Dynamic Management of Food Redistribution for 412 Food Rescue

Data description

Describe the data that will be gathered in the course of the research project, including whether the data should be preserved for long-term access.

412 Food Rescue is a non-profit that uses crowd-sourcing to redistribute food donations to distribution centers that serve the hungry. This project will develop scheduling and logistic technology to automate the efficent assignment of both distribution centers to food donations and transporters to relocate the donations to the assigned destinations. 412 already has an application that collects data on their food redistribution operations. The technology in this project will require 412 to communicate a subset of that data to the scheduling system, in particular, three types of data.

The first type of data will be for food opportunities, i.e., the timing, type, location (latitude/longitude coordinate), and amount of food being donated. The second type of data is for the distribution centers. A datum of this type will specify a numerical id for the center, the types of food it can receive, the type of food it would like to have, its location, its availability, and the number of people it serves. When an assignment of a distribution center to a food donation is made, a transportation request for moving a food donation to a distribution center will be generated and recorded. A request consists of the time of the request, the pick-up location, the types of food being donated, the quantity of the food, the state of the food (perishable vs. non-perishable), when the food is available for pick-up, the destination location, when food can be dropped off, and the deadline for when the food must be moved by. The third set of data applies to the group of people who can perform the request, i.e., those that can transport food donations to their prescribed destinations. A datum of this type will have a numerical identifier for the transporter, when the transporter is available to perform the request, and the transporter's carrying capabilities. When an assignment of a transporter to a request is made, that assignment and the computed timing of the assignment (scheduled pick-up and drop-off times) will be recorded.

This data will be used to drive the design and evaluate the performance of the proposed scheduling systems. During development of the technology, historical data over the last year will be used to generate test problems containing the types of data enumerated above in a JavaScript Object Notation (JSON) format. These problems will be used to test and evaluate the technology. A test suite of problems will be retained to do regression tests to guarantee that, as new features are added to the system, existing capabilities continue to perform. This test suite will be initiated and maintained over the lifetime of the project. This test suite will also be used to evaluate the system and show empirical proof of increased efficiency. The value of the data is most likely in comparing approaches to the optimization problems, but, as 412 owns the data, it will be their determination to who has access to them. We will share data only through 412.

Data format and metadata standards

Describe the standards and machine-readable formats that will be used in the course of the research project.

All data will be stored and maintain in JavaScript Object Notation (JSON) format, a public standard that is machine-independent. We will also maintain a document that will describe the model for the data, i.e., the JSON objects, their fields, and their semantics. The data and the document will be maintained in a version-controlled file system (subversion or git). To support reviewing data, we will develop a component to parse and generate instances of the data and store them in an in-memory database that supports sophisticated queries and

visualization using Google Charts.

Policies for access and sharing

Discuss the access policies that will apply to the data, so as to protect against the disclosure of identities, confidential business information, national security information, etc. and whether public use files may be generated from the data.

412 Food Rescue will own the data, so we will not extend access to the data. 412 will maintain the canonical copy of the data and will apply their policies to determine to whom the data may be shared. If 412 is willing to share data provided that it is deidentified, we will work with them to add noise to the locations and, if necessary, to the food donations and distribution center's requirements, depending on the concerns 412 has and what data they are willing to share. We will rely on 412's current processes for recording data and have flow-through permission from them to use the data.

Policies for re-use, redistribution, derivatives

Discuss the policies for re-use, re-distribution and derivative projects.

The data will be owned by 412 Food Rescue, and we will used them by their permission. Any requests for the data will be directed to them, and their policies will govern access to the data.

Plans for archiving and preservation

Outline the plans for archiving and preservation, specifying where research data will be deposited, and specify that data will be deposited at the time of initial publication of any related peer-reviewed journal article.

The data will be kept within our password-protected computing environment located within our laboratory at Carnegie Mellon University, which has limited access. Two developers will have access to the data and will be responsible for maintaining its security. We anticipate the data being kept throughout the project and any of its extensions, after which, they will be deleted.