Evaluating the Impact of Incorporating Information from Social Media Streams in Disaster Relief Routing

Ashlea Bennett Milburn, PhD
University of Arkansas
4207 Bell Engineering Center
Fayetteville, AR 72701
+1 (479) 575-3702
ashlea@uark.edu

Clarence L. Wardell III, PhD
CNA
4825 Mark Center Drive
Alexandria, VA 22311
+1 (703) 824-2970
wardellc@cna.org

ABSTRACT
In this paper, we describe a model that can be used to evaluate the impact of using imperfect information when routing supplies for disaster relief. Using two objectives, maximizing the population supported, and minimizing response time, we explore the potential tradeoffs (e.g. more information, but possibly less accurate) of using information from social media streams to inform routing and resource allocation decisions immediately after a disaster.

Categories and Subject Descriptors
A.0 [**General**]: c*onference proceedings*; G.1.6 [**Numerical Analysis**]: Optimization – *stochastic programming*; H.1.1 [**Models and Principles**]: Systems and Information Theory – value of information; H.4.2 [**Information Systems Applications**]: Types of Systems – *decision support, logistics*; I.6.5 [**Simulation and Modeling**]: Model Development.

General Terms
Algorithms, Performance, Experimentation, Theory, Verification.

Keywords
Vehicle routing problem, heuristic, imperfect information, social media, humanitarian logistics, disasters, emergency management.

1. INTRODUCTION
Information regarding the location and need of the population affected by a disaster is required in order to plan how to deliver support to that population. Traditionally, this information becomes available as flyovers and other forms of assessment are completed. However, critical hours pass while this information is collected, and certain critical needs may not be identified in this manner. A larger number of needs can possibly be identified in a shorter amount of time if data taken from social media is also used for disaster planning. Posts by affected populations to platforms such as Twitter, Facebook, and Ushahidi regarding where they are located and how much help they need could be useful in the response planning process, augmenting the data available through more traditional means [1]. As an example, following the Japanese earthquake and tsunami in March 2011, Twitter reported an event spike of up to 5,530 postings per second on their platform [2]. Because a large portion of this information is initially not verified, some of it may be inaccurate. Emergency managers in the United States have cited this concern as a major barrier to incorporating social media information in their decision-making [3]. If responders spend time traveling to locations where they think there is demand, only to see that no help is needed there, then the remaining population needing support could be adversely affected. Inaccurate information about needs has been identified as one of the primary impediments to the rapid delivery of goods immediately following a disaster [4].

WWW 2012 Companion, April 16–20, 2012, Lyon, France.
ACM 978-1-4503-1230-1/12/04.

In this paper, we examine the tradeoff between having more, but possibly inaccurate, data when routing vehicles for the delivery of relief supplies. We evaluate the tradeoff considering two separate objectives: maximizing the population supported, and minimizing response time. This work extends models in the disaster relief routing literature that address uncertainty in demand by considering the effects on decision-making when two distinct classes of information are taken into account [4]. More broadly, this work also contributes to the body of literature that addresses questions around the usefulness of information provided through citizen event reporting [5]. In Section 2, we formalize the problem statement. In Section 3, we discuss methodologies planned for the analysis of the impact of unverified data on disaster relief decision-making. Research questions to be addressed in future work are discussed in Section 4.

2. PROBLEM STATEMENT
A fleet of m homogenous vehicles, each having capacity u specifying the weight or volume that can be carried, are stationed at a depot and are available to serve customer requests. A set N of customer requests is known, where each request is a vector that represents a subgroup of the impacted population with a specific need at a known location x_i. The need may be for search and rescue, medical support, or relief supplies such as food, water, and shelter. In the remainder of this paper, we assume all requests specify a need for relief supplies but the ideas and methods described can be extended to accommodate other classes of needs in the future. The magnitude of need, denoted l_i, specifies the amount of relief supplies required by the subgroup, and generally is assumed to be proportional to the number of persons in the subgroup. The service time of the need, s_i, accounts for the time the vehicle must spend at the location delivering required goods before departing to serve the next request. Note the total time required for a vehicle to serve a request includes service time and additional travel time required to visit the customer.

We assume that a subset of requests have been previously verified, N_V, representing, for example, those demand points revealed by damage assessment teams. The remaining demand points, denoted N_S, represent those requests from a source such as social media. Requests in N_S have not been verified, and thus may be *inaccurate* according to some probability distribution. We define an *inaccurate request* as one where the magnitude l_i and service time s_i are other than indicated. The accuracy of a request is not known until the vehicle serving the request arrives at the associated location. Examples of inaccurate requests include needs that never actually existed (no subgroup is at the location), needs that exist at another location instead of the one indicated, needs that exist at the location but in a different quantity (more or less people in the subgroup), and needs that are served by another party between the time they are reported and the vehicle arrives. Denoting the time the vehicle arrives at location x_i as a_i, the accuracy of request $i \in N_S$ becomes known at time a_i. Then, magnitude of need l_i and service time s_i are updated according to the newly revealed (and newly verified) information.

In this setting, two separate problems are considered. In the first, which we refer to as *Minimize Arrival Time* (MAT), the problem is to assign each customer request in $N = N_V \cup N_S$ to exactly one vehicle while obeying vehicle capacity constraints, and create routes for each vehicle that begin and end at the depot and serve assigned requests. The objective is to minimize average vehicle arrival time for all requests. In the second, referred to as *Maximize Requests Served* (MRS), the problem is to assign each customer request in $N = N_V \cup N_S$ to at most one vehicle while obeying vehicle capacity constraints and create routes for each vehicle that begin at the depot, serve all assigned requests, and return to the depot by a fixed time limit T. Note that not all customer requests will be served in MRS. The objective is to maximize the number of requests that are served.

To evaluate the impact of incorporating the new class of unverified requests from social media in disaster relief planning, we compare solutions to MAT and MRS to the case where only requests in N_V are considered when planning vehicle routes. We denote the modified problems at MAT-V and MRS-V, respectively. In the next section, we present models and insights for the described problems.

3. MODELS AND INSIGHTS

All above problems (MAT, MRS, MAT-V, MRS-V) can be modeled as variants of the classical vehicle routing problem (VRP) defined on graph $G=(N',A)$ where node set N' is comprised of the depot location plus the locations x_i for all i in the appropriate request set (either $N_V \cup N_S$ or only N_V), and A is comprised of arcs linking depot and request locations. In our study, arcs are assumed to be reliable with deterministic travel times t_{jk} on each arc (j,k). MAT-V and MRS-V are modeled as deterministic VRP variants, while MAT and MRS are modeled as VRP variants where one class of requests is deterministic (N_V) and one is stochastic (N_S). Rather than provide these well known formulations here, we refer the reader to [6] and [7] for comprehensive reviews of deterministic and stochastic VRP variants, respectively. In what follows, we present basic insights from two extreme cases for the problems being studied: (i) every unverified request in N_S is inaccurate, and (ii) every unverified request in N_S is accurate. Furthermore, we restrict the types of inaccurate requests considered to those representing needs that never actually existed.

3.1 Minimize arrival time (MAT/MAT-V)

In MAT and MAT-V, all requests are served, and the objective is to minimize the average arrival time at request locations. Suppose all requests in N_S are inaccurate, so that all true need requests are expressed in N_V. Clearly, solving MAT-V will yield the optimal solution, while solving MAT will result in average arrival times that are at least as great as the solution provided to MAT-V. If instead all requests in N_S are accurate, the opposite is true. Solving MAT yields the true optimal solution. Solving MAT-V would leave all requests in N_S unserved for the time being. While these needs may eventually be discovered by more traditional means and serviced at some point, the arrival time of the relief support will be much later.

3.2 Maximize requests served (MRS/MRS-V)

In these problems, the objective is to serve as many requests as possible by a fixed time limit T. If all requests in N_S are inaccurate, then an upper bound on the number of requests that can be served when solving both MRS and MRS-V is $|N_V|$. If an instance is tightly constrained with respect to T, fewer requests will be served. Denoting optimal solutions to these problems as MRS* and MRS-V*, it is clear that MRS* <= MRS-V*, because time spent driving to and evaluating inaccurate requests in N_S will consume valuable time that could be spent serving those in N_V. If instead all requests in N_S are accurate, then an upper bound on the number of requests that can be served in both problems is $|N_V+N_S|$. This bound can only possibly achieved when solving MRS, and in this case, we find that MRS* >= MRS-V*.

4. RESEARCH QUESTIONS

Our future work will expand on the presented models and attempt to answer several of the many questions that remain around the usefulness of data obtained from social media platforms. These include evaluation of 1) the impact on the planning process of varying the spatial distributions of N_S and N_V requests, and 2) the impact on solution quality when varying the proportion (or distribution) of inaccurate requests in N_S from the extreme cases already presented. Additional extensions will include evaluation when arc quality is stochastic and analysis of policy implications when social media usage demographics are incorporated.

5. REFERENCES

[1] Harvard Humanitarian Initiative. 2011. *Disaster Relief 2.0: The Future of Information Sharing in Humanitarian Emergencies*. Washington, D.C. and Berkshire, UK: UN Foundation & Vodafone Foundation Technology Partnership.

[2] Hagan, J. Tweet science. *New York Magazine.*(Oct. 10 2011)

[3] Wardell, C., San Su, Y. 2011. *2011 Social Media + Emergency Management Camp: Transforming the response enterprise.* Technical Report. CNA. Alexandria, VA.

[4] de la Torre, L. E., Dolinskaya, I. S., and Smilowitz, K.R. 2012. Disaster relief routing: Integrating research and practice. *Socio-Economic Planning Sciences.* 46,1(Mar. 2012), 88-97.

[5] Kramer, M., Costello, R., and Griffith, J. 2009. Investigating the force multiplier effect of citizen event reporting by social simulation. *Mind & Society.* 8, 2 (Dec. 2009), 209-221.

[6] Toth, P. and Vigo, D. 2002. *The Vehicle Routing Problem*. SIAM Monographs on Discrete Mathematics and Applications, SIAM Publishing, 2002.

[7] Gendreau, M., Laporte, G., and Seguin, R. 1996. Stochastic vehicle routing. *European Journal of Operational Research* 88, 1(Jan. 1996), 3-12.