Towards a Large-Scale Intelligent Mobile-Argumentation and Discovering Arguments, Controversial Topics and Topic-Oriented Focal Sets in Cyber-Argumentation

Najla Althuniyan
University of Arkansas, Fayetteville

Follow this and additional works at: https://scholarworks.uark.edu/etd

Part of the Graphics and Human Computer Interfaces Commons, Mass Communication Commons, Programming Languages and Compilers Commons, Social Media Commons, and the Software Engineering Commons

Citation

This Dissertation is brought to you for free and open access by ScholarWorks@UARK. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of ScholarWorks@UARK. For more information, please contact scholar@uark.edu.
Towards a Large-Scale Intelligent Mobile-Argumentation and Discovering Arguments, Controversial Topics and Topic-Oriented Focal Sets in Cyber-Argumentation

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Computer Sciences

by

Najla Althuniyan
King Saud University
Bachelor of Science in Computer Science, 2006
California State University, Fullerton
Master of Science in Computer Science, 2013

July 2021
University of Arkansas

This dissertation is approved for recommendation to the Graduate Council.

______________________________  ________________________________
Xiaoqing “Frank” Liu, Ph.D.     Brajendra Nath Panda, Ph.D.
Dissertation Director          Dissertation Co-Director

______________________________  ________________________________
Justin Zhan, Ph.D.             Douglas Adams, Ph.D.
Committee Member              Committee Member

______________________________
Qinghua Li, Ph.D.
Committee Member
Abstract

User-generated content (UGC) platforms host different forms of information, such as audio, video, pictures, and text. They have many online applications, such as social media, blogs, photo and video sharing, customer reviews, debate, and deliberation platforms. Usually, the content of these platforms is provided and consumed by users. Most of these platforms, mainly social media and blogs, are often used for online discussion. These platforms offer tools for users to share and express opinions. Commonly, people from different backgrounds and origins discuss opinions about various issues over the Internet. Furthermore, discussions among users contain substantial information from which knowledge about collective intelligence can be extracted. Collective Intelligence is wisdom and knowledge that grows when a group works together collectively or cooperatively. In this dissertation, strides in the cyber-argumentation field are made. The body of this work in this dissertation revolves around different areas: (1) bringing the intelligent cyber-argumentation into the handheld device space and showing the effectiveness of bringing large-scale cyber-argumentation into handheld devices, (2) constructing the argument discovery framework and identifying the arguments attributes, (3) modeling the controversial degree of cyber-argumentation discussions using well-known measures and, (4) discovering topic-oriented focal sets in cyber-argumentation using link analysis, topic modeling and social roles. This dissertation is concluded by discussing the challenging technical implications of this emerging research area and proposing future work avenues.
Acknowledgment

I would like first to thank my dissertation committee and especially Dr. Xiaoqing Liu. Each committee member has given me helpful feedback and has been a great professor who has prepared me to get to this place in my academic life. This dissertation would not be possible without their help.

Funding-wise, this work was fully sponsored by the government of Saudi Arabia, represented by the Saudi Arabian Cultural Mission in the US, SACM. I would like to express my sincere gratitude to SACM. Any opinions, findings, and conclusions, or recommendations expressed in this material are those of the author and do not necessarily reflect the sponsor’s opinion.

I want to acknowledge my colleagues from the software engineering laboratory: Joseph Sirrianni, Md Mahfuzer Rahman, S M Nahian Al Sunny, and Md Rakib Shahriar, for their excellent collaboration. I appreciate pre-meeting and post-meeting discussions. It was a pleasure working with you.

From the bottom of my heart, I would like to say big thanks to my husband; Ahmed has been incredibly supportive of me throughout this entire process and has made endless sacrifices to help me get my degree. My children, Sima, Dima, and Rima. I was going to school just like you, had homework just like you, and had to do my homework just like you.

I would like to express my gratitude and appreciation for my extended family, friends, and neighbors. They deserve special thanks for their continued support and encouragement. I doubt that I would be in this place today without such a team behind me.

Finally, I would like to thank the University of Arkansas for providing such an excellent environment to earn my Ph.D. degree.
Dedication

To my late father: we will celebrate together in heaven!
# Table of Contents

Chapter 1.  Introduction ........................................................................................................... 1

1.1  Reference .......................................................................................................................... 5

Chapter 2.  Background of Cyber-Argumentation and Data Collection .......................... 7

2.1  Background of Argumentation Systems .......................................................................... 7

2.1.1  ICAS .............................................................................................................................. 8

2.2  Empirical Data Collection ............................................................................................... 12

2.2.1  Issue1: Same-sex couples and Adoption ................................................................... 12

2.2.2  Issue4: Government and Healthcare ......................................................................... 13

2.3  References ....................................................................................................................... 14

Chapter 3.  Design of Mobile Service of Intelligent Large-Scale Cyber Argumentation for Analysis and Prediction of Collective Opinions ......................................................... 16

3.1  Introduction ..................................................................................................................... 16

3.2  Related Work ................................................................................................................... 20

3.3  The Mobile Application of Cyber-Argumentation ......................................................... 22

3.3.1  The Mobile Application Architecture of Cyber-Argumentation ............................... 22

3.3.2  The Back-end ............................................................................................................... 28

3.3.3  Front-end and Back-end Communications ................................................................. 28

3.4  Discussion and Analysis .................................................................................................. 29

3.4.1  Example: Issue: Guns on Campus ............................................................................ 29

3.5  System Usability Study .................................................................................................... 31

3.5.1  Methodology ............................................................................................................... 32

3.5.2  RESULTS ..................................................................................................................... 32

3.6  Conclusion and Future Work .......................................................................................... 33
Chapter 5.  Modeling and Analyzing Controversy in Large-Scale Cyber-Argumentation...... 63

5.1  Introduction ............................................................................................................................. 63

5.2  Background ............................................................................................................................... 64

5.3  Controversy Measures ................................................................................................................ 65

5.3.1  Controversy Definition ............................................................................................................ 65

5.3.2  Controversy Attributes .......................................................................................................... 66

5.3.3  Data Collection ....................................................................................................................... 67

5.3.4  Estimating Controversy with Primitive measures ..................................................................... 67

5.3.5  Estimating Controversy Using Distance-Based Measure ...................................................... 68

5.3.6  Estimating Controversy Using Distribution-Based Measure ................................................. 69

5.3.7  Estimating Controversy Using Agreement Measurements ................................................... 71

5.4  Discussion and Results .............................................................................................................. 72

5.5  Conclusion And Future Work .................................................................................................... 83

5.6  References ..................................................................................................................................... 84

Chapter 6.  Discovering topic-oriented focal sets in cyber-argumentation using link analysis, topic modeling and social roles.................................................................................................................. 87

6.1  Introduction .................................................................................................................................. 87

6.2  Related Work ............................................................................................................................... 88

6.2.1  Community Detection ............................................................................................................. 88

6.2.2  Topic Modeling ......................................................................................................................... 91

6.2.3  Social Roles in UGC Platforms .............................................................................................. 91

6.3  Proposed Framework .................................................................................................................. 93

6.3.1  Data Collection ....................................................................................................................... 93
6.3.2 User Opinion Vector (UOV) ................................. 94
6.3.3 Focal Sets and Participants Partitioning Using Link Analysis ............................. 94
6.3.4 Focal Sets Opinion Vector .................................................. 95
6.3.5 Issue Topics by Topic Molding ........................................ 96
6.3.6 Individuals and Focal Sets and Topics Similarities ..................................... 98
6.3.7 Social Roles and Topic Intensity Degree in ICAS ...................................... 99

6.4 Discussion and Results .......................................................... 100
6.4.1 The FSA Algorithm with ICAS Discussions ........................................... 100
6.4.2 Topic Modeling with ICAS Discussion .................................................. 101
6.4.3 Individuals and Focal Sets and Topics Similarities ...................................... 101
6.4.4 Individual Opinion Intensity and Social Roles ......................................... 102
6.4.5 Focal Sets Opinion Intensity and Social Roles ......................................... 105
6.4.6 Illustration of Issue 1 ...................................................................... 106

6.5 Conclusion ............................................................................. 110

6.6 References ............................................................................. 111

Chapter 7. Conclusion and Future work .............................................. 115
# Table of Figures

- **FIGURE 1-1** DISSERTATION FRAMEWORK .......................................................... 2
- **FIGURE 2-1** ISSUE, POSITION AND ARGUMENT TREE – FROM [3] .................. 8
- **FIGURE 2-2** POSITION SUBTREE BEFORE AND AFTER REDUCTION .............. 10
- **FIGURE 3-1** APP ARCHITECTURE AND COMPONENTS ................................... 23
- **FIGURE 3-2** MAIN PAGE ................................................................................ 26
- **FIGURE 3-3** SELECTED ISSUE ....................................................................... 26
- **FIGURE 3-4** SELECTED POSITION ................................................................. 26
- **FIGURE 3-5** COLLECTIVE ANALYTICS .......................................................... 26
- **FIGURE 3-6** PERSONAL ANALYTICS .............................................................. 26
- **FIGURE 3-7** THE SLIDING BAR .................................................................... 26
- **FIGURE 3-8** FRONT-END ANF BACK-END COMMUNICATIONS .................. 29
- **FIGURE 4-1** USER/POSITIONS/ARGUMENT TREE ........................................ 42
- **FIGURE 4-2** USER INTERACTION GRAPH ...................................................... 50
- **FIGURE 4-3** PEARSON CORRELATIONS BETWEEN INDICATORS ............ 51
- **FIGURE 4-4** ARGUMENT DISCOVERY SCORING MODEL ............................ 53
- **FIGURE 4-5** THE ARGUMENT DISCOVERY FRAMEWORK .......................... 55
- **FIGURE 5-1** USER DISTRIBUTION FOR P0, P1, P2 AND P3 UNDER ISSUE 1 .... 73
- **FIGURE 5-2** USER DISTRIBUTION FOR P4, P5, P6 AND P7 UNDER ISSUE 2 .... 73
- **FIGURE 5-3** USER DISTRIBUTION FOR P8, P9, P10 AND P11 UNDER ISSUE 3 ..... 78
- **FIGURE 5-4** USER DISTRIBUTION FOR P11, P12, P14 AND P15 UNDER ISSUE 4 ... 80
- **FIGURE 6-1** DISCOVERING TOPIC-ORIENTED INDIVIDUALS & FOCAL SETS ...... 93
- **FIGURE 6-2** THE OPINION INTENSITY BAR ................................................. 102
- **FIGURE 6-3** USER 1 EGOCENTRIC GRAPH .................................................. 109
- **FIGURE 6-4** USER 11 EGOCENTRIC NETWORK ........................................... 109
<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>TABLE 2-1</td>
<td>ISSUES, POSITIONS, AND USERS DETAILS</td>
<td>13</td>
</tr>
<tr>
<td>TABLE 3-1</td>
<td>POSITIONS STATISTICS TO THE GUN ON CAMPUS ISSUE</td>
<td>30</td>
</tr>
<tr>
<td>TABLE 3-2</td>
<td>SUS QUESTIONS AND RESPONSES</td>
<td>34</td>
</tr>
<tr>
<td>TABLE 4-1</td>
<td>DISCOVERING ARGUMENTS WITH COOPERATIVE INDICATORS</td>
<td>56</td>
</tr>
<tr>
<td>TABLE 4-2</td>
<td>DISCOVERING ARGUMENTS WITH MUTUALLY EXCLUSIVE INDICATORS</td>
<td>57</td>
</tr>
<tr>
<td>TABLE 4-3</td>
<td>DISCOVERING ARGUMENTS WITH UN-RECOMMENDED OPERATOR BETWEEN INDICATORS</td>
<td>58</td>
</tr>
<tr>
<td>TABLE 4-4</td>
<td>DISCOVERING ARGUMENTS BY THE ARGUMENT DISCOVERY MODEL</td>
<td>59</td>
</tr>
<tr>
<td>TABLE 5-1</td>
<td>ISSUES, POSITIONS, AND USERS DETAILS</td>
<td>74</td>
</tr>
<tr>
<td>TABLE 5-2</td>
<td>RESULTS FROM ALL CONTROVERSY MEASURES</td>
<td>75</td>
</tr>
<tr>
<td>TABLE 5-3</td>
<td>CORRELATIONS BETWEEN DISCUSSION ATTRIBUTES AND CONTROVERSY MEASURES</td>
<td>81</td>
</tr>
<tr>
<td>TABLE 6-1</td>
<td>RESULTS FROM APPLYING THE FSA ALGORITHM</td>
<td>100</td>
</tr>
<tr>
<td>TABLE 6-2</td>
<td>TOPIC MODELING EVALUATION</td>
<td>101</td>
</tr>
<tr>
<td>TABLE 6-3</td>
<td>STATISTICS OF INDIVIDUAL SOCIAL ROLES AND OPINION INTENSITY IN ICAS DISCUSSIONS</td>
<td>103</td>
</tr>
<tr>
<td>TABLE 6-4</td>
<td>STATISTICS OF FOCAL SET OPINION INTENSITY IN ICAS DISCUSSIONS</td>
<td>106</td>
</tr>
<tr>
<td>TABLE 6-5</td>
<td>TOPICS BY NMF AND LDA</td>
<td>108</td>
</tr>
</tbody>
</table>
List of Published Papers

Chapter 2:


Chapter 3:


Chapter 4:


Chapter 5:

N Althuniyan, “Modeling and Analyzing Controversy in Large-Scale Cyber-Argumentation” submitted to the International Conference on Transdisciplinary AI for review and publication.

Chapter 6:

Chapter 1. INTRODUCTION

UGC platforms, such as social media, blogs, photo sharing, and websites allow substantial discussion and user participation. Usually, UGC content is created by participants, and it does not get filtered as content in traditional media outlets. Thus, UGC content creators are rewarded by receiving recognition from content consumers, who use UGC platforms for information or entertainment. As a result, UGC applications create an attractive user environment and adapt AI models to help users be more active, be more creative, and develop new personal and business opportunities. Also, UGC platforms are used for reasons other than entertainment, such as politics, education, marketing, and malicious activities. Cyber-argumentation is an example of UGC platforms.

This dissertation work is based on cyber-argumentation. Figure 1-1 shows the framework of this dissertation. This research has used the intelligent cyber-argumentation system, ICAS, detailed in Chapter 2. ICAS hosts structured discussions. Each discussion is represented as a weighted argumentation graph. This system uses the fuzzy logic engine [1, 2] to drive the user’s opinion vector from the argumentation graph. This dissertation is a progress work for the field of cyber-argumentation by introducing additional models and work to facilitate and analyze online deliberation effectively, as in Figure 1-1. The green rectangles in Figure 1-1 are the work discussed in this dissertation.

Cyber-argumentation and debate platforms are designed to enable online deliberations with large-scale, in-depth argumentation for productive discussions over the internet. These platforms host structured argumentation that allow complex analytical models to mine the argumentation for collective intelligence. However, not all of those argumentation platforms have mobile-platform versions, and if they exist, they have fundamental capabilities, such as basic statistics. In this
research, the design of a mobile application for cyber-argumentation is presented, such as reporting basic statistics and complex analytics. This mobile application supports intelligent cyber-argumentation and large-scale discussions and provides meaningful analytics on mobile devices. The platform has incorporated several analytical models to capture collective opinions, detect opinion polarization, and predict missing user opinions with the limitation of mobile device’s screen size. An example is used to illustrate our design and models in the mobile space, and a system usability study of our application is presented.

Each discussion in cyber-argumentation platforms is represented by a tree-structure. The discussion massively grows as users join the discussion. However, a study from our lab showed that users only view, on average, 3% of the discussion content. Therefore, not all opinions are constructive or worth more discussion. Moreover, the screen size of a mobile application makes it hard for users to keep track of the discussion or identify which opinions are constructive. Thus, there is a need for an opinion discovery method in cyber-argumentation. In social media and blog platforms, opinions are discovered by engagement information, impact score, or reverse-chronological order. On the other hand, some UGC platforms discover constructive opinions based on textual features. However, there are no pre-identified feature sets for the opinions that have been set for searching and discovering opinions in academia or public-debate platforms.

FIGURE 1-1 DISSERTATION FRAMEWORK
Identifying features for locating constructive opinions helps improve the discourse quality and provides an attractive online discussion platform. In this research, a new framework for opinion selection and discovery is proposed. It locates constructive opinions based on four unique features: engagement, recentness, controversy, and author influence. Therefore, it provides an attractive and dynamic discourse incorporating opinions features based on users’ preferences. First, these features are defined in the cyber-argumentation space. Consequently, a new framework that combines those features for opinion search and discovery is developed. Finally, an application example on the ICAS dataset demonstrates the proposed frameworks effectiveness in discovering and searching for constructive opinions.

This framework guided the dissertation direction into two additional tasks: the first task is to analyze and model the controversial degree of cyber-argumentation discussions. The second task is to discover topic-oriented focal sets in cyber-argumentation using link analysis, topic modeling, and social roles.

For the first task, certain phenomena, such as controversy, often appear in online argumentation, making the discussion between participants heated. Heated discussions can be used to extract new knowledge. Therefore, detecting the presence of controversy is an essential task to determine if collective intelligence can be extracted from online discussions. This research uses existing measures for estimating controversy quantitatively in cyber-argumentation. First, it defines controversy in different fields and then identifies the attributes of controversy in online discussions. Then, the distributions of user opinions and the distance between opinions are used to calculate the controversial degree. Finally, the results from each controversy measure are discussed and analyzed using an empirical study generated by ICAS. This work is an improvement over the
existing measurements because it does not require ground-truth data or specific settings and can be adapted to distribution-based or distance-based opinions.

For the second task, constructive opinions are often provided by users with specific profiles and have implicit topics. Users may interact with each other cooperatively or collectively to support, attack, or deliver agendas due to similar interests. Much research has been conducted to detect groups who shared similar interests using link analysis techniques, text analysis techniques, and other data mining techniques, or a combination of different methods to discover or detect communities or hidden communities with specific characteristics in social media and blogs. However, most of the work that has been done has focused on improving the community, or hidden community, detection algorithms concerning the research challenges. Cyber-argumentation platforms are full of valuable hidden features worth further study and research due to the extensive discussion they contain. This research proposes a new framework to discover groups, or focal sets, with similar interests in topics using focal structure analysis algorithm [3] and topic modeling techniques. Then, the social roles of the focal set members are investigated. By combining these techniques, the groups and individuals behind specific topics in the discussion are discovered and studied. Moreover, the topic intensity degree among participants and groups is measured.

Our work can be useful for UGC platforms to identify groups with similar interests using their discussion and interaction information. Mainly, this framework helps discover those groups, identifying group members' social roles instead of finding influential individuals and measuring topic intensity degrees. Since UGC platforms are not only for entertainment, the proposed framework can leverage new knowledge or malicious activities from the collective discussions. It estimates individual interests and power in the identified groups. Besides that, it should help develop new measures and policies with the associated risks or opportunities in cyberspace.
In this chapter, the introduction of the dissertation is presented. Chapter 2 provides the background information of cyber-argumentation systems used as a foundation of this dissertation. It also describes the dataset used in this dissertation. Some of the work in this chapter was published in a conference paper in IEEE-Second International Conference on Transdisciplinary AI (TransAI). Chapter 3 presents my contribution to this dissertation by designing a mobile app of intelligent large-scale cyber argumentation to analyze and predict collective opinions. This chapter is published in Artificial Intelligence and Mobile Services – AIMS 2019 and IEEE-Second International Conference on Transdisciplinary AI (TransAI). Chapter 4 presents my work in developing the Opinion Discovery Framework: Toward a Quality Opinion-Centric Platform, a conference paper presented in the International Conference on Computing and Data Engineering (ICCDE2020) and published in the Journal of Advances in Information Technology, 2020. Chapter 5 presents my work on developing a model for analyzing the argumentation controversy in large-scale cyber-argumentation. Chapter 6 presents a framework for discovering topic-oriented focal sets in cyber-argumentation platforms using link analysis, topic modeling, and social roles. The work in chapter 5 and 6 are to be submitted for publication. Chapter 7 concludes this dissertation and directs it for future work.

1.1 REFERENCES


Chapter 2. BACKGROUND OF CYBER-ARGUMENTATION AND DATA COLLECTION

2.1 BACKGROUND OF ARGUMENTATION SYSTEMS

Cyber-argumentation platforms are outlets that deal with contentious debate and deliberation from which these systems draw conclusions. Most of these platforms are built on formal and informal argumentation frameworks. The formal argumentation framework is an abstract argumentation created by Dung [1], known as the Dung Abstract Framework. It defines the argumentation systems as a set of arguments and a set of defeasibility relations. The system is viewed as an oriented graph, whose nodes are the different arguments, and the edges represent the defeasibility relationship between them. The formal argumentation platform was the fundamental framework for many informal argumentation frameworks, such as logic-based argumentation frameworks, value-based argumentation frameworks, and assumption-based argumentation (ABA) frameworks. However, the Issue-Based Information System by Kunz and Rittel [2], known as IBIS, was invented earlier than Dung’s argumentation framework. It has been used widely as an informal argumentation framework to support coordination and solving problems that involve multiple stakeholders. Many argumentation tools have been built using IBIS.

For over a decade, Prof. Liu and his research group have investigated and developed large-scale cyber-argumentation platforms [3-16]. As a starting point, Liu et al. [3] implemented an intelligent system for collaborative engineering design and conflict resolution. This system has been updated and developed over time to accommodate different research goals. The main objective of these tools developed and led by Prof. Liu is to capture decision reasoning by structuring online discussions and publishing many studies on cyber-argumentation for practical analysis [6-9]. For example, the studies from [10-13] are to support decision-making, [14] is to analyze large-scale discussions and collective opinion effectively, [12] is to analyze the credibility of arguments, [10]
is to detect conflicting opinions, [13] is to identify outlier opinions, [14] is to identify opinion factions, [15] is to measure opinion polarity in online-discussion, and [16] is to predict the collective and individual opinions. The current version is the Intelligent Cyber Argumentation System (ICAS). ICAS is an online web-based argumentation tool developed by Prof. Liu's research group and is used as a foundation for this dissertation.

2.1.1 ICAS

![Argumentation Tree Diagram]

**FIGURE 2-1 ISSUE, POSITION AND ARGUMENT TREE – FROM [3]**

One argumentation system that uses the IBIS structure is the ICAS. Each issue and related nodes are modeled into a single tree, as shown in Figure 2-1. This section explains the major components of ICAS and the collective analytical models adapted to help users understand and analyze the lengthy discussions.

**Issues** are the unsolved problems or questions open for discussion. They usually start a new structured discussion. Therefore, the issue is placed at the root of the tree. All related nodes connected to the issue-tree are part of the discussion.

**Positions** are the stances or possible solutions for an issue. They represent the spectrum of perspectives on the issue. Each position responds to one and only one issue. Positions are usually neutral and not for supporting or attacking the issues; they serve as suggestions or assertions responding to an issue. A position’s discussion is made up of all arguments under it.
**Arguments and Reactions** are users’ opinions for or against positions or other arguments, but not for or against issues. An argument is made of text and an agreement value. The text is the place for rationale, opinion, facts, data, rhetoric, etc. The argument text describes the user’s argument in words and can be up to 2500 characters. The agreement value is a user-defined value that indicates the user’s agreement or disagreement on the parent node. Arguments and reactions can be made to positions (hence, called first-level arguments) or to arguments, (hence, called counter-arguments). The agreement value is a real number between -1 and 1. The sign indicates that the user agrees or disagrees with the parent argument, and the value represents the intensity of their agreement/disagreement. There are 11 levels, which indicate different levels of agreement. These levels are in one interval [-1, 1] and are separated by 0.2 interval length. The sections are: Completely Disagree (-1.0), Strongly Disagree (-0.8), Moderately Disagree (-0.6), Weakly Disagree (-0.4), Slightly Disagree (-0.2), Indifferent (0), Slightly Agree (+0.2), Weakly Agree (+0.4), Moderately Agree (+0.4), Strongly Agree (+0.8), and Completely Agree (+1). For example, a user may post (-0.4) as an agreement value. Since the sign is negative, it is a disagreement. The 0.4 indicates that the user is moderately disagreeing. If the agreement value is 0, it means the user is neutral (indifferent) about the argument.

Users can contribute to the system without writing any additional argument by using a reaction functionality. Reactions are statements of agreement/disagreement without any text and used to support or attack arguments or positions explicitly. This helps reduce redundant arguments in the system but still allows users to stay engaged in the discussion.

**The Collective Analytical Models**

To understand the discussion’s direction and participants’ attitudes, users need to read every argument in the issue tree. This is time-consuming and most likely, users cannot comprehend the
whole discussion while it grows massively over time. Therefore, the ICAS has incorporated three analytics models to make the app intelligent and informative for users. The models are the collective intelligent index introduced by [3], the polarization index developed by [15], and the prediction of collective opinions on the position level presented by [16]. These models use analytical techniques to help users understand the discussion, even without participating. In the next subsections, there are more details about these models in ICAS.

The Collective Intelligence Index

Each reaction or argument to a node has an agreement value. This agreement value ranges from -1 and 1. On a position level, the position or argument’s overall agreement value is calculated as:

$$Position_{overall\_agreement} = \frac{\sum_{i=1}^{N} AN_i}{N}$$

\(AN_i\) is a reduced argument or a reaction to a position.

For analytics, determining the user’s agreement towards a position is needed. Some users make direct reactions to positions, while others make reactions to arguments. Reactions that are made to positions directly have an agreement value that is easy to calculate. However, it is not clear how the argument is related to the parent position for the arguments further down the tree. The argument’s agreement level with its parent position is needed to be driven. For example, position P, as shown in Figure 2-2 (left), is not known exactly the argument B’s agreement value to position P.
P. To solve this problem, the application uses a fuzzy logic engine to reduce arguments agreement values to its parent position. In the argument reduction procedure, the fuzzy inference engine takes in two inputs and produces one output. The inputs are the agreement values of an argument to be and its parent. The output will be the newly reduced agreement value for the argument to be from the fuzzy inference engine. The fuzzy inference engine uses fuzzy membership functions to quantify the linguistics agreement terms used in the application, as listed in the previous section. The fuzzy membership function used in the application is the piecewise linear trapezoidal function. Membership functions are defined by using four vertices of the trapezoids to define the fuzzy sets.

**The Polarization Index**

The polarization index, developed by Sirrianni et al. [15], measures the amount of polarization between the individual users in a given position. This index is a distribution-based index that returns a value between 0 and 1, where 0 indicates no polarization and 1 indicates complete polarization. Polarization takes different kinds of formats in online deliberations. If polarization is detected between participants, then consensus cannot be reached. This causes the quality of the discourse to plummet. This index helps users know to which degree the discussion in a position is polarized. Please refer to Sirrianni et al. [15] for a full reference on the polarization index.

**Prediction of Collective Opinions**

In large-scale cyber argumentation and deliberation, discussions are often incomplete. Not all users explicitly share their full positions. Therefore, it is difficult to determine the collective opinions in large-scale argumentation. If the agreement value for non-participating users can be predicted, it will help us to understand the overall user attitude towards the issue or position. The model uses the cosine similarity with the position correlation collaborative filtering (CSCCF) model from Rahman et al. [16] to predict the collective agreement value for all non-participating users in a
particular position. To predict position p, for users who did not participate in a discussion of this position, the model calculates their prediction value using the CSCCF model. It averages all the prediction values as a collective prediction on this position.

2.2 Empirical Data Collection

The dataset used in this research is from a study produced by students from an introductory level sociology class who participated in an online discussion using ICAS. Mr. Joseph Sirrianni, a Ph.D. candidate in the CSCE department of the University of Arkansas, led this study in collaboration with Dr. Douglas Adams, an associate professor from the Sociology and Criminology Department. This study was conducted in the spring of 2018 and lasted for twenty-six days. There are four issues discussed in this study. Each issue has four positions. Therefore, ICAS participants have discussed sixteen different positions heavily between them. The issues and the positions were predetermined before the discussion. Table 2-1 gives us more details about users’ participation, arguments, and reactions for each position in ICAS.

2.2.1 Issue1: Same-sex Couples and Adoption

This issue posed the following question: “Should same-sex married couples be allowed to adopt children?” and contained the following four positions for discussion:

- (P0) No, same-sex couples should not be allowed to legally adopt children in any circumstances.
- (P1) Yes, but adoption should be limited to only blood relatives of the couple, such as nieces/nephews.
- (P2) Yes, but same-sex couples should have special vetting to ensure that they can provide as much as a heterosexual couple.
• (P3) Yes, same-sex couples should be treated the same as heterosexual couples and be allowed to adopt via the standard process.

<table>
<thead>
<tr>
<th>Issues</th>
<th>Position #</th>
<th># Arguments</th>
<th># Reactions</th>
<th># Distinct Users</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Issue 1</strong></td>
<td>P0</td>
<td>1017</td>
<td>726</td>
<td>289</td>
</tr>
<tr>
<td></td>
<td>P1</td>
<td>576</td>
<td>330</td>
<td>263</td>
</tr>
<tr>
<td></td>
<td>P2</td>
<td>705</td>
<td>409</td>
<td>279</td>
</tr>
<tr>
<td></td>
<td>P3</td>
<td>591</td>
<td>347</td>
<td>259</td>
</tr>
<tr>
<td><strong>Issue 2</strong></td>
<td>P4</td>
<td>782</td>
<td>259</td>
<td>269</td>
</tr>
<tr>
<td></td>
<td>P5</td>
<td>593</td>
<td>203</td>
<td>252</td>
</tr>
<tr>
<td></td>
<td>P6</td>
<td>533</td>
<td>193</td>
<td>249</td>
</tr>
<tr>
<td></td>
<td>P7</td>
<td>620</td>
<td>229</td>
<td>253</td>
</tr>
<tr>
<td><strong>Issue 3</strong></td>
<td>P8</td>
<td>883</td>
<td>296</td>
<td>274</td>
</tr>
<tr>
<td></td>
<td>P9</td>
<td>593</td>
<td>238</td>
<td>255</td>
</tr>
<tr>
<td></td>
<td>P10</td>
<td>581</td>
<td>243</td>
<td>257</td>
</tr>
<tr>
<td></td>
<td>P11</td>
<td>636</td>
<td>226</td>
<td>253</td>
</tr>
<tr>
<td><strong>Issue 4</strong></td>
<td>P12</td>
<td>747</td>
<td>202</td>
<td>258</td>
</tr>
<tr>
<td></td>
<td>P13</td>
<td>623</td>
<td>215</td>
<td>252</td>
</tr>
<tr>
<td></td>
<td>P14</td>
<td>547</td>
<td>151</td>
<td>244</td>
</tr>
<tr>
<td></td>
<td>P15</td>
<td>556</td>
<td>157</td>
<td>238</td>
</tr>
</tbody>
</table>

2.2.2 **Issue 4: Government and Healthcare**

This issue posed the following question: “Should individuals be required by the government to have health insurance?” and contained the following four positions for discussion:

• (P12) No, the government should not require health insurance.

• (P13) No, but the government should provide help paying for health insurance.
● (P14) Yes, the government should require health insurance and should punish anyone who does not have it.
● (P15) Yes, the government should require health insurance and guarantee health coverage for everyone.

2.3 REFERENCES


Chapter 3. Design of Mobile Service of Intelligent Large-Scale Cyber Argumentation for Analysis and Prediction of Collective Opinions

3.1 Introduction

People are eager to discuss local and global issues and share their opinions and thoughts with others. Online discussion allows for massive participation from users with different personal backgrounds and perspectives. Large-scale cyber-argumentation platforms facilitate an extensive online discussion where users can post issues and viewpoints, and then others respond to them. In addition, advanced integrated models can analyze the argumentation data and capture the crowd wisdom from these discussions. These analytics help users understand extensive discussions without having to go through the whole discussion. ICAS is an example of a large-scale cyber-argumentation platform with an advanced model to analyze discussions. However, it does not support mobile platforms. People mostly use social media and networking services for large-scale issue-based discussions because of their popularity and accessibility. However, social media and networking services, such as Facebook [1], Twitter [2], and LinkedIn [3], are not explicitly designed to handle large-scale cyber-argumentation. They are not structured for substantial discussions because they are user and connection-centric, not issue-centric. Discussions on the same issue are fragmented; they occur in many users’ local pages and/or are separated into numerous different threads, which are hard to follow. It is easy to lose track of the conversation.

Instead of using social media services for argumentation, some people use online debate services. These debate services, such as Kialo [4], Contra [5], Debate [6], Arguly [7], Debate Map [8] and Ppl’s View [9] are designed to handle argumentation in a user-friendly interface. However, these services have limited capabilities. Typically, they present only two sides of the issues, forcing users to agree or disagree completely with one side. Some issues have neutral positions that need
more discussion. These kinds of platforms do not support neutral positions. However, Kialo [4] has updated the issue structure to have either a single thesis or multi theses for an issue. Then for each thesis, users can submit the pros and cons. The structure has improved a little bit by having more than two-sides on an issue, but still, it requires users to come to a complete agreement or disagreement with a side. Furthermore, in Kialo [4], any author can update their issue after having reviews. These updates may be as minor as spelling mistakes or as major as changing the issue itself. If these edits are significant, as in most cases, not all reviews and replies are valid anymore. Moreover, some of these services require users to reply or comment to express their opinions on an issue. In some cases, users would like to express their agreement or disagreement without the need to post a new comment or reply. For example, if a user sees that someone else has made the argument they were going to make, they should express their partial or complete agreement without making a redundant post of the same reasoning. In addition, they lack analytics to understand the discussion.

Debate Map [8] uses the tree structure for deliberation. When the discussion grows, it is hard to follow as the tree expands and has more depth. While these debates services facilitate argumentation and deliberation better than social media services, they lack support for mobile applications. Only Contra [5] and Ppl’s View [9] have mobile applications of the services listed above. Both applications have no web-based platform and lack analytical models and present modest numbers for users. Ppl’s View [9] differs from Contra [5] by using the hashtag format as a title for issues. Thus, these services lack the analytical power of cyber-argumentation systems and only rarely have robust mobile applications. In the services mentioned above, users cannot explicitly state the degree of agreement or disagreement on a solution for an issue or other
arguments. Therefore, users feel restricted and limited in those kinds of discussions. Moreover, many of these services do not provide analytics for facilitating argumentation effectively.

Other available services, such as Quora [10] and Reddit [11], are open for public questioning and discussion. Quora [10] is a question and answering service, and Reddit [11] is a social news aggregation and discussion service. Those services have rich conversations and attract tremendous user engagement. Many people use Quora [10] as a resource for research, information, and general interest. Some users use Quora [10] to add and build their social networks. However, Quora [10] heavily emphasizes the quality of questions and, in particular, answers. Reddit [11] displays the most upvoted content as "good" content. Therefore, it gets more visibility than "bad" content. The downvote and upvote features for answers, but these features are intended as a judgment of quality, not a judgment of agreement or disagreement. They have a similar structure to the issue tree, but they cannot effectively facilitate many types of discussions, like debates. For example, neither service has a way for users to express explicit agreement or disagreement outside of writing it in their text. This means that the only way to tell if an idea is being supported or not is to read the text replies from every user, which is very cumbersome for the reader. In addition, they provide fundamental analytics to users. They are not designed to handle and process online argumentation.

Researchers have developed a few online debate platforms for argumentation, where users can have many more solutions to a single issue. These tools (Pührer [12], Bex et al. [13]; Iandoli et al. [14]) can handle argumentation better than the above-listed services, but they are not widely available and supported after their research is finished. In addition, the number of participants in these tools is small. Moreover, there are no statistics or measures regarding collective opinions about the posted issues or arguments hosted by these platforms. Most of these platforms do not provide mobile applications, and if they have, they have limited functionality.
Developing an application for mobile-based online argumentation is challenging due to the limited mobile screen space and the potentially massive size of the issue tree. Thus, an intelligent mobile application with a native application for online argumentation in the mobile space is designed to resolve this problem. This application represents the issue tree in a user-friendly way. The issue tree breaks down the substantially complicated question into smaller solvable ones. The root of the tree is the issue, question or unsettled problem or situation under consideration to be solved. The internal nodes and leaves of the tree are suggested views or solutions to the root. In this mobile application, users can easily navigate, participate, and follow the issue-based discussion and understand what is going on in discussions. It allows users to post issues, viewpoints, or arguments pertaining to different topics. All those functionalities are adapted successfully with the consideration of the limited mobile devices screen.

Unlike other argumentation services, which provided users a choice between distinct mutually exclusive positions, our application allows users to define positions that overlap or are alternatives to other positions. Everyone can react to each position or other arguments with different levels of agreement or disagreement. Moreover, users have the choice to add more arguments or counterarguments for deliberation. The resulting discussions between users highlight various aspects of the positions. The application adapted several analytical models to provide meaningful statistics for users, such as the collective opinions prediction and each position’s polarization index. Therefore, users stay aware and well informed about the collective interactions with the system. Finally, a system usability study has been conducted on fifteen users to measure the app usability on mobile devices and results are reported.
3.2 RELATED WORK

There has been some academic research done on both web-based and mobile-based argumentation platforms. Pührer [12] and Bex et al. [13] are examples of web-based argumentation platforms, while Iandoli et al. [14], Hansson and Ekenberg [15] are examples of a mobile-based argumentation platform. In the mobile-based argumentation services, ArgueApply from Pührer [12] is a mobile-based argumentation app that uses the Grappa framework Brewka and Woltran [16] to evaluate different viewpoints on a specific debate. In this app, users can post statements to an ongoing discussion with five agreement or disagreement levels. Then, the application reports all posted viewpoints and the extent to which users have accepted and rejected each viewpoint. The result is displayed as who is right in the discussion. In debates and deliberations, rightness and wrongness are typically not evaluated based on user-feedback but based on truth and logical reasoning. Moreover, if a user is unhappy with the discussion results, they could add unquestionable links that significantly influence the outcome. The discussion in cyber-argumentation systems should capture crowd thinking and not be biased based on individual tendencies. This application presents modest analytics for users, such as who is viewing or joining the discussion. Another example of an argumentation mobile-based platform is from Hansson and Ekenberg [15]. It allows users to submit issues, develop and vote on it. Then others can comment, edit, and add additional documents. It is also structured to have sub-options such as rating, pros, and cons. However, this tool is not designed for public use. The issue’s initiator invites people to participate in the issue. They can set a time limit for the users to contribute to the issue. This application has more functionalities and features than other argumentation platforms. However, this application is meant to be used for professional use or within a closed community where participants know each other. Moreover, opinion analytics and stats in this application are
measured in the user score, which measures both user activity and how much the following activity this activity creates. These stats are modest and straightforward.

In web-based argumentation services, there are different design choices and presentations adapted to handle cyber-argumentation. ArguBlogging from Bex et al. [13] is designed to let users share opinions across blogs. It allows users to post arguments with two responses: agree or disagree. However, it gives users the option to provide more details (links) to support or attack the argument as evidence. It only supports two kinds of blogging sites: Tumblr and Blogger. Still, it is limited to binary sides as responses to posted arguments and no analytics or statistics are provided to the user. Iandoli et al. [14] developed another example of web-based argumentation. They use a collaborative computer-supported argument visualization platform (CCSAV) with moderation. CCSAV platforms are a representation-centric process to consolidate the collective efforts for the creation of a shared argument map based on the Issue-Based Information System (IBIS) argumentation web-framework by Kunz and Rittel, 1970 [17]. The argument map is presented to the users as a tree or network filled with arguments and counterarguments, such as issues, claims, premises, and evidence, with different relationships, such as support, attack, warrant, etc. This kind of platform requires skilled users because it expects users to form the argument map accurately, which for novice users can be challenging to follow. It has a better visual representation of the discussion. However, it requires moderation. If a user does not map the argument accurately, the created argument becomes pending until a moderator validates the issue. They estimated 1:20 as a moderator to active participants ratio, according to Klein [18]. In a real-time setting, it is not feasible to hire the same moderators-users ratio to process all pending posts manually. There was no system usability study conducted to measure product usability for all the research mentioned above products.
The argumentation map in Klein [18] had more than 200 users and more than 5000 postings (ideas, arguments, and comments) in three weeks. They claimed to be large-scale deliberation. In our empirical study, there are more than 300 users and 10000 arguments in five weeks. Please refer to section 2.2 for more information on our empirical data. There is potential for it to be used by a larger user population.

According to Statista [19], there was about a 13% increase in mobile app downloads in 2018 than in 2017. Therefore, mobile phones are taking more focus and use from people. The mobile market growth is expecting mobile users using smartphones to be more than 50% of the total users by 2018. Therefore, developing a mobile-based argumentation application is a significant achievement to bring a large-scale cyber-argumentation to everyone worldwide. This research paper contributes with a mobile application design to support large-scale cyber-argumentation with a simple design and intelligent metrics and analysis capabilities. This design allows users to view issue trees and create new issues, positions, and arguments. Moreover, it enables users to react and reply to each other based on their opinions. Finally, the application produces collective and personal statistics and additional useful information to users using the system, such as predicting collective opinions index and polarization index on the position level.

3.3 THE MOBILE APPLICATION OF CYBER-ARGUMENTATION

3.3.1 THE MOBILE APPLICATION ARCHITECTURE OF CYBER-ARGUMENTATION

This mobile application of cyber argumentation is built on the ICAS platform. An overview of the environment used to develop this application is given in Figure 3-1. The framework of this mobile application consists of two major components: the front-end and the back-end.
3.3.1.1 The front-end

The front-end is an Android application developed by the Cordova framework and written using JQuery and HTML 5. Android is a well-known open-source architecture for mobile application development. Android includes operating systems, application framework, Linux kernel and a set of API libraries. This application requires a minimum SDK level of 28 (Android 9) as a front-end. Android 9 has introduced a new set of features for power management and battery saver improvement. This feature is essential to our app because of the requests and connections between the front and back ends. Cordova is an open-source mobile development framework that uses standard techniques such as JQuery and HTML 5 for cross-platform development. The JQuery DOM Cache was used to increase our application's performance by referencing cached DOM elements.

![Diagram of APP ARCHITECTURE AND COMPONENTS](image)

**FIGURE 3-1 APP ARCHITECTURE AND COMPONENTS**

This application uses a simple technical environment to demonstrate the effectiveness of bringing multi-sided argumentation to handheld devices. Also, this application allows how users can navigate easily between the issues and arguments in the system.
3.3.1.2 User Interface

Although the mobile screen is limited, and the issue tree can be expanded drastically, this application considers simplicity and ease of use. The main components of the app are issues, position, arguments, reactions, and analytics functionalities. Below are the details of each component. Additionally, like most mobile applications, this app requires registration before using it. The registration process is similar to any mobile app registration.

Issues

The main page is loaded when a user is logged in to the application, as in Figure 3-2. The main page displays a list of issues. Every user has the choice either to add a new issue or view the existing issues for debate. If the user is adding a new issue, he/she can add an issue header, description, and a link to any web address (optional data). Moreover, a user can select any issue to view all related information. All related positions are displayed when a user selects an issue, as in Figure 3-3. Users can navigate to the issues list by tabbing in the home bar. Unlike most of the user-generated content, the selected issue description is fixed on the top of the screen. Therefore, users can scroll up and down between the positions and still know which issue they are viewing. There is a high chance of having similar issues posted in the system, depending on trending topics or events.

Positions

Like issues, to engage in, or the user can select to add a new position, as in Figure 3-3. When a user selects a position, all related arguments are displayed. The number of replies, the number of reactions, and the position's overall agreement value are shown; see Figure 3-4. Due to the screen size limitation, small icons are used to label these numbers. These numbers summarize the discussion size and the degree to which participants have agreed/disagreed with the position. From
the position tab, users can reply or react to any position. Users can navigate or go back to the selected issue by tabbing on the issue text at the top of the screen. Like issues, and for the same reasons, the position description is fixed under the issue description. It also allows users to view the discussion analytics under the selected issues based using the analytics. More details are provided later.

**ARGUMENTS AND REACTIONS**

Users can add arguments by using the reply button. Users can explicitly support or attack arguments or positions using the react button, as in Figure 3-4. Levels of agreements/disagreements are selected using the sliding bar, as in Figure 3-7. This sliding bar is divided into 11 levels indicating different levels of agreement. These levels are in one interval $[-1, 1]$ and are separated with 0.2. They also can use the reply button to add an argument to a desired argument or position. If the argument has a counter-argument, a button with a plus sign appears next to it. It helps the user to know which arguments are expandable and which are not. It also summarizes the discussion under the expandable argument as in positions. To maintain the discussion’s integrity, the application service does not allow standard users to delete or modify issues, positions, and arguments after posting them. Instead, users can post updates to arguments and reactions they authored. An update creates a new argument with the user’s modification and is linked to the old argument. The old argument is still visible in the application but marked as an updated one. The application will keep both values because the old argument may have some children that are calculated in the collective intelligence index. This update’s restriction is made because to restrict the author from changing the issue, positions or arguments to something contradictory or unrelated to the original post. This restriction ensures transparency and encourages honesty for all users.
FIGURE 3-7
MAIN PAGE

FIGURE 3-7
SELECTED ISSUE

FIGURE 3-7
SELECTED POSITION

FIGURE 3-7
THE SLIDING BAR

FIGURE 3-7
COLLECTIVE

FIGURE 3-7
PERSONAL
**The Discussion Analytics**

There is two discussion analytics in the application. The first one is collective intelligence analytics, as in Figure 3-5. It presents meaningful numbers about the collective efforts on each position under the selected issue. It contains three collective measurements. The first analytics is the collective intelligence index. This index represents the overall discussion agreement for the selected position. The second index is the collective opinion index. This index reflects the users’ total agreement value, including the users who have not participated in a selected position but have participated in other discussion positions. The third index is the polarization index. This index tells the user how polarized the discussion is under a selected position—more information about these indexes is discussed in section 3.5. The second analytic functionality is personal intelligence analytics, as in Figure 3-6. It summarizes the number of reactions and replies the user makes for each position under the selected issue. It also calculates the overall individual agreement value based on arguments and or reactions made by the user.

Each analytical function illustrates results for all positions under the selected issue on one page. All indices values are posted to the end-users with some formatting. Each index has two attributes: a value and a color. If the index is posted along with red color, most of the participants are against the selected position. If it is posted with green, most of the participants are for the selected position, e.g., the collective index, the Prediction of Collective Opinions index, and the number of arguments and reactions made by the user. However, in some situations, such as totals and the polarization index, the numbers are black since they do not represent any supporting or attacking scenarios.
3.3.2 **THE BACK-END**

The back-end consists of three major components: The Apache server which hosts the web-based application, the ICAS server, and the Neo4j database, as seen in Figure 3-1. All analytics in the application are handled in the back end, ICAS. The Apache server connects to the ICAS server using the WebSocket connection. The ICAS server connects to the Neo4j database using the Bolt connection. The current back-end is a new version of ICAS. There are many functionalities and features added to support argumentation and deliberation.

3.3.3 **FRONT-END AND BACK-END COMMUNICATIONS**

This application is connected to other layers like most mobile applications, as shown in Figure 3-8. At first, the mobile application sends the user logging information to the ICAS server. The ICAS server communicates with the Neo4j database to validate the result. If the user is found in the database, the ICAS server returns a message to the application that contains the issue list. If the user is not found, the ICAS server sends a message to the application, including an error message. When the user selects an issue for browsing, the mobile application requests the ICAS server to retrieve all issue details. The ICAS server forwards the request to the Neo4j database. The Neo4j database returns all issue information to the ICAS server, and then the server delivers the information to the mobile application in the correct format. The same process is applied if the user selects a position or an expandable argument under the chosen issue. For collective and personal analytics, these analytics require an enormous amount of calculations. Therefore, these calculations are not part of the issue details. If the user decides to view these analytics, the mobile application sends requests to the ICAS server to retrieve all analytics details. The ICAS server forwards the request to the Neo4j database. The Neo4j database returns all analytics information to the ICAS server, and then the server delivers the information to the mobile application in the
correct format. These calculations are done on the issue-level for all positions at once. There are many other communications between the mobile application, ICAS server, and Neo4j database. However, only the high levels of communications between the front-end and the back-end are reported.

FIGURE 3-8 FRONT-END ANF BACK-END COMMUNICATIONS

3.4 DISCUSSION AND ANALYSIS

This mobile application is capable of facilitating large-scale cyber-argumentation in a limited handheld screen size. Here is a detailed example of an issue in the application.

3.4.1 EXAMPLE: ISSUE: GUNS ON CAMPUS

The discussion from Issue 3 is summarized in Table 3-1. The number of participants among all positions is about the same except for position 8. There are many reasons for this exception. One reason is being a safety concern due to the fact people feel the need to carry guns on campuses in some cases. However, this position is against carrying guns on campuses regardless of specific reasons or special cases where guns on campuses need to be permitted. The same applies to the number of arguments made in each position.
TABLE 3-1 POSITIONS STATISTICS TO THE GUN ON CAMPUS ISSUE

<table>
<thead>
<tr>
<th>Position #</th>
<th>Position 8</th>
<th>Position 9</th>
<th>Position 10</th>
<th>Position 11</th>
</tr>
</thead>
<tbody>
<tr>
<td># participants</td>
<td>164</td>
<td>129</td>
<td>132</td>
<td>130</td>
</tr>
<tr>
<td># first-level arguments</td>
<td>508</td>
<td>386</td>
<td>387</td>
<td>355</td>
</tr>
<tr>
<td>Overall Collective Agreement</td>
<td>0.2</td>
<td>0.13</td>
<td>-0.36</td>
<td>0.15</td>
</tr>
<tr>
<td>Argument Collective Agreement</td>
<td>0.19</td>
<td>0.09</td>
<td>-0.34</td>
<td>0.14</td>
</tr>
<tr>
<td>React Collective Agreement</td>
<td>0.07</td>
<td>0.07</td>
<td>-0.12</td>
<td>0.06</td>
</tr>
<tr>
<td>Polarization Index</td>
<td>0.30</td>
<td>0.23</td>
<td>0.22</td>
<td>0.26</td>
</tr>
<tr>
<td>Collective predication Index</td>
<td>0.31</td>
<td>0.00</td>
<td>-0.35</td>
<td>0.36</td>
</tr>
</tbody>
</table>

The table has the overall agreement values of arguments and reactions. The reaction agreement is the collective average for users reactions towards the position without a text. The user's reaction is the user agreement degree on the posted position. The argument agreement is the collective average users’ reactions towards a position with a text. In this application, the user can react only or react and reply to a position or an argument. The overall collective agreement is the average users’ reactions to the position and all related arguments regardless of the text. These measurements help non-participants understand participants’ reactions to each position under the selected issue. It also informs the user about how many users have participated and the arguments made so far under the selected issue. For example, position 8 seems to be more popular than position 11 because it has more participants and arguments than position number 11. People who
like to join hot topic discussions may favor position 8 over position 11. For position 10, it seems that people are not in favor of additional training to carry a gun on campus. This position has an agreement value of -0.36.

Regarding the polarization index, the positions under the Guns on Campus issue have similar polarization values. The polarization index for all the positions is less than or equal to 0.30. All of the positions are slightly polarized but not to a significant degree. This index helps users comprehend the distribution of the participant’s agreement for a particular position. It leads to understanding the discussion better.

Finally, the application adapted the collective opinions prediction index to predict the future of a discussion about a position. In this application, about 94 participants out of 336 have responded to all positions. Only about one-fourth of the users have participated in all discussions in the application. Incomplete discussions are the typical case for argumentation and deliberation. However, users may want to know the collective opinions prediction, which predicts the collective user agreement if everyone had participated. The collective prediction index calculates the collective opinion for non-participants based on similar participants’ arguments, reactions to positions, and correlations between positions. In Position 8, less than half of the users did not participate in the conversation. If they join in the future, the application would predict that they might slightly agree on the position with an agreement value 0.31.

### 3.5 System Usability Study

Usability is a multidimensional aspect of a product. It merely describes how a product is usable in its context. However, when measuring mobile device system usability, the screen size should be considered an additional aspect. The screen size of handheld devices is the primary design factor for designing mobile applications user interface.
Commonly, some variables need to be defined to measure system usability. There are three main variables for the application usability of this mobile application: users, goals, and context. The target users of this application are adults without age limitations. Thus, from defining the users, the goals can be identified. This application should allow users to discuss any issue and view related arguments and analytics. This application is designed to bring multi-sided deliberation and discussions in handheld devices.

### 3.5.1 Methodology

In this application, attempts are made to measure the system usability for our application. Since the data in ICAS was collected using the web-based version of ICAS, fifteen graduate students as users are recruited to evaluate this mobile application. Each of the users had used the mobile version of ICAS for discussions and evaluation. The SUS questions from Usability.gov are used. Researchers have modified the questions to accommodate our mobile application. Each question has five possible answers. The possible answers range from one to five, where one means strongly disagree, and five means strongly agree.

### 3.5.2 Results

Table 3-2 contains the SUS questions and our users’ evaluations for our application. From Table 3-2, the majority of the users agreed on using the app frequently. Surprisingly, users were split equally about their feelings toward the app’s complexity. However, most of the users found the app easy to use. Furthermore, few users thought they needed technical support to be able to use this application. Regarding the app functionally, most of the users found that this app has integrated many functions well. However, about 27% of the users found some inconsistencies in this application. 74% of the users thought that most people would learn to use the application quickly, but the minority of users found the app to be cumbersome. About 87% of users felt confident to
use the app, and 80% needed to learn nothing before using the app. The researchers can conclude that this application is an initial step to bring multi-sided argumentation and deliberation into handheld devices. The final application score based on SUS scores and measures is 68, which is average, according to Usability.gov [20]. In addition, according to [21], the demonstrated application can be labeled as a “Good” application; it got a SUS mean score equal to 68, which is closer to 71.4 than 50.9. Moreover, the 70% percentile is calculated, and it was 72.5. According to [22], our application scores as “B” using the percentile rank. It can be concluded this application has the potential in bringing multi-sided large-scale cyber-argumentation into the limited screen size platforms.

3.6 Conclusion and Future Work

In this paper, an innovative large-scale intelligent cyber-argumentation mobile service is developed. This mobile application represents argumentation and facilitates structured discussion of multiple positions on an issue intelligently in handheld mobile devices. Participants can add issues, positions, and arguments. They can explicitly provide degrees of their agreements/disagreements when reacting to positions and arguments. Discourses and conversations are structured and easy to navigate and understand. This application uses intelligent analytic models such as polarization index, collective opinion prediction, and analytics about the collective and personal participation. These models allow participants to focus on and be well informed about the ongoing conversations. This mobile application substantially improves argumentation in the mobile space over the existing social media and debate service. An example and system usability study of our application is presented in this paper.
Currently, this app is being updated to include social networking services for deliberation, mobile notifications, and automatic issue generation. In addition, this app is being enhanced for its usability based on the users’ feedback.

**TABLE 3-2 SUS QUESTIONS AND RESPONSES**

<table>
<thead>
<tr>
<th>Question #</th>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Neutral</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1. I think that I would like to use this application frequently.</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>Q2. I found the application unnecessarily complex.</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Q3. I thought the application was easy to use.</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>Q4. I think that I would need the support of a technical person to be able to use this application.</td>
<td>4</td>
<td>6</td>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Q5. I found the various functions in this application were well integrated.</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td>Q6. I thought there was too much inconsistency in this application.</td>
<td>8</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Q7. I would imagine that most people would learn to use this application very quickly.</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>Q8. I found the application very cumbersome to use.</td>
<td>8</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Q9. I felt very confident using the application.</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>5</td>
</tr>
<tr>
<td>Q10. I needed to learn many things before I could get going with this application.</td>
<td>3</td>
<td>9</td>
<td>0</td>
<td>3</td>
<td>0</td>
</tr>
</tbody>
</table>

3.7 **REFERENCES**

[1] Facebook: www.facebook.com
[3] LinkedIn: www.linkedin.com
[8] Debate Map: from debatemap.live
[9] Ppls’ View: from Google Play Store
[20 Usability from: www.usability.gov
Chapter 4. **OPINION DISCOVERY FRAMEWORK: TOWARD A QUALITY OPINION-CENTRIC PLATFORM**

4.1 **INTRODUCTION**

User-generated content (UGC) platforms allow extensive discussion and user participation, such as social media, blogs, photo sharing, and websites. Generally, UGC content is created by regular people and mostly unsanitized as traditional media outlets. Therefore, UGC content suppliers are rewarded by receiving recognition from content consumers, who use UGC platforms for information or entertainment. As a result, UGC applications create an attractive user environment and adapted AI models to help users be more active, creative and develop new personal and business opportunities. Cyber-argumentation, an AI sub-field, is an example of a UGC platform. Online-Argumentation platforms allow vast discourse between participants as well as understanding the discussions. Cyber argumentation addresses issues by creating well-defined structures for deliberation and has exhibited the ability to evaluate the discussion on large-scale platforms and in different contexts [1-4].

While large-scale deliberation systems host massive opinions, not all opinions are worth further discussion. Opinions in cyber-argumentation systems hold rich information beyond its text. Commonly, in public deliberation systems [5, 6], opinions can be discovered by engagement information, impact score, or reverse-chronological order. These discovering methods omit important features an opinion may exhibit. Mostly, they degrade some constructive opinions using only one feature to search and discover opinions. On the other hand, constructive opinions, or critiques, usually hold good intentions, motives for improvements, and positive feedback that makes a particular situation better. Therefore, identifying features of constructive opinions for searching and discovering opinions is needed. However, it is a challenging task because there are
no clear definitions and measurements for opinion features. Some opinions are informative, related, controversial, or none of the previous. For instance, constructive opinions receive more attention over time than unconstructive opinions. These opinions create a controversial state over the written argument, attracting the audience to react and engage. Thus, constructive opinions retain some unique characteristics. Recognizing the different attributes of worthwhile opinions provides signals to what degree each opinion should be promoted and presented to users. Discovering and promoting opinions using different opinions features will lead to higher quality discussions because various dimensions of searching opinions are considered instead of only one dimension. Constructive discourse results in strengthening online communities and uniting users [7]. Therefore, users reach their best-reasoned judgment to solve a problem, increase participation in the discussion, and achieve the argumentation platform goals.

This research presents a novel selection framework for searching and discovering constructive opinions in cyber-argumentation platforms. This framework searches and discovers opinions based on four non-textual opinion features used to seek a dynamic opinion-centric platform. This framework is a user-customized search and uses four opinion-distinguishing features: engagements, recentness, author influence, and controversy. First, these features are defined in the cyber-argumentation field, analyze the relationship between them, and introduce our new framework. The used dataset is the dataset collected from our argumentation tool and our framework to create a dynamic opinion discovery system.

4.2 RELATED WORK

Limited research has been done on argument search and discovery. Nguyen et al. [8] has developed an end-end process in argument discovery to minimize the crowd cost and maximize the quality of crowd answers using argument text. Mishne [9] identified blog opinions by scoring posts based
on various aspects associated with an opinion about a topic. It included shallow sentiment analysis, spam detection, and link-based authority estimation. Furthermore, Amgoud & Ben-Naim [10] proposed a new semantics family, which searches and rank-orders arguments from the most acceptable to the weakest. Thus, their approach offers a theoretical framework for comparing semantics. While Bonzon et al. [11] have proposed six new ranking-based semantics for search results based on the propagation of the weights of the arguments, their approach gave a higher weight to non-attacked arguments. Moreover, Eirinaki at el. [12] presented an algorithm that analyzed the overall sentiment of a document/review and identified the semantic orientation of the review-specific components that lead to a particular sentiment. Finally, Pu et al. [13] defined new ranking-based semantics, called categoriser-based ranking semantics, for abstract argumentation framework. All the research work mentioned above focuses on discovering arguments using textual information as the foundation for argument discovery methods.

Moreover, there are some public debate platforms, such as Kialo [5] and Debate.com [6], where opinions are discovered differently. For example, in Kialo, opinions under the thesis statement are searched and listed using the impact score. The opinions are listed in decreasing order based on their impact score and the impact scores are calculated using the users’ ratings; a user can rate an opinion impact on a scale [1 - 5] and other information. Likewise, in Debate.com, users can choose from different discovery methods to sort out the opinions, such as by date, most agreed, most disagreed and unanswered. Under each opinion, the counter-opinions are listed based on the number of users’ interactions with these opinions, such as replies and likes. In all the platforms mentioned above, the temporal information has been omitted. In addition to that, the opinion author’s status has not been considered for scoring or discovering those opinions. Furthermore, there was no consideration for controversy in their discovery methods.
On other UGC platforms, there are many aspects to consider for scoring or searching posts. In particular, social media applications have rich content from users. This content is processed to bring the most attractive and related content to users. For instance, Facebook [14], the most popular platform for social connections, has developed a personalized news feed algorithm to rank users' stories that matter most to users every time they visit Facebook. It searches and evaluates each story based on who posted it, what kind of media it contains, and interactions made so far to each story. The Facebook story ranking algorithm scores each story based on that information and places the ones that matter the most to users on the top of the feed. Another UGC example is Twitter [15]. Twitter is a popular microblogging platform. It categorizes each user’s timeline into three sections. The first section displays the top tweets using tweets engagement, user’s connection, and user preferences information, while the final section lists the remaining tweets and events in reverse-chronological order. Each UGC platform has its design and attributes. Considering those attributes to evaluate content is a significant reason for platform success and attracting more users.

Cyber-argumentation platforms are becoming widely available and supported with the expansion of UGC platforms and advanced AI techniques. Cyber-argumentation platforms differ from other UGC platforms by capturing the crowd wisdom and collective opinions dynamically. There is a need to develop an opinion discovery framework toward a quality opinion-centric platform to achieve this goal. This paper proposes a new opinion discovery framework that uses features of different non-textual opinions to create a quality opinion-centric platform.

4.3 Problem Description

ICAS has more than 10,600 arguments spread among 16 positions. On average, there are 408 first-level arguments for each position. From our study, on average, users view only 3% of the content.
Users do not explore all arguments due to the massive content and time limitation. Therefore, there is the need to select and present the most constructive arguments for users. However, this is a challenging task for many reasons. First, this dataset has large-scale discussions, which stresses the need for an argument discovery framework. Therefore, users have no time to go through all posted arguments to find constructive arguments. Secondly, users have no choice to select or express their preferences. Finally, there are no clear argument features set that can be used to build the arguments list. Arguments can be old, new, engaging, not engaging, short, long, reasonable, sound, valid, controversial, docile, etc.

Arguments features are either explicit, such as created date, number of reactions, agreement value, etc., or implicit, such as the degree of engagement, controversy, etc. ICAS searches and sorts arguments by reverse-chronological order, which is the common argument discovery method for deliberation platforms. Therefore, all new arguments will always be at the top of the list, pushing old arguments down the list. This searching method is reasonable and commonly used, but it has limitations. First, it is not capable of recognizing the old arguments with recent user engagement. Therefore, these arguments receive less or no user engagement over time. Secondly, ordering the arguments list by date cannot capture the other significant arguments’ features. However, not all arguments are engaging. In our dataset, more than half of the arguments received no engagement from users. Discovering arguments by engagement information gives more exposure to arguments with higher user engagement regardless of the other features. Therefore, the new arguments are penalized for appearing further down on the search list. Failing to recognize an argument's temporal aspect can negatively affect the user experience and make the arguments list appear unchanged. Another vital feature of arguments is the degree of controversy. Some arguments are controversial, making users reacting with different levels of agreement or disagreement.
Identifying the controversial degree for arguments is the key component of the argumentation systems.

Finally, the last feature that plays an essential role in argument discovery is the argument’s author’s influence. People come to the platform with different knowledge and skills. Arguments’ authors directly or indirectly impact readers' thoughts, feelings, and actions. Failing to recognize the authors’ influences as an aspect of an argument can undermine its significance. Up to our knowledge, no research has been done so far that considers the arguments mentioned above features for argument search and discovery. Hence, the researchers are referring to these features as indicators for the rest of this research.

### 4.4 Proposed Framework

According to the multi-faceted concept of argument discovery, a new framework is proposed to search and discover the most constructive arguments in a position tree based on four non-textual indicators: Argument’s Engagement, Argument’s Recentness, Argument’s Degree of Controversy, and Arguments Author’s Influence Degree. In this research, the proposed framework is named the Argument Discovery Framework. The researchers define constructive arguments as new engaging arguments that create a controversial state written by an influential user. Therefore, this research focuses only on the indicators mentioned above to discover the most constructive opinions in online discussions.

This section first defines the argument indicators: engagement, recentness, controversy, and author influence, and methods to quantify the cyber-argumentation space. Then, the correlations between the indicators are analyzed to apply the recommended scoring method. Finally, the aggregated argument discovery framework is introduced.
An argument is made up of related arguments and reactions. An argument is defined in ICAS as a tuple:

\[ <a, p, u> \]

To create an argument, there are three entities involved while submitting a new argument process:

- \( a \in A \): is the argument to be made. It has some features such as ID, text, agreement value, created date, last activity time.
- \( p \in \{ A, P \} \): is the parent node for the made argument. \( p \) could be an argument \( \in A \), or a position \( \in P \). It has the same features as the argument.
- \( u \in U \): is the user who authored the argument.

Similarly, the reaction is defined in ICAS as a tuple:

\[ <r, p, u> \]

To create a reaction, there are three entities involved while submitting a new argument process:
● $r \in R$: is the reaction to be made. It has some features such as ID, agreement value, created date.

● $p \in \{A, P\}$: is the parent node for the made argument. $p$ could be an argument $\in A$ or a position $\in P$. It has the same features as the argument.

● $u \in U$: is the user who posted the reaction.

It is intuitive to construct a graph to illustrate the involved factors in argumentation behavior. The graph is the argumentation graph, as shown in Figure 4-1. Now, a formal definition of the argumentation graph is given as follow:

**Definition 1. (Argumentation Graph)** argumentation behavior can be represented as a graph $AG = (V, E)$ where:

- $V = A \cup P \cup R \cup U$. Four types of entities are involved: A arguments, P positions, R reactions and U users.
- $E = \langle u, p \rangle \cup \langle u, a \rangle \cup \langle u, r \rangle \cup \langle a, p \rangle \cup \langle r, p \rangle \cup \langle a, a \rangle \cup \langle a, r \rangle$. $\langle u, p \rangle$ represents $u$ posted a position $p$. $\langle u, a \rangle$ represents $u$ posted an argument $a$. $\langle u, r \rangle$ represents $u$ posted a reaction $r$ to an argument or position. $\langle a, p \rangle$ represents an argument $a$ posted to a position $p$. $\langle r, p \rangle$ represents $r$ a reaction made to a position $p$. $\langle a, a \rangle$ represents an argument $a$ posted to an argument $a$. $\langle r, a \rangle$ represents $r$ a reaction made to an argument $a$.

The argumentation graph serves as the baseline graph for this research work. This graph is used and modified in many ways to extract extra information used in this research.
4.4.2 Engagement

Usually, user engagement is defined as the quality of user experience with technology [16]. In ICAS, users can engage with the system by viewing or adding issues, positions, or arguments, update an argument and react to an argument or position and other functions. All user’s interactions made to issues, positions or arguments are considered as an engagement. Therefore, an entity's engagement is measured by the amount and interactions made by users to that entity. Users can react or reply to an argument or position. A user’s reaction is determining the level of agreement/disagreement with the parent entity without text. A user’s reply is determining the level of agreement/disagreement with the parent entity with text. Therefore, replies have more weight than reactions in an entity engagement measurement because they contain additional information. To measure the degree of engagement for each argument, some data sources from the AG are used as follow:

- $V = A \cup P \cup R$. Three types of entities are involved: A arguments, P positions, and R reactions.
- $E = \langle a, p \rangle \cup \langle r, p \rangle \cup \langle a, a \rangle \cup \langle r, a \rangle$. $\langle a, p \rangle$ represents an argument a posted to a position p. $\langle r, p \rangle$ represents a reaction made to a position p. $\langle a, a \rangle$ represents an argument a posted to an argument a. $\langle r, a \rangle$ represents a reaction made to an argument a.

The researchers assume that engagement information contributes to:

- Node Weights: can be interpreted as the sum of all edges’ weights connected to this node.
Edge Weight: is obtained from the relation between entities. If the relation is created between two arguments, that edge weighs two. If the relation is created between a reaction and an argument, that edge weighs one.

According to the above assumptions, the argument is defined as a tuple:

\(<a, p, e_{ap}>\).

Similarly, the reaction is defined as a tuple:

\(<r, a, p, e_{rp}>\).

To calculate the argument engagement information in AG, there are two entities involved while submitting a new argument process:

- a, r, a, p: the same as in the AG.
- e_{ap}: is the engagement weight assigned to the created relation between the argument a and every ancestor in the path of argument a. The weight is equal to two because each argument is made up of an agreement value and a text.
- e_{rp}: is the engagement weight assigned to the created relation between the reaction r and every ancestor in the path of reaction r. The weight is equal to one because each reaction is made up of an agreement value only.

The total engagement score \(TE(a_i)\) can be calculated for each argument as:

\[
TE(a_i) = \sum_{r_j, a_j \rightarrow a_i} w_{ij}
\]  

(4.1)

Where

\[
w_{ij} = \begin{cases} 
2, & \text{if } a_j \text{ has a relation with } a_i \text{ in AG} \\
1, & \text{if } r_j \text{ has a relation with } a_i \text{ in AG} \\
0, & \text{else}
\end{cases}
\]  

(4.2)
To measure the degree of engagement for argument, the total engagement of argument \( a \) is divided by the sum of all \( a \)'s siblings' total engagement at the same level. The engagement score for argument \( a \) is calculated as:

\[
E_{score}(a_i) = \frac{TE(a_i)}{\max(\sum_{a_j \in A} TE(a_j), 1)}
\]  

(4.3)

To avoid dividing by zero, the maximum function is added to choose between the maximum total engagement information and one. This smoothing to take into account arguments or positions without children. The formula (4.3) returns a double value \([0, 1]\), where 0 means the argument has no engagement and 1 means the argument is the most engaging.

### 4.4.3 Recentness

Recentness can be interpreted in different ways in ICAS. For specific temporal actions, such as posting new arguments, this is considered a recent activity. On the other hand, partial updates to arguments such as updating content and receiving new counter-arguments or reactions are considered recent updates. For both situations, the argument time information has changed. Specifically, in the second situation, the argument tree has expanded and received more information, support, or an attack. Therefore, each entity's recentness can be measured by passing on its last activity time to all the entity’s ancestors. The same data sources are used to measure the engagement information to measure the degree of recentness for each entity, argument or position.

To calculate the argument recentness information in AG, there are two entities involved while submitting a new argument process:

- **a, r, p:** the same as in the AG.
- **\( r_{ap} \):** is the recentness weight assigned to the created relation between argument \( a \) and every ancestor in the path of argument \( a \).
- \( r_{r_r} \): is the recentness weight assigned to the created relation between reaction \( r \) and every ancestor in the path of reaction \( r \).

To calculate the recentness score for argument \( a \), the recentness information are needed to be found for each argument and the maximum recentness information among its siblings. However, the recentness information for arguments is composed of date and time information. Therefore, the last activity time normalization is needed for all arguments into an interval of time units, such as minutes, hours or days, to find the maximum recentness. This normalization will provide more information about an interval of time units. The interval’s endpoints can be set to thresholds. For example, the upper endpoint can be set to the current time, and the lower endpoint can be set to a few hours or days early. In this study, days are used as interval units. The argument’s last activity time (date only) is normalized to an interval of integers. The normalized start date of the study is used as the lower endpoint for the interval. The normalized last date of the study is used as the upper endpoint for the interval. The recentness information for each argument is calculated as follows:

\[
R(a_i) = \text{normalized (last activity time}(a_i)) \quad (4.4)
\]

The maximum recentness information is the maximum value of recentness information from all arguments with the same level as \( a_i \). After that, calculate the recentness score for each argument as:

\[
R_{\text{score}}(a_i) = \frac{R(a_i)}{\text{Max } (R_{a_j \in A}(a_j), 1)} \quad (4.5)
\]

Similar to the engagement score, smoothing is applied for the recentness score. The formula (4.5) returns a double value \([0,1]\), where 0 means the argument is an old argument and 1 means the argument is new.
4.4.4 Controversy

The controversy is a major phenomenon in cyber-argumentation. However, there is no unified definition for controversy in UGC platforms. Dori-Hacohen [17] quantifies controversy as the degree of disagreement among large groups of people in discussion or issues. Garimella et. al. [18] quantifies controversy based on the topic’s and user’s levels using different methods. In ICAS, the argument’s controversy measures the degree of users’ agreement or disagreement for a position or an argument. Argument \( a \) with constant support means that users agreed on supporting and accepting this argument. Argument \( a \) with constant attack means that users agreed on disagreement or rejection of this argument. However, argument \( a \) with different levels of support and attack is more controversial. Therefore, to measure the degree of controversy for argument \( a \), some data sources from AG are used, the same sources used to define arguments’ engagement and recentness. The agreement value of an argument or reaction is used as the base information. Arguments with a range from -1 to +1. (-1) means a strong disagreement or attack, and (+1) means strong agreement or support. The standard deviation is used to measure the degree of controversy for an argument or a position. Usually, the standard deviation of data points is frequently used to measure the volatility of those points. If all data points are closer to the mean, the standard deviation is low. If the data points are spread out over a wider range of values, then the standard deviation is high. Therefore, an argument with constant attack or support will have a low standard deviation. It means that the supporters or attackers agree. However, arguments with mixed attacks or supports will have a higher standard deviation. This disagreement can be captured by calculating the standard deviation on the agreement values made to an argument. The controversy degree for argument \( a \) is calculated for all direct arguments’ and reactions’ agreement values made to argument \( a \) as:
\[ C_{\text{score}}(a) = \sqrt{\frac{1}{N - 1} \sum_{i=1}^{N} (x_i - \bar{x})^2} \] (4.6)

\( x_i \) is the agreement value of a direct child for argument \( a \), \( \bar{x} \) is the mean value of argument \( a \) direct children agreement values, \( N \) is the number of direct children to argument \( a \). For siblings arguments, the \( C_{\text{score}}(a) \) gets normalized to \([0,1]\) where 0 means the argument has a low controversial degree and 1 means the argument has a high controversial degree.

4.4.5 Author’s Influence

This indicator determines a user’s influence on the discussion. Users in the discussion are authors, readers, or both. This research is interested in the author’s influence. When a user posts an argument and gets many replies and reactions, he influences the responders. This could happen in a different context. A user may have a strong influence in one or more positions and no influence on the other positions. Therefore, the degree of author influence on the position level is calculated. The AG is used to create the User Interaction Graph (UIG) in Figure 4-2. Now, the formal definition of the User Interaction Graph is as below.

Definition 2. (User Interaction Graph) users’ interaction behavior can be represented as a graph \( \text{UIG} = (V, E) \) where:

- \( V = U \), only one type of entity, which is users \( U \).
- \( E = \langle u_i, u_j \rangle \) represents that \( u_i \) has replied or reacted to an augment written by \( u_j \).

The researchers assume that author influence information contributes to:

- Node Weights: can be interpreted as the degree of the user’s influence on the discussion.
- Edge Weight: is obtained from the relation between users. If $u_i$ has replied to an augment written by $u_j$, that edge weighs two. If $u_i$ has reacted to an augment written by $u_j$, that edge weighs one.

The UIG information as the data sources is used to calculate the author influence's degree in the discussion. PageRank Algorithm [19] has been widely used to identify influentials in different scenarios [20, 21]. To measure the author’s influence for an argument $a$, the weighted PageRank Algorithm [19] is used.

$$AI_{Score}(u_i) = PR(u_i) = (1 - d) + d \sum_{u_j \rightarrow u_i} PR(u_j) W^{in}_{(u_j, u_i)} W^{out}_{(u_j, u_i)} \quad (4.7)$$

$W^{in}_{(u_j, u_i)}$ is the weight of link $(u_j, u_i)$ calculated based on the number of in links of user $u_i$ and the number of all references users of $u_j$, $W^{out}_{(u_j, u_i)}$ is the weight of link $(u_j, u_i)$ calculated based on the number of out links of user $u_i$ and the number of all references users of user $u_j$. The $PR(u_i)$ score range is $[0, 1]$, where 0 mean user $u_i$ has a low influence degree and 1 means user $u_i$ has a high influence degree as an author.

![Figure 4-2 USER INTERACTION GRAPH](image_url)

4.4.6 Correlations between Indicators

The PEARSON correlation [22] between the indicators mentioned above for all discussions of the positions in ICAS is performed. Figure 4-3 shows us the correlations between the four indicators.
According to [22], the recentness indicator (R_Score) does not correlate with the controversy indicator (C_Score) and has a very weak correlation with the other indicators. This is because, over time, users lose interest in viewing and interacting with old postings. However, the engagement indicator (E_Score) has a moderate correlation with the controversy indicator (C_Score) and a strong correlation with the author's influence indicator (AI_Score). The reason behind those correlations is that these scoring methods use the same sources of data to score the arguments for those indicators. Although the controversy indicator (C_Score) and the author influence indicator (AI_Score) use the same data source for scoring arguments, they have a very weak correlation with each other. In other words, a controversial argument does not have to be written by an influential author and vice versa. To find the relation between any indicators, the PEARSON correlation [22] between them is performed. Then, based on the R-value, the appropriate operator can be used to combine those indicators. Therefore, the relationships between the indicators are analyzed to help build the argument discovery framework.

4.4.7 Relationships among Indicators

The researchers found that some of the indicators strongly correlate, such as the engagement indicator (E_Score) and the controversy indicator (C_Score). However, moderate, weak, or no
correlations exist between other indicators, as shown in Figure 4-3. Thus, the relationships between
the indicators are categorized as in [23] into conflicting, cooperative, and mutually exclusive.

4.4.7.1 Conflicting Indicators (⊗)

According to Liu et al. [23], two indicators are said to be conflicting if the correlation value r is
less than 0. If there is an increase in one indicator scoring value, it always decreases in the other
indicator scoring value. Therefore, they are completely conflicting. In our setting, there are no
conflicting indicators. However, if this framework gets expanded or modified in the future, it may
have conflicting indicators. In this situation, for each argument a, the fuzzy compromise operator
⊗ is applied on the argument indicators to score argument a. Consider the following set of
indicators scores for argument a: \{I_1(a), I_2(a), \ldots I_n(a)\}. This operator combines all included
indicators and uses the average function to tradeoff between those conflicting indicators as follows:

\[ I_1(a) \otimes I_2(a) = \frac{I_1(a) + I_2(a)}{2} \]  \hspace{1cm} (4.8)

4.4.7.2 Cooperative Indicators (⊕)

According to Liu et al. [23], two indicators are said to be cooperative if the correlation value r is
greater than 0. If there is an increase in one indicator scoring value, it always increases the other
indicator scoring value. Therefore, they are satisfied simultaneously. An example of this
correlation is the correlation between the engagement (E_Score) and the controversy (C_Score)
indicators. In this situation, for each argument a, the fuzzy conjunction operator ⊕ is applied to
score arguments. Consider the following set of indicators scores for an argument a: \{I_1(a), I_2(a),
\ldots I_n(a)\}. This operator combines all included indicators and uses the MIN function [23] for
combining those cooperative indicators to get services at the lowest price. However, the MIN
function is not suitable because it would degrade the argument score in our context. Because the
researchers are trying to find the arguments with the highest score and the indicators are
cooperative, the indicators scores are multiplied to calculate the final score of argument a as follows:

\[ I_1(a) \oplus I_2(a) = I_1(a) \times I_2(a) \]  
(4.9)

### 4.4.7.3 Mutually Exclusive Indicators (⊙)

According to Liu et al. [23], two indicators are mutually exclusive if there are no correlations between them. Therefore, they cannot be satisfied at the same time to the highest degree. An example of this correlation is between the recentness (R_Score) and the controversy (C_Score) indicators. In this situation, for each argument a, apply the fuzzy disjunction operator ⊙ is applied to score arguments. Consider the following set of indicators scores for an argument a: \{I_1(a), I_2(a), \ldots I_n(a)\}. This operator combines all included indicators and uses the MAX function to combine those mutually exclusive indicators as follows:

\[ I_1(a) \odot I_2(a) = Max(I_1(a), I_2(a)) \]  
(4.9)

### 4.4.8 The Argument Discovery Scoring Model

Not all users are willing to choose the argument discovery method. Moreover, not all users are expected to understand the argument indicators and their correlations, especially when the recentness indicator has different correlations values with the other indicator. As a result, users may not be able to build the recommended argument discovery method. Yet, the researchers want
to discover the most constructive arguments for users. Therefore, the argument discovery model is presented. This model aggregates the argument indicators while paying equal attention to each of the four indicators and uses a structured equation modeling (SEM) technique [24] to build a linear model to synthesize the argument’s indicators. SEM is frequently used to evaluate and assess unobservable ‘latent’ constructs using one or more observed variables. The argument discovery score method is formulated using SEM as follows:

\[
S(a) = (\alpha \times R\_score(a)) + (\beta \times E\_score(a)) + (\gamma \times PR(u_a)) + (\delta \times C\_score(a))
\]  \hspace{1cm} (4.10)

In our experiments, the parameters \(\alpha, \beta, \gamma, \delta\) were set to 0.05, 0.2, 0.72, 0.03, respectively. At first, each indicator is computed for each argument separately using the equations mentioned above (4.3, 4.5, 4.6, 4.7). Then, those four indicators are combined to calculate the argument score for discovery (4.11). The arguments’ scores are in \([0, 1]\). The arguments are displayed in the arguments list based on the arguments’ score in decreasing order. As a position or an argument is made of a tree with different levels, this framework is applied recursively for each indicator for each argument from the leaves to the tree root.

### 4.4.9 The Argument Discovery Framework

In Section 4.5.7, the argument discovery scoring method is introduced. In some situations, users would like to focus on one or more indicators, so they may need to select the argument discovery method. Therefore, an argument discovery framework is built to accommodate users’ preferences. Figure 4-5 depicts an overview of the argument discovery framework. The proposed framework is straightforward, as follows:

1. The user selects position P, retrieve all P’s arguments.
2. For each retrieved argument $a$, calculate the indicators scores ($4.3, 4.5, 4.6, 4.7$) and the argument discovery method using ($4.10$).

3. List all arguments to the user by the argument discovery method score ($4.10$) in decreasing order.

4. Ask the user to build the preferred discovery method:
   
   • Use the indicators score for each argument from step 1.
   
   • Apply the selected discovery method based on the type of relationship between indicators. Calculate the argument score.

5. List all arguments to the user by the selected argument discovery method score in decreasing order.

FIGURE 4-5 THE ARGUMENT DISCOVERY FRAMEWORK

4.5 APPLICATION EXAMPLE

In this section, an application example with different scenarios is presented to illustrate the proposed framework on the same dataset from section 2.2. In this example, different argument discovery requests are used to search and recommend the most relevant arguments based on a user’s preferences. Each example demonstrates a scenario from the proposed framework. All the data demonstrated in this example are from Position P0.
4.5.1 Scenário 1

A user submits an argument discovery request of two cooperative indicators with the recommended operator. For example, if the request is “Engagement ⊕ Controversy”, the top 5 relevant arguments are in Table 4-1.

From Table 4-1, all arguments discovered by this discovery method are engaging but not necessarily recent. For example, the last time a30 received attention from users was on day 1 of the study. Likewise, the last time a39 received attention from users was on day 2 of the study. However, the user did not specify the recentness in the request. All arguments in the above table had scored above the 90th percentile in terms of controversy.

<table>
<thead>
<tr>
<th>Argument No.</th>
<th>Author Name</th>
<th>Created Date</th>
<th>No. Reactions</th>
<th>No. Arguments</th>
</tr>
</thead>
<tbody>
<tr>
<td>a30</td>
<td>user53</td>
<td>1</td>
<td>0</td>
<td>11</td>
</tr>
<tr>
<td>a39</td>
<td>user106</td>
<td>2</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>a369</td>
<td>user58</td>
<td>9</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>a785</td>
<td>user303</td>
<td>23</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>a161</td>
<td>user87</td>
<td>4</td>
<td>2</td>
<td>9</td>
</tr>
</tbody>
</table>

4.5.2 Scenário 2

A user submits an argument discovery request of two mutually exclusive indicators with the recommended operator. For example, if the request is “Recentness ⊓ Controversy”, the top 5 relevant arguments are in Table 4-2.
All arguments discovered by this discovery method have scored above the 90th percentile in terms of controversy. However, those arguments are not the newest. For an argument to generate a controversial situation in the discussion, it requires more time to be seen and discussed by users. For example, a813 was created on day 24, and the study lasted for 26 days. Other arguments scored higher in terms of controversy, but they did not score high according to the user’s request. Because these indicators are mutually exclusive, it is hard to find arguments that simultaneously scored the highest for all indicators.

4.5.3 Scenario 3

A user submits an argument discovery request of two conflicting indicators without the recommended operator. For example, if the request is “Controversy ⊗ Recentness”, the top 5 relevant arguments are as in Table 4-3.

Arguments discovered by this discovery method are new, but some more recent arguments have not been recommended. In terms of controversy, only two arguments have scored above the 90th percentile. Since those indicators are mutually exclusive, the arguments order is different from Scenario 2. Therefore, using the not recommended operator between the indicators leads to
different results or no results in some cases. Combining indicators without knowing the relationship between the indicators could lead to unsatisfying results. Therefore, the argument discovery method is recommended for users who do not know or do not understand how to combine arguments with the recommended operator.

**TABLE 4-3 DISCOVERING ARGUMENTS WITH UN-RECOMMENDED OPERATOR BETWEEN INDICATORS**

<table>
<thead>
<tr>
<th>Argument No.</th>
<th>Author Name</th>
<th>Created Date</th>
<th>No. Reactions</th>
<th>No. Arguments</th>
</tr>
</thead>
<tbody>
<tr>
<td>a947</td>
<td>user29</td>
<td>25</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>a943</td>
<td>user226</td>
<td>25</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>a818</td>
<td>user249</td>
<td>24</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>a936</td>
<td>user19</td>
<td>25</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>a813</td>
<td>user249</td>
<td>24</td>
<td>0</td>
<td>4</td>
</tr>
</tbody>
</table>

**4.5.4 SCENARIO 4**

In this scenario, the results of applying the argument discovery model (4.11) are presented. The top 5 relevant arguments are as table 4-4.

Arguments in the above table are engaging and recent. Moreover, some of those arguments had scored above the 90th percentile in terms of all of the argument indicators. This model does not require users to understand the argument indicators nor the mechanism of different argument discovery methods. Furthermore, this mechanism does not require additional processing since it is automatically applied after selecting a position.
TABLE 4-4 DISCOVERING ARGUMENTS BY THE ARGUMENT DISCOVERY MODEL

<table>
<thead>
<tr>
<th>Argument No.</th>
<th>Author Name</th>
<th>Last Activity Day</th>
<th>No. Reactions</th>
<th>No. Arguments</th>
</tr>
</thead>
<tbody>
<tr>
<td>a943</td>
<td>user226</td>
<td>25</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>a947</td>
<td>user29</td>
<td>25</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>a818</td>
<td>user249</td>
<td>24</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>a813</td>
<td>user249</td>
<td>24</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>a936</td>
<td>user19</td>
<td>25</td>
<td>0</td>
<td>2</td>
</tr>
</tbody>
</table>

4.6 EVALUATION OF THE FRAMEWORK

The framework implantation is simple and straightforward. The number of arguments under the selected position is bounded upon the user selecting a position to view. Additionally, the number of argument’s indications is specified, too. Then, the score of each indicator for each argument is computed. Finally, the user request is aggregated with an upper bound, which is the number of arguments under the selected position. Thus, the entire argument discovery framework is linear. The bottlenecks are the number of arguments under the selected position and the number of arguments indicators chosen by the user under consideration.

Since ICAS is a discussion platform, the number of arguments posted is high and keeps growing over time. A parallel implementation of the argument discovery framework could be performed to overcome this bottleneck because all arguments at the same level under the selected position are independent. The parallel pattern implementation is based on the divide and conquer approach. The divide step splits arguments under the chosen position into subsets based on the number of available processors. Each processor executes the argument discovery framework on its assigned subset and then returns the discovered arguments results. Finally, the conquer step combines each
of the results obtained from each processor to obtain the final results. The execution time of the
argument discovery framework will be reduced based on the number of multicore processors used.
In our setting, the highest number of arguments per position did not exceed 1500 arguments.
Therefore, the parallel implementation was not performed on our framework. However, the
parallel implementation of the argument discovery framework is essential as the number of
arguments grows drastically. The researchers expected the system to scale linearly with the
increasing number of arguments coming to the discussion using parallel implementation fashion.

4.7 Conclusion and Future Work

In this research, a novel framework is proposed for discovering constructive opinions in cyber-
argumentation platforms. This framework discovers opinions based on four non-textual features.
It identifies and measures the degree of recentness, engagement, controversy, and author influence
for each opinion. Then, it combines those indicators for discovering constructive opinions and
recommends them to users. Furthermore, it allows users to select and build their argument
discovery methods.

Unlike other cyber-argumentation platforms, arguments promoted by our framework are
meaningful because they are discovered using a list of non-textual features while paying equal
attention to each of the four essential features of opinions. Moreover, this framework provides
users with chances to specify their preferences. Thus, it encourages constructive discourse between
users that improves the quality of the discussion. Our framework has many potential applications
in the context of opinions search and discovery. It can be adapted and customized by many UGC
applications, such as promoting posts on social media, reviews in online retailers, and replies on
online news platforms.
This work focused on discovering opinions on the position and the argument level. However, the researchers did not perform the discovery method on the issue level. Issues may have different settings and measurements on the issue level. Alternatively, issues may exhibit other features set for issues discovery. This is left for future work.

4.8 REFERENCES


Chapter 5. Modeling and Analyzing Controversy in Large-Scale Cyber-Argumentation

5.1 Introduction

Discussions and debates can sometimes become heated when about controversial issues. A group of participants presents much greater intelligence than individuals in discussions. Controversy exists in in-person or online discussions [1 - 11, 13, 19 - 26], regardless of the medium. Controversy takes place in cyber-argumentation platforms too. Therefore, to make sense of group structures and interactions in large-scale argumentation, measuring and analyzing the controversial degree in online discussions is necessary. There has been some work done to measure the controversy in different contexts. For example, [1-8] have analyzed and quantified controversy in social media platforms, such as Facebook and Twitter. [1, 9-14] have analyzed and quantified controversy in discussion platforms such as Wikipedia, Reddit, and Slashdot. [15, 16] have addressed the controversy in deliberation and argumentation by mining disagreements and conflicts between participants. [17] estimated two types of controversy using movie ratings. In online discussions, people work together, cooperatively or collectively. They discuss the same issue from different perspectives. At the same time, most people choose a side for or against an issue. Cyber-argumentation platforms facilitate massive participation from users and structure discussions in an effective way. This leads to a better understanding of group interactions and benefits from collective intelligence. When people have different opinions, with different intensities, controversy or conflicts in people’s opinions exist. A much richer understanding can be obtained by identifying the disagreements and conflict in a discussion [15]. Measuring and quantifying controversy in cyber-argumentation adds more capabilities and functionalities of analyzing online discussions that allow constructive discourse between participants.
In this research, controversy within cyber-argumentation is defined based on the controversy definitions from different academic areas. The attributes of controversy are identified. Based on the definitions and the attributes, existing measurements are used to quantify the degree of controversy in cyber-argumentation. More findings and analytics are discussed.

5.2 BACKGROUND

Controversy phenomena have been studied for a long time in different domains. For example, in Cambridge Dictionary [18], controversy is defined as “a disagreement, often a public one, that involves different ideas or opinions about something.” However, there are many definitions of controversy in social science research. For instance, Johnson et al. [19] stated that controversy exists when ideas, information, conclusions, theories, or opinions are incompatible with another person's, and the two people seek to reach an agreement. Stradling et al. [20] identified controversial issues as “those problems and disputes that divide society and for which significant groups within society offer conflicting explanations and solutions based on alternative values.” Hess [21] defined the Controversial Public Issue as an unresolved public policy question that sparks significant disagreement. Harwood & Hahn [22] defined the discussion of a controversial issue as a reflective dialogue among students, or between students and teachers, about an issue on which there is disagreement. Malikow [23] expressed that controversy exists when a strong intellectual argument can be made for two or more conflicting positions. The issue involves two or more parties with equal and competing interests. Andersson [24] said, “controversy could be defined as a consistent conflict or a residual difference regarding an issue, such as financial situations, subject knowledge, religion, morals, etc.” In [19-24], these studies focused on defining controversial issues and when and how to teach them in classrooms, which is not the scope of this research. These definitions are useful for identifying the controversy problem and its attributes in
online discussions. Additionally, there are some existing controversy definitions in the computer science literature. For example, Sznajder et al. [25] stated that the definition in which concepts are controversial is itself controversial. Dori-Hacohen [1] defined controversial topics as those that generate strong disagreement among large groups of people. Zielinski et al.[9] said that controversy is composed of the object, a group of people, and their opinion about the given object. From the controversy definitions above, all definitions are consistent with that controversy exists when a disagreement happens between two or more parties. The parties’ opinions or ideas have more than two positions or two mutual solutions. Finally, the disagreement between parties is about an object, mostly the user’s opinions about an object. However, these definitions disagreed on the degree or the sensitivity of the disagreement and the existence of conflicts between parties. This is an important area to be researched and explored. However, in this research, defining controversy in online discussions and identify the attributes of controversy in online discussions. Finally, existing measures are used to estimate the controversial degree of online discussions. This approach helps determine the controversial degree of online discussions without ground-truth data to validate results.

5.3 Controversy Measures

5.3.1 Controversy Definition

Given the definitions of controversy in section 5.2, controversy can be identified as a disagreement between groups of people about an object. Therefore, the degree of disagreement can be measured through people’s opinions, and then the degree of controversy can be estimated. However, the agreement measures in the literature are more popular than the disagreement measures. The agreement measures can be used to estimate agreement or consensus in a population. And yet, they can also be used as disagreement measures. For example, Whitworth & Felton [27] measured
disagreement as an inverse measure of agreement. In addition, Akiyama et al. [28] viewed consensus as to the complement of disagreement. Finally, Tastle & Wierman [29] stated that the mirror image of consensus (agreement) is dissension (disagreement). Before estimating the controversial degree in cyber-argumentation as a disagreement or dissension problem, it is necessary to identify the important attributes that make up the argumentation controversy for extracting the crowd wisdom.

5.3.2 Controversy Attributes

Given the above-mentioned controversy definitions, the most controversial situation happens when participants are split equally among all possible opinions. Therefore, there is no consensus among the participants. In this case, users are at their maximum possible disagreement. Thus, the controversy degree is at its highest level. On the other hand, there is no controversy between users if all users have the same opinion. In this case, all users will wind up on the same pole, regardless of if users are for or against a position. If the standard deviation measure is taken for all user opinions, it will result in zero. The in-between case exists when users are in different poles with different opinions. Therefore, users have different in-group populations along with varying intensities of opinions. There are a theoretical minimum and maximum for estimating controversy in cyber-argumentation from all the cases mentioned above. Therefore, the controversy attributes are:

- A discussion about an object.
- The number of participants in a discussion.
- The participants’ opinions.

From the attributes mentioned above, the controversy degree can be estimated as distance-based measures and distribution-based measures.
5.3.3 Data Collection

The same data in section 2.2 is used in this research. However, this research is applied to the position level. Therefore, all arguments and positions under the position are considered in this research.

5.3.4 Estimating Controversy with Primitive Measures

Two primitive measures can be used to measure the controversial degree in cyber-argumentation discussions: variance and standard deviation. Variance and standard deviation are widely used to measure the dispersion of data across the range interval. Variance is the average squared deviation of values from the mean. It can measure the variance in user opinions in cyber-argumentation as:

\[ \sigma^2(P) = \frac{\sum_{i=1}^{n}(a_i - \mu)^2}{n} \]  

(5.1)

\(a_i\) is the averaged agreement value for user \(u_i\), \(\mu\) is the averaged agreement value for all users who participated in the discussion rooted by \(P\), and \(n\) is the number of participants in the discussion rooted by \(P\). If users’ opinions are very close to \(\mu\), then the variance will be small and vice versa.

Standard deviation is a well-known measurement of the amount of variation or dispersion of a dataset. It can measure the variance in user opinions in cyber-argumentation as:

\[ \sigma(P) = \sqrt{\frac{\sum_{i=1}^{n}(a_i - \mu)^2}{n}} \]  

(5.2)

\(a_i\) is the averaged agreement value for user \(u_i\), \(\mu\) is the averaged agreement value for all users who participated in the discussion rooted by \(P\), and \(n\) is the number of participants in the discussion rooted by \(P\). Like variance, a low standard deviation indicates that the users’ opinions in \(P\)’s discussion tend to be very similar to the mean \(\mu\); a high standard deviation indicates that the users’ opinions in \(P\)’s discussion are spread out across the agreement spectrum and far from the mean \(\mu\).
5.3.5 Estimating Controversy Using Distance-Based Measure

In ICAS, users can express their opinion with an agreement value. There are 11 agreement levels, which indicate different levels of agreement, as in section 2.1.1. However, a user opinion can have a different numerical value than the values mentioned above if the user’s argument value or reaction has reduced. To this point, each user will have a single averaged agreement value. This value can be used to measure the opinion distance between two users using Euclidean distance. This measure is inspired by [30]. If there are two users: \( u_i, u_j \) with two agreement values: \( a_i, a_j \), the distance between the users’ opinion can be calculated as:

\[
D(u_i, u_j) = \sqrt{ \frac{(a_i - a_j)^2}{d} } \quad (5.3)
\]

The distance between two users is normalized by dividing by \( d \). \( d \) is the maximum difference between opinions, which is equal to 2. Therefore, the distance result is bounded between 0 and 1. 0 indicates that both users have the same opinion, which indicates no controversy. On the other hand, 1 indicates full controversy, which means both users have opposite opinions. The next step is to identify the opinion distance between user \( u_i \) and \( n \) users under position \( P \). This distance is measured as follows:

\[
D(u_i, P) = \frac{1}{n-1} \sum_{j=1, j \neq i}^{n} \sqrt{ \frac{(a_i - a_j)^2}{d} } \quad (5.4)
\]

Equation (5.4) is a pair-wise comparison between two users' opinions. It reduces the effect of outliers and is normalized for the range of maximum disagreement on any construct and the number of constructs considered. Therefore, this measurement is not biased by a single outlier.

The final step is to identify the opinion distance between all users under position \( P \). This distance is measured as follows:
\[ D(P) = \frac{1}{n(n-1)} \sum_{i=1}^{n} \sum_{j=1, j \neq i}^{n} \sqrt{(a_i - a_j)^2} \]  \hspace{1cm} (5.5)

5.3.6 **Estimating Controversy Using Distribution-Based Measure**

Numerical values may not always express users’ opinions. Instead, in most discussion platforms, users express their opinions using categorical or Likert-scale options. For example, Facebook users use emoticons to interact with posts. These types of opinions still exhibit controversy definitions and attributes. The controversy degree can be estimated from the distributions’ perspective if disagreement and dissension measures are used. Therefore, the complement of group agreement or consensus measures is used to determine controversy in ICAS discussions. This section uses three different distribution-based measures to estimate controversy: the dissention measure, the soft and hard controversy measures.

5.3.6.1 The Dissention Measure

This measure is a Shannon entropy-based measure developed by [29]. Initially, the authors developed the consensus measure, which takes a probability distribution over a discrete set of choices and produces a single value in the range \([0, 1]\), where 0 indicates complete disagreement, and 1 indicates complete agreement. The dissension equation is calculated as:

\[ Dnt(P) = - \sum_{i=1}^{k} p_i \left( 1 - \frac{|a_i - \mu|}{d} \right) \]  \hspace{1cm} (5.6) by [29]

\( a_i \) is the averaged agreement value for user \( u_i \), \( \mu \) is the averaged agreement value for all users who participated in the discussion rooted by \( P \), \( k \) is the number of possible agreement values on the spectrum, \( d \) is the distance between the maximum and minimum possible choices in the range interval, and \( p_i \) is the probability (relative frequency) of the agreement value \( a_i \).
5.3.6.2 Soft and Hard Controversy

[17] has identified two kinds of controversy measurements. The first type is hard controversy \( H \), which occurs when the debate is polarized at the two extremes of the spectrum: a “love-hate movie.” This measure is often used more for polarization than controversy. This measure is a normalized standard deviation and is calculated as:

\[
H(P) = \frac{1}{C_H} \left[ \sum_{i=0}^{k} p_i(a_i - \mu)^2 \right]^{1/2}
\]  

(5.7) by [17]

\( a_i \) is the averaged agreement value for user \( u_i \), \( \mu \) is the averaged agreement value for all users who participated in the discussion rooted by \( P \), \( k \) is the number of possible agreement values in the spectrum, \( p_i \) is the probability (relative frequency) of the agreement value \( a_i \), and \( C_H \) is the highest possible variance. High \( H \) can indicate polarization or bimodality, even if the two peaks are not located at the boundaries of the distribution. This measure will be in \([0, 1]\), where 0 indicates no polarization, and 1 indicates perfect polarization.

The second type is the soft controversy \( S \), which occurs when users’ opinions distribute evenly across the spectrum. It can be calculated as:

\[
S(P) = 1 - \frac{1}{C_S} \left[ \sum_{i=0}^{k} (p_i - 0.1)^2 \right]^{1/2}
\]

(5.8) by [17]

\( k \) is the number of possible agreement values on the spectrum, and \( p_i \) is the probability (relative frequency) of the agreement value \( a_i \) and \( C_S = \sqrt{0.9} \). This measure will be in \([0, 1]\), where 0 indicates total agreement and 1 represents complete disagreement.
5.3.7 Estimating Controversy Using Agreement Measurements

Several indices have been developed and used to measure the agreement between participants in different contexts [31 - 33]. However, in this research, only three methods are considered: Finn’s Index of agreement is widely used for inter-rater agreement and reliability, $a_{wg}$ method is used as an improved agreement method over Finn’s Index, and Krippendorff’s Index is another multiple interrater agreement. These methods can assess the agreement between multiple raters with missing values.

5.3.7.1 Finn’s Index of Agreement

Finn’s Index is based on the assumption that when raters are in total agreement with each other, the observed variance will be 0. However, when the scoring is random, the variance will be larger. It can be calculated as:

$$FI(P) = 1 - \frac{S_0^2}{S_M^2}$$ (5.9) by [31]

$S_0^2$ is the observed variance in the participants’ agreement values, and $S_M^2$ is the maximum possible variance in the agreement values, which is equal to $\frac{A^2 - 1}{12}$, and $A$ is the number of agreement values or response options. This measure results in $[0, 1]$, where 1 indicates total agreement and 0 indicates total disagreement.

5.3.7.2 The $a_{wg}$ Method

Within-group index, $a_{wg}$, uses the mean of the ratings rather than the population-based ratings. Thus, it avoids the assumption that the data is uniformly distributed. It is calculated as:

$$a_{wg}(P) = 1 - \frac{2(S_0^2)}{[(H + L)M - (M^2) - (H \cdot L)]\left[\frac{J}{J-1}\right]}$$ (5.10) by [32]
$S_0^2$ is the observed variance in the participates’ agreement values, $H$ is the maximum possible agreement value, $L$ is the minimum possible agreement value, $J$ is the number of participants, and $M$ is the mean agreement for all the participants in $P$’s discussion. This measure will result in $[0, 1]$, where 1 indicates total agreement and 0 indicates total disagreement.

5.3.7.3 Krippendorff’s Alpha

Krippendorff’s Alpha is another measure of interrater agreement. It is designed to assess the agreement among multiple raters. Krippendorff’s Alpha can work with any number of participants and incomplete data. Those two conditions are the common scenarios in the discussions. This measure can be calculated as:

$$a_k(P) = 1 - \frac{D_o}{D_E}$$  \hspace{1cm} (5.11) \text{ by [33]}

$D_o$ is the observed level of disagreement among agreement values and $D_E$ is the level of disagreement among agreement values due to chance or error. The result of this measure will have different values: $\alpha = 1$ indicates perfect agreement, and $\alpha = 0$ indicates the absence of agreement (it is disagreement in this case), and $\alpha < 0$ is when disagreements exceed what can be expected by chance or error.

For all agreement measures, the mirror value can be used to estimate the disagreement values [27-29]. These disagreement values can be used as controversy estimates for online discussions.

5.4 DISCUSSION AND RESULTS

Table 5.1 shows the discussion attributes, such as number of posts (the total number of arguments and reactions), the normal distributions per posts, the number of distinct users joined the discussion, the normal distributions of users, number of poles, the maximum distance between poles and the overall agreement values for each position. These numbers vary significantly between positions. However, users who have participated in positions ranged from 238 to 289
users. The normal distribution of users is used to calculate whether poles exist or not. For each agreement level, if the number of users is greater than the normal distribution for users, then a pole is formed at this agreement level. Then, the distance between the furthest poles is calculated. The number of poles ranges from 3 to 6 poles, and the distance between poles ranges from 0.4 and 2; the maximum possible distance between poles is 2. For P0, P2, P7, and P11, the poles are in alignment. On the other hand, the poles in P10 are spread across the agreement value spectrum. The rest of the positions are in between the previous cases. The agreement value of each position is the normalized collective reduced agreement values of users’ posts.

Table 5.2 shows the different controversy values for all positions calculated by different controversy measures (5.1-5.11).

![FIGURE 5-2 USER DISTRIBUTION FOR P0, P1, P2 AND P3 UNDER ISSUE 1](image1.png)

![FIGURE 5-1 USER DISTRIBUTION FOR P4, P5, P6 AND P7 UNDER ISSUE 2](image2.png)
### TABLE 5-1 ISSUES, POSITIONS, AND USERS DETAILS

<table>
<thead>
<tr>
<th>Issues</th>
<th>Position #</th>
<th>Arguments #</th>
<th>Reactions #</th>
<th># Distinct Users</th>
<th>Normal Distribution of Post</th>
<th>Normal Distribution of Users</th>
<th># of Poles</th>
<th>Max Distance between Poles</th>
<th>Overall Collective Agreement Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Issue 1</td>
<td>P0</td>
<td>1017</td>
<td>726</td>
<td>289</td>
<td>159.5455</td>
<td>26.27273</td>
<td>3</td>
<td>0.8</td>
<td>-0.289</td>
</tr>
<tr>
<td></td>
<td>P1</td>
<td>576</td>
<td>330</td>
<td>263</td>
<td>92.27273</td>
<td>24.09091</td>
<td>5</td>
<td>1.2</td>
<td>0.041576</td>
</tr>
<tr>
<td></td>
<td>P2</td>
<td>705</td>
<td>409</td>
<td>279</td>
<td>102.9091</td>
<td>25.36364</td>
<td>3</td>
<td>0.4</td>
<td>0.453004</td>
</tr>
<tr>
<td></td>
<td>P3</td>
<td>591</td>
<td>347</td>
<td>259</td>
<td>82</td>
<td>23.54545</td>
<td>4</td>
<td>1</td>
<td>-0.23836</td>
</tr>
<tr>
<td>Issue 2</td>
<td>P4</td>
<td>782</td>
<td>259</td>
<td>269</td>
<td>87.81818</td>
<td>24.36364</td>
<td>5</td>
<td>1.4</td>
<td>-0.14203</td>
</tr>
<tr>
<td></td>
<td>P5</td>
<td>593</td>
<td>203</td>
<td>252</td>
<td>68.27273</td>
<td>22.90909</td>
<td>6</td>
<td>1</td>
<td>0.235686</td>
</tr>
<tr>
<td></td>
<td>P6</td>
<td>533</td>
<td>193</td>
<td>249</td>
<td>61.90909</td>
<td>22.18182</td>
<td>6</td>
<td>1.6</td>
<td>-0.00793</td>
</tr>
<tr>
<td></td>
<td>P7</td>
<td>620</td>
<td>229</td>
<td>253</td>
<td>68.45455</td>
<td>23.09091</td>
<td>4</td>
<td>0.6</td>
<td>0.491102</td>
</tr>
<tr>
<td>Issue 3</td>
<td>P8</td>
<td>883</td>
<td>296</td>
<td>274</td>
<td>95.45455</td>
<td>24.90909</td>
<td>4</td>
<td>1.6</td>
<td>0.16381</td>
</tr>
<tr>
<td></td>
<td>P9</td>
<td>593</td>
<td>238</td>
<td>255</td>
<td>72</td>
<td>23.18182</td>
<td>4</td>
<td>1.2</td>
<td>0.140152</td>
</tr>
<tr>
<td></td>
<td>P10</td>
<td>581</td>
<td>243</td>
<td>257</td>
<td>72.27273</td>
<td>23.45455</td>
<td>5</td>
<td>2</td>
<td>0.149686</td>
</tr>
<tr>
<td></td>
<td>P11</td>
<td>636</td>
<td>226</td>
<td>253</td>
<td>69.36364</td>
<td>23.09091</td>
<td>3</td>
<td>0.4</td>
<td>-0.37641</td>
</tr>
<tr>
<td>Issue 4</td>
<td>P12</td>
<td>747</td>
<td>202</td>
<td>258</td>
<td>76.45455</td>
<td>23.54545</td>
<td>5</td>
<td>1</td>
<td>0.102259</td>
</tr>
<tr>
<td></td>
<td>P13</td>
<td>623</td>
<td>215</td>
<td>252</td>
<td>84.36364</td>
<td>22.90909</td>
<td>6</td>
<td>1.6</td>
<td>0.226293</td>
</tr>
<tr>
<td></td>
<td>P14</td>
<td>547</td>
<td>151</td>
<td>244</td>
<td>58.54545</td>
<td>22.27273</td>
<td>6</td>
<td>1.4</td>
<td>-0.20155</td>
</tr>
<tr>
<td></td>
<td>P15</td>
<td>556</td>
<td>157</td>
<td>238</td>
<td>57.45455</td>
<td>21.63636</td>
<td>5</td>
<td>1</td>
<td>0.008861</td>
</tr>
<tr>
<td>Position #</td>
<td>$\sigma^2(P)$</td>
<td>$\sigma(P)$</td>
<td>Dnt($P$)</td>
<td>S($P$)</td>
<td>H($P$)</td>
<td>D($P$)</td>
<td>FI($P$)</td>
<td>$a_{wg}(P)$</td>
<td>$\alpha_k(P)$</td>
</tr>
<tr>
<td>------------</td>
<td>----------------</td>
<td>-------------</td>
<td>---------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>-------------</td>
<td>-------------</td>
</tr>
<tr>
<td>P0</td>
<td>0.228093</td>
<td>0.680668</td>
<td>0.554738</td>
<td>0.764483</td>
<td>0.123287</td>
<td>0.728487</td>
<td>0.022487</td>
<td>0.511621</td>
<td>-0.00347</td>
</tr>
<tr>
<td>P1</td>
<td>0.238672</td>
<td>0.605898</td>
<td>0.466246</td>
<td>0.900971</td>
<td>0.033236</td>
<td>0.468766</td>
<td>0.024226</td>
<td>0.483502</td>
<td>-0.00379</td>
</tr>
<tr>
<td>P2</td>
<td>0.263939</td>
<td>0.64575</td>
<td>0.508252</td>
<td>0.664185</td>
<td>0.143376</td>
<td>0.268642</td>
<td>0.035065</td>
<td>0.698105</td>
<td>-0.0036</td>
</tr>
<tr>
<td>P3</td>
<td>0.258221</td>
<td>0.585955</td>
<td>0.450379</td>
<td>0.849046</td>
<td>0.03929</td>
<td>0.320268</td>
<td>0.027132</td>
<td>0.562477</td>
<td>-0.00388</td>
</tr>
<tr>
<td>P4</td>
<td>0.23598</td>
<td>0.58474</td>
<td>0.44456</td>
<td>0.859842</td>
<td>0.034855</td>
<td>0.270999</td>
<td>0.032967</td>
<td>0.5055</td>
<td>-0.00375</td>
</tr>
<tr>
<td>P5</td>
<td>0.204719</td>
<td>0.558383</td>
<td>0.413863</td>
<td>0.844203</td>
<td>0.036181</td>
<td>0.275362</td>
<td>0.026530</td>
<td>0.44833</td>
<td>-0.00398</td>
</tr>
<tr>
<td>P6</td>
<td>0.278917</td>
<td>0.595766</td>
<td>0.457229</td>
<td>0.878648</td>
<td>0.033298</td>
<td>0.293191</td>
<td>0.022218</td>
<td>0.572054</td>
<td>-0.00412</td>
</tr>
<tr>
<td>P7</td>
<td>0.205904</td>
<td>0.549988</td>
<td>0.406315</td>
<td>0.694234</td>
<td>0.065437</td>
<td>0.293723</td>
<td>0.021995</td>
<td>0.564656</td>
<td>-0.00395</td>
</tr>
<tr>
<td>P8</td>
<td>0.350837</td>
<td>0.664262</td>
<td>0.536066</td>
<td>0.856457</td>
<td>0.044463</td>
<td>0.318553</td>
<td>0.021069</td>
<td>0.727424</td>
<td>-0.00366</td>
</tr>
<tr>
<td>P9</td>
<td>0.264844</td>
<td>0.575175</td>
<td>0.435885</td>
<td>0.860156</td>
<td>0.032981</td>
<td>0.333486</td>
<td>0.031646</td>
<td>0.54637</td>
<td>-0.00394</td>
</tr>
<tr>
<td>P10</td>
<td>0.325281</td>
<td>0.646155</td>
<td>0.512324</td>
<td>0.871512</td>
<td>0.04056</td>
<td>0.310409</td>
<td>0.026731</td>
<td>0.663248</td>
<td>-0.00389</td>
</tr>
<tr>
<td>P11</td>
<td>0.265159</td>
<td>0.603984</td>
<td>0.463362</td>
<td>0.741847</td>
<td>0.054042</td>
<td>0.34617</td>
<td>0.025827</td>
<td>0.645699</td>
<td>-0.00395</td>
</tr>
<tr>
<td>P12</td>
<td>0.223941</td>
<td>0.591292</td>
<td>0.451567</td>
<td>0.85748</td>
<td>0.033747</td>
<td>0.244658</td>
<td>0.024648</td>
<td>0.447453</td>
<td>-0.00388</td>
</tr>
<tr>
<td>P13</td>
<td>0.216527</td>
<td>0.562471</td>
<td>0.417062</td>
<td>0.84354</td>
<td>0.036175</td>
<td>0.282604</td>
<td>0.021099</td>
<td>0.460247</td>
<td>-0.00398</td>
</tr>
<tr>
<td>P14</td>
<td>0.209687</td>
<td>0.540777</td>
<td>0.401721</td>
<td>0.812892</td>
<td>0.035687</td>
<td>0.335611</td>
<td>0.028713</td>
<td>0.435563</td>
<td>-0.0041</td>
</tr>
<tr>
<td>P15</td>
<td>0.312642</td>
<td>0.636291</td>
<td>0.502012</td>
<td>0.882989</td>
<td>0.036938</td>
<td>0.352526</td>
<td>0.021123</td>
<td>0.630697</td>
<td>-0.00422</td>
</tr>
</tbody>
</table>
For Issue 1, in Figure 5-1, P0 is rejected by most of the participants with different intensities. The number of participants mostly decreases as the agreement value increases. The poles formed by users have a maximum distance of 0.8; therefore, they are not in alignment. However, all poles are placed on the agreement value spectrum's negative side. Thus, the overall collective agreement value of this position is -0.289. Because the number of posts is high and the number of poles is less than the other positions under this issue, this position reports the highest degree of the $Dnt(p)$, $H(p)$, and $D(p)$ measures, as in Table 5-2. As for P1, participants do not distribute uniformly across the agreement value spectrum. However, users form poles around the center of the spectrum. Participants have low intestines when supporting or opposing this position. P1 has more poles than the other positions; however, the number of users in each pole is less than the number of users in a pole of the other positions. Since the poles formed in P1 are closer to each other and centered around the middle, this position reported the highest $S(p)$ and the lowest $H(p)$ measures of all positions in this study. The overall collective agreement value of P1 is 0.041576. P2 received a great deal of support from users and very little opposition. About one-forth of the participants agree completely with this position. Indeed, the number of participants increases as the agreement value increases. Thus, it has an overall agreement value of 0.453004. Because the poles in this position are in alignment and the majority of the participants agreed on this position, this position reported the highest $H(p)$ and $Fl(p)$ and lowest in $S(p)$ and $D(p)$ measures. In this position, participants are more polarized in their opinion than in the other positions. Finally, P3 is rejected with a great deal of opposition with little support. However, most participants are unsure about their opinion (agreement value > -0.25 and < +0.25). The poles formed by participants are not in alignment and have a distance of 0.8. This position scored in the middle of all controversy measures for all the positions in this study. The overall collective agreement value is -0.23836.
For Issue 2, in Figure 5-2, P4 received almost an equal amount of supports and attacks by participants. However, most participants were unsure about their opinion (agreement value > -0.5 and < +0.5). The poles formed by participants are not in alignment and have a distance of 1.4. Most of these poles are formed at the center of the agreement value spectrum. As a result, this position scored low in the $D(p)$ measure because this distance between participants' opinions is low and scored high in the $Fl(p)$ measure. All scores show that most of the participants are being neutral about P4, which indicates some polarization. Thus, the overall collective agreement value of this position is -0.14203. As for P5, participants do not distribute uniformly across the agreement value spectrum. This position received more support than the opposition by participants. However, users form poles around the positive side of the agreement value spectrum. The majority of the participants are supporting this position with low intensity (agreement value > 0 and < +0.50). It has a larger number of poles than P4; however, the poles are parallel. Therefore, this position reported the low scores in the $Dnt(p)$, $D(p)$, and $a_{wg}(p)$ measures because the majority of the users are in agreement with supporting this position with different intensities. Thus, the overall collective agreement value of this position is 0.235686. About P6, participants do not distribute uniformly across the agreement value spectrum; however, users form poles around the center of the spectrum. Participants have low intensities when supporting or opposing this position. P6 has a larger number of poles than P4 and P7. However, the number of participants in each pole in P6 is less than the number of participants in any pole in P4 and P7. Since the poles formed in P6 were more in alignment and centered around the middle, this position reported the highest $S(p)$ measure and the lowest in the $H(p)$ measure. The overall collective agreement value of this position is -0.00793. Finally, P7 is supported by users and very little opposition. About one-fourth of the participants agree completely with this position. The number of participants in this position
increases as the agreement value increases. Thus, it has an overall agreement value of 0.491102. Because the poles in this position are in alignment and the majority of the participants agree on this position, this position reported high in \( H(p) \) and low in \( S(p) \) and \( D(p) \). In this position, participants are more polarized in their opinion than the other positions under Issue 2.

For Issue 3, as in Figure 5-3, P8 is supported by a moderate deal of support from users and little opposition. However, many participants were unsure about their opinion (agreement value > -0.25 and < +0.25). Thus, it has an overall agreement value of 0.16381. The poles in this position are not in alignment; the distance between the furthest poles is 1.4. Therefore, participants are not on the same page in this position. For this reason, this position reported high in \( Dnt(p) \) and low in \( FI(p) \) measures. However, P8 scored high in the \( awg(p) \) because most participants support this position with different intensities. Therefore, participants are not in agreement with this position.

As for P9, participants do not distribute uniformly across the agreement value spectrum and users form poles toward the right side of the agreement values spectrum. Participants have low intensities when supporting or opposing this position. However, the majority of participants were unsure about this position. It has more poles than others under the same issue. This position reported the high in the \( H(p) \) and the \( FI(p) \) measures because users are polarized around the center of the agreement value spectrum, indicating some agreement level. The overall collective agreement value of this position is 0.140152. P10 follows P8’s pattern. P10 is supported by a

![Figure 5-3 User Distribution for P8, P9, P10 and P11 Under Issue 3](image)

**FIGURE 5-3 USER DISTRIBUTION FOR P8, P9, P10 AND P11 UNDER ISSUE 3**
moderate deal of support from users and little opposition. The number of participants mostly increases as the agreement value increases. Thus, it has an overall agreement value of 0.149686. The poles in this position are not in alignment; the distance between the furthest poles is 2, which is equal to the maximum possible distance between any poles. Therefore, the participants’ opinions vary in this position. For this reason, this position reported the high in $D_{nt}(p)$ measure. Nevertheless, P10 scored high in $a_{wg}(p)$ measure due to the majority of the participants are supporting this position. Finally, P11 is rejected by most of the participants with different intensities. The number of participants mostly decreases as the agreement value increases. The poles formed are in alignment, and participants seemed to be polarized around the negative side of the agreement value spectrum but not significantly. For this reason, this position scored high in the $a_{wg}(p)$, $H(p)$, and $D(p)$ measures and low in the $S(p)$ measures, as in Table 5-2. This position has an overall collective agreement value of -0.37641.

For Issue 4, as in Figure 5-4, P12 is supported by a moderate deal of support from participants and little opposition. However, most participants are unsure about their opinion (agreement value > -0.25 and < +0.25). Thus, it has an overall agreement value of 0.102259. The poles in this position are in alignment; the distance between the furthest poles is 1. For this reason, this position reported the low in $D(p)$ measure, which indicates fewer differences in opinions in this discussion than in the other discussions. P13 is supported by a moderate deal of support from users and little opposition. The number of participants mostly increases around the center of the agreement value spectrum. Since many participants are unsure about their opinion, it has an overall agreement value of 0.226293. The poles in this position are not in alignment; the distance between the furthest poles is 1.6. Therefore, the participants’ opinions vary in this position. This position scored low in $D(p)$ and $FI(p)$ because most participants are unsure or agreed on this position. This position scored low
in all measures relative to the other positions. P14 follows P12’s pattern; it is supported by a moderate deal of support from participants and little opposition. However, most participants are unsure about their opinion (agreement value > -0.5 and < +0.5). Thus, it has an overall agreement value of -0.20155. The poles in this position are in alignment; the distance between the furthest poles is 1.6. Thus, this position reported the low in Dnt(p) because most users were unsure about this position. It also scored low in $a_{wg}(p)$ because users’ opinions have different supporting or opposing discussions intensively. Finally, P15 is supported by a great deal of support from users and little opposition. However, most participants are unsure about their opinion. Thus, it has an overall agreement value of 0.008861. Because the poles in this position are in alignment and the majority of the participants are unsure about this position, this position reported a high in S(p) and D(p) and very low in FI(p) measures. Participants seemed to have different opinions about this position.

In this work, several controversy measures are used and proposed. The Pearson correlation is used to assess these measures against the discussion attributes, shown in Table 5.3. Pearson correlation is a statistic that measures the linear correlation between two variables [34].

The Euclidian distance method, $D(P)$ as a distance-based measure, shows a strong correlation with the number of posts and a moderate correlation with users' numbers across positions. However, it
has a weak negative correlation with the number of poles and the distance between the poles. Therefore, the distance between user’s opinions decreases as the number of poles increases.

TABLE 5-3 CORRELATIONS BETWEEN DISCUSSION ATTRIBUTES AND CONTROVERSY MEASURES

<table>
<thead>
<tr>
<th></th>
<th>Dnt(P)</th>
<th>S(P)</th>
<th>H(P)</th>
<th>D(P)</th>
<th>FI(P)</th>
<th>awg(P)</th>
<th>ak(P)</th>
</tr>
</thead>
<tbody>
<tr>
<td># Posts</td>
<td>0.6241</td>
<td>-0.2803</td>
<td>0.6948</td>
<td>0.7544</td>
<td>0.1108</td>
<td>-0.0324</td>
<td>0.8877</td>
</tr>
<tr>
<td># Participants</td>
<td>0.6146</td>
<td>-0.3584</td>
<td>0.6967</td>
<td>0.5047</td>
<td>-0.0055</td>
<td>-0.2095</td>
<td>0.9981</td>
</tr>
<tr>
<td># Poles</td>
<td>-0.5365</td>
<td>0.6314</td>
<td>-0.6870</td>
<td>-0.3843</td>
<td>0.1767</td>
<td>0.5599</td>
<td>-0.6319</td>
</tr>
<tr>
<td>Max Distance between Poles</td>
<td>-0.1238</td>
<td>0.7588</td>
<td>-0.6892</td>
<td>-0.3975</td>
<td>-0.2370</td>
<td>0.1851</td>
<td>-0.3009</td>
</tr>
<tr>
<td>Overall Agreement Value</td>
<td>-0.1353</td>
<td>-0.2407</td>
<td>0.1615</td>
<td>-0.4501</td>
<td>-0.0177</td>
<td>-0.1686</td>
<td>0.0139</td>
</tr>
<tr>
<td>Variance</td>
<td>0.6703</td>
<td>0.2822</td>
<td>-0.0666</td>
<td>-0.0732</td>
<td>-0.9964</td>
<td>-0.8326</td>
<td>0.0280</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>0.9982</td>
<td>-0.0523</td>
<td>0.5231</td>
<td>0.5014</td>
<td>-0.6342</td>
<td>-0.6443</td>
<td>0.5876</td>
</tr>
</tbody>
</table>

The other distribution-based controversy measures have different settings. For example, the Dnt(P) measure from (5.6) strongly correlates with the number of posts and the number of users. It also has a moderate negative correlation with the number of poles. However, this measure has a very strong correlation with the standard deviation measure. On the other hand, the S(P) measure, from (5.8), has a weak negative correlation with the number of posts and users and a strong correlation with the number of poles and the distance between the furthest poles. This measure is sensitive to opinion distributions. If all posts appear poles that are spread across the agreement spectrum, as in
P0, P2, P7, and P11, the controversy score will decrease significantly. If the posts are equally distributed on the all agreement value’s spectrum, the controversy score will increase. Unlike the $S(P)$, the $H(P)$ measure (5.7) reports the opposite correlations with discussion attributes. This measure is strongly correlated with the number of users and posts, but it has a strong negative correlation with the number of poles and the distance between them. This measure can capture the so-called “support-attack” position. Therefore, when a position receives a high number of supports and a high number of attacks simultaneously, this measure's score increases. This explains the high correlation between this measure and the standard deviation of polarized opinions. This measure is more suitable as a polarization measure [17] than as a controversy measure.

The agreement measures are used to estimate the controversy in the ICAS discussions and mirrored (disagreement $= 1 –$ agreement) [27-29] to calculate the disagreement level between participants. The $FI(P)$, calculated by (5.9), reported less than 1% for all positions, as in Table 5.2. However, Finn’s index assumes that the variance in users’ opinions or agreement values follows a uniform distribution, which is unlikely in practice [35]. Despite the differences in all the discussion attributes, this index reported similar scores [0.21-0.35] for all positions and has a strong negative correlation with the standard deviation and variance measures. It reports weak or no correlations with the other discussion attributes. The $awg(P)$ measure, calculated as in (5.10), reports different scores for all discussions. It has a robust negative correlation with the standard deviation and variance measures and has a moderate correlation with the number of poles formed. Finally, the $a_{sd}(P)$ measure, as in (5.11), has reported values less than zero, which indicates that disagreement in ICAS discussions exceeds what can be expected by chance or error. It has different kinds of correlations with the discussion attributes. It is strongly correlated with the number of posts and users. Also, it has a strong negative correlation with the number of poles and a weak negative
correlation with the maximum distance between poles. Surprisingly, this measure does not correlate with the variance measure and has a moderate correlation with the standard deviation measure, even if the standard deviation and variance measures are highly correlated.

In this study, several existing measures from different contexts are used to quantify the controversial discussions. Each measure is a package of strengths and weaknesses. These measures are easy to adapt and calculate if the data is preprocessed in the right format. Unlike other controversy measures, the measures used in this research have been widely used in literature in different contexts and can easily be adapted to multiple platforms. These measures can be used to compare the controversial degree for discussions, especially with the absence of ground-truth data for validation. Moreover, they do not require human-in-the-loop or validation to estimate the controversial degree of online discussions. The right measure will depend on platform attributes and needs. For example, like ICAS, the distance-based measure can be used because a numerical value expresses user opinions. However, in a particular setting, such as reactions on Facebook, distribution-based measures can be used to estimate the controversial degree of online discussions.

5.5 CONCLUSION AND FUTURE WORK

Online discussions are full of kinds of opinions that could be similar, different, or conflicting. People exchange opinions and sometimes get excited about things from different perspectives. They also reveal the controversy degree between participants in a discussion. Controversy is an important concept to analyze when conducting online argumentation and deliberation. In this study, controversy has been redefined in the cyber-argumentation field. The controversy definitions and attributes are used to measure the controversial degree of online discussions using existing measures. Results show different correlations between controversy estimates and
discussion attributes using pre-existing measures. These measures do not require ground-truth data or specific settings and can be adapted to distribution-based or distance-based opinions.

This work focused on measuring controversy at the position level. However, many questions still need to be answered. For example, will the controversy measure for the issue-level exhibit different definitions, attributes, and settings? Measuring the controversial degree on the issue level would allow for the exploration of the different types of participant attitudes. The other questions are: What is the relationship between polarization and controversy in online discussions? What can be done to increase or decrease the controversy in online discussions? These questions are left for future work.

5.6 References


84


Chapter 6. DISCOVERING TOPIC-ORIENTED FOCAL SETS IN CYBER-ARGUMENTATION

USING LINK ANALYSIS, TOPIC MODELING AND SOCIAL ROLES

6.1 INTRODUCTION

According to Nielsen [1], users’ participation in online social networks is unequal. In fact, in all large-scale user-generated content (UGC) platforms, user participation follows the 1-9-90 rule. Whereas 1% of users account for most of the content, 9% of users contribute from time to time, and 90% of users are lurkers who do not participate much in discussions. Therefore, there is a need to investigate users: who is talking, what they are talking about and how they are connected to extract new knowledge from UGC platforms? Users do not work alone; they are connected with others directly, such as social connections, or indirectly, such as user interaction connections like a “reply-to” or “like” on social media platforms. Besides, users may work with each other cooperatively or collectively to support, attack, or deliver agendas due to similar interests. These connections represent some similarities that can be used for community detection or group identification. These types of connections are studied extensively in research to detect communities or hidden communities with similar information or interests.

Community or hidden community detection can be performed using link analysis techniques, such as community detections or graph algorithms [2 - 11]; text analysis techniques, such as topic modeling [12, 13] or sentiment analysis [14]; or other data mining techniques, such as clustering, classification deep learning [15], using game-theoretic modeling [16] or using node attributes and edge structure [17]. Some studies have combined different techniques to discover or detect communities or hidden communities with certain similarities in social media and blogs. However, most of the work that has been conducted focused on improving the community or hidden community detection and structure concerning the research challenges. Furthermore, few
researchers have tried to find representative users in the detected communities to analyze these users. To our knowledge, minimal work has been done on cyber-argumentation platforms. Mostly, community detection has been studied extensively in social media platforms such as Facebook and Twitter.

In this research, a new framework of existing work is proposed to discover topic-oriented focal sets using the Focal Structure Analysis Algorithm (FSA) [18] and the topic modeling in the Intelligent Cyber-Argumentation System platform (ICAS). First, topic-modeling techniques are performed at the issue-level to identify most topics discussed by users under the parent issue. Second, the users’ interaction information is used to apply the FSA algorithm to identify sub-communities and focal structures. A pairwise similarity is then performed to measure the similarities between identified topics and community topics to find the users and focal sets behind specific issues. Finally, further experiments are conducted to study users’ roles and the intensity of their opinion in online dissuasion. This work adds more significance to cyber-argumentation platforms by discovering which communities or focal structures are behind specific topics and discovering users’ roles and opinion intensity.

6.2 RELATED WORK

6.2.1 COMMUNITY DETECTION

Community detection [11], influential nodes [19, 20], and topic modeling [20] are well-studied topics in academia. In the last few years, researchers were focused on identifying hidden communities using different research methodologies [2, 3, 5, 7, 8, 10, 14]. Tang et al. [2] used a novel integration scheme based on structural features to improve conventional community detection methods from one-dimensional to multi-dimensional networks. Hajdu et al. [3] discovered the passenger communities and most frequent trips in the transfer network using graph
information and the community detection algorithm. Wang et al. [5] proposed community kernel detection to uncover the hidden community structure in large social networks and discovered influential users. He et al. [8] presented a new approach: Hlidden C0mmunity Detection (HICODE), which identifies hidden communities and dominant community structures by weakening the strength of the dominant structure to uncover the hidden community structure beneath. Prem and Blei [21] used a Bayesian model of networks that allows communities to overlap. A corresponding algorithm naturally interleaves subsampling from the network and updating an estimate of its communities in massive networks. Peng et al. [7] developed an unsupervised learning method to discover implicit communities hidden in tweet datasets. Fortunato [6] has studied well community detection algorithms in graphs. Fortunato and Barthelemy [11] found that modularity optimization might fail to identify modules with smaller communities compared to the massive network size. Lin et al. [22] used a distant-based modularity method for community detection in incomplete network information to discover hierarchical and overlapped communities. Behera et al. [23] used a parallel programming framework to reduce running time for uncovering the hidden communities in a social network.

All the researches mentioned above have focused on link-based community or hidden community detection. However, some work has been done through community detection using additional preprocessing or additional techniques such as text analysis. For example, Fu et al. [14] used topic identification to identify the target participants, then applied sentiment analysis and opinion mining for users with similar topic interests. Finally, they applied multi-level community detection to find communities constructed by the users who have a consistent opinion. Zhao et al. [13] addressed the semantics problem shared by people in community detection by introducing topical clustering as an additional step to strengthen or weaken the community connections. Dang and Nguyen [24]
proposed a new approach for topic modeling using community findings in dynamic networks. They used topic modeling to refine the network based on the topics to reveal the structures and communities in dynamic social networks. Abdelbary [15] proposed multi-layer community detection by applying Gaussian Restricted Boltzmann Machine on users’ posts to identify their topics of interest and then construct communities.

Influential nodes are significant in different contexts. Two well-known algorithms are used for influential nodes in graphs. The PageRank Algorithm [25] simply counts the number and the weight of the links to a node in a graph. The underlying assumption of the PageRank algorithm is: the more critical the node in the graph, the more likely to receive more links over time. On the other hand, the HITS Algorithm [26] assigns two scores for each node: authority, which estimates the value of the node's content, and hub, which estimates the value of the links of that node to the other nodes. Both algorithms have been used, optimized, and improved in many contexts. For example, Chen et al. [27] proposed a new metric to identify influential nodes in a network by trading-off between the low-relevant degree centrality and other time-consuming measures. Kempe et al. [28] implemented the Decreasing Cascade Model to choose the active set of nodes for behavior spreading. Wang et al. [29] proposed a new algorithm that detects communities in social networks with the consideration of information diffusion and a dynamic programming algorithm for selecting communities to find influential nodes.

The FSA algorithm developed by Sen et al. [18] detects disease-release structure in a network context. It uses the Protein-Protein Interaction (PPI) networks to identify smaller and more relevant focal structures instead of identifying large clusters or communities by applying the modularity algorithm [30] recursively. This algorithm has been used to determine the focal sets in organizing mass protests on social media [31]. Unlike the traditional influential finding algorithms, this
algorithm was able to identify a set of influential nodes in a network that forms a compelling power. In this research, a community as a group or focal set is used to reference a set of individuals who shared links in the argumentation graph.

### 6.2.2 Topic Modeling

Topic modeling and sentiment analysis attract researchers’ attention due to the massive text generated by UGC, which helps to extract crowd wisdom. LDA [32], Latent Dirichlet Allocation, is a well-known algorithm in the topic modeling field. It is designed as a multi-level Bayesian to model items as a collection of a finite mixture over an underlying set of topics. Xu et al. [33] used non-negative matrix factorization for document clustering in a given corpus. These algorithms are used and improved enormously in academia to fulfill research needs. For example, Jelodar et al. [34] reviewed scholarly articles published from 2003 to 2016 related to LDA-based topic modeling to discover the research development, current trends, and intellectual structure of topic modeling. Debortoli et al. [35] combined LDA with a logistic regression model to explain user satisfaction with an IT artifact by analyzing more than 12,000 online customer reviews.

### 6.2.3 Social Roles in UGC Platforms

Users in UGC platforms hold different roles; users may switch roles based on their contributions. Different researchers have studied and investigated the social roles of users in various domains. For example, Stuetzer et al. [39] have studied brokering behavior in online learning communities by examining role patterns and information between learners and educators using social network analysis. Chan et al. [40] used nine different role features to profile the user roles in discussion forums. Then, they used two-stage clustering to describe the forums based on their role composition. White et al. [38] used a mixed membership formulation to cluster users with similar egocentric network structures based on the profile models' network statistics. Welser et al. [39]
used editing patterns and egocentric network visualizations to develop “structural signatures” as quantitative indicators of role adoption. In another similar domain, Wikipedia talk pages are community-oriented pages. Gleave et al. [40] have standardized the “social role” in online communities as a blend of psychological, social structural and behavioral properties. They measured and analyzed strategies for identifying social roles in Wikipedia and Usenet. Fisher et al. [41] have used social network analysis or SNA, to characterize authors in Usenet newsgroups. They found that second-degree egocentric networks provide obvious differences between different types of authors and newsgroups. On another platform, Reddit.com, Buntain and Golbeck [42] confirmed the existence of the “Answer-Person” role in Reddit and provided an automated method for identifying this role based on user interaction. Mantzaris and Higham [43] have proposed a new model that quantifies nodes' communication data in social networks, known as dynamic communicators, using standard centrality measures.

Social roles have been studied in the knowledge management domain. For example, Davidson et al. [44] have examined different roles in online communities and developed Reader to Leader framework to utilize online forum users' role evolution. Cranefield et al. [45] have studied lurking behavior in online communities. They found four different key roles in transferring knowledge in two sets of activities: monitoring the knowledge agenda and monitoring or being monitored. Akar et.al [46] identified user roles in an online community using structural role theory, SNA, and community members’ contribution behavior. Finally, social roles have been used in Enterprise Social Networks (ESN). Hacker et al. [47] have determined knowledge actions and knowledge worker roles to characterize ESN user behavior.
6.3 Proposed Framework

Discussions in cyber-argumentation contain more than text and opinions. This proposed framework uses the FSA algorithm to discover topic-oriented hidden communities under a selected issue.

![Diagram of Discussion Framework]

**Figure 6-1 Discovering Topic-Oriented Individuals & Focal Sets**

This framework is divided into key modules, as shown in Figure 6-1. First, the data is collected from the ICAS platform. Then, topic modeling is performed on the text collected and focal set identifications on the participants using link analysis. Subsequently, a pairwise similarity is conducted between the identified topics, individuals, and focal sets’ posts. Finally, the participants' social role and opinion intensity are used to analyze the discussion dynamic in ICAS discussions. In the next subsections, more information about the technical details of the framework.

6.3.1 Data Collection

The same data in Section 2.2 is used. However, this framework is applied on the issue level. Therefore, all arguments and positions under the issue are considered in this research.
6.3.2 User Opinion Vector (UOV)

In ICAS, an issue, as in Figure 2-1, consists of positions and arguments at different tree levels. To derive the user opinion vector, all arguments and reactions made at any tree level need to be a direct child of its position parent. To do so, the fuzzy reduction engine from section 2.1.1 is used. Then, all the reduced agreement values from all arguments and reactions per position for the user are averaged as in formula 2.1. If there are j positions under issue I, then the same process for all positions under the selected tree is applied. In case the user has not participated in one or more positions, the missing values are imputed with zeros. Finally, the user vector opinion is derived as

\[ UOV(u, I) = (AV_{p1}^u, AV_{p2}^u, AV_{p3}^u, \ldots, AV_{pj}^u) \] (6.1)

6.3.3 Focal Sets and Participants Partitioning Using Link Analysis

In ICAS, it is assumed that participants do not have any explicit relationships or social connections during the discussion. It is also assumed that users work with each other unintentionally or collectively. With these assumptions, some users share similar interests and behaviors. The FSA algorithm [18] is applied to ICAS discussions to discover sub-communities and focal structures in this section. However, the issue tree, as shown in Figure 2-1 represents the discussion in ICAS. Therefore, a graph transformation is needed to apply the FSA algorithms. The user interaction graph (UIG) is created from the discussion tree, as in section 4.4.2. The resulting graph consists of nodes representing users in the discussion and edges representing reply-to and react-to relationships between users. As the graph is created, the FSA algorithm is applied to partition the UIG into smaller graphs. As a result, different communities are formed at different levels of the UIG.
6.3.4 Focal Sets Opinion Vector

This section studies the users’ opinions within the focal set. Users who shared the same topic do not necessarily have to share the same opinion about the topic. Therefore, it needs more processing to investigate users’ opinions within the same focal set. Two concepts are introduced for this purpose: Average User Opinion and Average Focal Set Opinion—details in the next two subsections.

6.3.4.1 Average User Opinion (AUO)

On the issue level, to determine the average users' opinions, users’ opinions for all positions under the issue need to be considered as follow:

\[
AUO(I) = \left( \frac{\sum_{i=1}^{n} AV_{i p_1}}{n}, \frac{\sum_{i=1}^{n} AV_{i p_2}}{n}, \frac{\sum_{i=1}^{n} AV_{i p_3}}{n}, \ldots, \frac{\sum_{i=1}^{n} AV_{i p_j}}{n} \right) \tag{6.2}
\]

In (6.2), n is the number of participants in issue I, j is the total number of positions for I and \( AV_{i p_1} \) is the average user i opinion on position p1.

The average user opinion has the same size as the user opinion vector. Therefore, each user opinion vector can be compared with AUO for a particular issue.

6.3.4.2 Focal Set Average Opinion

On the issue-level, to determine the average focal set opinion for a particular community, users’ opinions for all positions under the issue within the same focal set need to be considered.

\[
FSAO(C) = \left( \frac{\sum_{i=1}^{n} AV_{i p_1}}{n}, \frac{\sum_{i=1}^{n} AV_{i p_2}}{n}, \frac{\sum_{i=1}^{n} AV_{i p_3}}{n}, \ldots, \frac{\sum_{i=1}^{n} AV_{i p_j}}{n} \right) \tag{6.3}
\]

In (6.3), n is the number of community C participants, j is the total number of positions for I and \( AV_{i p_1} \) is the average user i opinion on position p1.

The focal set average opinion has the same size as the user opinion vector. Therefore, FSAO with AUO can be compared for a particular issue as the norm.


6.3.5 Issue Topics by Topic Molding

6.3.5.1 Corpus Preparing

Although an issue has explicit positions, it may have other implicit topics or positions as the discussion goes on between users. Moreover, in the ICAS design, positions do not have to be distinct mutually exclusive positions; they overlap or are alternatives to other positions. Thus, there is a need to identify the main topics that are discussed under an issue. However, the corpus needs to be prepared and preprocessed for topic modeling. Therefore, each user’s arguments are combined under the selected issue, tokenized and stop words, and low-frequency words are removed, and finally, stemmed. Each user has a single document of his processed text. Finally, the corpus is made up of all the users’ documents.

Topic modeling is done in different ways. A well-known algorithm is Latent Dirichlet Allocation (LDA) by Beli [32]. Another topic modeling method uses non-negative matrix factorization (NMF) [33]. The next subsections introduce both methods in the proposed framework.

6.3.5.2 Topic Modeling by LDA

LDA is a generative probabilistic model of a corpus [32]. LDA determines each word $w$ in each document $d$ and comes from a topic $t$ from a per-document distribution over topics. The LDA model converts the whole corpus (document-word) matrix to two lower matrices $\alpha$, $\beta$. Matrix $\alpha$ is a document-topic matrix and matrix $\beta$ is a topic-word matrix. To do so, it uses two probability values:

- The probability distribution of words in topics: $\Phi_{wt} = P(w|d)$.
- The probability distribution of topics in documents: $\Theta_{td} = P(t|d)$.

The probability of a word in a document is:

$$P(w|d) = \sum_{t \in T} p(t, d) p(t|d)$$

(6.4) by [32]
Where \( T \) is the total number of topics. If conditional independence is assumed, then:

\[
P(w|t, d) = P(w|t)
\]

(6.5) by [32]

Therefore,

\[
P(w|d) = \sum_{t=1}^{T} p(t)p(t|d)
\]

(6.6) by [32]

The generative process of the LDA model can be described as the joint probability distribution, and the likelihood of generating the whole corpus \( D \) is:

\[
P(D|\alpha, \beta) = \prod_{d=1}^{M} p(\theta_d | \alpha) \left( \prod_{n=1}^{N_d} \sum_{t_{d,n}} p(t_{d,n}|\theta_d) p(w_{d,n}|t_{d,n}, \beta) \right) d\theta_d
\]

(6.7) by [32]

\( \alpha \) is a matrix where each row is a user document \( d \), and each column represents a topic \( t \). A value in row \( d \) and column \( t \) represents how likely a user document \( d \) contains topic \( t \), whether the distribution is symmetric or asymmetric. \( \beta \) is a matrix where each row represents a topic \( t \) and each column represents a word \( w \). Therefore, \( \theta_d \) is the topic proportion for document \( d \). A value in row \( t \) and column \( w \) represents how likely a topic \( t \) contains word \( w \). Therefore, \( t_{d,n} \) is the topic assignment for word \( n \) in a user document \( d \), \( w_{dn} \) is the observed word in document \( d \).

**6.3.5.3 Topic modeling by NMF**

Non-negative Matrix Factorization is a dimension reduction method that factors high-dimensional vectors into a low-dimensionality representation. If a corpus \( D \) consist of \( n \) words and \( m \) documents, the \( D \) matrix can be factored into matrix \( W \) (\( n \) words and \( k \) topics) and matrix \( H \) (\( k \) topics and \( m \) documents). The \( W \) matrix holds the topics discovered from the documents, and \( H \) holds the membership weights for the topics in each document. However, topic modeling by NMF usually requires a corpus \( D \) to be normalized by TF-IDF. \( W \) and \( H \) matrices are optimized over an objective function as
\[
\frac{1}{2} \|D - WH\|_F^2 = \sum_{i=1}^{n} \sum_{j=1}^{m} (D_{ij} - (WH)_{ij})^2 \quad (6.8) \text{ by [33]}
\]

Using the objective function, the update rules for matrices \(W\) and \(H\) are:

\[
W_{ic} \leftarrow W_{ic} \frac{(AH)_{ic}}{(WHH)_{ic}} \quad (6.9) \text{ by [33]}
\]

\[
H_{cj} \leftarrow H_{cj} \frac{(WA)_{cj}}{(WWH)_{cj}} \quad (6.10) \text{ by [33]}
\]

The updated values are calculated in parallel operations. Then, the new matrices are used, \(W\) and \(H\), to calculate the reconstruction error. This process is repeated until convergence.

6.3.6 Individuals and Focal Sets and Topics Similarities

A pairwise similarity is used between identified topics and individuals and focal sets’ text to measure the contributions from individuals and focal sets’ text. The similarity between individual texts and identified topics is a simple task to calculate. However, the similarity between the focal set’s text and identified topics is not a straightforward calculation. It requires more processing. Thus, the following algorithm determines the focal set contribution to each topic.

**Algorithm 1 Focal Set - Topic Similarity**

**Input:** list of focal sets and list of topics

**Output:** matrix of the similarity between focal sets content and topics

- For each focal set:
  - For each topic:
    - For each user in the focal set:
      - User contribution = pairwise similarity between user content and topic
      - focal set contribution += User contribution
    - focal set topic similarity = focal set contribution / focal set size
Finally, the focal sets are sorted based on their similarity score in decreasing order for each identified topic. By setting a similarity threshold, communities that have contributed the most to a particular topic can be identified.

6.3.7 **Social Roles and Topic Intensity Degree in ICAS**

Most research on social roles [37 - 47] has used graph information to identify participants' social roles. They have mainly used the node position in a graph where the node in-degree and out-degree information determines the participant roles. The number and description of roles in the research [37 - 47] have varied concerning the associated platform. Since ICAS is a discussion and argumentation platform, the UIG from [48] is used. In addition, the node attributes from the UIG are used to determine the user’s role in the discussion.

[40-41] have identified three social role signatures based on the egocentric networks for participants: Answer Person, Discussion Person, and Discussion Catalyst. An “Answer Person” does more posting than receiving in a discussion. Therefore, the out-degree for this user in a graph is much higher than their in-degree. A “Discussion Person” has frequent reciprocal exchanges with other participants. As a result, this user has similar in-degree and out-degree scores. Finally, the “Discussion Catalyst” is responsible for posting messages that initiate long threads. This role may not have higher out-degree scores. Users with this role do not have more postings than others, but they are likely to attract others to engage in their threads. In this research, the roles mentioned above are used to label users in discussions.
6.4 DISCUSSION AND RESULTS

6.4.1 THE FSA ALGORITHM WITH ICAS DISCUSSIONS

The FSA algorithm is used as a link-based algorithm to partition users in each discussion into focal sets. Table 6-1 shows the FSA algorithm results applied to the UIG for each discussion hosted by ICAS.

The modularity, clustering coefficients, and transitivity scores are averaged using the FSA algorithm for multiple rounds. In this research, the number of nodes belongs to a focal set in the first round of using the FSA Algorithm ranges 30-40; in the last two rounds, the focal set size ranges from 2-4 nodes.

<table>
<thead>
<tr>
<th>Issue</th>
<th># Users</th>
<th># Connections</th>
<th># Focal Sets</th>
<th># Levels</th>
<th>Modularity</th>
<th>Clustering Coefficient</th>
<th>Transitivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Issue 1</td>
<td>305</td>
<td>4701</td>
<td>83</td>
<td>10</td>
<td>0.609477</td>
<td>0.097111</td>
<td>0.087107</td>
</tr>
<tr>
<td>Issue 2</td>
<td>291</td>
<td>3412</td>
<td>79</td>
<td>9</td>
<td>0.681429</td>
<td>0.169482</td>
<td>0.196258</td>
</tr>
<tr>
<td>Issue 3</td>
<td>297</td>
<td>3696</td>
<td>84</td>
<td>10</td>
<td>0.633422035</td>
<td>0.139165273</td>
<td>0.130571023</td>
</tr>
<tr>
<td>Issue 4</td>
<td>280</td>
<td>3198</td>
<td>88</td>
<td>11</td>
<td>0.635557726</td>
<td>0.097178521</td>
<td>0.17741009</td>
</tr>
</tbody>
</table>

Issue 1 has the highest number of participants and connections. The FSA algorithm looped 10 times and resulted in 85 focal sets. The UIG was sparse due to the high number of users participating in this discussion. As a result, the UIG had the lowest modularity, cluster coefficient, and transitivity scores. Issue 2 has 79 focal sets from applying the FSA algorithm within 9 rounds. However, this issue reported the highest modularity, cluster coefficient, and transitivity scores. This issue has a denser UIG graph than the other issues’ UIGs. Issue 3 and Issue 4 are in between Issue 1 and Issue 2, as shown in Table 6-1.
6.4.2 **Topic Modeling with ICAS Discussion**

LDA and NMF models are used for topic modeling on the issue level. Table 6.2 reports the evaluation result from applying both models.

<table>
<thead>
<tr>
<th>Issues</th>
<th>PMI - NMF</th>
<th>PMI - LDA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Issue 1</td>
<td>0.05277667</td>
<td>-0.0382568</td>
</tr>
<tr>
<td>Issue 2</td>
<td>0.011041</td>
<td>-0.007667</td>
</tr>
<tr>
<td>Issue 3</td>
<td>0.1094469</td>
<td>0.021647</td>
</tr>
<tr>
<td>Issue 4</td>
<td>0.0495284</td>
<td>-0.005781</td>
</tr>
</tbody>
</table>

In each issue, the number of topics is set to 4, with 10 tokens in each topic. For both models in all issues, the topics were largely dissimilar. With the limited corpus from each issue discussion, NMF reported better topic results, evaluated by pointwise mutual information (PMI), than LDA. The reason behind this result is that NMF normalizes the corpus using the TF-IDF. However, Topic modeling with NMF model requires more run time than the LDA model.

6.4.3 **Individuals and Focal Sets and Topics Similarities**

For each discussion topic, the topic is passed as a search query for all user documents. Based on the similarity scores between the user’s document and the topic, users’ documents are sorted in decreasing order to find users who contribute most to a topic.

Algorithm 1 is used for the focal sets to find the similarities between topics and focal set documents. Due to the space limitation, the results of the focal sets documents and topic similarities are not included in this dissertation, but they are available upon request. However, one significant finding in this area is that the similarity between topics and focal sets documents increases as focal set size decreases. This correlation is because some users contribute to the discussion but without any text, which negatively affects the topic similarity.
6.4.4 **INDIVIDUAL OPINION INTENSITY AND SOCIAL ROLES**

The intensity measure usually is split into three levels: low, medium, and high. In this research, the same levels of intensity are used to determine user opinion intensity. Each user opinion vector (UOV) is calculated as (6.1) and then is compared with AUO as (6.2) to label the user. This distance between UOV and AUO determines the user opinion intensity level, as shown in Figure 3. For each user, if the UOV is within 0.33 distance or less from the AUO, then this user opinion is labeled as a low-intensity opinion. If the UOV is within a distance between 0.34 and 0.67 from the AUO, then this user opinion is labeled as medium-intensity opinion. If the UOV has a distance of more than 0.67 from the AUO, the user’s opinion is labeled as high-intensity.

![Figure 6-3 THE OPINION INTENSITY BAR](image-url)
TABLE 6-3 STATISTICS OF INDIVIDUAL SOCIAL ROLES AND OPINION INTENSITY IN ICAS DISCUSSIONS

<table>
<thead>
<tr>
<th>Issues</th>
<th>Social Roles</th>
<th>Low-Intensity</th>
<th>Medium-Intensity</th>
<th>High-Intensity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Answer Person</td>
<td>35</td>
<td>56</td>
<td>9</td>
</tr>
<tr>
<td>Issue1</td>
<td>Discussion Person</td>
<td>29</td>
<td>104</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>Discussion Catalyst</td>
<td>11</td>
<td>22</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>Answer Person</td>
<td>40</td>
<td>45</td>
<td>8</td>
</tr>
<tr>
<td>Issue2</td>
<td>Discussion Person</td>
<td>53</td>
<td>101</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>Discussion Catalyst</td>
<td>6</td>
<td>25</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Answer Person</td>
<td>21</td>
<td>55</td>
<td>11</td>
</tr>
<tr>
<td>Issue3</td>
<td>Discussion Person</td>
<td>22</td>
<td>121</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td>Discussion Catalyst</td>
<td>3</td>
<td>23</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Answer Person</td>
<td>35</td>
<td>47</td>
<td>7</td>
</tr>
<tr>
<td>Issue4</td>
<td>Discussion Person</td>
<td>49</td>
<td>92</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td>Discussion Catalyst</td>
<td>3</td>
<td>12</td>
<td>10</td>
</tr>
</tbody>
</table>

To investigate the user’s social role, the egocentric network information is used to label the user as Answer Person, Discussion Person, or Discussion Catalyst, as previously mentioned in section 6-4.

Table 6-4 presents the number of labels for participants according to their opinion intensity and social roles. Most of the participants are labeled with either a “Discussion person,” “Medium-Intensity,” or both at the same time across discussions. Users with “Answer Person” and “Discussion Person” roles occur more frequently with the opinion of “Low-Intensity” than an opinion of “High-Intensity,” except for Issue 3. However, users with the “Discussion Catalyst” role have a different pattern. For Issue 1, users with the opinion of “High-Intensity” occur equally
to users with the opinion of “Low-Intensity.” However, in Issue 4, users with the opinion of “High-Intensity” are more than users with the opinion of “Low-Intensity,” while for Issue 2 and Issue 3; users with the opinion of “Low-Intensity” occur more than users with the opinion of “High-Intensity.”

The total number of participants across all discussions is 310 users. Of these users, 5% of the participants are users with an “Answer Person” role across all discussions, 25% of the participants have a “Discussion Person” role across all discussions, and only one user found who had the “Discussion Catalyst” across all discussions. On the other hand, users seem to switch social roles among discussions. For example, 72% of the participants alternated between the “Answer person” and “Discussion Person” roles across all discussions, and 38% of the participants had either “Discussion Person” or “Discussion Catalyst” roles across all discussions. However, no user was found with the “Answer person” and “Discussion Catalyst” roles in all discussions. These statistics show that users may switch roles from the “Answer person” role to the “Discussion Person” role or from the “Discussion Person” role or “Discussion Catalyst” role but not from the “Answer person” role to the “Discussion Catalyst” role according to the user behavioral postings.

The PageRank Algorithm [25] has been applied to each issue discussion. The users were sorted in decreasing order based on their PageRank score. Of the top 10%, most of the users were labeled as “Discussion Catalyst,” and a few of them were labeled as “Discussion Person.” Moreover, according to the PageRank Algorithm scores, two users were found to be influential in all discussions, six users found in three discussions, and twenty-three users found in two ICAS discussions.

Similarly, opinion intensity degree dynamics for users have an identical pattern to the user’s social role dynamics. 30% of the participants have an opinion with a “Low-Intensity” degree, 14% of the
participants have an opinion with a “Medium-Intensity” degree, and only one participant has an opinion with a “High-Intensity” degree across all issues. Of the total users, 50% either have opinions of “Low-Intensity” or “Medium-Intensity” degree, 40% have opinions of “High-Intensity” or “Medium-Intensity” degree, and 1% have opinions of “Low-Intensity” or “High-Intensity” degree across all discussions. Like social role dynamics, users may change their opinion intensity degree from “Low-Intensity” to “Medium-Intensity” or “Medium-Intensity” to “High-Intensity” degree but rarely from “Low-Intensity” to “High-Intensity” degree.

6.4.5 Focal Sets Opinion Intensity and Social Roles

Like individual opinion intensity, the focal set opinion intensity is calculated using (6.3) and then is compared with AUO as (6.2) to label the focal sets as in the previous section. Table 6-4 shows us the focal set labels in all ICAS discussions. Most of the focal sets in all issue discussions have an opinion with a “Low-Intensity” degree; fewer focal sets have an opinion with a “Medium-Intensity” degree and no focal sets with a “High-Intensity” degree. With the users being grouped in focal sets, the opinion intensity degree reduces.

The focal sets found in all ICAS discussions have different dynamics. Users with the “Answer Person” role appeared in 80% of the focal set members and made up 34%. Participants with the “Discussion Person” role appeared in 92% of the focal sets and made up 53% of the focal set members. However, users with the “Discussion Catalyst” role appeared in 55% of the focal sets and made up 12% of the focal set members. At least one user with a “Discussion Person” or “Discussion Catalyst” role exists in each identified focal set. The researcher found no focal set where all the members were labeled with the “Answer Person” role.
TABLE 6-4 STATISTICS OF FOCAL SET OPINION INTENSITY IN ICAS DISCUSSIONS

<table>
<thead>
<tr>
<th>Issues</th>
<th>Low-Intensity</th>
<th>Medium-Intensity</th>
<th>High-Intensity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Issue 1</td>
<td>71</td>
<td>12</td>
<td>0</td>
</tr>
<tr>
<td>Issue 2</td>
<td>44</td>
<td>35</td>
<td>0</td>
</tr>
<tr>
<td>Issue 3</td>
<td>66</td>
<td>17</td>
<td>0</td>
</tr>
<tr>
<td>Issue 4</td>
<td>73</td>
<td>13</td>
<td>0</td>
</tr>
</tbody>
</table>

As per the opinion intensity in focal sets, users with “High-intensity” opinion appeared in 39% of the focal set and made up 8% of the focal set members. Participants with “Medium-intensity” opinion appeared in 99% of the focal set and made up 75% of the focal set members. Finally, users with “Low-intensity” opinion appeared in 58% of the focal set and made up 17% of the focal set members.

6.4.6 ILLUSTRATION OF ISSUE 1

The graph partitioning using the link analysis algorithm and applying the FSA algorithm to Issue 1 resulted in 83 focal sets. Users may appear in one or more focal sets, depending on the argumentation connections made by users to others under Issue 1. This discussion has about 305 participants and 4701 argumentation connections (from arguments and reactions). This algorithm looped for ten times. The averaged modularity score from applying the FSA algorithm on the argumentation graph of Issue 1 is 0.609477, the averaged cluster coefficient is 0.097111, and the transitivity score is 0.087107. In this discussion and all ICAS discussions in this study, the system admin only posts issues and positions.
In contrast, arguments and reactions are posted by participants only, not the system admin. Therefore, as participants join the discussion, everyone replies or reacts to the admin’s post, which is the scenario’s position. However, the system admin never responds to any post, argument or reaction, in the discussion. About 66% of participants’ posts are made directly to the admin’s post, and the rest are made between participants. Thus, the averaged cluster coefficient and transitivity scores are very low in this study.

The other module in this framework is topic modeling by NMF and LDA. In table 6-5, the extracted topics can be seen by both models: NMF and LDA. Topics by LDA are not as meaningful as topics by NMF. For example, T1a and T2a do not have a clear idea. T1c and T1d are similar and have more of the discussion’s stop words. On the other hand, topics by NMF are much better than topics by LDA. For example, T1b is about necessary conditions in the adopting home: care, love, home, and family. T2b is about what the Bible thinks about straight and gay parents and people. T3b is about family qualifications for adopting a child. T4b is about what Figures a child needs in a family.

Regarding the pairwise similarity between users’ content and identified topics, if a user has a similarity with a topic from Table 6-5 of 70% and more, the researcher considers this user as having discussed this topic in his content. For the topics modeled by NMF, T1b is discussed by 84 users. Only one user was labeled as “Discussion Catalyst” with a “High-Intensity” opinion. The majority of the users have a “Medium-Intensity” opinion with different social roles. Yet, most of the users with “Low-Intensity” opinions are labeled as “Answer Person.” On the other hand, only 11 users discuss T2b. Only 60% of these users have been labeled as “Discussion Person,” a “Medium-Intensity” opinion. Similarly, only 10 users have discussed T3b. Lastly, only 11 users have discussed T4b. Still, the majority of these users have a “High-Intensity” opinion. Likewise,
topics that LDA models have a similar pattern. For example, T1a was discussed by 47 users, T2a was discussed by 55 users, T3a was discussed by 44 users, and T4a was discussed by 94 users. The majority of users in these topics are labeled as “Answer Person” or “Discussion Person” and with “Medium-Intensity” or “Low-Intensity” opinions. T3a has no users with a “High-Intensity” opinion.

The majority of users in these topics are labeled as “Answer Person” or “Discussion Person” and with “Medium-Intensity” or “Low-Intensity” opinions. T3a has no users with a “High-Intensity” opinion.

To measure the similarity between a focal set’s text and a topic, Algorithm 1 from Section 6.3.6 was used. The similarity between the focal sets’ text and topics in Table 6-5 is significantly less than the similarity between a user’s text and a topic. Some of the users in the focal set do not contribute much to the discussion by replies. They contribute with reactions only. For instance,
user1, user11, user2, user30 and user5 are members of many focal sets with a high similarity between their texts and T1b. However, user2 has a similarity of 40% with T1b. This user provides little content to the discussion but engages with others by reactions.

In the same way, user2, user17, user30 and user50 have appeared in many focal sets that have high similarity between their texts and T2b; however, their text contribution to T2b is not significant.

User1, user11, user14 and user3 have appeared in many focal sets with a high similarity between their texts and T3b. However, only user1 has a significant contribution to T3b. T4b has a similar pattern to T2b. User4, user8 and user17 have appeared in many focal sets that have high similarity between their texts and T4b; however, their text contribution to T4b is not significant. All users who do not significantly contribute to the focal set’s text are labeled as “Answer Person.” Users with the “Answer Person” label may not add much to the content, but they help promote discussion content. For example, user 1 posts to others more than they receive, as in Figure 6-3. On the other hand, user11 receives incoming links more than he posts, as in Figure 6-4.

The FSA algorithm and topic modeling techniques are useful tools to extract knowledge within their domain. Combining these techniques leads to more valuable information hosted by UGC.
platforms. It leads to discovering and profiting hidden communities and users based on their discussion and contribution by aggregating individuals' data. These algorithms are beneficial for identifying the social roles and identifying the opinion intensity for participants in cyber-argumentation platforms.

6.5 Conclusion

Identifying topic-oriented communities is an essential task in online discussion platforms. Moreover, it is important to measure the opinion intensity and identify the social roles for users who make up the focal set to understand discussions and participants better. A new framework is proposed in this research. This framework can measure the users’ opinion in a discussion and compare it to the average user’s opinion vector as well as identify focal sets using the FSA algorithm and measure their opinion intensity degree. Additionally, the framework applies a pairwise similarity between topics discussed by individuals and focal sets to leverage topic-oriented individuals and focal sets in cyber-argumentation. Finally, it analyzed the individuals and focal set dynamics using users’ opinions and social roles.

This framework can be used to identify similar focal sets and individuals who are behind certain topics but not connected. It can also be used to blend communities and individuals of polarized opinions in an online discussion. This balances the focal sets and individuals in a discussion and draws out the crowd wisdom in the cyber-argumentation platform.

There are many options to expand this model. For example, since the individual and the focal set opinion intensity degree are calculated, is it possible to calculate the topic intensity degree according to the participants' attitudes? If yes, how to increase or decrease the topic intensity degree? Another task that could be done is to investigate non-textual users who appeared in
multiple focal sets but have not contributed to the discussions with replies. This is left for future work.

6.6 References


Chapter 7. CONCLUSION AND FUTURE WORK

User-generated content (UGC) platforms host different forms of information, such as audio, video, pictures, and text. They have many online applications – i.e., social media, blogs, photos and video sharing, customer reviews, and debate sites. Usually, the UGC platform content is uploaded and consumed by users. Most of these platforms, mainly social media and blogs, are often used for an online discussion and provide tools for users to share and express opinions. Commonly, people from different backgrounds and origins discuss opinions about various issues over the Internet. Furthermore, discussions among users contain substantial information from which researcher can extract knowledge about collective intelligence. Collective Intelligence grows when a group works together collectively or cooperatively.

Cyber-argumentation and debate platforms are designed to enable online deliberations with large-scale, in-depth cyber-argumentation for productive discussions over the Internet. These platforms host structured argumentation networks that allow complex analytical models to mine the argumentation for collective intelligence. However, not all of the argumentation platforms have mobile-platform versions and — if these exist; they have primitive capabilities — such as basic statistics. In this dissertation, the design of a mobile application for cyber-argumentation is presented via reporting basic and complex stats and analytics. This mobile application supports intelligent cyber-argumentation and large-scale discussions and provides meaningful analytics on mobile devices. The platform has incorporated several analytical models to capture collective opinions, detect opinion polarizations, and predict missing user opinions. It also considers the limitation of mobile devices' screen size to allow argumentation. An example is used to illustrate our design and models in the mobile space, and a system usability study of our application is presented.
A tree structure represents each discussion in cyber-argumentation platforms. The discussion massively grows as users join the discussion. However, a study from our lab showed that users only view, on average, 3% of the discussion content. Therefore, not all posted opinions are constructive or worth more discussion. Moreover, the screen size of a mobile application makes it hard for users to keep track of the discussion or identify which opinions are constructive. Thus, there is a need for an opinion discovery method in cyber-argumentation. In social media and blog platforms, opinions are discovered by engagement information, impact score or reverse-chronological order. On the other hand, some UGC platforms discover constructive opinions based on textual features. However, there are no pre-identified feature sets for the opinions that have been set for searching and discovering opinions in academia or public-debate platforms. Identifying features for locating constructive opinions helps improve the discourse quality and provides an attractive online discussion platform. In this research, a new framework for opinion selection and discovery is proposed. This framework locates constructive opinions based on four unique features: engagement, recentness, controversy, and author influence. Therefore, it provides an attractive and dynamic discourse, incorporating the opinions feature based on users’ preferences. First, these features are defined in the cyber-argumentation space. Then, a new framework that combines these features for opinion search and discovery is developed. Finally, examples on a deliberation dataset demonstrate the proposed framework's effectiveness in discovering and searching for constructive opinions. This application indicates the need for further investigation. First: how to measure and model the controversial degree of a particular discussion or argument? Second: how to discover the topical focal sets in discussions using additional information from the argumentation graph and topic modeling techniques? These two questions lead to new directions for this dissertation.
Certain phenomena, such as controversy, often appear in online argumentation, making the discussion heated between participants. Heated discussions can be used to extract new knowledge. Therefore, detecting the presence of controversy is an essential task to determine if collective intelligence can be extracted from online discussions. This research uses existing measures for estimating controversy quantitatively in cyber-argumentation. First, it defines controversy from the controversy definitions in different fields, and then it identifies the attributes of controversy in online discussions. The distributions of user opinions and the distance between opinions are used to calculate the degree of online discussions' controversiality. Finally, each controversy measure results are discussed and analyzed using an empirical study generated by ICAS. These measures are improvements over existing measurements because it does not require ground-truth data or specific settings and can be adapted to distribution-based or distance-based opinions. Constructive opinions are often provided by users with specific profiles and have implicit topics. Users may work together cooperatively or collectively to support, attack, or deliver agendas surrounding similar interests. Much research has been conducted to detect groups or individuals who shared similar interests using link analysis, text analysis, and other data mining techniques, or a combination of different methods to discover communities or hidden communities with specific characteristics in social media and blogs. However, most of the previous work that has been done has focused on improving the community, or hidden community, detection algorithms concerning the research challenges. Cyber-argumentation platforms are full of valuable hidden features worth further study due to the extensive discussion they contain. This research proposed a new framework for discovering groups, or focal sets, with similar interests in topics using the focal structure analysis algorithm and topic modeling. Then, the social roles of the focal set members are investigated. By combining these techniques, groups and individuals behind specific
topics in online discussions are discovered and studied. This work can be useful for UGC platforms to identify groups with similar interests using their discussion and interaction information. It mainly helps discover these groups and identify group members' social roles and measure opinion intensity. Since UGC platforms are not used for entertainment only, the proposed framework can leverage new knowledge or malicious activities from the collective discussions. It estimates individual interests and power in the identified groups. Beyond this, it should help researchers and policymakers develop new measures and policies with cyberspace's associated risks as well as opportunities.

As UGC platforms are growing significantly over time, cyber-argumentation platforms are full of valuable information. The current research work and the work in this dissertation need more study. There are many open questions to explore in each area of this dissertation:

Chapter 3: the first question is, with the massive amount of connections between handheld devices and back-end servers required to process online discussion data, how to optimize these connections and still facilitate the end-user with all needed information? The other question is: how to connect this application to existing social media platforms to exchange user-based and discussion-based information? How to relate discussions in social media and cyber-argumentation?

Chapter 4: the second question, in the opinion discovery framework, if the length of arguments and the number of views were included in this opinion discovery framework, would it lead to better results? How can they be incorporated with the other argument indicators to find constructive opinions?

Chapter 5: the third question is, in analyzing and modeling controversy in cyber-argumentation, what is the relationship between controversy and polarization in collective intelligence development? How to lower or increase the controversial degree as the discussion augments?
Chapter 6: the fourth question is, this work presents a hybrid model to identify topical focal sets in cyber-argumentation. This study was performed on a closed and mentored community. However, in the real setting, bots and trolls exist. How to identify them and prevent them from posting scams or irrelevant posts?

The questions mentioned above describe the work left to do. They lead the research into areas where further work is needed.