Modeling Civil Conflict and Aid Delivery in Uganda

Peter J. Bull
Institute for Applied Computational Science
Harvard University
52 Oxford Street
Cambridge, MA 02139
bull@fas.harvard.edu

Isaac M. Slavitt
Institute for Applied Computational Science
Harvard University
52 Oxford Street
Cambridge, MA 02139
slavitt@fas.harvard.edu

Abstract

Using MCMC techniques, we model civil conflict in Uganda and optimize the limited relief aid that can be provided in these scenarios. We describe a method to simulate civil conflict events in space and time given historical data about these events. We also optimize the distribution of limited aid resources to refugees from these crises. Modeling aid delivery as a combination of the traveling salesman problem and the knapsack algorithm — two NP-hard problems — we find acceptable solutions using stochaastic metaheuristics.

The Data

The data comes from ACLED (Armed Conflict Location and Event Data Project), which is a dataset with locations, dates, fatalities, motivation, actors involved, and other information about civil conflicts in Africa. Their collection of data on Uganda covers 1997-2013, and they have a real-time tracker of events reported in 2014.[1] The need for an understanding of these patterns of conflict is clear, as ACLED notes:

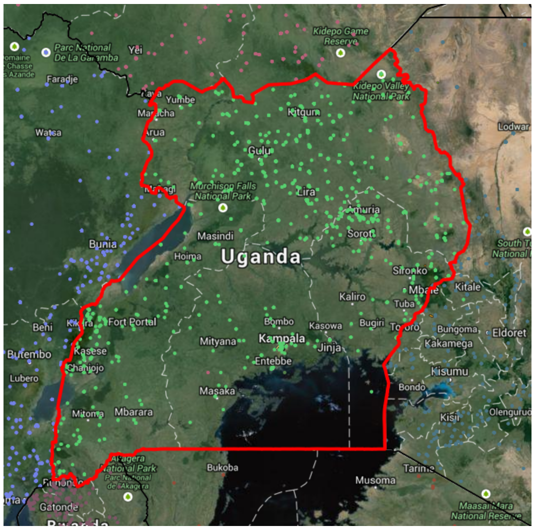
> This dataset codes the dates and locations of all reported political violence events in over 50 developing countries. Political violence includes events that occur within civil wars or periods of instability. Although civil war occurrence is decreasing across African countries, new forms of political violence are becoming more common.

For Uganda, the dataset contains around 4,500 observations of civil violence. Each observation includes the data, location, an quantified estimate for how precise these measures are, a number of fatalities for the event, and the actors involved. Figure 1a shows these conflicts scattered on a political map of Uganda.

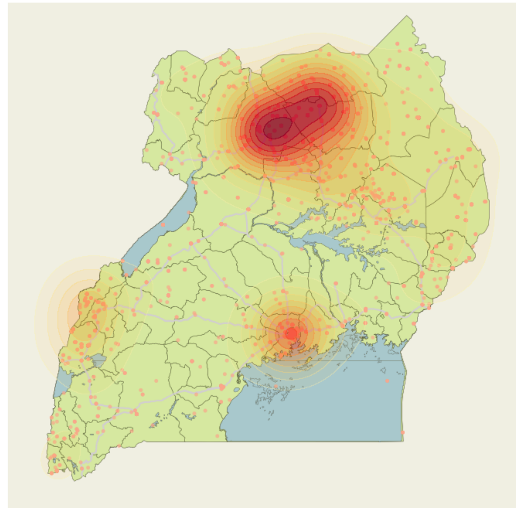
Modeling civil conflict events in space

While we could have trivially fit a Gaussian Process (GP) using the given data, we thought it would be more interesting to apply more directly the MCMC sampling methods learned in this class. From the MCMC standpoint, it was simple to treat the entire country of Uganda as a probability distribution from which geospatial conflict events could be sampled. We took historical conflict location data from the entire ACLED data set and smoothed it using a Matrn covariance function. Figure 1b shows this smoothing applied to the same conflicts depicted in 1a.

We then discretized that smooth function, which left us with a 2D matrix from which it was easy to sample with any number of samples. Although we experimented with many different methods, we ultimately settled on slice sampling for its favorable, efficient properties. Figure 2 shows this matrix.



(a) Civil conflicts in Uganda 1997-2013.



(b) Conflicts with Matrn smoothing.

Figure 1: Civil conflicts in Uganda.

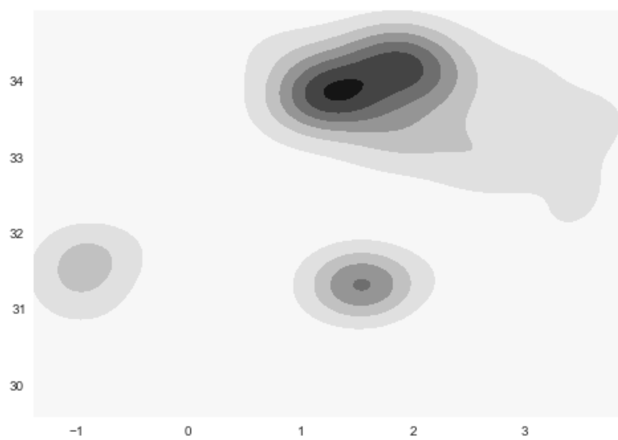


Figure 2: Discetized matrix for sampling.

Optimizing Resource Placement and Allocation

In the second part of this project, we use the first section as both inspiration for the aid delivery analogy and also as a source of randomly sampled data points representing geospatially distributed conflicts.

The Traveling Salesman Problem

One question of particular interest that we can tackle using techniques from the class is how to route emergency aid to locations where it is needed. For concreteness, let's postulate a Red Cross medical or food supply caravan that originates from the organization's in-country headquarters. This caravan wishes to visit all n emergent locations in order to deliver needed supplies. They wish to do so in the most efficient manner possible.

This is fundamentally an optimization problem, and one that is well known – it was first described in 1932 by Karl Menger (shortly after his year here at Harvard as a visiting lecturer) and has been studied extensively ever since.[2] Here is the traditional convex optimization specification of the problem:[3]

$$
\begin{aligned}
&\min \sum_{i=0}^{n} \sum_{j\neq i, j=0}^{n} c_{ij} x_{ij} \\
&\text{s.t.} \\
&\quad x_{ij} \in \{0,1\} && i,j = 0,\cdots,n \\
&\quad \sum_{i=0, i\neq j}^{n} x_{ij} = 1 && j = 0,\cdots,n \\
&\quad \sum_{j=0, j\neq i}^{n} x_{ij} = 1 && i = 0,\cdots,n \\
&\quad u_i - u_j + n x_{ij} \leq n-1 && 1 \leq i \neq j \leq n
\end{aligned}
$$

As is clear from the specification, this is an integer linear program (ILP) where:

- x_{ij} is a binary decision variable indicating whether we go from location i to location j.
- c_{ij} is the distance between location i and location j.
- The objective function is the sum of the distances for routes that we decide to take.
- The final constraint ensures that all locations are visited once and only once.

Answering complex, realistic questions

Packing the aid truck - adding the Knapsack Problem

We extend the TSP into a multi-objective optimization problem where *the contents of the aid trucks* also have an optimization component. Therein lies the knapsack problem: subject to a volume or weight constraint, and given that different locations might have very different needs such as food, vaccinations, or emergent medical supplies, *which supplies do we pack on the trucks*?

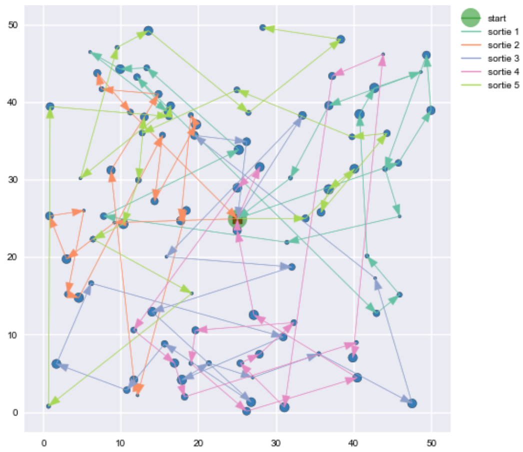
Here's the unbounded[1] version of the knapsack problem:

$$
\begin{aligned}
&\max \sum_{i=1}^{n} v_i x_i \\
&\text{s.t.} \\
&\quad x_i \in \mathbb{Z} \\
&\quad x_i \geq 0 \\
&\quad \sum_{i=1}^{n} w_i x_i \leq W
\end{aligned}
$$

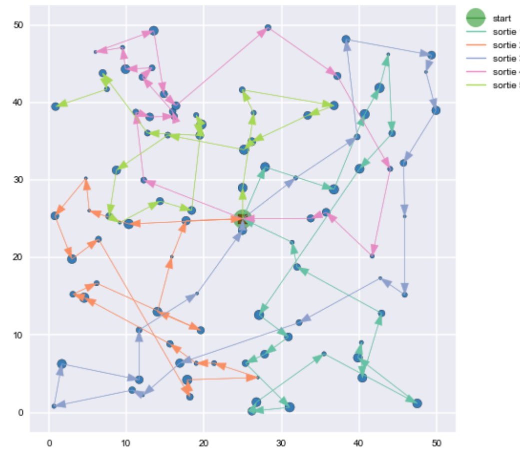
In this formulation:

- x_i is a zero or positive integer decision variable indicating how many units of item i we load on the truck.
- v_i is the utility we get from bringing along item i.
- w_i is the weight of item i.
- W is the maximum weight the truck can carry.

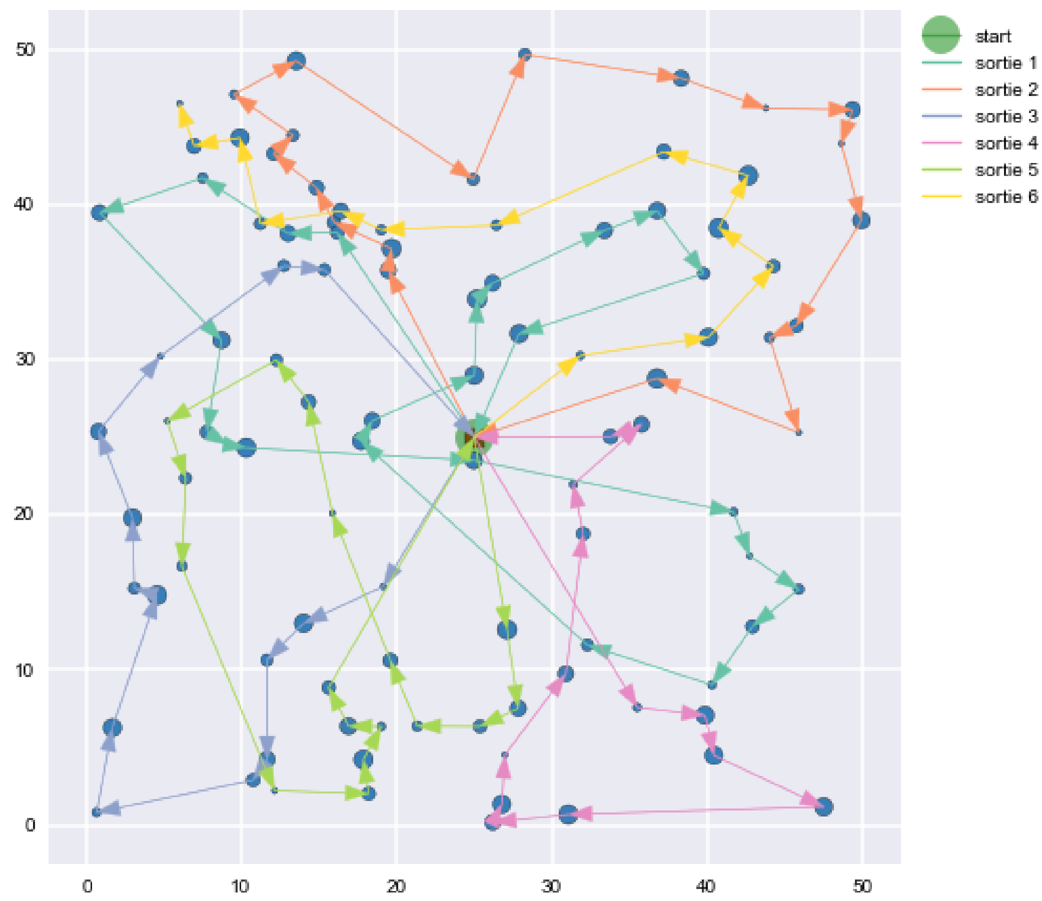
[1] Often, this problem is formulated such that you can only bring one of each item, but that doesn't make sense here. We want to be able to bring as many types of each type of aid as we think necessary, and we'll assume that as many as desired are available to load on the trucks before starting out from HQ.



(a) After 5,000 iterations.



(b) After 20,000 iterations.



(c) After 100,000 iterations.

Figure 3: Routing for the TSP/Knapsack hybrid.

The problem, of course, is that brute force solution of the TSP is $\mathcal{O}(n!)$. Traditional, deterministic algorithm approaches such as branch-and-bound or branch-and-cut are still impractical for larger numbers of nodes. In many cases, exhaustive search for global optimality is not even particularly helpful as long as the solution found is good enough. We will use simulated annealing (SA) to get acceptable solutions to the TSP (c.f. the class lectures and homework problem).

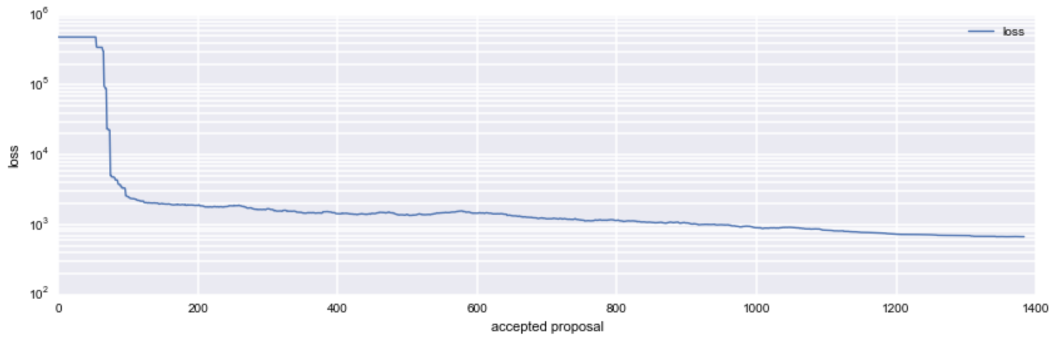


Figure 4: Loss function acceptances over 100,000 iterations.

Finding the optimal site for the resupply location

Lorem ipsum dolor sit amet, consectetur adipisicing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Factoring in travel time and loss from service delays

Lorem ipsum dolor sit amet, consectetur adipisicing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

References

[1] C. Raleigh, A. Linke, H. Hegre, and J. Karlsen, "Introducing ACLED-armed conflict location and event data," *Journal of Peace Research*, vol. 47, no. 5, pp. 1–10, 2010.

[2] K. Menger, "Das botenproblem," *Ergebnisse eines Mathematischen Kolloquiums*, vol. 2, pp. 11–12, 1932.

[3] W. Winston, *Operations Research: Applications and Algorithms*. Thomson Brooks/Cole, 2004.