

Technical Disclosure Commons

Defensive Publications Series

November 2023

IDENTIFYING ORIGINATORS/SPEAKERS OF STATEMENTS/IDEAS IN VIRTUAL MEETING ENVIRONMENTS

Hazim Dahir

Zaheer Aziz

Subha Dhesikan

Follow this and additional works at: https://www.tdcommons.org/dpubs_series

Recommended Citation

Dahir, Hazim; Aziz, Zaheer; and Dhesikan, Subha, "IDENTIFYING ORIGINATORS/SPEAKERS OF STATEMENTS/IDEAS IN VIRTUAL MEETING ENVIRONMENTS", Technical Disclosure Commons, (November 19, 2023)

https://www.tdcommons.org/dpubs_series/6431



This work is licensed under a [Creative Commons Attribution 4.0 License](https://creativecommons.org/licenses/by/4.0/).

This Article is brought to you for free and open access by Technical Disclosure Commons. It has been accepted for inclusion in Defensive Publications Series by an authorized administrator of Technical Disclosure Commons.

IDENTIFYING ORIGINATORS/SPEAKERS OF STATEMENTS/IDEAS IN VIRTUAL MEETING ENVIRONMENTS

AUTHORS:

Hazim Dahir

Zaheer Aziz

Subha Dhesikan

ABSTRACT

In technology, entertainment, business, and/or legal proceedings, especially in scenarios in which significant monetary and/or legal rights are at stake, it can be important to determine when a statement was made and by whom the statement was made. As a majority of business is currently conducted virtually or as a hybrid (e.g., partial virtual presence), it would be desirable to be able to determine, with high certainty, an originator of an "idea", phrase, formula, etc. This proposal provides techniques that may be utilized to enable speaker attribution for virtual meeting environments.

DETAILED DESCRIPTION

Presented herein are techniques through which speaker determination and/or attribution may be provided for virtual meeting environments in order to determine/attribute an idea discussed during a meeting to a particular meeting participant. Specifically, a system may be provided that generates, for a meeting, a dynamic meeting attribution list that identifies, tracks, and records an attribution for all ideas that may be expressed in conversation occurring during the meeting. The system may perform real-time comparisons of topics in order to find similar and dissimilar topics, which can be utilized to form 'themes'. Further, a graphical representation of the same information can be provided in a conference display, which can assist participants and help to channel conversations or provided later on as part of meeting notes.

Consider Figure 1, below, which is a diagram illustrating a system flow that may be utilized to facilitate speaker determination and attribution for topics that may be discussed in a virtual/hybrid meeting in accordance with the techniques of this proposal.

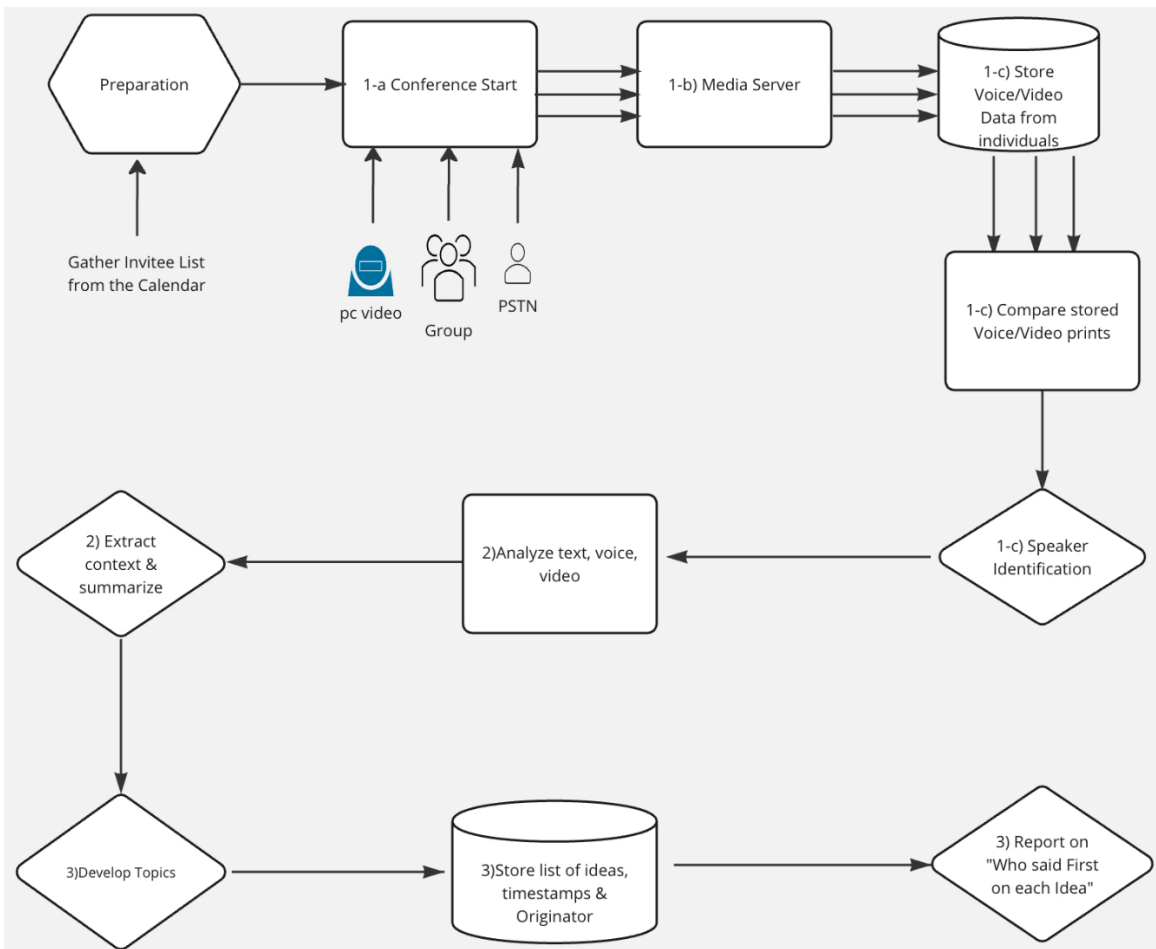


Figure 1: Speaker Determination System Flow

With reference to Figure 1, the speaker determination and attribution system flow may involve four main steps that may be performed: 1) speaker recognition; 2) capture and summarization of natural language; 3) identification and development of topics/themes; and 4) generating/outputting information/reports identifying speakers of different ideas discussed during a meeting.

For the first step involving speaker recognition, a logic engine can be utilized to identify a speaker with a high degree of confidence. For example, consider that a conference or virtual event is in progress with multiple participants. The participants could be of various types:

- I. Individual participants (with or without a camera on);
- II. Multiple participants behind a single End-Node (with or without a camera on);

- III. Participants connected through a Public Switched Telephone Network (PSTN) (audio only); or
- IV. Any combination of all three.

The media server, regardless of the process through which each of the participants joins the meeting, can gather the list of participants as individual channels or virtual nodes, and can then identify the channels and individuals. In some instances, individuals can be identified through their login process, such as identification that can be facilitated through a login or the end node through which an individual logs into a meeting. In some instances, individuals may be identified through the use of stored voice prints, such that received voice streams can be compared to pre-stored voice data (e.g., voice prints) in order to identify individuals. In still some instances, real-time voice biometrics (e.g., speech-to-text (STT), text extraction from an attendee's introduction, and/or stored voice prints) can be used to identify individuals. In still some instances, facial recognition, which may be available for participants utilizing cameras, may be used to identify individuals. For example, at a room level, facial recognition may be provided using people focus capabilities and "tiling" of individual participants.

For the second step involving the capture and summarization of natural language, the media server has access to every stream or channel (virtual node) in which a participant or group of participants are together in the meeting. For this proposal, the continuous set of spoken words by a single person is considered a "stream". Along with the stream, various other pieces of metadata are also captured, importantly, the speaker of a given stream as mentioned in step 1, above (Speaker Recognition).

In accordance with techniques of this proposal, every time a person speaks, a stream is captured and STT operations can be performed on the stream in order to generate a text version. Natural language processing (NLP) can be performed on the text to clean the data, for example, to remove stop words, and to tokenize and lemmatize the data. Python provides various libraries (e.g., nltk, etc.) that can be utilized to facilitate such operations.

From the clean data of each speaker's stream, a summary can be formed, context extracted and text summarization can be generated. There are various algorithms that can be employed to analyze a stream and extract the foundational ideas and the surrounding

context. For example, in some instances, extractive summarization algorithms can be utilized to identify significant sentences and add them to a summary of a meeting. In some instances, abstractive summarization algorithms may also be utilized, as different people may have different ways through which to convey a similar idea or concept, such that abstractive summarization may be used to derive/interpret concepts and develop a summary therefrom.

Based on one or more summaries that can be developed, latent semantic analysis can be utilized to facilitate 'singular value decomposition', which may provide for the ability to extract relationships between words. Large language model (LLM) processing based on Bidirectional Encoder Representations from Transformers (BERT) modeling, Generative Pre-trained Transformer (GPT) modeling, or the like may then be utilized to provide speaker identity and relative timestamps to one or more summaries for a meeting in order to establish the original contributor for each of one or more ideas expressed during the meeting.

The third step involves identification of topics and themes in the conversation. A meeting summary that is populated with corresponding speaker identities/timestamps can be broken down into a number of topics/ideas with their entire context such as the originator and a summary of what was said by each speaker, along with relative timestamps.

NLP techniques such as Term Frequency Inverse Document Frequency (TF-IDF) (between streams), TextRank, Yet Another Keyword Extractor (YAKE), or the like can be leveraged in order to facilitate identifying topics and generating themes for a meeting based on one or more summaries generated for the meeting.

Topics between speakers can be compared to determine whether similar or dissimilar topics are discussed. Semantic analysis tools can be utilized to identify similarities or dissimilarities of topics in two different streams spoken by two different speakers. With these tools, the topics from different streams can be associated with each other based on similarity in order to create themes and identify the original contributor for each theme.

These techniques can be done during the meeting and at the meeting continues and/or at the end of the meeting. If done during the meeting, a graphical representation of

such information may be provided in a conference display that may assist the participants and/or channel conversations. For example, such topics and attribution can be displayed in the corner of the display screen, such as a mind map, word cloud, flow chart, hierarchy of texts, etc., so that every participant can be aware of the topics, sub-topics, and attributes during the meeting. This may also facilitate the participants to add to the conversation and the topics.

Likewise, the summary, topics and its attribution can be recorded for use after the meeting such that the attribution can be available as the meeting progresses and/or after the meeting. Imagine, for example, that every stream from every participant is depicted as a word cloud. For example, if there are five participants in a meeting, then five word clouds could be provided with every topic/word highlighted and time-stamped, in which the topics/words could then be analyzed utilizing, the aforementioned statistical analysis to determine originators, attributions, and summaries.

Techniques proposed herein can leverage any combination of machine learning (ML) and/or artificial intelligence (AI) AI libraries, tools, and/or the like to achieve speaker identification/attribution for topics/ideas discussed during a meeting. Apart from the ones mentioned above, any of the following tools/libraries can be used in specific scenarios to determine topics, outcomes, engagement, attribution, or even sentimentality in accordance with techniques proposed herein:

- Named Entity Recognition (NER): Identifying and categorizing entities, such as people, organizations, and locations, in text data to attribute credit to individuals who make significant contributions;
- Part-of-Speech (POS) Tagging: Tagging each word in a text data with its part of speech to identify the verbs or nouns associated with particular ideas or contributions and attribute credit to the person who first introduced or developed them;
- Sentiment Analysis: Identifying a sentiment or tone of text data to attribute credit to individuals who consistently express positive or negative opinions about particular ideas or contributions; and/or

- Topic Modeling: Identifying the underlying topics or themes in text data to attribute credit to individuals who make significant contributions to these topics or themes.

The proposal described herein has a variety of uses. For example: while different people may have different ways of conveying an idea, such different ways of speech should not prevent the said people from being acknowledged as having originated the idea. Thus, techniques of this proposal provide for summarizing and developing topics based on each individual stream and then attaching other subsequent speakers to topics with their contributions to the topics.

Lack of identification of ideas is a problem acutely felt by the female gender or by non-native English speakers, whose ideas may get less traction when presented, as compared to others who presents the same idea. While some of this lack of identification may be attributed to bias, many issues that may be created by a lack of identification of ideas can be addressed by techniques of this proposal. Consider, for example, a scenario as shown below in Figure 2.

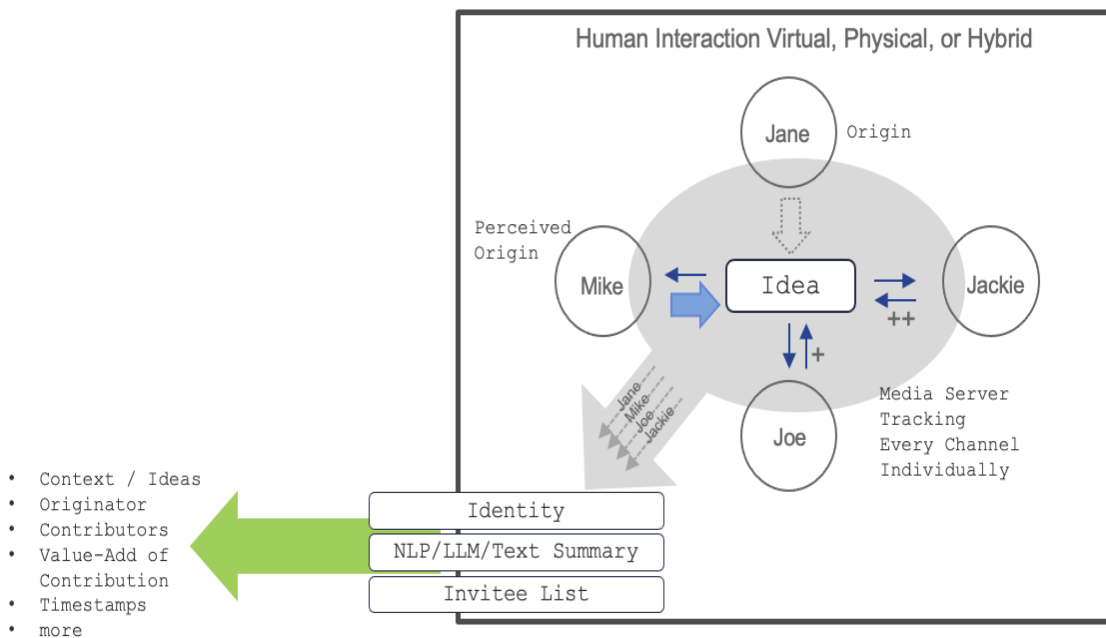


Figure 2: Example Scenario

As shown in Figure 2, consider that for a particular meeting, a participant, Jane, was an active participant and the originator of the “idea” or “topic” but was not noticed

(e.g., due to noise of the meeting, voice volume, speaking style etc.). Shortly after and during the course of the meeting/conversation, another participant, Mike, mentions the idea and takes credit for it. This could have legal, business or diversity implications. However, such an issue can be addressed by techniques of this proposal such that, if not confidently, then at least with a high level of certainty, Jane can be correctly identified as the source of the idea or topic.

As a summary, the techniques proposed herein may facilitate generating dynamic meeting summary and attribution list that identifies, tracks, and records all topics, themes and attributions in a given conversation. Real-time comparisons can be performed for topics in order to determine similar and dissimilar topics to form themes. In some instances, a graphical representation of such information can be provided in a conference display to assist the participants and channel the conversation. Such topic/idea identification/attribution may find applicability in an array of scenarios, such as fair attribution, measuring employment engagement, diversity inclusion and many more.