Validating Safe Intersection Crossing Technology in the Field

Data Collection

What data will you collect or create?

This project aims to consolidates recent advances to PedPal, a mobile app designed to assist pedestrians with disabilities in safely crossing signalized intersections, and produce a hardware software solution that is validated through deployment at a local intersection close to the CMU campus. As such, data collection is not an explicit end goal of the project. However, to validate the efficacy of the hardware/software solution that is produced, crossing trials at the target intersection will be performed and data will be collected for purposes of quantifying performance. In particular, an attempt will be made to measure both the effectiveness of ultra-wide band beacons for localization and the responsiveness of the PedPal cloud server and intersection manager that must be developed to interface with the controller at the intersection. Specifically, the following data will be collected for each experimental crossing trial at the target intersection:

- crossing trial number
- crossing distance
- start time of the cross
- end-time of the cross
- detected position (location) of the pedestrian after each second of the cross as provided by the app.
- actual (ground truth) position of the pedestrian after each second of the cross (as derived from a timed video of the cross with street markings)
- accuracy of each detected location relative to ground truth.
- number of other PedPal users simultaneously moving through the intersection (0 to 2)
- average roundtrip communication time required per message by DSRC message type (MAPP, SPaT, SRM, SSM) during crossing

How will the data be collected or created?

To collect the information related to the detected locations over time during crossing trials, we will instrument the PedPal app to log this information and subsequently retrieve it from the logs. To obtain the actual (ground truth) locations over time during crossing trials, we will record a timed video of each trial, and derive the actual locations from this information. To facilitate this process, we will introduce temporary markings in the intersection (e.g., marking tape) at regular intervals. From this baseline information, accuracy information will be derived.

To collect message roundtrip information by command, we will further instrument the PedPal app to log this information and subsequently retrieve it from the logs.

Documentation and Metadata

What documentation and metadata will accompany the data?

A data dictionary will be provided with the file containing the data in comma separated form and header information, to explain the semantics of the values in each field (column).

Ethics and Legal Compliance

How will you manage any ethical issues?

Our intent is that all trials will be performed by project personnel, strictly for purposes of quantifying localization accuracy and system responsiveness. It is not our intent to conduct a study of prospective users of the PedPal Lite system. Hence, there are no ethical issues to manage to our understanding.

How will you manage copyright and Intellectual Property Rights (IP/IPR) issues?

CMU will retain IP rights to all data collected and created over the project lifetime. It is not anticipated that there will be any 3rd party interest in this data. However, we expect to publish the results of these experimental trials, and if there is an interest in the data for purposes of comparison with other similar technologies, then a mechanism for making the data available will be developed.

Storage and Backup

How will the data be stored and backed up during the research?

Access to this data is provided only to project personnel through controlled accounts which are issued by the PI.

Data collected in the field will be downloaded from the app and extracted from the video using local project member laptops and then entered into GitHub.

How will you manage access and security?

GitHub encrypts all data in transit, all login information and credentials are always protected. GitHub stores a one-way hash of all user passwords using bcrypt. Your account login is protected from brute force attack with rate limiting.

Selection and Preservation

Which data are of long-term value and should be retained, shared, and/or preserved?

All of the data mentioned above will be considered sharable once the project has completed. As mentioned earlier, our intent is to publish results following from the experimental trials and the data

will be preserved to support this activity. At the end of the project we will move all data collected to Carnegie Mellon University's data archive for permanent preservation.

Although we don't anticipate it, we will make this data available in as is condition, should another third party desire access to the data for future research purposes.

What is the long-term preservation plan for the dataset?

As indicated above, the data will be moved to CMU's data archive facility at the completion of the project (or to another permanent data archive if this is deemed necessary).

Data Sharing

How will you share the data?

We do not anticipate any external party being interested in this data. But we are willing to share it for various educational and/or research purposes. We will take a low overhead approach to sharing, where an interested party contacts the PI and the PI then responds by forwarding the data by email in its planned comma separated values format.

Are any restrictions on data sharing required?

No.

Responsibilities and Resources

Who will be responsible for data management?

The Intelligent Coordination and Logistics Laboratory will be responsible for implementing the data management plan, until which time it is deemed necessary to enlist an independent repository service for indefinite preservation.

What resources will you require to deliver your plan?

No additional software, hardware or staff training will be necessary.
