

A Framework of a Multi-User Voice-Driven BIM-Based Navigation System for Fire Emergency Response

Hui Zhou, Mun On Wong, Huaquan Ying, Sanghoon Lee*
Department of Civil Engineering, The University of Hong Kong, Pok Fu Lam Road, Hong Kong
shlee1@hku.hk (*Corresponding author)

Abstract. Navigation support is of significant importance for fire evacuation and rescue due to the complexity of building indoor structures and the uncertainty of fire emergency. This paper presents a framework of a multi-user voice-driven building information model (BIM)-based navigation system for fire emergency response. Classes of the navigation system is first defined being consistent with the open BIM data standard (i.e. Industry Foundation Classes, IFC). A string-matching method is then developed to generate a navigation query from each voice navigation request based on the Levenshtein distance and Burkhard and Keller (BK)-tree of a fire navigation associated lexicon. With the semantic information of the location in the navigation query, the spatial geometric information of the location is extracted from the BIM model and visibility graph-based route plans for multiple users are generated. To deliver the route planning to building users in an intuitive and direct manner, patterns of different voice prompts will be designed to automatically broadcast the navigation route step by step. Finally, the proposed navigation system will be validated with virtual reality (VR) based experiment.

1. Introduction

Delayed evacuation service is more likely to cause a high proportion of emergency casualties and property losses, especially in residential and high-rise buildings (Purser and Bensilum, 2001). Navigation support is a crucial link in the chain of evacuation services to prepare routing plans and provide path-finding guidance for fire evacuation, rescuing trapped people and relocating dangerous goods. In practice, a static 2D route map is commonly used for navigation planning in case of a fire. However, it is not reliable in terms of effective and efficient navigation in fire emergencies with high uncertainty. Actually, extensive information about building facilities, occupants, fire and dangerous materials are required by first emergency responders for decision-making and response operations (Gökdemir, 2011). Building information model (BIM), as a comprehensive repository with structured building data, has been used in existing indoor navigation systems (Ivanov, 2017). For example, Lertlakkhanakul et al. (2009) developed a BIM-based system called “GongPath” for indoor pedestrian navigation. Lin et al. (2013) retrieved IFC BIM data to create a grid graph to support path planning in 3D indoor spaces, using the fast-marching method. However, their navigation systems or methods were generally designed for a given pair of locations. For the navigation in fire emergency response, a more comprehensive system including the determination of location information for the pair of origin and destination in each navigation query needs to be built.

Wang et al. (2015) built a web-based prototype of equipment maintenance in fire emergencies. The prototype was delivered to users by mobile devices where they can input their location information by finding and clicking the corresponding building components, which in turn may make an eye-off-road distraction for users. Another commonly used personal localization mechanism within existing navigation systems is based on sensing technologies. Yoon et al. (2016) developed a Wi-Fi-based system to track the movement of victims for localization and assessment. Cheng et al. (2017) created a Bluetooth-based sensor network to monitor information of temperature, smoke and signal strength for providing early alarm services and evacuation tracking. Rueppel and Stuebbe (2008) integrated wireless LAN (WLAN), ultra-wide-band (UWB), and radio frequency identification (RFID) techniques to develop a hybrid

system to support indoor navigation in complex buildings. In such mechanisms, users are required to use mobile devices to scan the sensing tags or capture the signals delivered by position sensors. However, it may also cause dangers of eye-off-road distraction for responders on the path (Zhang et al., 2017). Moreover, sensor-based devices may be lacking in current buildings or be easily broken by fires or high temperatures, and the field of vision to find a sensor tag may be blurred by smoke. In addition, it is ignored that responders under emergencies can take an initiative to send evacuation or rescue queries with their location information using direct and prompt natural language speech. In regard to voice-based localization, Ivanov (2017) applied natural voice-based queries to retrieve relevant location data to develop a navigation system for visually impaired people. The navigation system was implemented with simulations in a 2D environment where a 2D escape route was generated as the only navigation assistance. However, such a route representation is not convenient enough for fire responders, especially for trapped occupants who may not be able to keep rational cognition and judgment under emergencies. To address these issues, voice-driven navigation queries and commands can be applied as a potential method because they are more natural and intuitive for users under fire emergencies. Moreover, scenarios for collaborative fire response among multiple users were seldom discussed in traditional fire emergency navigation systems.

Therefore, this paper aims to propose a framework for a multi-user BIM-based voice-driven navigation system for fire emergency response. Virtual Reality (VR) technology will be adopted to validate the proposed systems since it enables people to be immersed in the virtual environment (Zou et al., 2017). The rest of this paper is organized as follows. Classes of the proposed navigation system are defined in Section 2. Section 3 elaborates the two modules of the navigation system. Specifically, Section 3.1 presents the solutions to voice recognition and navigation query generation for voice navigation requests. Section 3.2 shows the developed approach of the BIM-based navigation model and the potential ways for voice navigation command generation among multiple users. Section 4 introduces the probable design of VR-based experimental validation for the proposed system. Finally, the paper is concluded in Section 5 and future research work is also presented.

2. Classes of the Proposed Navigation System

The classes of the proposed navigation system are composed of two aspects: users and facilities (as presented in Figure 1), for which objects are identified based on the information requirements for each module of the navigation system. The class of users contains objects of residential occupants, firefighters, and facility managers. For facilities, classes for building elements, transport elements, firefighting devices, hazardous materials, and embedded sensors are defined. The defined classes are identified by the attribute “ID” and have the general attribute “Location” or “Geometry Info”, which refers to a 3D coordinate data in the global coordinate system of a building. Meanwhile, specific attributes are required for some classes. For the space and transport element classes, the attribute “Accessibility” is used to capture the information whether a space or transport element is accessible or not under emergency situations. The composite material information of each building element (e.g. fire resistance rating of composite materials) are specified in their attributes. For firefighting devices, the device type and quantities of devices are required. For hazardous materials, material information such as material type and toxicity, and the storage volume is defined within the system.

The semantics and topology relationships between the defined classes are consistent with those in the Industry Foundation Classes (IFC) schema, which is an open standard BIM data format, in order to use an IFC BIM model as a data hub of the proposed navigation system. In this paper, IFC4 is used as the base specification. Figure 2 presents the IFC data structure that

supports the required information extraction for the defined classes. It is noted that most of the defined classes can be mapped with the IFC instances, with reference to IfcBuilding, IfcBuildingStorey, IfcSpace, IfcBuildingElement, IfcTransportElement, and IfcDistributionElement. However, the remaining classes such as the fire extinguisher in the firefighting device class and the hazardous material class, have no corresponding definitions in the IFC schema. In this case, the missing information is either stored in the database or added as an extension of IFC entity.

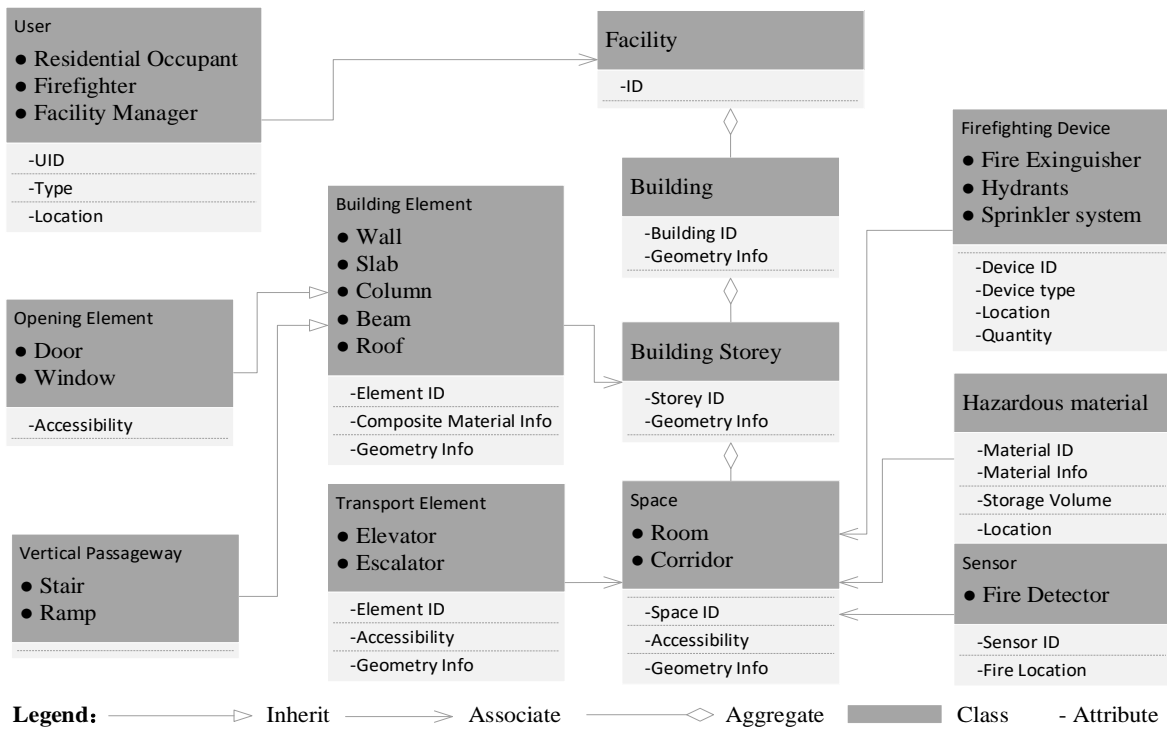


Figure 1: Classes definition of the proposed navigation system

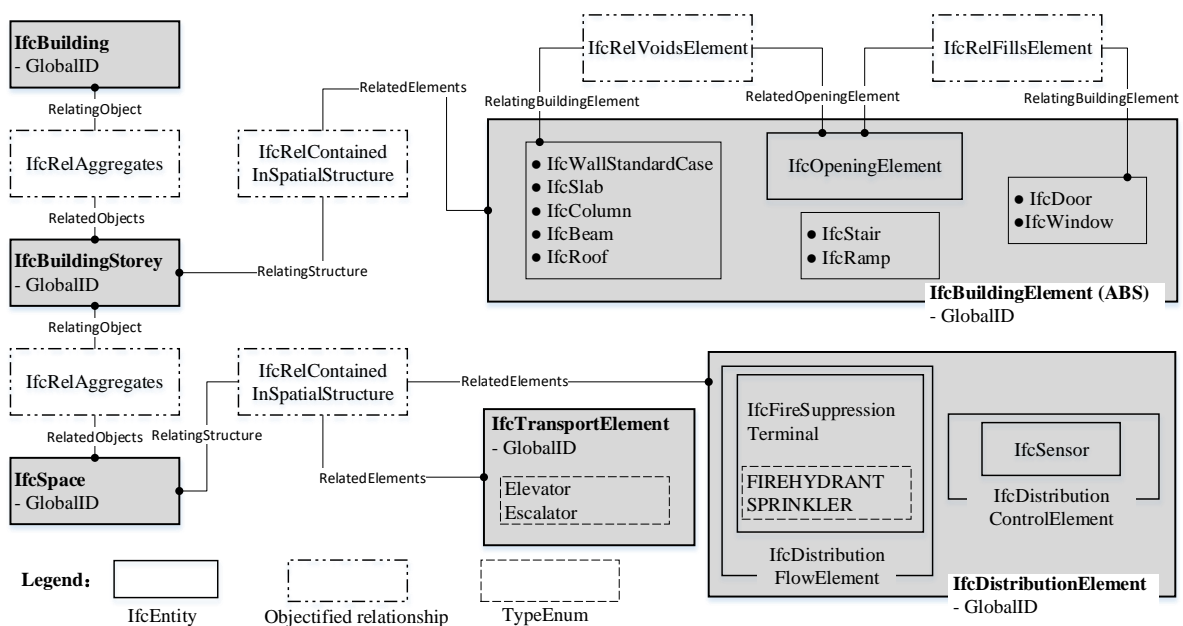


Figure 2: IFC data structure to support classes information extraction

3. The Framework of the Multi-User Voice-Driven BIM-Based Navigation System

The presented framework consists of two modules, as shown in Figure 3. The first module aims to recognize users' voice inputs and generate corresponding navigation queries. The second module is designed to retrieve the information required by each query from the BIM-based data hub and generate navigation models for different users. Details for each module are elaborated in the following sub-sections.

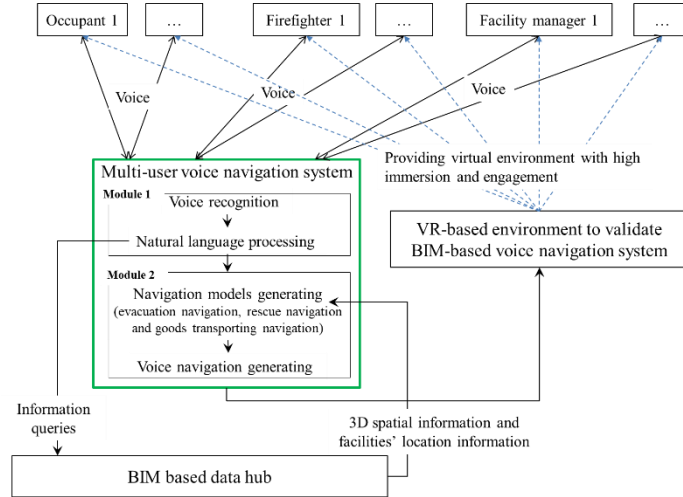


Figure 3: The framework of the multi-user voice-driven BIM-based navigation system

3.1 Module 1: Voice Recognition and Navigation Query Generation

3.1.1 Voice Recognition

Once a fire emergency occurs, users can communicate with the navigation system in natural English speech through mobile electronics which can record and play audio. The Google speech-to-text (STT) API (<https://cloud.google.com/speech-to-text/>) is then adopted to recognize and convert the audio streams received from users into sequences of words/sentences automatically. Based on its cloud storage, the Google STT API can provide synchronous voice recognition with significant accuracy (Harsur and Chitra, 2017; Kępuska, and Bohouta, 2017). The converted words/sentences are saved as .txt files and further used for navigation associated information extraction in Section 3.1.2. In this paper, it is hypothesized that effective network connectivity is available all the time to ensure real-time human-to-machine interaction.

3.1.2 Navigation Query Generation

To generate navigation queries from the recognized speeches, a fire emergency navigation-oriented lexicon is first constructed as a dataset for textual information processing. The predetermined vocabulary includes four categories: users, locations, request details, and interaction dialogs. Each category is composed of one to several sub-categories, which are further labelled under each category with a prefix of the category index, as shown in Table 1. Particularly, the location vocabulary represents the semantics of building components and is labelled sequentially according to the building components' relationships, which helps to enable more accurate indoor localization by organizing the location vocabulary from coarse to fine (e.g. from a building storey level to a space level).

Based on the constructed navigation-oriented lexicon, converted sentences are then input for detecting and extracting navigation-associated keywords using the string-matching approach based on "Levenshtein distance" (Levenshtein, 1966; Salehinejad et al., 2017). To reduce the

search scope for each target word, the predefined navigation associated vocabulary is organized in the structure of a BK-tree (Burkhard and Keller, 1973). Concepts of the Levenshtein distance and BK-tree are briefly introduced as follows.

Table 1: Vocabulary context for the fire emergency navigation-oriented lexicon

Category (Index)	Label: ID	Example words/phrases
User (U)	-User Type: U01	Residential occupant, firefighter, facility manager
Location (L)	-Building storey: L01	Floor, level
	-Space: L02	Space, room, toilet
	-Building structure element: L03	Wall, column, slab, beam, roof
	-Opening element: L04	Door, window, exit
	-Vertical passageway: L05	Stair, ramp
	-Transport element: L06	Elevator, escalator, lift
	-Firefighting devices: L07	Fire extinguisher, hydrant, hose reel, sprinkler
	-Hazard material: L08	Oil, gas
	-Sensor: L09	Fire sensor, smoke sensor
	-Landmark: L10	Desk, statue, reception
Request Detail (RD)	-Action command: RD01	Walk, go, approach
	-The verb for request: RD02	Find, look for, search, relocate
	-Where adverb: RD03	On, in, inside, outside, at
	-Preposition for the position: RD04	From, to
Interaction Dialog (InD)	-Confirmation: InD01	Yes, no, yep, sure, okay

For two strings S_1 and S_2 with the lengths $|S_1|$ and $|S_2|$, respectively, $L(S_1, S_2)$ indicates the distance between the first i character of S_1 and the first j character of S_2 , such that

$$L(S_1^i, S_2^j) = \begin{cases} L(S_1^{i-1}, S_2^{j-1}), & \text{if } S_1^i = S_2^j \\ 1 + L, & \text{otherwise} \end{cases} \quad (1)$$

Where

$$L = \min(L(S_1^{i-1}, S_2^j), L(S_1^i, S_2^{j-1}), L(S_1^{i-1}, S_2^{j-1})) \quad (2)$$

Specifically, $L(S_1^i, S_2^0) = i$ and $L(S_1^0, S_2^j) = j$, while S_k^0 represents an empty string. $L(S_1^i, S_2^j)$ is the minimum number of single-character operations (i.e. insertions, deletions, or substitutions) needed to match S_1^i and S_2^j .

An arbitrary element a is selected as the root of BK-Tree. $d(a, b)$ is defined as a distance function for each pair of elements of a set of objects. The m -th sub-tree branch is recursively built of all elements b such that $d(a, b) = m$. Figure 4(b) shows the constructed BK-tree of a set of navigation associated vocabularies with the word “floor” as the root. The number on edges indicates the Levenshtein distance between two strings. The root word can be determined more reasonably in accordance with the frequency of words used for indoor navigation.

A two-step approach is developed for the navigation query generation, based on principles of the Levenshtein distance and BK-tree. As shown in Figure 4(a), the first step extracts the associated information from converted sentences for navigation purpose. In detail, the sentences are first tokenized with word tags which represent their part of speech (POS) (e.g., noun, verb, preposition, etc.). Subsequently, noun phrases are extracted by predefined patterns (e.g. noun phrase: determiner + adjective + noun). Next step executes string matching based on the BK-tree word dictionary for nouns inside the extracted noun phrases. For a given word query, the distance D from the current root is first calculated. Then the words located in the D th sub-tree branch are determined as the candidate list, while the sub-tree root is changed into the current root. Progressively, the target word can be detected when the distance from the query word to

obstacles exist, the connections between the vertices of obstacles or buffers and the doors are created. The accessibility of the graph nodes is changed in accordance with the dynamic information in fire emergencies, such as people distribution, fire and smoke spread trends, etc. Figure 6 presents an example of a BIM-based navigation for fire response in a sample building based on the proposed method.

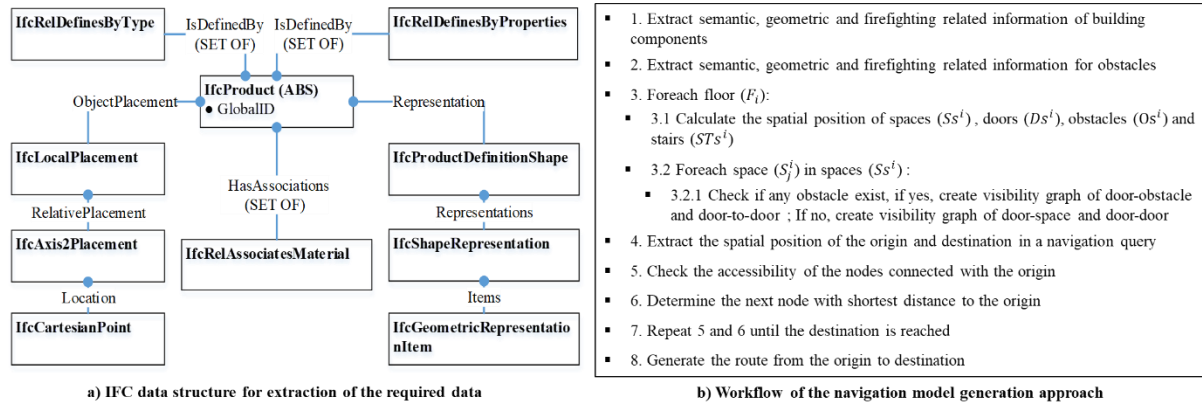


Figure 5: Diagram of the navigation model generation

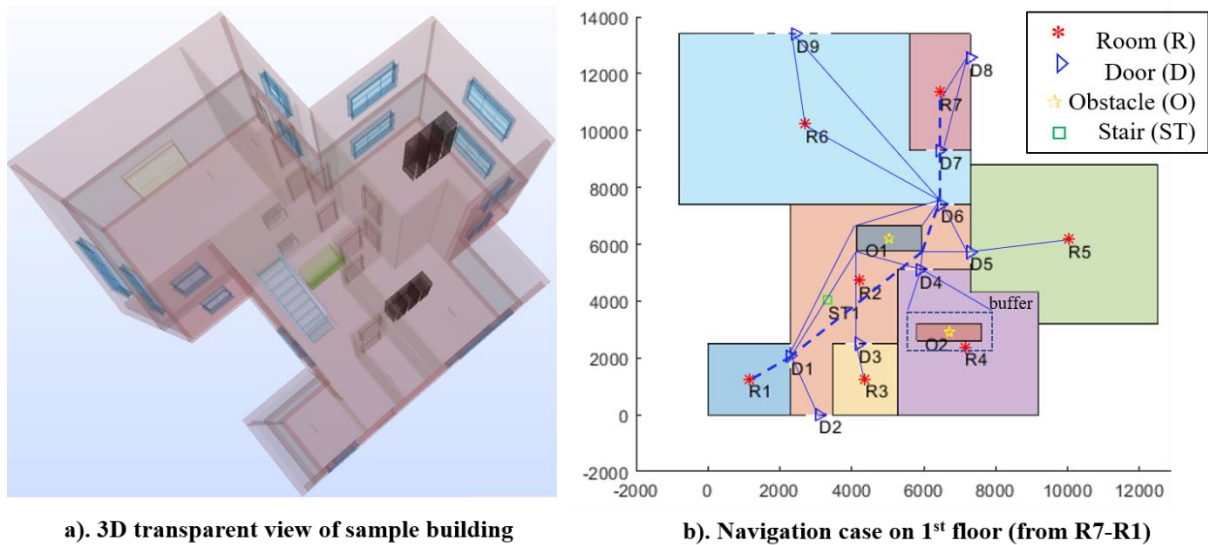


Figure 6: BIM-based navigation for fire response in a sample building

3.2.2 Voice Navigation Command Generation

Patterns of voice prompts for navigation to the route nodes are designed based on the types of graph nodes and their connections in the visibility graph (i.e. door-to-space, door-to-door, and door-to-obstacle). As shown in Figure 7, the voice prompt for navigation to a door could be “GO THROUGH THE” i th “DOOR ON THE LEFT” or “FRONT DOOR”. The pattern is automatically determined for each route node, as the uniform word vocabulary for the semantic information of the location has been defined with unique labels in Module 1. The patterns for voice navigation to other locations such as spaces, obstacles, and stairs are also defined in the similar manner. Based on the designed navigation patterns, the optimal route from the origin to the destination is automatically translated into textual navigation commands with the support from the developed vocabulary in Module 1. Then, the text is input into a voice synthesizer tool called “eSpeak” to generate voice navigation commands.

As the navigation model is initially developed to enable multi-user navigation services, a communication mechanism between firefighters, facility managers, and residential occupants

need to be predefined. Potential interactions could be rescue requests from the trapped people, hazardous materials relocation, and properties protection, etc. Figure 8 explains the potential interactions between multiple users with two navigation scenarios. The first scenario is that a trapped occupant sends a help request signal with his/her location to the navigation system which then delivers the message to firefighters and further plan the shortest route for rescue. Also, facility managers and residential occupants inform the navigation system of the locations of potential hazards and important properties in the building. This helps to update the building information in the navigation system and further plan shorter routes for users. Within these interactions, the navigation system requests users to confirm the query for the current location and ultimate destination, and it sends a message delivery confirmation to users. For such cases, sentences of general dialogs (e.g. your message is sent to firefighters) are defined and sent at the appropriate time.

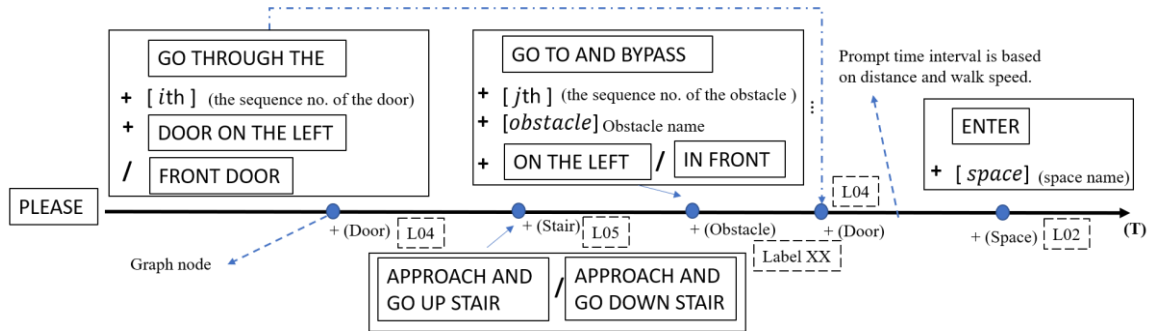


Figure 7: Voice prompt pattern for voice navigation command generation

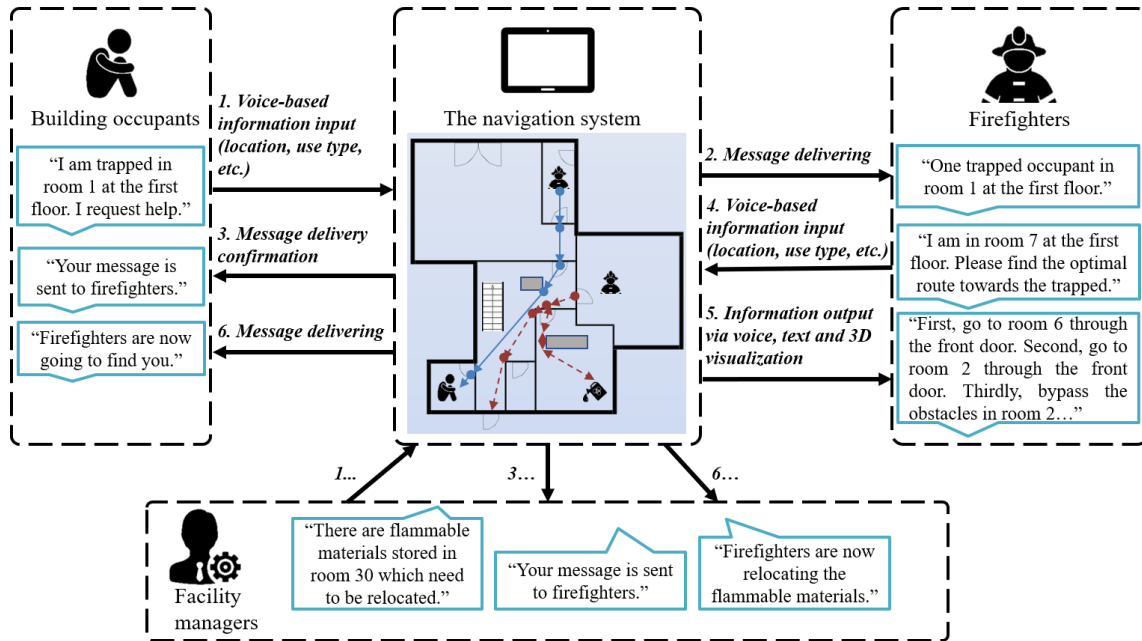


Figure 8: Multi-user voice-driven navigation cases with the assistance of the proposed system

4. VR-based Experimental Validation

Based on Modules 1 and 2, a multi-user BIM-based voice-oriented prototype for fire emergency navigation will be developed. Unity, a high-performance game development platform, will then be adopted to build a VR-based environment where people can be engaged in the fire emergency scenarios and interact with the navigation system as building users, as shown in Figure 9. To present full-scale immersive fire emergency scenarios, the VR-based environment will further be implemented in a CAVE system, as presented in Figure 9 (a).

The scenarios will be designed based on a preliminary study of potential reasons for several disastrous fire emergency cases in Hong Kong. Two groups of participants, namely the experimental group and the control group will be invited to finish designed tasks within the comparative experiment, assisted by the proposed navigation system or not, respectively. To realize voice navigation for multiple users, a multi-player High-Level API (HLAPI) as a Unity plug-in will be adopted for building the network of users' interaction. A Unity package named Tridify (<https://www.tridify.com/>) will also be used to ensure uniform coordinates between the VR environment and BIM models. Thus, the navigation system tested in the VR environment will further obtain the accurate position of participants in buildings through the tracking sensor.

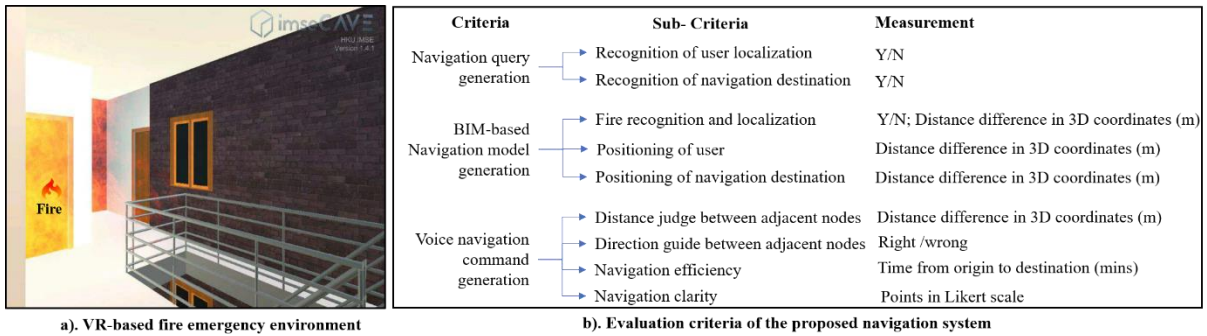


Figure 9: VR-based experimental validation of the proposed navigation system

To evaluate the performance of the proposed navigation system, assessment criteria will be built in terms of key module functions. Specifically, for navigation query generation, it can be assessed by whether the navigation system can recognize the semantic information of the users' locations and the navigation destinations. With regard to the BIM-based navigation model generation for multiple users, the proposed system will be tested to check whether it can recognize and localize the fire, users and navigation destinations based on the semantic information of locations in the navigation queries. For voice navigation command generation, the distance and direction guide between two adjacent nodes on the generated navigation route will be examined. Meanwhile, navigation efficiency and clarity are essential to the safe evacuation and rescue under fire emergencies. Figure 9 (b) illustrates the hierarchy of the evaluation criteria of the proposed navigation system. The corresponding measurement for each sub-criterion is also specified in Figure 9 (b). For instance, the time from the origin to the destination consumed by the experiment participants will be recorded to evaluate the navigation efficiency of the system.

5. Conclusions and Future Work

This paper proposes a framework of a multi-user voice-driven BIM-based navigation system for fire emergency response. This framework is expected to extend traditional BIM-based navigation approaches with intuitive voice-based localization and navigation prompt delivery. Specifically, a string-matching method is introduced to generate the navigation query from each voice navigation request based on the Levenshtein distance and BK-tree of a fire navigation associated lexicon. Approaches for dynamic route planning and intuitive navigation command generation for multiple users are proposed with the multi-user interaction definition and patterns of voice prompts. Therefore, the multi-user navigation system will be not only a navigation tool for planning the shortest route, but also an information sharing platform for involved participants to collaborate with each other. However, the approaches or algorithms involved in each module needs to be implemented with more fire emergency scenarios in real-world buildings. In addition, to avoid invalid voice input, the proposed navigation system hypothesizes that building users will only input navigation requests. For more general cases, sentiment analysis of the recognized voice information should be required. Moreover, deep

learning methods for the navigation query generation and the voice navigation command generation can be integrated with the proposed system to ensure great efficiency in navigation services.

6. Acknowledgments

The work described in this paper was supported by a grant from Graduate Collaborative Research Awards funded by Universitas 21.

References

- Burkhard, W.A. and Keller, R.M. (1973). Some approaches to best-match file searching. *Communications of the ACM*, 16(4), pp.230-236.
- Cheng, M.Y., Chiu, K.C., Hsieh, Y.M., Yang, I.T., Chou, J.S. and Wu, Y.W. (2017). BIM integrated smart monitoring technique for building fire prevention and disaster relief. *Automation in Construction*, 84, pp.14-30.
- Cloud Speech-to-Text. Accessed 30 March 2019, <<https://cloud.google.com/speech-to-text/>>.
- Cheng, J.C., Tan, Y., Song, Y., Mei, Z., Gan, V.J. and Wang, X. (2018). Developing an evacuation evaluation model for offshore oil and gas platforms using BIM and agent-based model. *Automation in Construction*, 89, pp.214-224.
- Gökdemir, N. (2011). Identification and representation of information items required for vulnerability assessment and multi-hazard emergency response operations. Middle East Technical University.
- Harsur, A. and Chitra, M. (2017). Voice based navigation system for blind people using ultrasonic sensor. *IJRITCC*, 3, pp.4117-4122.
- Ivanov, R. (2017). An approach for developing indoor navigation systems for visually impaired people using Building Information Modeling. *Journal of Ambient Intelligence and Smart Environments*, 9(4), pp.449-467.
- Këpuska, V. and Bohouta, G. (2017). Comparing speech recognition systems (Microsoft API, Google API and CMU Sphinx). *International Journal of Engineering Research and Applications*, 7, pp.20-24.
- Lertlakhanakul, J., Li, Y., Choi, J. and Bu, S. (2009). GongPath: Development of BIM based indoor pedestrian navigation system, NCM 2009 - 5th International Joint Conference on INC, IMS, and IDC, pp. 382–388.
- Levenshtein, V.I. (1966), February. Binary codes capable of correcting deletions, insertions, and reversals. In *Soviet physics doklady*, 10(8), pp. 707-710).
- Li, N., Yang, Z., Ghahramani, A., Becerik-Gerber, B. and Soibelman, L. (2014). Situational awareness for supporting building fire emergency response: Information needs, information sources, and implementation requirements. *Fire safety journal*, 63, pp.17-28.
- Lin, Y.H., Liu, Y.S., Gao, G., Han, X.G., Lai, C.Y. and Gu, M. (2013). The IFC-based path planning for 3D indoor spaces. *Advanced Engineering Informatics*, 27(2), pp.189-205.
- Purser, D.A. and Bensilum, M. (2001). Quantification of behaviour for engineering design standards and escape time calculations. *Safety science*, 38(2), pp.157-182.
- Rueppel, U. and Stuebbe, K.M. (2008). BIM-based indoor-emergency-navigation-system for complex buildings. *Tsinghua science and technology*, 13(S1), pp.362-367.
- Salehinejad, H., Barfett, J., Aarabi, P., Valaee, S., Colak, E., Gray, B. and Dowdell, T. (2017), October. A convolutional neural network for search term detection. In 2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC) (pp. 1-6). IEEE.
- Tashakkori, H., Rajabifard, A. and Kalantari, M. (2015). A new 3D indoor/outdoor spatial model for indoor emergency response facilitation. *Building and Environment*, 89, pp.170-182.
- Tridify. Accessed 30 March 2019, <<https://www.tridify.com/>>.
- Wang, S.H., Wang, W.C., Wang, K.C. and Shih, S.Y. (2015). Applying building information modeling to support fire safety management. *Automation in Construction*, 59, pp.158-167.
- Yoon, H., Shiftehfar, R., Cho, S., Spencer Jr, B.F., Nelson, M.E. and Agha, G. (2016). Victim Localization and Assessment System for Emergency Responders. *Journal of Computing in Civil Engineering*, 30(2), pp. 1–14.
- Zheng, Y., Liu, Y. and Hansen, J.H. (2017), June. Navigation-orientated natural spoken language understanding for intelligent vehicle dialogue. In *Intelligent Vehicles Symposium (IV)*, 2017 IEEE (pp. 559-564). IEEE.
- Zou, H., Li, N. and Cao, L. (2017). Emotional response-based approach for assessing the sense of presence of subjects in virtual building evacuation studies. *Journal of Computing in Civil Engineering*, 31(5), p.04017028.