
Ricky J. Sethi

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.

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.
that White House Press

other attribute dimensions are amalgamated and weighted in
government institutions).

sources or from sources with specified properties (e.g.,
users could query the system for assertions from certain
Dublin Core metadata (http://dublincore.org). In this way,
Source could be fully described, for example using the
can support multiple evidence/premise nodes.

not directly with Claim components. In fact, multiple sources
Sources are only associated with Evidence components and
when there is no individual Evidence component for it; if
Sources). A Claim can be directly supported by a Source

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 (some-
times 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 con-
stituent components (SourceRating, ContentRating, Quest-
ionRating, etc.) on a MouseOver event, displaying details
of Users’ Ratings, Source Ratings, Expert Ratings, etc.

Our proposed framework is not just a system for argu-
mentation 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, in-
cluding 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.”

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].

Indeed, [21], [37] emphasize the need to employ online
collaboration and online communities, which include both
learners and experts creating knowledge together and evalu-
ating complex issues using multiple perspectives to think crit-
ically 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].

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. 2. Main landing screen.

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
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:

\[
V = \{v_{11}(\text{evidence}_1), v_{21}(\text{claim}_1), v_{31}(\text{source}_1), v_{42}(\text{source}_2), v_{52}(\text{claim}_2)\}
\]

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 \( v_b \):

- \( e_1 : v_{11}(\text{evidence}_1) \rightarrow v_{21}(\text{claim}_1) \)
- \( e_2 : v_{31}(\text{source}_1) \rightarrow v_{11}(\text{evidence}_1) \)
- \( e_3 : v_{41}(\text{source}_1) \rightarrow v_{22}(\text{evidence}_2) \)
- \( e_4 : v_{12}(\text{evidence}_2) \rightarrow v_{52}(\text{claim}_2) \)
- \( e_5 : v_{42}(\text{source}_2) \rightarrow v_{52}(\text{evidence}_2) \)

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.
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 WikiTribune1.

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 RaghuruRam Rangaraju for the development and to especially thank Yolanda Gil for the invaluable discussion and support.

REFERENCES


1https://www.wikitribune.com/