June 2023

ENDPOINT POLICY ENFORCEMENT OPTIMIZATION AND DYNAMIC ENVIRONMENT ASSESSMENT DURING VIDEO CONFERENCES

Derek Engi

Gonzalo A Salgueiro

Julie Allen

Pascale Delaunay

ENDPOINT POLICY ENFORCEMENT OPTIMIZATION AND DYNAMIC ENVIRONMENT ASSESSMENT DURING VIDEO CONFERENCES

AUTHORS:
Derek Engi
Gonzalo A Salgueiro
Julie Allen
Pascale Delaunay

ABSTRACT

Techniques are presented herein that support an audio and video policy framework that utilizes available sensors (encompassing both true Internet of things (IoT) sensors as well as the identification of beacons from mobile phones, consumer laptops, etc.) in conjunction with historical meeting data to suggest and enforce policies during teleconference meetings. By taking stock of the immediate environment surrounding a meeting participant, such a framework can enforce the audio and video options that are available to the user to align with secure practices and confidentiality protections. Aspects of the presented techniques support the identification of "safe spaces" through the use of Global Positioning System (GPS) coordinates, the detection of an increased presence of consumer devices for the purpose of suggesting and applying audio and visual security profiling, and the ability to send a notification to a telepresence user in the event that predicted meeting content may be too sensitive based on the environment of the meeting participant.

DETAILED DESCRIPTION

Techniques are presented herein that leverage consumer and industrial sensors for the purpose of suggesting and enforcing policies in a teleconferencing environment. Based on the historical usage patterns of a user, audio and visual input and output options, and the dynamics of the environment that is surrounding a user, the presented techniques propose optimal and secure policies that may ensure privacy and limit distractions for a user in a teleconferencing environment.

The notion of being able to identify beaconing devices and evaluate the landscape of sensor usage in an environment is well known. The techniques presented herein extend

beyond that capability by utilizing a changing environment to dynamically apply an audio and visual policy within a teleconferencing context based on multiple factors, including the evaluation of an environment.

The presented techniques support a method under which dynamic environmental factors and historical meeting information may play a role in the suggestion and enforcement of audio and visual profiles during a teleconference. Based on the level of sensor activity in a given area, the physical location of a meeting participant, and meeting insights (such as participants and content), the presented techniques may enforce a policy at the endpoint level to ensure privacy and secure communications.

Figure 1, below, illustrates a process that is possible under the techniques presented herein through which environmental factors (such as the physical location of a participant) as well as surrounding sensor data may establish an environmental profile for a meeting participant in their current location.

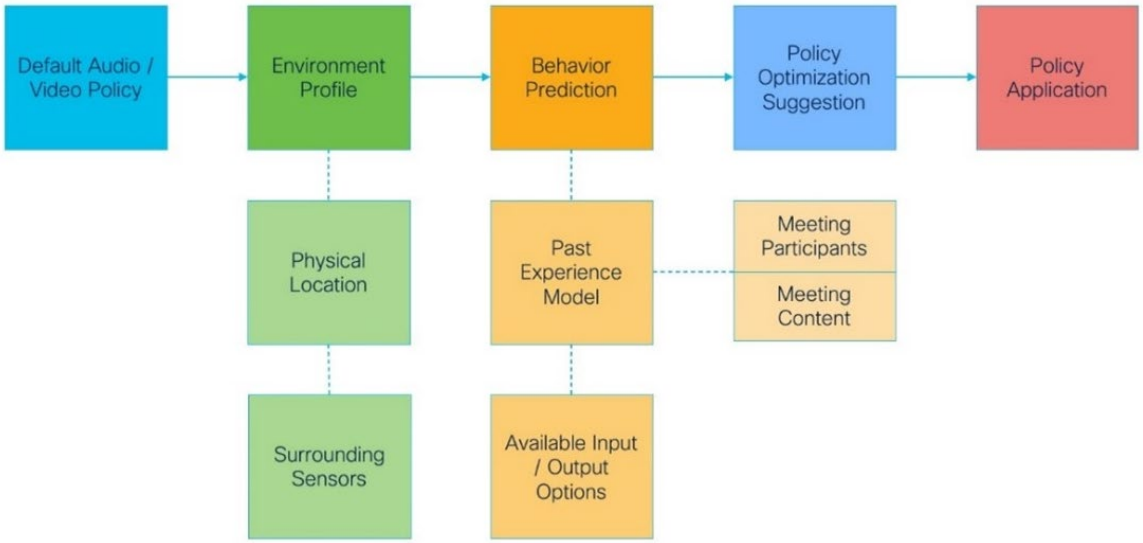

Figure 1: Determination of Policy Application Based on Sensor and Experience Models

Within the context of the techniques presented herein, the term sensor may encompass both true Internet of things (IoT) sensors as well as the identification of beacons from mobile phones, consumer laptops, etc. The presented techniques utilize the dynamic nature of a local sensor environment for the purpose of identifying when a suggestion or a

decision may be applied to better optimize the user experience of a primary meeting participant and limit distractions to other meeting participants.

Behavior prediction under the presented techniques may evaluate the previous experience that a user had for a given meeting based on the participants of, and the content for, the meeting. As an example, if it is common for the primary meeting participant to have a morning stand-up meeting while taking their dog for a walk, the behavior model may suggest an audio and visual policy comprising an increased weighting for the use of a headset and the disabling of video.

Additionally, the policy optimization suggestion and application models of the presented techniques may recommend a policy profile to a user or enforce an audio and visual policy based on criteria that are set by a teleconference administrator.

The techniques presented herein can be further understood through several illustrative scenarios. The different scenarios depict how a change in a sensor or environmental landscape can affect the policy recommendation and enforcement actions.

For a first scenario, Figure 2, below, depicts a home office user who has pre-registered their environment's geolocation data as a "safe space."

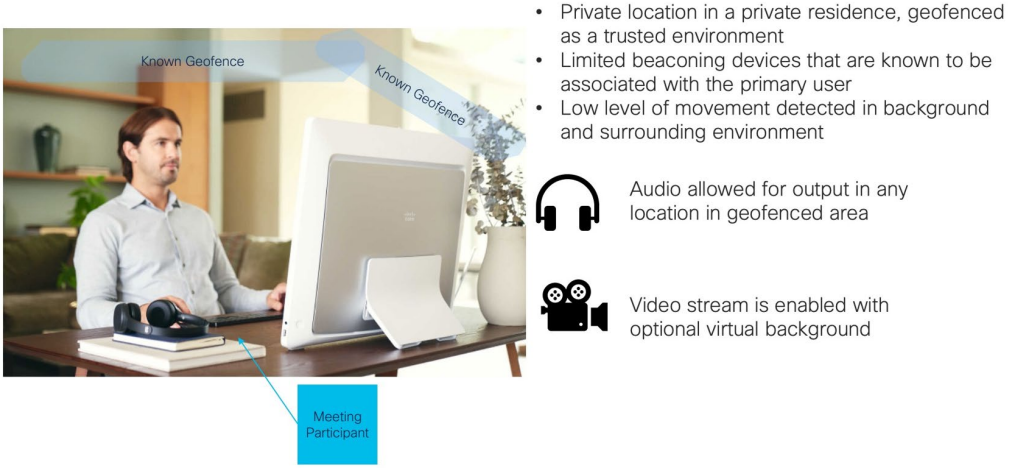

Figure 2: User in a Home Office

Under this scenario, during an initial profiling activity the sensors that are in the instant environment may be cataloged as known devices that are in the proximity of the user. Since there is a low level of predicted movement and a high level of known devices

in the immediate area, the suggested audio and visual policy for the user may be less restrictive based on the previous meeting conditions and content.

For a second scenario, Figure 3, below, illustrates the movement of a meeting participant into a semi-private area in a public space that is outside of a known geofenced area that has been previously registered as a "safe space."

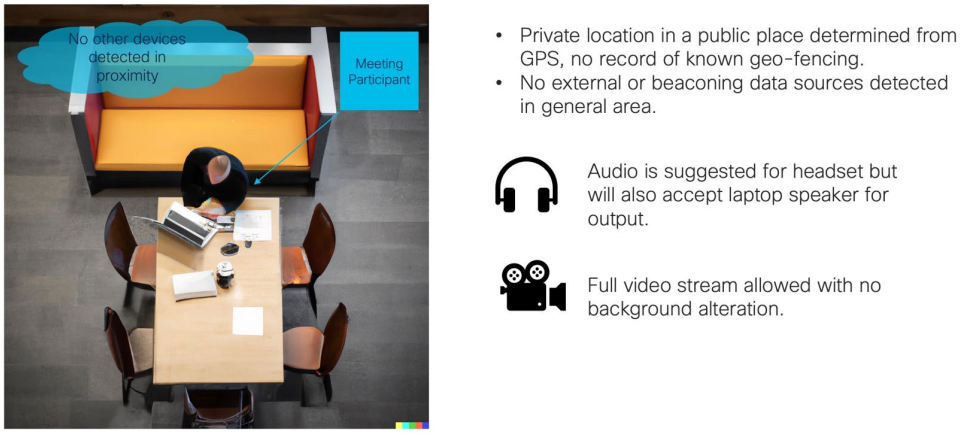

Figure 3: Meeting Participant in an Empty Coffee Shop

Under this scenario, while the user is in a public place the lack of sensor and environmental data may allow for a semi-restricted policy. The user may be encouraged to utilize a headset instead of a speaker for audio and any video streaming of the user may be allowed based on the expected meeting participants and content.

A third scenario, which provides an example of the dynamic assessment of a policy profile, encompasses a coffee shop that has an increasing number of patrons, as shown in Figure 4, below.

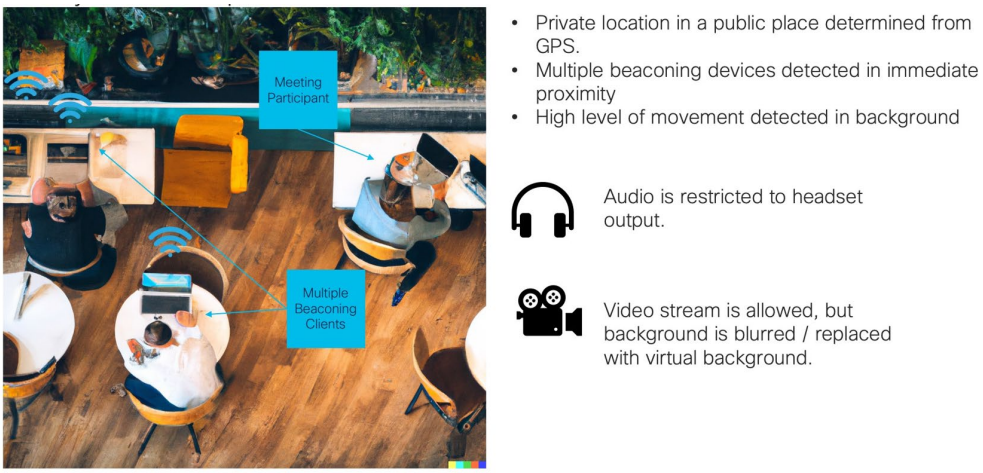

Figure 4: Meeting Participant in a Busy Coffee Shop

Under this scenario, as more sensors and beaconing devices are detected in the local proximity of the original meeting participant, the audio and visual profile may be enforced to redirect audio to the user's headset and restrict video streaming, or enforce a virtual background so as not to distract the other meeting participants. Additional audio optimization features (such as background noise removal, automatic muting, and the gating of the participants audio) may be enabled to account for the uptick in activity in the area around the user.

Furthermore, if there is historical information as part of the behavior model that indicates that this particular meeting is about to cover sensitive content, a warning or a notification may be displayed to the user indicating that the current location may not be ideal for discussing sensitive topics.

The techniques that are presented herein encompass a number of novel elements, several of which will be described below.

Under a first element, aspects of the presented techniques dynamically and continuously assess an environment for the presence of sensors and devices that could influence a decision regarding the discussion of sensitive topics during a teleconference meeting.

A second element encompasses the use of a surrounding environment and sensor information to suggest and enforce an audio and visual policy during a teleconference

meeting while a third element encompasses the use of historical meeting information (such as meeting participants and content) to suggest and determine the optimal audio and visual policies for a teleconference meeting.

A fourth element encompasses the ability to determine "safe spaces" through the use of Global Positioning System (GPS) coordinates and the ability to identify an increase presence of consumer devices for the purpose of suggesting and applying audio and visual security profiling.

And a fifth element encompasses the ability to send a notification to a telepresence user in the event that predicted meeting content maybe too sensitive based on the environment of the meeting participant.

It is important to note that some existing solutions may recognize where a person is based on their location through GPS coordinates and, if they are closer to a desk phone that may be used as the termination point for providing the audio and visual elements for a meeting, be able to transition to one audio profile over another based on such a location determination. However, none of those solutions encompass using consumer devices in the surrounding environment to justify the enforcement of, or the suggestion of, a particular audio and visual policy. The techniques presented herein may employ location details as a high-level method of determining the physical location of a meeting participant and identifying a public place from a private residence, but it is important to note that those techniques go well beyond GPS and geofencing to determine the potential security implications that could be raised by using certain audio or visual profiles during a meeting.

As described and illustrated in the above narrative, an audio and video profile may, according to the presented techniques, be defined as comprising the options that a meeting participant has available to them to actually participate in a meeting. Those options may include, for example, streaming video, a virtual background, no video, audio through laptop speakers, audio through headphones, no audio, etc.

For purposes of further illustration, Figure 5, below, depicts elements of a meeting where a user is working remotely while in a coffee shop.

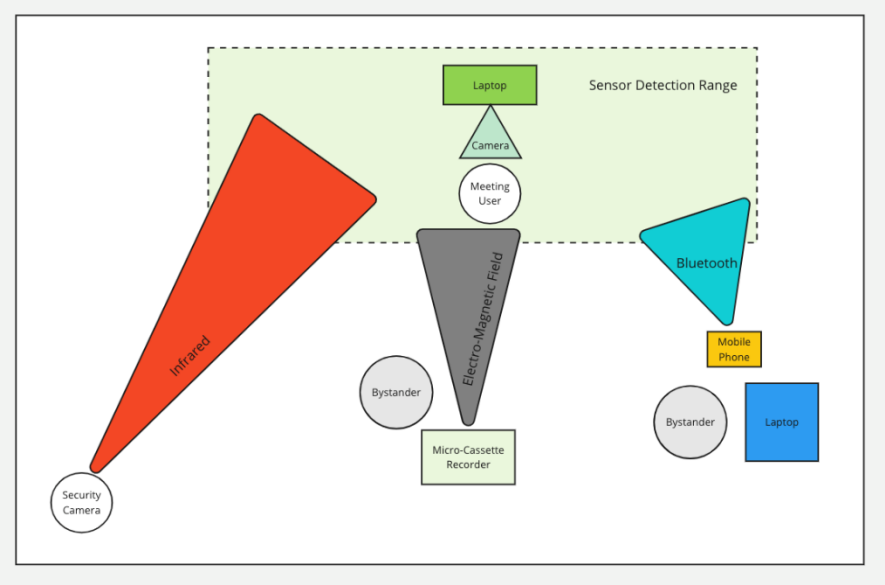

Figure 5: Meeting Participant Working Remotely

The techniques presented herein support the use of sensor detection methods to determine if consumer devices could potentially eavesdrop on the meeting that is depicted in Figure 5, above, and assess the level of security risk by evaluating the content that was previously discussed with the meeting participants. Such a capability is a novel aspect of the presented techniques. By leveraging previous meeting information (as described above) the presented techniques may determine if a meeting content was general chit-chat or did it potentially contain confidential information (where the techniques may invoke the digital equivalent of "closing the door of an office to further a discussion").

Furthermore, the presented techniques may evaluate the area that is surrounding a meeting participant and, in response, dynamically adjust the options for enforcing audio and visual policies. With reference to Figure 5, above, such an activity may encompass a number of elements.

First, the presence of infrared radiation (as detected by, for example, the laptop of a meeting participant) may result in a limiting of content sharing or the dimming of the screen of the participant to avoid the potential exposure of sensitive data to a seemingly passive observer.

Second, the beaconing of a coffee shop patron's cell phone, or the magnetic field detection of a cassette recorder, may limit the audio output options of a user to force them

to employ a headset so that the audio from other meeting participants cannot be captured by seemingly passive observers.

Third, a fluctuation in the signal strength of Bluetooth devices in an area may confirm that individuals are walking around the coffee shop and the background of the meeting participant's video may need to be blurred or stopped all together to reduce any distraction to the other meeting participants.

When used in conjunction with historical content and the behavior of a meeting's participants, physical location, and the items in the immediate environment, the presented techniques look to maximize privacy by taking precautions to limit the exposure of potentially sensitive material. Such precautions may be enforced through an ability level that is allowed for an end user regarding their sharing or receiving of audio and visual data and the permitted output methods.

In summary, techniques have been presented herein that support an audio and video policy framework that utilizes available sensors (encompassing both true IoT sensors as well as the identification of beacons from mobile phones, consumer laptops, etc.) in conjunction with historical meeting data to suggest and enforce policies during teleconference meetings. By taking stock of the immediate environment surrounding a meeting participant, such a framework can enforce the audio and video options that are available to the user to align with secure practices and confidentiality protections. Aspects of the presented techniques support the identification of "safe spaces" through the use of GPS coordinates, the detection of an increased presence of consumer devices for the purpose of suggesting and applying audio and visual security profiling, and the ability to send a notification to a telepresence user in the event that predicted meeting content may be too sensitive based on the environment of the meeting participant.