Spotting Fake News: A Social Argumentation Framework for Scrutinizing Alternative Facts

Ricky J. Sethi[1]

Abstract— **The proliferation of fake news in today's digital world has moved beyond a specific election cycle and now commands headlines globally. In this paper, we propose countering the spread of fake news on social networks by leveraging these crowds to instead help verify alternative facts. We present a prototype social argumentation framework to verify the validity of proposed alternative facts to help curb the propagation of fake news. We utilize fundamental argumentation ideas in a graph-theoretic framework that also incorporates semantic web and linked data principles. The argumentation structure is crowdsourced and mediated by expert moderators in a virtual community.**

I. Motivation

The proliferation of fake news in today's digital world has moved beyond a specific election cycle and now commands headlines globally. Alternative facts are shared on social networks and spread like wildfire across all sorts of social media. Being able to distinguish credible information from alternative facts is essential to curbing the propagation and amplification of such misinformation.

In this paper, we propose countering the spread of fake news on social networks by leveraging these crowds to instead help verify alternative facts. Computational approaches for addressing fake news usually focus on automated tools for detection. These tools flag previously identified hoaxes; or automatically detect fake news articles using natural language processing techniques with pre-existing ground truth; or track the viral-like transmission of hoaxes [23], [6], [17], [19], [29]. None of the existing approaches, however, deal with verification of the alternative facts which constitute the semantic content of such articles.

However, reliance upon misinformation and propaganda has been prevalent in much of human history [10], [20]. Critical thinking and evidence-based reasoning are essential for countering propaganda and misinformation intended to manipulate public opinion [28], [24]. In particular, argumentation has been shown to be a natural, substantiated approach for analyzing the veracity and reliability of assertions and claims [18], [8]. In fact, in considering how to assess critical thinking, [8] asserts the need to identify conclusions, reasons, and assumptions as well as judging the quality of arguments and developing positions on an issue. Using this sort of evidence based reasoning not only has the potential to identify fake news to a greater extent but also to imbibe users with the critical thinking ability to navigate future fake news articles.

Just as with propaganda and misinformation, these so-called ”alternative facts” have to be examined critically using evidence-based reasoning in the age of the World Wide Web. Collaborative interaction is seen as one of the keys for developing critical thinking and evidence-based reasoning in online forums [16], [36], [15]. Some [3] also identify the need to provide argumentation structures to create deeper personalized knowledge and go beyond a simpler social construction of collective knowledge in collaborative online forums using computer supported argumentation [22], [27].

In this paper, we present a prototype system that uses social argumentation to verify the validity of proposed alternative facts and help with fake news detection. We utilize fundamental argumentation principles in a graph-theoretic framework that also incorporates semantic web and linked open data principles [2], [14], [11]. The argumentation structure is crowdsourced and mediated by expert moderators in a virtual community. To the best of our knowledge, our novel computational approach is the only one to address the verification of alternative facts and fake news.

II. Background on Argumentation

Argumentation can be described as a ”kind of discourse through which knowledge claims are individually and collaboratively constructed and evaluated” based on evidence [9]. Building upon the central components of an argument, namely claim, data, warrant [32], Toulmin's ideas have influenced and been refined, extended, and formalized in artificial intelligence and other disciplines [33]. An argument can be seen as a set of premises and a conclusion supported by evidence [12], [35], [34] and this kind of argumentation plays a central role in the building of explanations, models, and theories [9], [30].

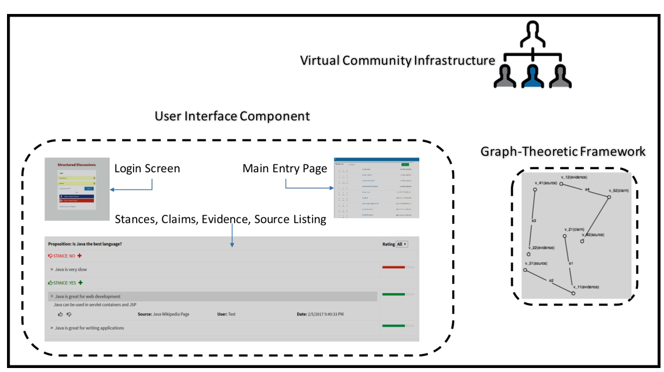

Fig. 1. Overview of our System Architecture.

[1]Department of Computer Science, Fitchburg State University, 160 Pearl St, Fitchburg, MA 01420 USA `rickys@sethi.org`

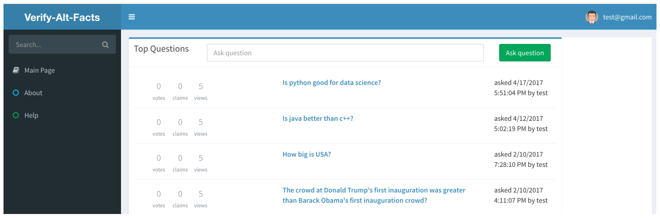

Fig. 2. Main landing screen.

Indeed, [21], [37] emphasize the need to employ online collaboration and online communities, which include both learners and experts creating knowledge together and evaluating complex issues using multiple perspectives to think critically about the topics being discussed [3], [37]. In addition, [13], [36] emphasize the benefit of online facilitators that are instructors or experts to help increase critical thinking and depth of knowledge in learners by guiding threaded discussions, especially to avoid early termination of threads. In fact, [4], [7] even show that simply observing the learning process of another learner engaged with an expert results in information gain via observational learning [1].

III. Overview of Approach

We build upon the Argument Interchange Format [22], [35], which models an argument as a network of connected nodes of information (claims and datum which we model as premises and evidence) and schemes (warrants or rules of inference which we model as a particular conclusion or stance). Our graph-theoretic approach also keeps track of provenance in argumentation schemes [30], [31].

We define an argument as being composed of Stances, Claims, and Evidence, where both Claims and Evidence are supported by Sources, typically web documents. A Claim is either an inference or a conclusion while Evidence (sometimes called a Premise) provides the support for that Claim.

A Stance is the final conclusion composed of Claims and Evidence, and their associated Sources. Stances are fundamental stands on a topic and can be mutually exclusive, should have cohesive sub-structures, and are composed of atomic argumentation components (Claims, Evidence, and Sources). A Claim can be directly supported by a Source when there is no individual Evidence component for it; if more Evidence components are added for a Claim, then Sources are only associated with Evidence components and not directly with Claim components. In fact, multiple sources can support multiple evidence/premise nodes.

The Sources themselves have their own properties. A Source could be fully described, for example using the Dublin Core metadata (http://dublincore.org). In this way, users could query the system for assertions from certain sources or from sources with specified properties (e.g., government institutions).

Our methodology also incorporates Ratings for each Source and user in the system. Different trust, authority, and other attribute dimensions are amalgamated and weighted in a Summary Rating; these compound ratings reveal their constituent components (SourceRating, ContentRating, QuestionRating, etc.) on a MouseOver event, displaying details of Users' Ratings, Source Ratings, Expert Ratings, etc.

Our proposed framework is not just a system for argumentation structure; instead, we organize the community and system to work together synergistically to support learning via critical thinking. Members of this virtual community can take three major roles: 1) Users, who are the information seekers submitting the queries; 2) Responders, who have some degree of expertise or background to add Claim, Evidence, and Source nodes; and 3) Moderators, who are contributors that guide the question and answer flow, including triaging incoming questions, matching experts to new questions, evaluating answers for quality assurance, etc. These roles are dynamic as they may evolve over time, and may be multi-faceted with different functions and capabilities.

An example exchange is shown in the screenshot in Figure 3; the example portrayed is based on the origin of the term, "alternative facts." A fact is usually defined as a piece of information used as evidence or as part of a report or article. The term, "alternative facts," originated when Kellyanne Conway, the Counselor to the President, stated in a January 22, 2017 interview on Meet the Press that White House Press Secretary Sean Spicer was giving "alternative facts" when trying to defend Sean Spicer's claim about the attendance numbers at President Donald Trump's inauguration. This usage is distinct from the legal term, *alternative facts*, used "to describe inconsistent sets of facts put forth by the same party in a court."

IV. System Architecture

We developed our system as a web-based application with a responsive interface that allows for viewing on desktops, tablets or mobile phones. It consists of three main components: the Graph-Theoretic Framework, the User Interface Component, and the Virtual Community to support Crowdsourcing. We describe each of these components in detail below.

A. Graph-Theoretic Framework

We create an *Argumentation Graph*, $G_A = (V, E, f)$, composed of a set of vertices, V, edges, E, and a function, f, which maps each element of E to an unordered pair of vertices in V. Each fundamental Claim, Evidence, or Source in an argument thus constitutes an atomic argumentation component, v_a, and is embedded as a vertex in the graph such that $v_a \in V$. The vertices contain not just the component's semantic content, but also the ratings, authority, trust, and other attribute dimensions of each atomic argumentation component. The edges $e \in E$ contain weights along the various dimensions of trust and authority as well as a pro/con designation for the connection (e.g., if an Evidence node is pro or con for a connected Claim node), while the function f maps how they're connected. Depending on the context of the argument, this graph can be undirected or directed, where

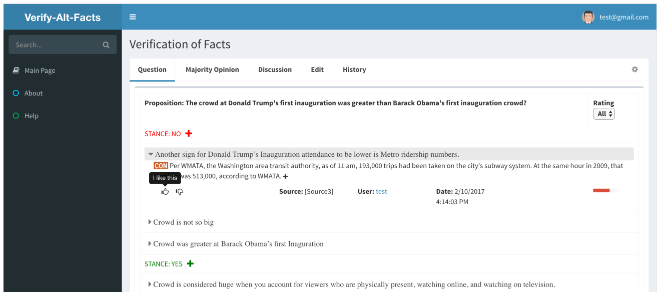

Fig. 3. Screenshot of main proposition screen.

the temporal component gives the direction to the directed graph.

In terms of a graph, we therefore see the set of vertices V as the set of Claims, Evidence, and Sources; the set of edges E as a set of links that may connect any two vertices (which can be a Source, Evidence, or Claim). Each subgraph or path traversal that can be obtained from a graph results in a Stance.

For example, we might have the following set of vertices and edges in a sub-graph:

$$V = \{v_{11}(evidence_1), v_{21}(claim_1), v_{31}(source_1), \\ v_{42}(source_2), v_{52}(claim_2)\}$$

and $E = \{e_1, e_2, e_3, e_4, e_5\}$. Through the function f, we get the following edges between two vertices, v_a and v_b:

- $e_1 : v_{11}(evidence_1) \rightarrow v_{21}(claim_1)$
- $e_2 : v_{31}(source_1) \rightarrow v_{11}(evidence_1)$
- $e_3 : v_{41}(source_1) \rightarrow v_{22}(evidence_2)$
- $e_4 : v_{12}(evidence_2) \rightarrow v_{52}(claim_2)$
- $e_5 : v_{42}(source_2) \rightarrow v_{52}(evidence_2)$

We can represent this in the form of an adjacency matrix; please note: the second index on the vertices represents the ordinal position of the kind of node it represents (e.g., Claim, Evidence, or Source) for illustrative purposes only.

There are two ways to represent the stances: one way is by making the Stance another node in G_A that is added by the moderators in a top-down manner. The other is to designate each sub-graph as a different Stance. Once the G_A is formed, we can form sub-graphs which represent the different stances we can infer from the argumentation graph where each sub-graph would be a separate Stance. Since a Stance is composed of Claims, Evidence, and their associated Sources, there is also a check on the minimum number of vertices a subgraph might contain. Our approach supports both ways of determining the various stances (what we call top-down vs bottom-up).

In this approach, a Stance is a sub-graph or tree of the argument, G_A. A particular path traversal would show the weights or quality of the Stance. Depending on the specific path taken through such an argumentation graph, the connections would allow atomic components to be incorporated in different Stances, with each Stance represented by some traversal of the graph.

B. User Interface Component

The User Interface Component of our system was developed as a web-based application and designed for viewing on desktops, tablets, or mobile phones. The front-end component was developed using HTML, CSS, Bootstrap, jQuery and JavaScript, while the back-end was developed using C# and Asp.net MVC 5 framework. The UI is comprised of three main components:

Database Component This back-end data collection component is responsible for the saving and retrieval of users' data, contributions, and activity and connects to the graph-theoretic framework using JSON objects and to the backend using an Object Relational Mapping (ORM) framework. It uses MS SQL as its relational database management system.

Data Collection Component This front-end component shows the questions to the users and gets questions. Other users can post their positive and negative claims based on their views and support their claim by providing evidence and sources. The algorithm in the backend sorts the claims based on filtering and sorting criteria like highest rating and source. When a user wants to add claims, evidence, or sources to a particular stance, they can click the plus button on respective stances which displays the modal popup for entry of claims, evidence, and sources. Users can rate the claim by clicking the upvote/downvote button.

Backend Component This backend component is handled by an Asp.net MVC5 architecture. The database component is accessed using the Entity framework, an Object Relational Mapping (ORM) framework that offers an automated mechanism for storing and accessing the data in the database. User input is validated to check for errors, SQL injection, etc. The backend component gets the users' data from the Data Collection Component and analyze the data using algorithm, then passed to the User Interface component. Data are sent in the form of hierarchical order using the backend algorithm based on weighted parameters for every claim given by the users. The data are returned in JSON form. JSON, often used for serializing and transmitting data over a network, is primarily used here to transmit data between the Graph-Theoretic Framework and the User Interface Component.

C. Virtual Community for Crowdsourcing

We provided the simplest possible argument structure that will enable understanding and participation as this was found to be an effective representation to enable volunteer contributors to create collaborative arguments [5]. These systems enable a community to collaboratively create answers to questions where many possible answers, or nuanced perspectives on a single answer, can be posited. We used a minimalistic argument structure to facilitate contributions and synthesis and designed the online community to best facilitate this interaction: we defined the nature of the community, clarified the contribution process, and then designed the initial system prototype. Our graph-theoretic framework can analyze arguments, which will enable the system to proactively relate viewpoints and derive source ratings.

We designed multiple community roles to support our collaborative argumentation system. Our community allows for a generalized five-pronged constituency: Questioners, Question Moderators, Experts, Contributors, and Answer Moderators. In terms of the overarching roles of Users, Responders, and Moderators (as outlined above), Questioners would be regular Users while Contributors would be rated Users of the site who are allowed to become Responders, along with the Experts. The Question and Answer Moderators would both fall under the Moderator role. These kinds of roles can be modified for communities like The Madsci Network [25] and WikiTribune[1].

These roles can also be generalized for formal settings in classrooms as well as expert discussions amongst domain experts. We have also developed an initial set of metrics to quantify the structure of these threaded discussions by measuring the redundancy of posts, the compactness of topics, and the degree of hierarchy in sub-threads which we will incorporate here as we open it up for widespread testing [26].

ACKNOWLEDGMENT

We would like to gratefully acknowledge support from the Amazon AWS Research Grant program. We would also like to thank RaghuRam Rangaraju for the development and to especially thank Yolanda Gil for the invaluable discussion and support.

REFERENCES

[1] Albert Bandura. *Social foundations of thought and action: A social cognitive theory*. Prentice-Hall, Englewood Cliffs, 1986.

[2] Christian Bizer, T Heath, and T Berners-Lee. Linked data-the story so far. *International journal on Semantic Web and Information Systems*, 5(3):1–22, 2009.

[3] DT Chen and David Hung. Personalised knowledge representations: the missing half of online discussions. *British Journal of Educational Technology*, 33(3):279–290, 2002.

[4] MTH Chi, M Roy, and RGM Hausmann. Observing tutorial dialogues collaboratively: Insights about human tutoring effectiveness from vicarious learning. *Cognitive Science*, 2008.

[5] Timothy Chklovski, V Ratnakar, and Yolanda Gil. User interfaces with semi-formal representations: a study of designing argumentation structures. *IUI*, 2005.

[6] Niall J Conroy, , Victoria L Rubin, and Yimin Chen. Automatic Deception Detection: Methods for Finding Fake News. In *ASIS&T*, 2015.

[7] Scotty D. Craig, Michelene T. H. Chi, and Kurt VanLehn. Improving classroom learning by collaboratively observing human tutoring videos while problem solving. *Journal of Educational Psychology*, 101(4):779–789, 2009.

[8] Robert H. Ennis. Critical thinking assessment, 1993.

[9] Sibel Erduran, Yasemin Ozdem, and Jee-Young Park. Research trends on argumentation in science education: a journal content analysis from 19982014. *International Journal of STEM Education*, 2(1):0–12, 2015.

[10] Stuart Ewen. *PR!: a social history of spin*. BasicBooks, 1996.

[11] Daniel Garijo and Yolanda Gil. A New Approach for Publishing Workflows : Abstractions , Standards , and Linked Data. In *SC WORKS*, 2011.

[12] Nancy L Green. Towards Intelligent Learning Environments for Scientific Argumentation. *Intelligent Tutoring Technlogies for ill-defined problems and ill-defined domains: Proceedings of the 4th internationial workshop on Intelligent tutoring systems and ill-defined domains held at the 10th International Conference on Intelligent Tutoring Sys*, (September 2015):29–36, 2010.

[13] Ying-Hua Guan, Chin-Chung Tsai, and Fu-Kwun Hwang. Content analysis of online discussion on a senior-high-school discussion forum of a virtual physics laboratory. *Instructional Science*, 34(4):279–311, jul 2006.

[14] Tom Heath and Christian Bizer. *Linked data: Evolving the Web into a global data space (1st edition)*, volume 1. 2011.

[15] Allan C Jeong. The Effects of Intellectual Openness and Gender on Critical Thinking Processes in Computer-Supported Collaborative Argumentation. *Argumentation*, 22(1):1–18, 2007.

[16] Heisawn Jeong and Michelene T. H. Chi. Knowledge convergence and collaborative learning. *Instructional Science*, 35(4):287–315, nov 2006.

[17] Zhiwei Jin, Juan Cao, Yu-Gang Jiang, and Yongdong Zhang. News Credibility Evaluation on Microblog with a Hierarchical Propagation Model. *2014 IEEE International Conference on Data Mining*, pages 230–239, 2014.

[18] Ralph Johnson. *The Rise of Informal Logic*. Windsor Studies in Argumentation, 1996.

[19] Adam Kucharski. Post-truth: Study epidemiology of fake news. *Nature*, 540(7634):525–525, 2016.

[20] Walter Lippmann. The phantom public, 1925.

[21] Stacey Ludwig-Hardman and Stephanie Woolley. Online learning communities: Vehicles for collaboration and learning in online learning environments. *World Conference on Educational Multimedia, Hypermedia and Telecommunication*, pages 1998–2000, 2000.

[22] Chris Reed, D. Walton, and F. Macagno. Argument diagramming in logic, law and artificial intelligence. *The Knowledge Engineering Review*, 22(01):87–109, 2007.

[23] Victoria Rubin, Yimin Chen, and Niall J. Conroy. Deception Detection for News: Three Types of Fake News. *ASIS&T*, 2015.

[24] Megan Barnhart Sethi. Information, education, and indoctrination: The federation of american scientists and public communication strategies in the atomic age. *Historical Studies in the Natural Sciences*, 42(1):1–29, 2012.

[25] Ricky J. Sethi and Lynn Bry. The Madsci Network: Direct Communication of Science from Scientist to Layperson. In *International Conference on Computers in Education (ICCE)*, 2013.

[26] Ricky J. Sethi, Lorenzo A. Rossi, and Yolanda Gil. Measures of Threaded Discussion Properties. In *Intelligent Support for Learning in Groups at International Conference on Intelligent Tutoring Systems (ITS)*, 2012.

[27] S.B. Shum and Nick Hammond. Argumentation-based design rationale: what use at what cost. *International Journal of Human-Computer Studies*, 40(4):603–652, 1994.

[28] Michael J Sproule. *Propaganda and democracy : the American experience of media and mass persuasion*. Cambridge University Press, Cambridge, U.K. New York, NY, 1997.

[29] M. Tambuscio, G. Ruffo, A. Flammini, and F. Menczer. Fact-checking effect on viral hoaxes: A model of misinformation spread in social networks. *WWW*, pages 977–982, 2015.

[30] Alice Toniolo, Federico Cerutti, Nir Oren, Tj Norman, and Katia Sycara. Making Informed Decisions with Provenance and Argumentation Schemes. In *11th International Workshop on Argumentation in Multi-Agent Systems*, pages 1–20. Aamas2014.Lip6.Fr, 2014.

[31] Alice Toniolo, Timothy Dropps, Robin Wentao, and John a Allen. Argumentation-based collaborative intelligence analysis in CISpaces. In *COMMA*, pages 6–7, 2014.

[32] S Toulmin. *The Uses of Argument*, volume 70. 1958.

[33] Bart Verheij. The Toulmin Argument Model in Artificial Intelligence. In Iyad Rahwan and Guillermo R. Simari, editors, *Argumentation in Artificial Intelligence*, chapter 11, pages 219 – 238. 2009.

[34] Douglas Walton. Argumentation Theory: a Very Short Introduction. pages 1–21, 2009.

[35] Douglas Walton, Christopher Reed, and Fabrizio Macagno. *Argumentation Schemes*. Cambridge University Press, 2008.

[36] Qiyun Wang and Huay Lit Woo. Investigating students' critical thinking in weblogs: an exploratory study in a Singapore secondary school. *Asia Pacific Education Review*, 11(4):541–551, may 2010.

[37] Yu-Chu Yeh. Integrating e-learning into the Direct-instruction Model to enhance the effectiveness of critical-thinking instruction. *Instructional Science*, 37(2):185–203, nov 2007.

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