

Mobile Navigation Guide for the Visually Disabled

Bo Huang and Nan Liu

A location-aware navigation system has been developed and implemented for the visually disabled or visually impaired; the system is designed to improve individuals' independent mobility. This self-contained, portable system integrates several technologies, including mobile personal digital assistants, voice synthesis, a geographic information system (GIS), and a differential Global Positioning System (DGPS). The system is meant to augment the various sensory inputs available to the visually impaired user. It provides the user with navigation assistance, making use of voice cues iterating contextual building and feature information at regular intervals, through automatic GPS readings and a GIS database. To improve the efficiency of the retrieval of contextual information, an indexing method based on road segmentation was developed to replace the exhaustive search method. Experimental results show that the performance of the system on searching the buildings, landmarks, and other features around a road has been significantly improved by using this indexing method.

To find their way around, human beings use their senses to obtain inputs from the surrounding environment. They use vision to orient their position and use a process known as navigation. Orientation refers to one's spatial relationship to various features in the environment in terms of physical distance and direction; navigation refers to the continual reestablishment of one's position along a preplanned route or with respect to a particular goal.

Visually impaired individuals do not have the ability to gather visible contextual information along a travel route. Over the years, aids such as wheelchairs, canes, and guide dogs have been employed to compensate for this disadvantage. The prototype developed in this project seeks to complement these aids by providing contextual and route information to the user with a self-sufficient, portable device.

Other navigation aids for the blind, such as tactile maps and devices that output spoken directions, can be inconvenient to carry and can be customized only manually, and then only with difficulty. In addition, most cannot indicate a user's location. Tactile maps are of low resolution, are difficult to acquire, and are difficult to use when moving. Spoken directions are somewhat easier to acquire—either by recording a friend's voice or by using the output of text-to-speech software applied to text directions—and somewhat easier to use. Spoken directions are a feasible solution; however, once recorded, they cannot be customized, and the location-sensing systems with which they have been integrated are not very portable.

Department of Civil Engineering, National University of Singapore, 10 Kent Ridge Crescent, Singapore 119260. Current affiliation: B. Huang, Department of Geomatics Engineering, University of Calgary, Calgary, Alberta T2N 1N4, Canada.

Transportation Research Record: Journal of the Transportation Research Board, No. 1885, TRB, National Research Council, Washington, D.C., 2004, pp. 28–34.

Advances in mobile technology have made compact devices such as mobile phones, tablet PCs, and personal digital assistants (PDAs) capable of handling sophisticated user-oriented applications. The device used in this research is the iPAQ PocketPC H3630 running on the Windows CE operating system (OS), preloaded with digital maps and a spatial database that models a user's environment. The geographic information system (GIS) data were developed on a desktop computer using ArcPad and ArcView, provided by Environmental Systems Research Institute (ESRI), and subsequently migrated onto the iPAQ. The PDA serves as a self-contained repository of location-specific information; it does not require an ongoing online connection with a central database or server, thus alleviating bandwidth, connectivity, and subscription cost problems. This stand-alone system is integrated with a Global Positioning System (GPS) receiver and audio earphones to optimize listening in an urban environment.

The idea of using GPS to aid in the navigation of the visually impaired was proposed more than a decade ago, in 1985, by Loomis (J. M. Loomis, Digital Maps and Navigation System for the Visually Impaired, Department of Psychology, University of California–Santa Barbara). Today most GPS applications requiring real-time positioning accuracy better than 25 m use a differential Global Positioning System (DGPS). Correction signals from a GPS receiver at a known, fixed location are transmitted by radio link to the mobile receiver and allow the mobile receiver to determine its position with an absolute positional accuracy on the order of 1 m or better. This accuracy is vital to the safety of visually impaired pedestrians, who may have to travel on narrow walkways.

Aside from the design issues of hardware miniaturization and cost, this project emphasizes the use of a spatial indexing method to replace the exhaustive search method when contextual building and feature information is being dynamically retrieved. Because all spatial data are preloaded into the PDA, accessing data by exhaustively searching the database takes considerable time. The indexing method discussed later in this paper utilizes the balanced tree (B-tree) (*I*) and greatly improves the search time.

This paper discusses (*a*) related work; (*b*) system design, particularly hardware and software components; (*c*) implementation issues, particularly indexing and route-finding methods; and (*d*) conclusions and future anticipated work.

RELATED WORK

Some of the earliest research done on navigation systems for the visually disabled took place in the 1980s, when Loomis and colleagues proposed the use of GPS, GIS, and auditory displays to aid the navigation of the visually disabled. In the proposal, a GPS receiver and a fluxgate compass were used to determine the head position and orientation for use in the virtual display. Their prototype

system consisted of three modules: a DGPS module to determine the traveler's position, a GIS spatial database of their test site, and a user interface that employed spatialized sound from a virtual acoustic display to convey information about the surrounding environment to the blind traveler.

A related commercial product from Arkenstone, Strider (2), consists of GPS receivers plugged into a notebook computer. This product provides detailed maps that cover most of the regions in the United States. A talking interface allows for "virtual exploration" and helps users learn the details of streets and routes from the comfort of their homes. This product has been under continuous development and is now marketed under the name of Braille Note GPS.

MOBIC (3, 4) is a GPS-based travel aid for the blind and elderly. The prototype used for testing was a handheld computer preloaded with digital maps and having limited wireless capabilities to obtain information from a database. It employed speech synthesizers to recite preplanned routes. It also included DGPS corrections provided through a mobile phone link and a compass for orientation information. The main difficulty users faced was handling disconnections with the central database.

The People Sensor is an electronic travel aid designed to address two issues important to visually impaired people: (a) inadvertent cane contact with other pedestrians and objects and (b) speaking to a person who is no longer within hearing range. The device uses pyroelectric and ultrasound sensors to locate and differentiate between animate (human) and inanimate (nonhuman) obstructions in the detection path (5). The People Sensor focused on ergonomics as a primary design concern and avoided employing an auditory module that could interfere with the user's hearing perception of the environment. The People Sensor does not provide positioning information and is not meant to be a stand-alone system.

Another GPS-based system is being developed in Japan as part of a research project by Makino and colleagues. This system is unique; it uses a mobile phone to link the user and the GIS database (6). The mobile phone transmits the traveler's GPS coordinates to the GIS computer at a central facility. The GIS then outputs through synthetic speech that is transmitted back to the traveler, providing information on his or her position. Use of a mobile phone link has the advantages of minimizing the weight and computing power that must be carried by the traveler and of simplifying the update of the spatial database. However, the system suffers from a maximum GPS reading error of 16 m.

Drishti is a wireless pedestrian navigation prototype that integrates wearable computers, GIS, GPS, and voice technologies. It was developed by Helal et al. (7). The system guides an individual using both static and dynamic data. Drishti uses commercial off-the-shelf hardware and software, including Xybernaut MA IV, ESRI's ArcSDE, and IBM's ViaVoice. The system allows intelligence perceived by the blind user to be added to the central server hosting the database. If GPS tracking is lost, positioning will make use of dead reckoning methods and the user's average speed of walking. This system is much dependent on wireless connectivity.

A portable traveling support system using image processing was proposed by Kaluwahandi and Takodoro (8). This system combines a pedometer, a 16-bit single-board computer, three response switches, and a buzzer, all worn in a waist pouch; a laptop carried in a backpack; a camera attached to the waist; and a loudspeaker. The microcomputer calculates the traveling locus by using (a) the data of the traveling distance and direction measured by the pedometer and (b) image processing using the laptop computer and the camera; it then compares this information to the stored traveling route to the destination. In this

system, there may be an inaccuracy in determining the user's direction due to distortion of the terrestrial magnetism in the magnetic sensor.

Tyflos, another intelligent system, helps a visually impaired user to be partially independent and to walk and work in a 3-D dynamic environment. The Tyflos system carries two vision cameras that capture images from the surrounding 3-D environment, either at the user's command or on an ongoing basis (video); the system then converts these images into verbal descriptions of each image and relays them to the user (9). The main user issue with Tyflos is the amount of equipment carried by the user; this equipment includes glasses, vision cameras, laser range scanners, a microphone, a portable computer, speakers, and a communications system.

Although image processing and computer vision techniques demonstrate great potential for aiding real-time navigation, lighting, weather, obstructions, and other environmental constraints have to be addressed before a system can be effective. Generally, accurate position and camera orientation are necessary but can be difficult to achieve.

All of the devices discussed so far have attempted to tackle the problem of modeling the world, either through image processing techniques or using GIS spatial databases. These devices are often complex, and users normally need to undergo training or have some technological know-how to operate them successfully. Many techniques have been used to track a user's travel route and position; from the research, DGPS appears to offer the most precise results, although it only works outdoors. One issue that has not been adequately addressed in the work discussed above is the core technique or method by which information about the ambient environment (e.g., the surrounding buildings and landmarks) is retrieved. This is an important issue: efficient retrieval of spatial data leads to fast response time, which significantly affects the safety and comfort of the blind traveler.

SYSTEM DESIGN

The navigation prototype in this study was designed to suit the needs of visually disabled users precisely. Commercial hardware and software were used to develop the prototype; the focus was on hardware miniaturization, self-sufficiency for the user, and efficiency.

User Requirements

For general end-users, the navigation system had to meet several basic requirements. The system had to be context sensitive; that is, it had to be able to recognize the user's current location and retrieve relevant feature and landmark information. This was achieved accurately with DGPS. The system had to be user-friendly; no training should be required to use it successfully. The system also had to address several needs specific to visually impaired individuals; these needs were determined on the basis of preliminary investigations undertaken in the MOBIC project (3, 4) and on a paper by Bradley et al. Visually impaired individuals expressed the need to know the directions to a destination prior to a journey; this helps them in contingency situations, allowing them to refer to passers-by for manual guidance. Because the individuals suffer from a lack of sensory perception, there is a need to compensate with nonspeech and speech output; most individuals prefer a system that combines speech output with vibration alerts (7). In addition, on average, visually impaired individuals require three times more directional information and almost nine times more descriptive

information than seeing individuals do. For visually impaired individuals, vocal clues provide excellent navigation assistance. The prototype in this paper addresses these issues.

Hardware Components

Compaq iPAQ H3700

The mobile device used in this study is the iPAQ PocketPC running Windows CE 3.0 as the OS. The iPAQ has a processor running at 206 Mhz, has 64 MB of RAM, and weighs 6.3 oz. The OS is a multi-threading system designed to run on embedded and mobile devices. The user interface is similar to the Windows OS available for desktop computers. Files customized on a desktop computer can be conveniently migrated to the iPAQ. The iPAQ has an audio earphone attached. The specifications of the device are given in Figure 1.

DGPS Receivers

The GPS employed is the Handy Type GPS Receiver Turbo G2 from TOPCON Corporation. With its own internal camera, the GPS has a positioning accuracy of 1 m. The system is further equipped with the LEGANT-G external antenna, which enables a tracking accuracy of 50 cm. This precision is important for safety reasons; it is used to check whether the user is veering off pedestrian walkways onto roads or otherwise straying off course.

Software Components

The customization of the interface and the development of the prototype took place on a desktop PC running GIS applications provided by ESRI. These applications are described below.

ArcGIS Tools

Various GIS tools provided by ESRI were employed, namely, ArcView, ArcPad, and ArcPad Studio. These tools provide data visualization, query, and integration capabilities and the ability to create and edit geographic features. The map data files created were added and assembled in ArcMap to form a three-layer campus map of the National University of Singapore (NUS). The map was created in

ArcMap and had to be converted to the format supported by ArcPad for use on site. Most of ArcPad’s customization was performed in ArcPad Studio, in which applets, customized toolbars and forms, and default configuration files can be built. The fundamental feature of ArcPad’s customization environment is the ability to use scripts.

Voice Synthesis and Recognition

Commercial off-the-shelf software such as 2nd Speech Center and Coolspeaking was used to generate the voice IDs on the desktop computer before they were transferred to the mobile device. Subsequently, software such as IBM’s ViaVoice was used to provide a spoken dialogue interface. This software comes into play during the selection of a destination by the user.

IMPLEMENTATION

For this project, the area of study chosen was the NUS campus in Kent Ridge, which has an approximate land size of 150 ha. It was assumed that users would travel only on pedestrian walkways; thus, the NUS walkways were modeled into a network of links from which routes from source to destination could be ascertained. The safety of visually impaired pedestrians was further taken into account by allowing opposite walkways to be reachable only via zebra crossings (i.e., street crossings, often marked with stripes, where pedestrians have the right of way).

System Architecture

Figure 1 shows the prototype’s architecture. The maps of the study area were generated by combining various shapefile layers created and edited in ArcView. The map consists of three layers:

- Background.jpg,
- Walkway.shp, and
- Building.shp.

Data were collected on NUS’s buildings and features; the data included detailed textual descriptions, photographs, and voice IDs. The creation, editing of maps, and customization of user tools were performed on a desktop Pentium IV PC running Windows 2000 and using the GIS applications previously mentioned. All data were then

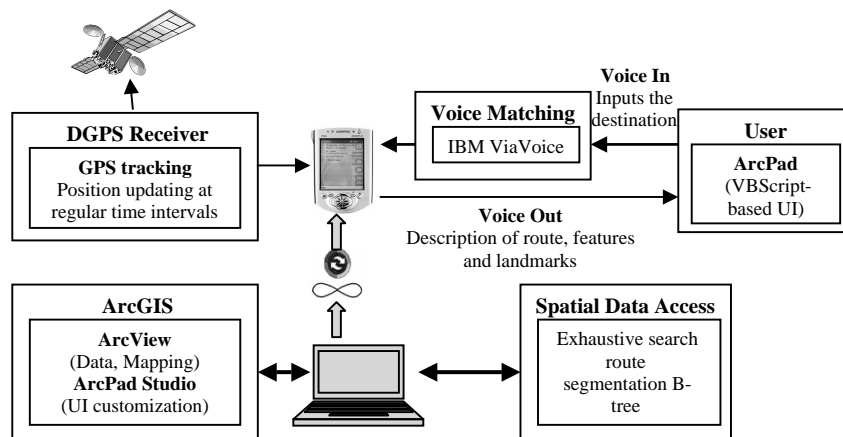


FIGURE 1 System architecture.

transferred from the desktop to the iPAQ so that the PDA could function as a stand-alone device.

The user communicates with the mobile device through speech. Each landmark or building contains a voice ID. When the system prompts the user for a destination, the user's response is matched to the voice ID via matching software. The routing algorithm then generates a route based on Dijkstra's algorithm to the destination, and

the route is recited to the user. Figure 2a shows the computed travel route, and Figure 2b shows an example of the directions that users hear through their earphones.

As the user travels along the intended route to the destination, the DGPS receiver tracks the location of the user at regular intervals, every 2 min or so. At these intervals, the system conducts a search based on the current coordinates of the user and returns a feature



FIGURE 2 PDA visualization: (a) iPAQ with route displayed on map, (b) pretravel route instructions, (c) voice cue example, and (d) alert message.

nearest to the user’s current position. Figure 2c shows the voice cues that the system would project to the user. Based on the voice IDs of the features stored in the database, the system states to the user the features that he or she is passing. The system also provides a vocal alert when the user nears or reaches road junctions. Figure 2d shows this warning.

Our system uses a simple method to help the user stay on course. If the user’s location deviates from the computed route by more than 1 m (the DGPS is at 0.5 m accuracy), the system will prompt the user to retrace his or her steps and return to the correct route.

Indexing Method

The primary goal of this effort is to create a stand-alone, compact system for the visually disabled user. One major area of concern is how efficiently information is retrieved when GPS tracks the user. Because the visually impaired individual needs frequent updates of contextual information, search time must be fast. When the database was first established, exhaustive search was used to find buildings and landmarks. Exhaustive search is a primitive brute-force method that does not yield optimal search times. The indexing method was proposed as a replacement. These two methods are discussed below.

Exhaustive Search

The spatial searching algorithm was incorporated in the VBScript source code that automated the customized tasks. All buildings and landmarks in NUS were stored as building features; therefore, when searching for a building that fulfilled given criteria, the system had to access every item in the building.shp layer, from the first record to the last record. Subsequently, it had to search every single coordinate of every vertex of building features to calculate the distance between the buildings and the user’s current location. For example, if there are *n* buildings, the time complexity of the search is *O(n)*.

Although a time complexity of *O(n)* does not appear to be terribly inefficient, it should be remembered that aside from the retrieval of data, there must be a compromise between the loading of voice IDs and the frequency of GPS tracking. A search method that takes too long will interfere with the accuracy of the features retrieved by the system. Therefore, an indexing method was proposed as a substitute for the exhaustive search method.

Route Segmentation B-Tree

Many spatial indexing methods have been developed, including B-trees, Quad-trees, and R-trees (10). All of these methods are

intended to accelerate the retrieval of an object based on its location. None of these techniques is clearly superior. Some form of hierarchical organization is generally advantageous, but performance depends significantly on the distribution patterns of the spatial data on the map.

The lack of inbuilt spatial indexing in ArcPad led to the introduction of an indexing method called Route Segmentation B-Tree. A B-tree may have a variable number of keys and children. The keys are stored in nondecreasing order. Each key has an associated child that is the root of a subtree containing all nodes with keys less than or equal to the key but greater than the preceding key. A node also has an additional rightmost child that is the root for a subtree containing all keys greater than any keys in the node. Because each node tends to have a large branching factor (a large number of children), it is typically necessary to traverse relatively few nodes before locating the desired key. A simple search operation with a B-tree is shown in Figure 3. In this example, buildings have been indexed by their distance from the user’s current location. The user wishes to search for a landmark that is 21 m away.

Instead of choosing between a left child and a right child, as a binary tree search would, a B-tree search must make an *n*-way choice. The correct child is chosen by performing a linear search of the values in the node. After finding the value greater than or equal to the desired value, the child pointer to the immediate left of that value is followed. If all values are less than the desired value, the rightmost child pointer is followed. Of course, the search can be terminated as soon as the desired node is found. In this case, all children to the left of 20 are less than 21; thus, the child pointer immediately to the right of the 20 is followed. Key 21 is found, returning the building that is required.

To apply the B-tree to the prototype system, the walkways were split into a certain number of segments (i.e., links, in this case), and indices were assigned to each walkway segment. Each of these route indices serves as roots of subtrees. Building feature IDs were assigned to particular routes based on proximity.

Whenever searching is initiated through automated GPS tracking, the system will find the ID of the link or walkway segment that the user is currently on and search only for building features belonging to that particular record, instead of searching all buildings in the building layer. Assuming all links have the same number of buildings allocated to them, it can be approximated that with a same total number of *n* buildings and *k* walkway segments, the time complexity of the search is *O(log n)*.

Before Route Segmentation B-Tree was developed, it was necessary first to build a data structure of the proper size. “Size” refers to the appropriate number of walkway segments, which determines the improvement in time complexity by using the indexing algorithm instead of exhaustive search. Time constraints prevented the determination of segment size from being studied in detail. In addition,

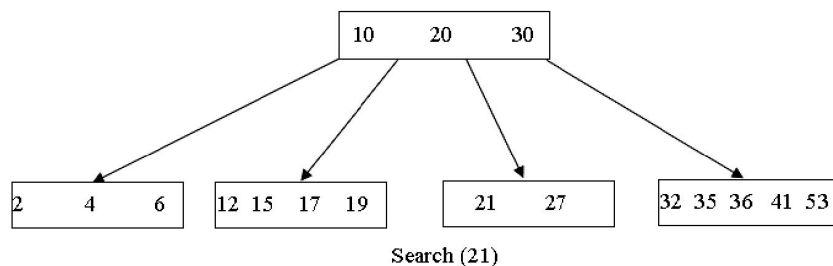


FIGURE 3 Search for key 21 in a B-tree.

because pedestrian crossing and traffic junctions were considered to be allowable paths, they were represented as walkway segments.

There are three layers in the NUS campus map, but only the walkway layer and building layer were involved in the Route Segmentation B-tree method. Within the ArcView and ArcMap application environment, a new field called "ID" was appended into the building.shp and walkway.shp layers. Each record in those layers was assigned a unique ID. A building was assigned to a link if any point of the building feature fell within a 60-m radius of the link. Figure 4 shows how building features were assigned to walkways, and Figure 5 illustrates the Route Segmentation B-Tree. The point at which two or more links converge is known as a node. All junctions of the walkway segments are nodes.

Improvement of Efficiency in Performing Spatial Queries

The built-in function Timer in VBScript was used in the source code to obtain the process time of the two algorithms. For exhaustive search, 30 points representing a user's locations were picked arbitrarily on the map; the process time was output in the form of a Message Box automated by the VBScript source code. For exhaustive search, the computed average time was 3.47 s.

For the indexing method, the maximum number of candidates in terms of buildings needed for distance checking was 34. Exhaustive search must retrieve location coordinates from 134 buildings; in contrast, the indexing method needs to access only about 10% of the candidates to perform the same spatial query.

Making use of the higher numerical precision available in the PC system clock, the results acquired were scaled to reflect the process time on the mobile device, which has a lower processor speed. By replacing exhaustive search with the indexing method, the process time has been reduced by 2.37 s or 68%, from 3.47 s to 1.1 s.

CONCLUSIONS AND FUTURE WORK

Over the past two decades, numerous systems designed to provide guidance to the visually impaired have been researched and developed. These systems have employed a range of technologies, from mobile wireless communications and image processing to computer vision and artificial intelligence. The prototype system discussed

here provides precise tracking of a user through DGPS, utilizes relevant context-aware information, is compact, provides user self-sufficiency, and complements traditional aids such as canes and guide dogs. The system employs the most current GPS receiver, which yields an accuracy of 50 cm. The contextual awareness available in the system is believed to enhance the navigational experience, especially for the blind user (6). Moreover, the presence of vocal cues in the system compensates for the user's lack of sensory perception. The system uses an indexing method called Route Segmentation B-Tree, which significantly improves data retrieval efficiency as compared to exhaustive search methods.

Because of the focus on the abovementioned areas, this research has abbreviated the portions on route computation. Although travel has been restricted to pedestrian walkways, which ensures a certain degree of safety, unforeseen difficulties and obstacles such as road constructions and overhanging branches can still pose a hazard. The route of least hazard, rather than the shortest route, may often be the one most desirable for the visually disabled user. Further research could work on developing an algorithm to map such a route. In addition, due to the dynamic nature of the environment, future systems can incorporate an intelligent agent that allows the user to input obstacles into the system, and the system to compute a new best route.

The system in this study uses simple voice-matching software to match the user's destination input to the voice IDs of the landmarks. Further work can be performed on developing an efficient way to iterate the available destinations to the user and to carry out fast matching or input to IDs.

The prototype was tested in the NUS campus by users wearing eye masks to ascertain how well it functions. Further tests involving users who are truly visually disabled will be carried out. Because the campus area is small, the PDA employed was able to store all the GIS data needed for use in the field. If the device needed to be used in a larger area, such as the whole of Singapore, storage problems might result. Using a central server may be helpful; work can be done on developing a system with persistent, reliable, inexpensive online wireless communications.

A final area of note: GPS is not a foolproof system. The GPS employed offers a great degree of precision; however, GPS signals can be lost near tall buildings and under tree canopies. When this happens, the user is forced to rely on traditional navigation aids. Further research could explore a contingency plan to address the loss of GPS signals so users will not become lost if their signal is lost. Furthermore,

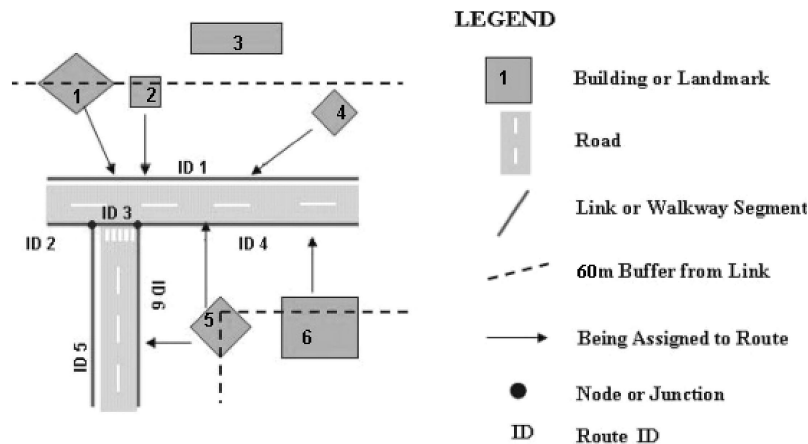


FIGURE 4 Example of how buildings are assigned to road segments.

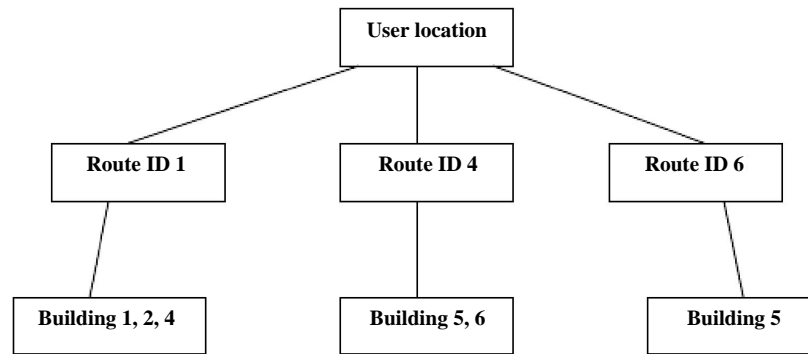


FIGURE 5 B-tree diagram of route assignment (from Figure 4).

GPS does not work indoors. Other tracking or positioning systems that can compensate for this limitation may need to be explored.

REFERENCES

1. Neubauer, P. *B-Trees: Balanced Tree Data Structures*. 1999. www.bluerwhite.org/btree. Accessed April 2003.
2. Busboom, M., and M. May. Mobile Navigation for the Blind: Making GPS and Other Commercial Products More Accessible. Presented at International Conference on Wearable Computing, Vienna, Austria, 1999.
3. Petrie, H., V. Johnson, T. Strothotte, T. Raab, S. Fritz, and R. Michel. MOBIC: Designing a Travel Aid for Blind and Elderly People. *Journal of Navigation*, Vol. 49, No. 1, 1996, pp. 45–52.
4. Strothotte, T., H. Petrie, V. Johnson, and L. Reichert. MOBIC: User Needs and Preliminary Design for a Mobility Aid for Blind and Elderly Travelers. Presented at Second TIDE Congress, Paris, La Villette, April 26–28, 1995.
5. Ram, S., and J. Sharf. The People Sensor: A Mobility Aid for the Visually Impaired. *Proc., Second International Symposium on Wearable Computers*, Pittsburgh, Pa., IEEE Computer Society, 1998, pp. 166–167.
6. Makino, H., I. Ishii, and M. Nakashizuka. Development of Navigation System for the Blind Using GPS and Mobile Phone. *Proc., 18th Annual International Conference of the IEEE*, Amsterdam, Netherlands, Vol. 2, IEEE, 1996, pp. 506–507.
7. Helal, A., S. E. Moore, and B. Ramachandran. Drishti: An Integrated Navigation System for Visually Impaired and Disabled. *Proc., Fifth International Symposium on Wearable Computers*, Zurich, Switzerland, IEEE Computer Society, 2001, pp. 149–156.
8. Kaluwahandi, S., and Y. Tadokoro. Portable Traveling Support System Using Image Processing for the Visually Impaired. *Proc., 2001 International Conference on Image Processing*, Vol. 1, Oct. 7–10, pp. 337–340.
9. Bourbakis, N. G., and D. Kavradi. An Intelligent Assistant for Navigation of Visually Impaired People. *Proc., 2nd IEEE Annual International Symposium on Bioinformatics and Bioengineering Conference*, Bethesda, Md., IEEE, 2001, pp. 230–235.
10. van Oosterom, P., and E. Claassen. Orientation Insensitive Indexing Methods for Geometric Objects. *Proc., Fourth International Symposium on Spatial Data Handling*, Zurich, Switzerland. International Geographical Union, Commission on Geographical Information Science, 1990, pp. 1016–1029.

Publication of this paper sponsored by Accessible Transportation and Mobility Committee.