 = Strong Liberal).

The histograms of each of the user overall agreement distributions for each of the positions in the dataset are shown in Fig 7.

TABLE III
The Polarization value for all of the positions for each polarization model.

Position	MAP	FM	MBLB
G1	0	0.2649	0.4289
G2	0.0032	0.1998	0.3616
G3	0.0027	0.2350	0.4135
G4	0.0034	0.2328	0.2059
H1	0.0148	0.1808	0.3493
H2	0.0239	0.1819	0.2558
H3	0.0210	0.1629	0.2567
H4	0.0001	0.2259	0.4601
R1	0.0264	0.1781	0.3001
R2	0.0297	0.1652	0.1994
R3	0.0002	0.2183	0.4662
R4	0.0048	0.1741	0.1496
S1	0	0.2437	0.2269
S2	0.0029	0.2102	0.2674
S3	0.0059	0.2133	0.4633
S4	0	0.2914	0.2032

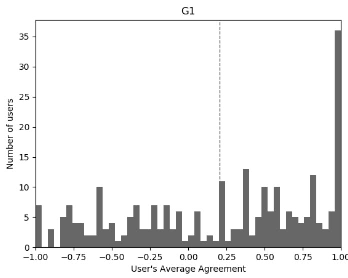
Acknowledgment

This an extended and revised version of a preliminary conference paper that was presented in IEEE ICCC 2018 [85].

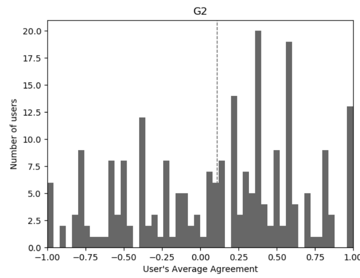
References

[1] J. Conklin and M. L. Begeman, "gIBIS: A tool for all reasons," *Journal of the American Society for Information Science*, vol. 40, no. 3, pp. 200–213, 1989. [Online]. Available: https://asistdl.onlinelibrary.wiley.com/doi/abs/10.1002/%28SICI%291097-4571%28198905%2940%3A3%3C200%3A%3AAID-ASI11%3E3.0.CO%3B2-U

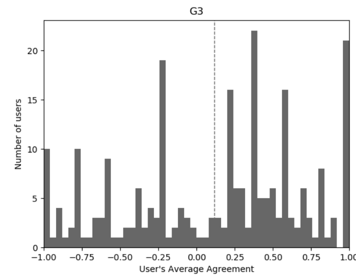
[2] N. Karacapilidis and D. Papadias, "Computer supported argumentation and collaborative decision making: the HERMES system," *Information Systems*, vol. 26, no. 4, pp. 259–277, 2001. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0306437901000205



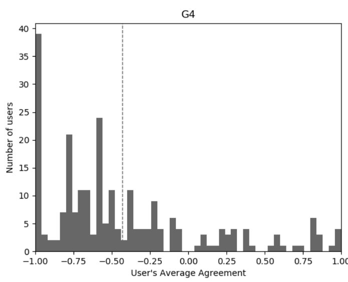
(a) Agreement Distribution for Position G1



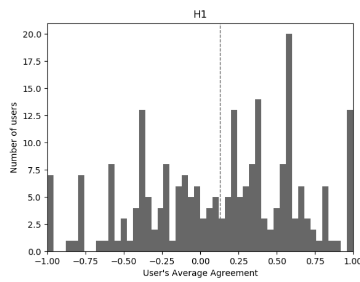
(b) Agreement Distribution for Position G2



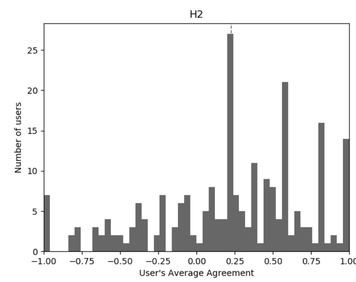
(c) Agreement Distribution for Position G3



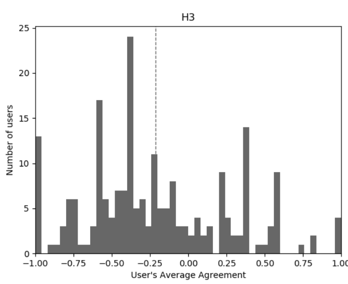
(d) Agreement Distribution for Position G4



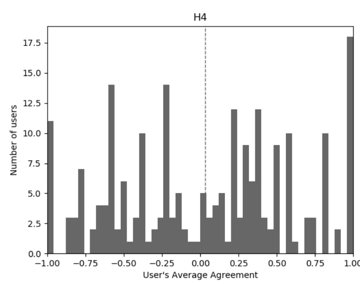
(e) Agreement Distribution for Position H1



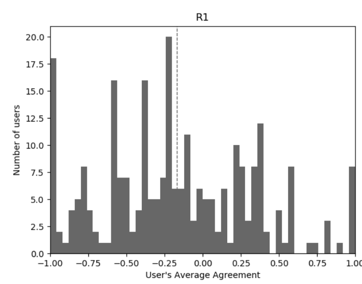
(f) Agreement Distribution for Position H2



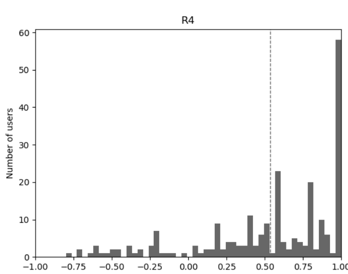
(g) Agreement Distribution for Position H3



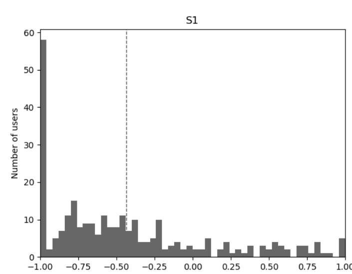
(h) Agreement Distribution for Position H4



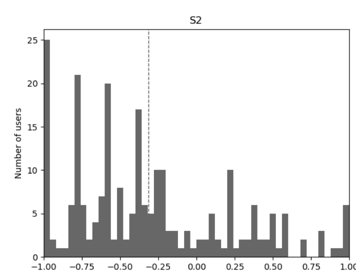
(i) Agreement Distribution for Position R1



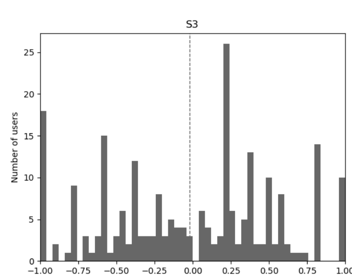
(j) Agreement Distribution for Position R4



(k) Agreement Distribution for Position S1



(l) Agreement Distribution for Position S2



(m) Agreement Distribution for Position S3

Fig. 7. Histograms of users by their overall average agreement for each remaining position.

[3] C. Reed and G. Rowe, "Araucaria: software for argument analysis, diagramming and representation," *International Journal on Artificial Intelligence Tools*, vol. 13, no. 4, pp. 961–979, 2004. [Online]. Available: https://www.worldscientific.com/doi/abs/10.1142/S0218213004001922

[4] M. Tzagarakis, G. Gkotsis, M. Hatzitaskos, N. Karousos, and N. Karacapilidis, "CoPe_it!: argumentative collaboration towards learning," in *Proceedings of the 9th international conference on Computer supported collaborative learning - CSCL'09*, vol. 2. Association for Computational Linguistics, 2009, pp. 126–128. [Online]. Available: http://portal.acm.org/citation.cfm?doid=1599503.1599546

[5] X. F. Liu, E. Khudkhudia, L. Wen, V. Sajja, and M. C. Leu, "An intelligent computational argumentation system for supporting collaborative software development decision making," in *Artificial Intelligence Applications for Improved Software Engineering Development: New Prospects*, ser. Advances in Computational Intelligence and Robotics, F. Meziane and S. Vadera, Eds. IGI Global, 2010, pp. 167 – 180. [Online]. Available: http://services.igi-global.com/resolvedoi/resolve.aspx?doi=10.4018/978-1-60566-758-4

[6] M. Klein, P. Spada, and R. Calabretta, "Enabling deliberations in a political party using large-scale argumentation: A preliminary report," in *Proceedings of the 10th international conference on the design of cooperative systems*, p. 17.

[7] A. C. B. Garcia and M. Klein, "Making sense of large-group discussion using automatically generated RST-based explanations," *SSRN Electronic Journal*, 2015. [Online]. Available: http://www.ssrn.com/abstract=2554838

[8] Z. Papacharissi, "The virtual sphere: The internet as a public sphere," *New Media & Society*, vol. 4, no. 1, pp. 9–27, 2002. [Online]. Available: https://doi.org/10.1177/14614440222226244

[9] E. Bakshy, S. Messing, and L. A. Adamic, "Exposure to ideologically diverse news and opinion on facebook," *Science*, vol. 348, no. 6239, pp. 1130–1132, 2015. [Online]. Available: http://science.sciencemag.org/content/348/6239/1130

[10] A. Bessi, F. Zollo, M. Del Vicario, M. Puliga, A. Scala, G. Caldarelli, B. Uzzi, and W. Quattrociocchi, "Users polarization on facebook and youtube," *PLoS ONE*, vol. 11, no. 8, 2016. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4994967/

[11] S. Flaxman, S. Goel, and J. M. Rao, "Filter bubbles, echo chambers, and online news consumption," *Public Opinion Quarterly*, vol. 80, pp. 298–320, 2016. [Online]. Available: https://academic.oup.com/poq/article/80/S1/298/2223402

[12] W. Quattrociocchi, A. Scala, and C. R. Sunstein, "Echo chambers on facebook," 2016. [Online]. Available: https://papers.ssrn.com/abstract=2795110

[13] S. Hong and S. H. Kim, "Political polarization on twitter: Implications for the use of social media in digital governments," *Government Information Quarterly*, vol. 33, no. 4, pp. 777–782, 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0740624X16300375

[14] C. R. Sunstein, *#Republic : Divided Democracy in the Age of Social Media*. Princeton University Press, 2017. [Online]. Available: http://0-search.ebscohost.com.library.uark.edu/login.aspx?direct=true&db=nlebk&AN=1431815&site=ehost-live&scope=site

[15] K. Coe, K. Kenski, and S. A. Rains, "Online and uncivil? patterns and determinants of incivility in newspaper website comments," *Journal of Communication*, vol. 64, no. 4, pp. 658–679, 2014. [Online]. Available: https://academic.oup.com/joc/article/64/4/658/4086037

[16] I. Rowe, "Deliberation 2.0: Comparing the deliberative quality of online news user comments across platforms," *Journal of Broadcasting & Electronic Media*, vol. 59, no. 4, pp. 539–555, 2015. [Online]. Available: http://0-search.ebscohost.com.library.uark.edu/login.aspx?direct=true&db=cms&AN=111289777&site=ehost-live&scope=site

[17] N. J. Stroud, J. M. Scacco, A. Muddiman, and A. L. Curry, "Changing deliberative norms on news organizations' facebook sites," *Journal of Computer-Mediated Communication*, vol. 20, no. 2, pp. 188–203, 2015. [Online]. Available: https://academic.oup.com/jcmc/article/20/2/188/4067559

[18] R. Medaglia and D. Zhu, "Public deliberation on government-managed social media: A study on weibo users in china," *Government Information Quarterly*, vol. 34, no. 3, pp. 533–544, 2017. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0740624X16301903

[19] K. Esau, D. Friess, and C. Eilders, "Design matters! an empirical analysis of online deliberation on different news platforms," *Policy & Internet*, vol. 9, no. 3, pp. 321–342, 2017. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/poi3.154

[20] D. Friess and C. Eilders, "A systematic review of online deliberation research," *Policy & Internet*, vol. 7, no. 3, pp. 319–339, 2015. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/poi3.95

[21] P. Aragón, V. Gómez, and A. Kaltenbrunner, "Detecting platform effects in online discussions," *Policy & Internet*, vol. 9, no. 4, pp. 420–443, 2017. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/poi3.158

[22] P. S. Hart and E. C. Nisbet, "Boomerang effects in science communication: How motivated reasoning and identity cues amplify opinion polarization about climate mitigation policies," *Communication Research*, vol. 39, no. 6, pp. 701–723, 2012. [Online]. Available: https://doi.org/10.1177/0093650211416646

[23] M. D. Conover, J. Ratkiewicz, M. Francisco, B. Goncalves, F. Menczer, and A. Flammini, "Political polarization on twitter," in *Fifth International AAAI Conference on Weblogs and Social Media*, 2011. [Online]. Available: https://www.aaai.org/ocs/index.php/ICWSM/ICWSM11/paper/view/2847

[24] S. Sobieraj and J. M. Berry, "From incivility to outrage: Political discourse in blogs, talk radio, and cable news," *Political Communication*, vol. 28, no. 1, pp. 19–41, 2011. [Online]. Available: https://doi.org/10.1080/10584609.2010.542360

[25] N. J. Stroud, "Polarization and partisan selective exposure," *Journal of Communication*, vol. 60, no. 3, pp. 556–576, 2010. [Online]. Available: https://academic.oup.com/joc/article/60/3/556/4098564

[26] J. K. Lee, J. Choi, C. Kim, and Y. Kim, "Social media, network heterogeneity, and opinion polarization," *Journal of Communication*, vol. 64, no. 4, pp. 702–722, 2014. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/jcom.12077

[27] A. J. Morales, J. Borondo, J. C. Losada, and R. M. Benito, "Measuring political polarization: Twitter shows the two sides of venezuela," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 25, no. 3, p. 033114, 2015. [Online]. Available: https://aip.scitation.org/doi/abs/10.1063/1.4913758

[28] C.-h. Chan and K.-w. Fu, "The relationship between cyberbalkanization and opinion polarization: Time-series analysis on facebook pages and opinion polls during the hong kong occupy movement and the associated debate on political reform," *Journal of Computer-Mediated Communication*, vol. 22, no. 5, pp. 266–283, 2017. [Online]. Available: https://academic.oup.com/jcmc/article/22/5/266/4666427

[29] C. A. Bail, L. P. Argyle, T. W. Brown, J. P. Bumpus, H. Chen, M. B. F. Hunzaker, J. Lee, M. Mann, F. Merhout, and A. Volfovsky, "Exposure to opposing views on social media can increase political polarization," *Proceedings of the National Academy of Sciences*, vol. 115, no. 37, pp. 9216–9221, 2018. [Online]. Available: https://www.pnas.org/content/115/37/9216

[30] D. Demszky, N. Garg, R. Voigt, J. Zou, M. Gentzkow, J. Shapiro, and D. Jurafsky, "Analyzing polarization in social media: Method and application to tweets on 21 mass shootings," in *17th Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)*, 2019. [Online]. Available: https://nlp.stanford.edu/pubs/demszky2019analyzing.pdf

[31] L. A. Adamic and N. Glance, "The political blogosphere and the 2004 u.s. election: Divided they blog," in *Proceedings of the 3rd International Workshop on Link Discovery*, ser. LinkKDD '05. ACM, 2005, pp. 36–43, event-place: Chicago, Illinois. [Online]. Available: http://doi.acm.org/10.1145/1134271.1134277

[32] P. C. Guerra, W. M. Jr, C. Cardie, and R. Kleinberg, "A measure of polarization on social media networks based on community boundaries," in *Seventh International AAAI Conference on Weblogs and Social Media*, 2013. [Online]. Available: https://www.aaai.org/ocs/index.php/ICWSM/ICWSM13/paper/view/6104

[33] V. Price, J. N. Cappella, and L. Nir, "Does disagreement contribute to more deliberative opinion?" *Political Communication*, vol. 19, no. 1, pp. 95–112, 2002. [Online]. Available: https://doi.org/10.1080/105846002317246506

[34] J. Barabas, "How deliberation affects policy opinions," *American Political Science Review*, vol. 98, no. 4, pp. 687–701, 2004. [Online]. Available: https://www.cambridge.org/core/journals/american-political-science-review/article/how-deliberation-affects-policy-opinions/60BBEAE885EB4EF99D81913375176743

[35] S. Iyengar, R. C. Luskin, and J. S. Fishkin, "Deliberative preferences in the presidential nomination campaign: Evidence from an online deliberative poll," 2005.

[36] D. C. Mutz and J. J. Mondak, "The workplace as a context for cross-cutting political discourse," *The Journal of Politics*, vol. 68, no. 1,

pp. 140–155, 2006. [Online]. Available: https://www.journals.uchicago.edu/doi/abs/10.1111/j.1468-2508.2006.00376.x

[37] K. Grönlund, K. Strandberg and, and S. Himmelroos, "The challenge of deliberative democracy online – a comparison of face-to-face and virtual experiments in citizen deliberation," *Information Polity*, vol. 14, no. 3, pp. 187–201, 2009. [Online]. Available: https://www.medra.org/servlet/aliasResolver?alias=iospress&doi=10.3233/IP-2009-0182

[38] J. S. Fishkin and R. C. Luskin, "Experimenting with a democratic ideal: Deliberative polling and public opinion," *Acta Politica*, vol. 40, no. 3, pp. 284–298, 2005. [Online]. Available: https://doi.org/10.1057/palgrave.ap.5500121

[39] V. R. K. Garimella and I. Weber, "A long-term analysis of polarization on twitter," in *Eleventh International AAAI Conference on Web and Social Media*, 2017. [Online]. Available: https://www.aaai.org/ocs/index.php/ICWSM/ICWSM17/paper/view/15592

[40] G. Olivares, J. P. Cárdenas, J. C. Losada, and J. Borondo, "Opinion polarization during a dichotomous electoral process," *Complexity*, vol. 2019, p. 5854037, 2019. [Online]. Available: https://doi.org/10.1155/2019/5854037

[41] R. S. Arvapally and X. F. Liu, "Polarisation assessment in an intelligent argumentation system using fuzzy clustering algorithm for collaborative decision support," *Argument & Computation*, vol. 4, no. 3, pp. 181–208, 2013. [Online]. Available: https://doi.org/10.1080/19462166.2013.794163

[42] M. Klein, "How to harvest collective wisdom on complex problems : An introduction to the MIT deliberatorium," 2011.

[43] J.-M. Esteban and D. Ray, "On the measurement of polarization," *Econometrica*, vol. 62, no. 4, pp. 819–851, 1994. [Online]. Available: https://www.jstor.org/stable/2951734

[44] A. Flache and M. W. Macy, "Small worlds and cultural polarization," *The Journal of Mathematical Sociology*, vol. 35, no. 1, pp. 146–176, 2011. [Online]. Available: https://doi.org/10.1080/0022250X.2010.532261

[45] W. Kunz and H. W. J. Rittel, "Issues as elements of information systems," in *Working Paper No. 131*, vol. 131, 1970.

[46] P. Schank and M. Ranney, "Improved reasoning with convince me," 1995, pp. 276–277.

[47] S. Vesic, M. Ianchuk, and A. Rubtsov, "The synergy: A platform for argumentation-based group decision making," 2012, pp. 501–502.

[48] L. Iandoli, M. Klein, and G. Zollo, "Enabling on-line deliberation and collective decision-making through large-scale argumentation: A new approach to the design of an internet-based mass collaboration platform," *International Journal of Decision Support System Technology (IJDSST)*, 2009. [Online]. Available: https://www.igi-global.com/article.aspx?ref=international-journal-decision-support-system&titleid=1745

[49] A. Gürkan, L. Iandoli, M. Klein, and G. Zollo, "Mediating debate through on-line large-scale argumentation: Evidence from the field," *Information Sciences*, vol. 180, no. 19, pp. 3686–3702, 2010. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0020025510002616

[50] F. Boltužić and J. Šnajder, "Back up your stance: Recognizing arguments in online discussions," in *Proceedings of the First Workshop on Argumentation Mining*. Association for Computational Linguistics, 2014, pp. 49–58. [Online]. Available: http://aclweb.org/anthology/W14-2107

[51] E. Cabrio and S. Villata, "A natural language bipolar argumentation approach to support users in online debate interactions†," *Argument & Computation*, vol. 4, no. 3, pp. 209–230, 2013. [Online]. Available: https://doi.org/10.1080/19462166.2013.862303

[52] S. E. Toulmin, *The Uses of Argument*. Cambridge University Press, 2003, google-Books-ID: 8UYgegaB1S0C.

[53] T. F. Gordon and N. Karacapilidis, "The zeno argumentation framework," in *Proceedings of the 6th international conference on Artificial intelligence and law*, ser. ICAIL '97. Association for Computing Machinery, 1997, pp. 10–18. [Online]. Available: https://doi.org/10.1145/261618.261622

[54] P. DiMaggio, J. Evans, and B. Bryson, "Have american's social attitudes become more polarized?" *American Journal of Sociology*, vol. 102, no. 3, pp. 690–755, 1996. [Online]. Available: https://www.journals.uchicago.edu/doi/abs/10.1086/230995

[55] D. Baldassarri and A. Gelman, "Partisans without constraint: Political polarization and trends in american public opinion," *American Journal of Sociology*, vol. 114, no. 2, pp. 408–446, 2008. [Online]. Available: https://www.journals.uchicago.edu/doi/abs/10.1086/590649

[56] S. Iyengar, G. Sood, and Y. Lelkes, "Affect, not Ideology a social identity perspective on polarization," *Public Opinion Quarterly*, vol. 76, no. 3, pp. 405–431, 2012. [Online]. Available: https://academic.oup.com/poq/article/76/3/405/1894274

[57] L. Mason, ""i disrespectfully agree": The differential effects of partisan sorting on social and issue polarization," *American Journal of Political Science*, vol. 59, no. 1, pp. 128–145, 2015. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/ajps.12089

[58] Y. Lelkes, "Mass polarization: Manifestations and measurements," *Public Opinion Quarterly*, vol. 80, pp. 392–410, 2016. [Online]. Available: https://academic.oup.com/poq/article/80/S1/392/2223374

[59] E. Suhay, E. Bello-Pardo, and B. Maurer, "The polarizing effects of online partisan criticism: Evidence from two experiments," *The International Journal of Press/Politics*, vol. 23, no. 1, pp. 95–115, 2018. [Online]. Available: https://doi.org/10.1177/1940161217740697

[60] A. Matakos, E. Terzi, and P. Tsaparas, "Measuring and moderating opinion polarization in social networks," *Data Mining and Knowledge Discovery*, vol. 31, no. 5, pp. 1480–1505, 2017. [Online]. Available: https://doi.org/10.1007/s10618-017-0527-9

[61] F. Schweitzer, T. Krivachy, and D. Garcia, "How emotions drive opinion polarization: An agent-based model," *arXiv:1908.11623 [nlin, physics:physics, q-bio]*, 2019. [Online]. Available: http://arxiv.org/abs/1908.11623

[62] S. Banisch and E. Olbrich, "Opinion polarization by learning from social feedback," *The Journal of Mathematical Sociology*, vol. 43, no. 2, pp. 76–103, 2019. [Online]. Available: https://doi.org/10.1080/0022250X.2018.1517761

[63] A. Gruzd and J. Roy, "Investigating political polarization on twitter: A canadian perspective," *Policy & Internet*, vol. 6, no. 1, pp. 28–45, 2014. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/1944-2866.POI354

[64] J. Schmitt, "How to measure ideological polarization in party systems," p. 31, 2016.

[65] M. Del Vicario, A. Scala, G. Caldarelli, H. E. Stanley, and W. Quattrociocchi, "Modeling confirmation bias and polarization," *Scientific Reports*, vol. 7, p. 40391, 2017. [Online]. Available: https://www.nature.com/articles/srep40391

[66] A. Bramson, P. Grim, D. J. Singer, S. Fisher, W. Berger, G. Sack, and C. Flocken, "Disambiguation of social polarization concepts and measures," *The Journal of Mathematical Sociology*, vol. 40, no. 2, pp. 80–111, 2016. [Online]. Available: http://www.tandfonline.com/doi/full/10.1080/0022250X.2016.1147443

[67] A. Bramson, P. Grim, D. J. Singer, W. J. Berger, G. Sack, S. Fisher, C. Flocken, and B. Holman, "Understanding polarization: Meanings, measures, and model evaluation," *Philosophy of Science*, vol. 84, no. 1, pp. 115–159, 2017, publisher: The University of Chicago Press. [Online]. Available: https://www.journals.uchicago.edu/doi/abs/10.1086/688938

[68] B. Apouey, "Measuring health polarization with self-assessed health data," *Health Economics*, vol. 16, no. 9, pp. 875–894, 2007. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/hec.1284

[69] T. S. Clark, "Measuring ideological polarization on the united states supreme court," *Political Research Quarterly*, vol. 62, no. 1, pp. 146–157, 2009. [Online]. Available: https://doi.org/10.1177/1065912908314652

[70] Z. Maoz and Z. Somer-Topcu, "Political polarization and cabinet stability in multiparty systems: A social networks analysis of european parliaments, 1945–98," *British Journal of Political Science*, vol. 40, no. 4, pp. 805–833, 2010.

[71] P. Rehm and T. Reilly, "United we stand: Constituency homogeneity and comparative party polarization," *Electoral Studies*, vol. 29, no. 1, pp. 40–53, 2010. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0261379409000419

[72] S. Gabbriellini and P. Torroni, "A new framework for ABMs based on argumentative reasoning," in *Advances in Social Simulation*, ser. Advances in Intelligent Systems and Computing, B. Kamiński and G. Koloch, Eds. Springer Berlin Heidelberg, 2014, pp. 25–36.

[73] R. M. Alvarez and J. Brehm, *Hard Choices, Easy Answers: Values, Information, and American Public Opinion*. Princeton University Press, 2002, google-Books-ID: dWh10zFdjNUC.

[74] T. J. Rudolph and E. Popp, "An information processing theory of ambivalence," *Political Psychology*, vol. 28, no. 5, pp. 563–585, 2007. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-9221.2007.00590.x

[75] L. Keele and J. Wolak, "Contextual sources of ambivalence," *Political Psychology*, vol. 29, no. 5, pp. 653–673, 2008. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-9221.2008.00659.x

[76] S. Yardi and D. Boyd, "Dynamic debates: An analysis of group polarization over time on twitter," *Bulletin of Science, Technology & Society*, vol. 30, no. 5, pp. 316–327, 2010. [Online]. Available: http://journals.sagepub.com/doi/10.1177/0270467610380011

[77] E. J. Olsson, “A bayesian simulation model of group deliberation and polarization,” in *Bayesian Argumentation*, F. Zenker, Ed. Springer Netherlands, 2013, pp. 113–133. [Online]. Available: http://link.springer.com/10.1007/978-94-007-5357-0_6

[78] M. D. Vicario, G. Vivaldo, A. Bessi, F. Zollo, A. Scala, G. Caldarelli, and W. Quattrociocchi, “Echo chambers: Emotional contagion and group polarization on facebook,” *Scientific Reports*, vol. 6, no. 1, pp. 1–12, 2016. [Online]. Available: https://www.nature.com/articles/srep37825

[79] C. Proietti, “Understanding group polarization with bipolar argumentation frameworks,” in *Frontiers in Artificial Intelligence and Applications*, ser. Frontiers in Artificial Intelligence and Applications, vol. 287. IOS Press, 2016, pp. 41 – 52.

[80] J. H. Evans, “Have americans’ attitudes become more polarized?—an update*,” *Social Science Quarterly*, vol. 84, no. 1, pp. 71–90, 2003. [Online]. Available: https://onlinelibrary.wiley.com/doi/full/10.1111/1540-6237.8401005

[81] D. Baldassarri and P. Bearman, “Dynamics of political polarization,” *American Sociological Review*, vol. 72, no. 5, pp. 784–811, 2007. [Online]. Available: https://doi.org/10.1177/000312240707200507

[82] M. McCombs, “The agenda-setting role of the mass media,” in *Proceedings of the 2002 Conference of Mass Media Economics*. London School of Economics, 2002.

[83] A. K. McCallum, “MALLET: A machine learning for language toolkit.” 2002. [Online]. Available: http://mallet.cs.umass.edu

[84] D. G. Myers and H. Lamm, “The group polarization phenomenon,” *Psychological Bulletin*, vol. 83, no. 4, pp. 602–627, 1976.

[85] J. Sirrianni, X. F. Liu, and D. Adams, “Quantitative modeling of polarization in online intelligent argumentation and deliberation for capturing collective intelligence,” in *2018 IEEE International Conference on Cognitive Computing (ICCC)*, 2018, pp. 57–64. [Online]. Available: doi.ieeecomputersociety.org/10.1109/ICCC.2018.00015

Joseph W Sirrianni received his B.S. in computer science from the University of Arkansas, Fayetteville, AR, USA in 2016. He is currently pursuing a Ph.D. degree from the University of Arkansas, Fayetteville, AR, USA.

His research focus is development of an intelligent cyber argumentation platform and modeling and analyzing various phenomena in cyber argumentation.

Dr. Xiaoging ”Frank” Liu is currently Dean of the College of Engineering and professor in both the School of Computing and School of Electrical, Computer, and Biomedical Engineering at the Southern Illinois University in Carbondale, Illinois. His current interests include software engineering, service computing, data analytics-based recommendation systems, cyber argumentation based social media and networking, cyber manufacturing, and applied artificial intelligence.

He served as a PI, Co-PI, or faculty participant of 29 funded research projects and published more than 150 referred papers in numerous journals and conferences, such as TWeb, TSC, JVR, SPIP, SQJ, ICSE, JSS, AAAI, ICWS, and CSCW. He received his PhD in computer science from the Texas A&M University in College Station in 1995.

Dr. Douglas Adams received his BA in Sociology in 1982 from Augsburg University (College) in Minneapolis, Minnesota. He earned his MA in Sociology in 1991, and his Ph.D. in Sociology in 1996 from the University of Arizona, Tucson.

He is currently an Associate Professor in the Department of Sociology and Criminology at the University of Arkansas. Since 2015, Dr. Adams has worked on developing a simple, cross-disciplinary, problem-based instructional process that is adaptive, and scalable. Since 2019, he has developed modifications to this process that blend both live, and on-line instructional discourse. Since 2017, Dr. Adams’ research has focused on the topic of Online Discourse; assisting in the development of a web-based discourse platform that facilitates social discourse (argumentation) on ”contested” (hot button) social issues.