Development of a Socially Cognizant Robotic Campus Guide

Benjamin Greenberg
Rutgers University
Department of Mechanical and Aerospace Engineering
Piscataway, New Jersey, USA
ben.greenberg@rutgers.edu

Daniel Nakhimovich
Rutgers University
Department of Computer Science
Piscataway, New Jersey, USA
d.nak@rutgers.edu

Richard Magnotti
Rutgers University
Department of Computer Science
Piscataway, New Jersey, USA
richard.magnotti@rutgers.edu

Hriday Purohit
Rutgers University
Department of Computer Science
Piscataway, New Jersey, USA
hriday.purohit@rutgers.edu

Sanskar Shah
Rutgers University
Department of Computer Science
Piscataway, New Jersey, USA
sanskar.shah@rutgers.edu

Aniket Satish Kulkarni
Rutgers University
Department of Computer Science
Piscataway, New Jersey, USA
aniket.kulkarni@rutgers.edu

Uriel Gonzalez-Bravo
Rutgers University
Department of Psychology
Piscataway, New Jersey, USA
uog3@scarletmail.rutgers.edu

Noah R. Carver
Rutgers University
Department of Computer Science
Piscataway, New Jersey, USA
nrc86@scarletmail.rutgers.edu

ABSTRACT

A robotic system to help lost students find their way around a college campus was designed, built, and tested. Socially cognizant design practices, including stakeholder engagement, and interdisciplinary team-building, were practiced. Users can interact with the robot through speech or touchscreen interfaces. The robot can provide verbal instructions on reaching a destination, or can guide the user to the destination, navigating in a socially conscious way. The speech, person detection, and navigation modules perform well in isolation and in concert. Future work includes technical improvements to the person detection and navigation systems, and evaluating social acceptance.

CCS CONCEPTS

• **Human-centered computing → Sound-based input / output**; **Interaction devices**; **Human computer interaction (HCI)**; **Social navigation**; **Social engineering (social sciences)**.

KEYWORDS

Robot, Social, Navigation, Speech, Voice, Human-Computer Interaction, Person Detection

ACM Reference Format:
Benjamin Greenberg, Daniel Nakhimovich, Richard Magnotti, Hriday Purohit, Sanskar Shah, Aniket Satish Kulkarni, Uriel Gonzalez-Bravo, and Noah R. Carver. 2024. Development of a Socially Cognizant Robotic Campus Guide. In *Companion of the 2024 ACM/IEEE International Conference on Human-Robot Interaction (HRI '24 Companion), March 11–14, 2024, Boulder, CO, USA.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3610978.3641263

HRI '24 Companion, March 11–14, 2024, Boulder, CO, USA

ACM ISBN 979-8-4007-0323-2/24/03
https://doi.org/10.1145/3610978.3641263

Figure 1: Campus Guide Robot

1 INTRODUCTION

Mobile robots are becoming increasingly common in public social settings, from restaurants to hospitals to airports. One of the challenges in designing these systems to operate effectively is to constrain their behavior to be socially acceptable. Social acceptability can mean different things to different people in different settings, but it is something that must be thoughtfully considered early in the design process to produce robots that best complement the people around them.

This project focuses on a particular social mobile robotic application: helping people who are lost on a college campus. College campuses can be sprawling complexes. For example, Rutgers University spans over six thousand acres and consists of over nine hundred buildings [1]. Navigating such a vast campus can be challenging for newcomers, especially when resources for easing this burden are difficult to find. A reasonable person might think that the lecture hall named PH-111 is in the Physics Building. That person would not only be wrong, but also confused, frustrated, and late for class (in the Pharmacy Building). This is a problem that many people have experienced. According to the Rutgers University Department of Student Orientation and Transition Programs (RUDSOTP), this is a common experience.

To address the problem of people being lost on a college campus, we have designed, built, and evaluated the performance of a robotic campus guide. We hope that some of the design principles and findings from this project may inform the design process of guide robots in other settings, from hospitals, to conference centers, to airports.

2 DESIGN CONTEXT

The context for this robotic application is a college campus. Every day, people set foot on college campuses for the first time. These people may enter the campus with different motivations depending on whether they are faculty members, parents, conference attendees, guest lecturers, contractors, couriers, sports fans, or students. However, all of these people share the experience of being in a new and unfamiliar place, with the aim of getting to a particular location.

2.1 Socially Cognizant Design

A goal of this project was to implement socially cognizant design principles. One of these principles is to build an interdisciplinary team. Our team consists of people from several STEM fields, as well as people from social sciences. This diverse set of perspectives enables us to anticipate and mitigate unintended consequences of our design choices.

Another principle of socially cognizant design is to consider the social impacts that the system being built will have on an individual level and a societal level.

Guiding a lost person around a college campus involves several social dynamics. Some of these are described below, along with the solutions that we devised to address social needs:

- Robot <-> User: How the person seeking guidance, and the robot interact with each other. There are numerous elements to this interaction. We have considered factors such as:
  - How does a user know what this robot is for, and how does the robot invite users into an interaction?
    - Guide robots should be placed near bus stops, entrances to buildings, or other places of congregation
    - The robot will have a screen that displays a message stating its purpose and inviting interaction
    - While waiting for a user to engage with it, the robot should periodically invite interaction audibly by saying "Ask me 'Take me to my classroom'"
  - How should the robot move once a person begins interacting with it?
    - We considered several options, including having the robot rotate to face the nearest person to it, and tilting its camera upward to indicate that the person has the attention of the robot. This has not been implemented.
  - Though many researchers have investigated social robot navigation in general, there is far less work on the social requirements for how a robot should guide a person. Factors such as speed, checking in, and maintaining engagement were discussed.
- Robot <-> Bystanders: How the robot interacts with other people in the environment whom aren't being guided.
  - The primary focus in this interaction is how the robot should navigate around other people in the environment. Significant research has been done by others in the field of social robot navigation, we have used that work as a starting point for ours.
  - Another aspect of this interaction could include robot vocalizations such as "excuse me" when passing near a bystander.
- User <-> Bystanders: How the person seeking guidance interacts with other people in the environment.
  - This mode of interaction is one that we have not spent significant time investigating. One speculation is that some users might be self-conscious about being seen using the robotic guide by other community members.

2.2 Stakeholder Engagement

In any complex project, it is crucial to identify and engage with as many stakeholders as possible. For this robotic campus guide project, we have identified several stakeholders including: students, campus safety officers, facilities management staff, and the Rutgers University Department of Student Orientation and Transition Programs (RUDSOTP).

The team from the RUDSOTP was gracious to invite us to one of their team meetings to discuss the issue of students struggling to find their classes, and the potential for a robotic campus guide to help solve the problem. They offered great insight into the pain points and needs of students and staff, and provided valuable input during our design process.

As briefly mentioned earlier, the RUDSOTP confirmed that there is a real problem with students struggling to find their classes. This is a serious enough problem that their team was motivated to build and maintain a website to serve as a central hub for information regarding finding classes. The team suggested that the robot have a QR code available to direct students to this page in case they don't need or want the full robot guide experience.

During the discussion, one of their team members raised a concern that the use of a robot campus guide would remove the human touch that can be an important part of making new community members feel welcomed. One idea that was raised was a potential video call feature utilizing the robot's screen and camera to connect the lost student to a staff member who could provide some guidance and a personal touch. This has not been implemented. Another team member brought up a contrary point, noting that some students can be intimidated by talking to strangers and would actually welcome the option to interact with an automated system.

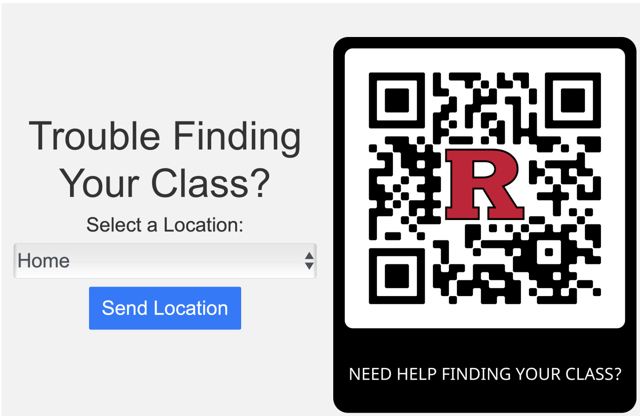

Figure 2: Touchscreen User Interface

One area of enthusiasm for an automated system was the potential to improve accessibility for the significant number of non-native English speakers on campus. Since most of the signage is in English, this can pose a challenge to non-native English speakers. The touchscreen interface of the robot offers the opportunity to toggle between different languages. This feature has not been implemented.

Other concerns that the RUDSOTP team raised were more pragmatic in nature. These included concerns about robots navigating the more congested parts of campus without being annoyances to the people in those areas, robustness against vandalism, and the complication of the buses that are often necessary to get from one part of the campus to another distant part. We did not focus our design on these issues, but it is extremely valuable to have them already brought to our attention as we continue improving our system.

3 SYSTEM OVERVIEW

This project leveraged a LoCoBot WX250S robot as the base platform. It was chosen because of its inbuilt navigation capabilities. We added a touchscreen as well as a microphone and speaker to serve as interfaces with people. The fully integrated system is pictured in Figure 1. Using these interfaces, a user can request help finding a location on campus. The touchscreen interface is pictured in Figure 2.

Depending on the request, the robot will either tell the user how to reach their goal, or will guide them to their goal.

4 SPEECH MODULE

The speech module for this system consists of an input and an output system. Using the microphone on the system, the robot is constantly waiting to be addressed. When a decibel threshold is exceeded, the robot listens to what is said. Depending on the input, it can enter Guidance mode or Instruction mode.

Guidance mode is triggered by key words like "take," "guide," and "bring." For example, if a person says "take me to room three," the robot will recognize that it is in guidance mode, and will check its database to see if the destination exists. If the robot is able to find the destination, then it responds by saying: "I'll take you to room three" or whatever destination was requested, and begins navigating to that location. If the destination is not part of the robot's database, it tells the user that it does not know where that room is.

Instruction mode is the default mode, that is, unless guidance key words are used, the robot will remain in Instruction mode. A user can ask the robot for instructions on how to get to a destination on campus. For example, one could ask: "How do I get to room three?" In this case, the robot uses a large language model (LLM) to generate a set of instructions describing the steps needed to reach the goal and reads them aloud. The LLM is provided with a written description of the environment as part of the mapping process. The combination of the written description and the question are sufficient for the robot to produce coherent instructions. To interface with a LLM we query OpenAI's API for GPT. GPT is a large language model pre-trained on 45 TB of scraped web data [2].

5 NAVIGATION MODULE

For the robot to guide a user to their destination, it must have a navigation system to plan paths, avoid collisions, and control the motors on the base. The baseline system built in to the robot uses the Navigation stack [3], which relies on Robot Operation System (ROS). The Navigation stack uses a 2D costmap to represent the environment, with obstacles represented as high-cost regions. A path planner is used to find the lowest cost path from the robot's current position to its goal. This results in robot motions that avoid obstacles and any other factors that cause high-cost areas.

This system can be considered an 'antisocial' navigation system because it has no distinction between people and any other type of obstacle. However, because the Navigation stack is plugin-based, it is possible to supplement this antisocial system with additional social layers. One such social layer was implemented by the team at Pablo de Olavide University, as an update to David Lu's original social navigation layer. This system [4] treats people differently from other obstacles, by adding additional high-cost regions to the costmap around people. This plugin adds costs around people in three cases:

- Circular cost region around stationary people
- Rounded-oblong cost region extending to the side of moving people to preference a passing side (i.e. pass people on the left or right)
- Rounded-oblong cost region extending in the direction of motion of moving people to avoid getting in the way of people

To improve upon the social layer, our team changed the way that costs are represented around moving people. We recognized that the existing shape, while it does deter a robot from planning a path directly in front of a moving person, does not take into account any uncertainty in the direction of motion. In other words, the existing model assumes that a person will continue moving straight. To account for the uncertainty about a person's future motion, we chose to represent the costs around the moving person as a sector of a circle, with the person being at the vertex of this sector. The velocity of the person is proportional to the radius of the sector (faster people cover more distance), and inversely proportional to the angle swept by the sector (faster people are less likely to change

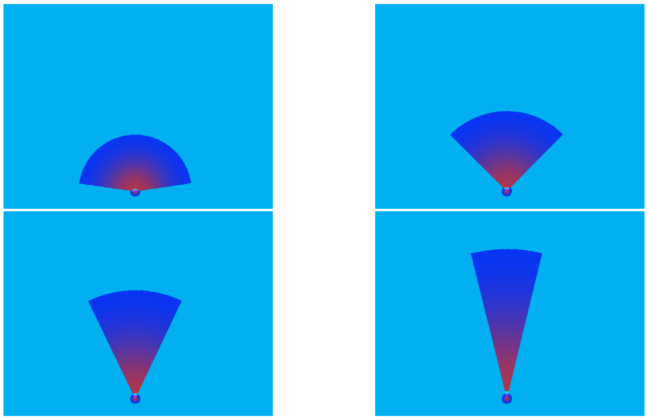

Figure 3: Costmap regions added for a person moving 0.2 m/s, 0.5 m/s, 1.0 m/s and 2.0 m/s, clockwise from top left

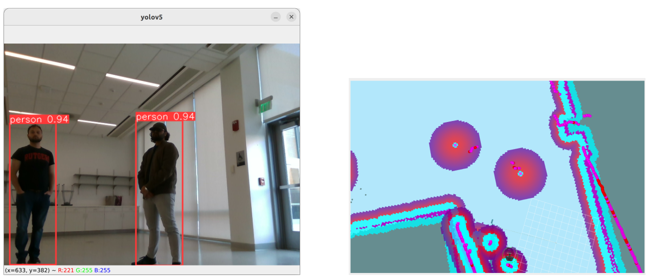

Figure 4: 2D image used to detect people (left) and 3D positions of these people projected onto map (right)

direction quickly). Costmap shapes for various person speeds are shown in Figure 3.

6 PERSON DETECTION MODULE

In order to use the social layer of the Navigation stack, it is necessary to detect where people are in the environment. We used the Intel Realsense D435 camera at our disposal to collect RGBD images continuously. We took advantage of the YOLOv5 system to detect people in the images and draw bounding boxes around them. We then used the depth data in the image to find the 3D position of the person. Several averaging approaches were attempted, but the most reliable method we found was to take the 25th percentile depth point in the bounding box. This system works well for detecting the locations of stationary people, but due to the lack of individuation of people it does not allow us to detect the velocities of people in the environment. A 2D image of two people, and the corresponding detected positions of them are shown in Figure 4.

7 EVALUATION AND RESULTS

System modules were evaluated independently, as well as in concert. This took place in a representative indoor space due to limitations on the mobility of the robot.

The person detection module was evaluated for accuracy in comparison to LIDAR data. We found that there were no significant differences between the performance of these systems when a person was within 4.4 meters of the robot. Our system's performance degraded beyond that point.

The speech detection module successfully understood the utterances of users for all of the conditions tested. However, it was observed to have difficulty understanding some non-American accents.

The navigation module successfully planned and executed motion to the desired location in four out of five trials. The robot was able to avoid collisions with static people in the environment. Because our person detection system does not provide velocities of people, we had to simulate this input to evaluate the changes we made to the navigation module. We found that the robot successfully avoided the moving person for each of its five trials.

The end-to-end process from speech or touch command to collision-free guidance to the desired destination was shown to be successful and reliable on each of the nine trials.

8 CONCLUSION AND FUTURE WORK

This project successfully implemented socially cognizant design principles to build a robot that serves a real need in our community. Our design was informed by technical as well as social constraints.

The individual modules and overall system exhibited high levels of performance during evaluation.

Moving forward, we will continue improving the robotic campus guide by developing a new person detection module that provides the velocities of people. This will enable real-world testing of our modified social costmap layer.

We also plan to get feedback from community members on how effectively this system addresses the problems identified. Through this iteration process we hope to build a more reliable and socially acceptable system that can begin helping people as soon as possible.

ACKNOWLEDGMENTS

This work has been supported by the National Science Foundation (NSF) National Research Traineeship (NRT) entitled "NRT-FWHTF: Socially Cognizant Robotics for a Technology Enhanced Society (SOCRATES)" Grant No. 2021628. We would like to thank Dr. Jacob Feldman for his guidance and leadership, as well as Lina Moe for her insights on public policy considerations.

REFERENCES

[1] [n. d.]. Facts and Figures: Rutgers By the Numbers | Rutgers University. https://www.rutgers.edu/about/by-the-numbers
[2] [n. d.]. What is GPT? https://aws.amazon.com/what-is/gpt/
[3] Steve Macenski, Francisco Martín, Ruffin White, and Jonatan Ginés Clavero. 2020. The Marathon 2: A Navigation System. In *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. https://github.com/ros-planning/navigation2
[4] Robotics-Upo. [n. d.]. Robotics-UPO/nav2_social_costmap_plugin: Plugin-based layer for social navigation in the NAV2 system of ROS2. https://github.com/robotics-upo/nav2_social_costmap_plugin