

Delivery Volume Improvement for Planning Medication Distribution

Jeffrey W. Herrmann, Sara Lu, Kristen Schalliol

Institute for Systems Research
University of Maryland
College Park, Maryland, USA
jwh2@umd.edu

Abstract—This paper considers the problem of planning the logistics of distributing medication to points of dispensing (PODs), which will give medication to the public. Previous work on a two-stage routing and scheduling approach showed that it can generate solutions with reasonable minimum slack. This paper presents a delivery volume improvement algorithm that can increase the minimum slack of a given solution.

Keywords—logistics and inventory, vehicle routing, scheduling, homeland security

I. INTRODUCTION

Improving emergency preparedness requires planning responses to bioterrorist attacks. In the case of a large scale bioterrorist event, such as the release of anthrax, public health officials may decide that mass dispensing of medication is needed. According to the Centers for Disease Control and Prevention, large cities and metropolitan areas need to dispense antibiotics to their entire identified population within 48 hours of the decision to do so [1]. Cities in every state are developing plans for opening points of dispensing (PODs) to give prophylactic medication to persons who are currently healthy but may have been exposed to a pathogen. PODs may be setup in schools, recreation centers, churches, and other non-medical facilities. Other modes of dispensing medication are being considered, but PODs are the primary focus of planning activities.

The proposed research is motivated by work with county public health departments in the state of Maryland who must plan the logistics for distributing medication to the PODs from a central location. We consider the problem at the state and local levels (not the national level). After the decision for mass dispensing is made, county public health departments will begin preparing to open multiple PODs simultaneously at a designated time. The state will request medication from the federal government, who will deliver an initial but limited supply of medication to a state receipt, storage, and stage (RSS) facility (which we call the “depot”). Contractors will deliver more medication to the depot, but the state will begin shipping medication from the depot to the PODs before everything arrives from the contractors. The deliveries to the depot arrive in batches that we call “waves.”

Poor medication distribution plans will delay the time that some PODs receive medication. This can delay the opening of

these PODs, and some residents may not get their medication in a timely manner, which increases their risk of death or illness. Clearly, there are many uncertainties in medication distribution, including the timing of shipments to the depot, the time needed to load and unload trucks, travel times, and the demand for medication at each POD. For this reason, planners need a robust plan. In particular, it is better if the plan calls for delivering medication to PODs much earlier than it is needed. This improves the likelihood that the PODs will open on-time, will not run out of medication during operations, and will dispense medication to the largest number of people in a timely manner.

The operations of firefighters, emergency medical services, and police departments have motivated research into location models [2-4] and dynamic vehicle routing models [5-7]. However, these models are not relevant to the medication distribution problem, which is more closely related to the inventory routing problem [8-11] and the production-distribution scheduling problem [12]. Still, the models used for those problem are also not directly relevant.

This paper addresses the single-product, deterministic problem. Inventory is treated as a continuous variable, but the number of pallets must be an integer. We measure the medication with the number of regimens. In mass dispensing, each person will get one predetermined regimen, which is a bottle with a specific number of pills. All PODs have the same hours of operation, and loading and unloading times are independent of the quantity. We are ignoring other resources such as the loading docks at the depot, the available drivers, and the number of available pallets. This paper presents a delivery volume improvement technique that can increase the minimum slack of a given solution. More details about the problem and the approach can be found in [13].

II. PROBLEM FORMULATION

A problem instance specifies the following information. Without loss of generality we let time $t = 0$ correspond to the first instant that the depot has medication. PODs will begin operating at time $t = T_1$ and continue to operate until time $t = T_2$. In practice, these times may be on the order of 12 to 48 hours.

There are n sites (which correspond to PODs). Each site ($i = 1, \dots, n$) has a dispensing rate of L_i regimens per time unit. This is the rate at which the site consumes medication. The site needs a total of $(T_2 - T_1)L_i$ regimens. There is a depot ($i = 0$) that receives medication from a supplier. Let $I(t)$ be the cumulative amount of medication delivered to the depot at time t . $I(t)$ is a discontinuous, non-decreasing function due to the batch deliveries that are made there.

The time spent at site i (to load or unload a vehicle) is p_i for $i = 0, \dots, n$. The time to go from site i to site j is c_{ij} . There are V vehicles. Vehicle v has a capacity of C_v pallets of material. At each site, a vehicle will deliver one or more pallets. A pallet can hold at most P regimens.

Given a problem instance, a solution specifies one or more routes for each vehicle and the quantity delivered at each site. The key decision variables are the sequences, start times, and delivery quantities. Let r_v be the number of routes that vehicle v makes. For the j -th route for vehicle v , $m(v, j)$ is the number of sites on the route, $\sigma_{vj} = \{i_1, \dots, i_{m(v, j)}\}$ is the sequence of sites that the vehicle visits, and t_{vj} is the start time at which the vehicle begins loading at the depot. Finally, q_{vjk} is the quantity delivered to each site $k \in \sigma_{vj}$. Note that $q_{vjk} = 0$ if and only if $k \notin \sigma_{vj}$.

Given a solution, we can evaluate its feasibility as follows. Let y_{vj} be the total duration of a route. Let w_{vjk} be the duration between the start of the route and the time that the delivery at site $k \in \sigma_{vj}$ is complete. Let h_{vjk} be the (integer) number of pallets required to deliver q_{vjk} regimens to site $k \in \sigma_{vj}$.

$$y_{vj} = p_0 + c_{0i_1} + p_{i_1} + c_{i_1i_2} + \dots + p_{i_{m(v,j)-1}i_m} + c_{i_mi_0} \quad (1)$$

$$w_{vjk} = p_0 + c_{0i_1} + p_{i_1} + c_{i_1i_2} + \dots + p_{i_{j-1}i_j} \quad (2)$$

Certain constraints must be satisfied for the solution to be feasible. The quantity shipped from the depot cannot exceed the amount delivered to the depot (Equation 3). A vehicle cannot begin a new route until it returns to the depot (Equation 4). The number of pallets used must be sufficient (Equation 5). The vehicle capacity cannot be exceeded on any route (Equation 6). Each and every site must receive all needed medication (Equation 7). All route start times must be non-negative (Equation 8).

$$\sum_{(a,b):t_{ab} \leq t_{vj}} \sum_{k \in \sigma_{ab}} q_{abk} \leq I(t_{vj}) \quad v = 1, \dots, V; j = 1, \dots, r_v \quad (3)$$

$$t_{vj} \geq t_{v,j-1} + y_{v,j-1} \quad v = 1, \dots, V; j = 2, \dots, r_v \quad (4)$$

$$0 \leq q_{vjk} \leq Ph_{vjk} \quad v = 1, \dots, V; j = 1, \dots, r_v \quad (5)$$

$$\sum_{k \in \sigma_{vj}} h_{vjk} \leq C_v \quad v = 1, \dots, V; j = 1, \dots, r_v \quad (6)$$

$$\sum_{v=1}^V \sum_{j=1}^{r_v} q_{vjk} = (T_2 - T_1)L_k \quad k = 1, \dots, n \quad (7)$$

$$t_{vj} \geq 0 \quad v = 1, \dots, V; j = 1, \dots, r_v \quad (8)$$

The problem is to find a feasible solution with the largest amount of minimum slack. Given a feasible solution, evaluating its minimum slack requires measuring the slack of each delivery. For each site k , order all of the routes (v, j) such that $k \in \sigma_{vj}$ by the delivery time $t_{vj} + w_{vjk}$. Then, let the set E_{vjk} be those routes (a, b) that occur before route (v, j) in this ordering, and let Q_{vjk} be the total quantity delivered to site k on these earlier routes:

$$Q_{vjk} = \sum_{(a,b) \in E_{vjk}} q_{abk} \quad (9)$$

Let s_{vjk} be the slack of the delivery to site k in route (v, j) . That is, if this delivery were delayed more than s_{vjk} time units and no more medication were delivered to the site, the site would run out of medication. The time at which site k would run out of medication is $T_1 + Q_{vjk} / L_k$. The minimum slack S of a solution is the minimum slack over all vehicles, routes, and sites.

$$s_{vjk} = T_1 + Q_{vjk} / L_k - (t_{vj} + w_{vjk}) \quad (10)$$

$$S = \min_{v=1, \dots, V} \min_{j=1, \dots, r_v} \min_{k \in \sigma_{vj}} \{s_{vjk}\} \quad (11)$$

III. EXAMPLE

Consider a five-site, three-vehicle problem instance. This example uses five POD locations and three trucks from the TourSolver templates. The PODs will operate for 10 hours (starting 10 hours after the start of deliveries). The dispensing rates of the five PODs are given in Table I. The total number of regimens needed is 676,190. These regimens are delivered to the depot in three waves of 200,000, 240,000, and 236,190 regimens (note that the second wave is the largest). The waves are 4 hours apart. The load and unload times p_i were all set to 10 minutes. $T_1 = 10$ hours = 600 minutes. $T_2 = 20$ hours. $P = 10,000$ regimens per pallet. $C_1 = 20$ pallets. TourSolver generates three routes. Truck 1 visits POD 2, Truck 2 visits PODs 3 and 4, and Truck 3 visits PODs 5 and 1.

The delivery quantities are proportional to the dispensing rates and are shown in Table I. Each vehicle starts its route at times $t_{v1} = 0$, $t_{v2} = 240$, and $t_{v3} = 480$ (corresponding to the waves delivered at times 0, 4, and 8 hours). Table II shows the delivery durations w_{vjk} and the slack for each POD in each wave, which is calculated using (10). The runout times are

600, 777, and 990 minutes in waves 1, 2, and 3. The minimum slack is 429 minutes (just over 7 hours); this occurs at POD 1 in the third wave. Note that POD 1 is the last POD to receive its delivery in each wave.

TABLE I. DELIVERY QUANTITIES FOR EXAMPLE. ALL QUANTITIES IN REGIMENS.

POD	L_k (regimens per hour)	Wave 1 quantity	Wave 2 quantity	Wave 3 quantity
1	10,985	32,491	38,989	38,370
2	11,957	35,366	42,439	41,765
3	14,322	42,361	50,833	50,026
4	14,516	42,934	51,522	50,704
5	15,839	46,848	56,217	55,325

TABLE II. SLACK CALCULATIONS FOR EXAMPLE. ALL TIMES IN MINUTES.

POD	Delivery duration w_{vjk}	Wave 1 slack	Wave 2 slack	Wave 3 slack
1	81	519	456	429
2	54	546	483	456
3	57	543	480	453
4	69	531	468	441
5	50	550	487	460

IV. SOLUTION APPROACH

Instead of attempting to solve the problem as a large integer program, we adopt a two-stage solution approach that separates the problem into two subproblems: (1) the “routing problem” assigns sites to vehicles and creates routes for each vehicle, and (2) the “scheduling problem” determines when the vehicles should start these routes and how much material should be delivered to each site on each trip. In this approach, each available vehicle will have exactly one route (sequence of sites). A vehicle may perform that route more than once with different delivery quantities each time. For a complete description of the entire two-stage solution approach, see [13].

In this section we describe the delivery volume improvement algorithm.

After some experience with these solutions, we discovered that carefully manipulating the delivery quantities can increase minimum slack significantly. For example, consider the example presented in Section III. As shown in Table II, in the third wave of deliveries, the slack at site 1 is only 429 minutes, while the slack of site 5 is 460 minutes. Adjusting the delivery quantities in the first two waves will modify the slack at these sites. By delivering more to site 1 (and less to site 5), we increase the slack at site 1 (and decrease the slack at site 5). The best we can do is to make them equal. This is the objective of the delivery volume improvement algorithm,

which uses the following variables: Let T_1 be the start of dispensing. Let K_N be the target slack for wave N (for $N > 1$). Let X_{Nk} be the time that site k will receive a delivery in wave N . Let $Q_{1k}, \dots, Q_{N-1,k}$ be the amount delivered to site j in waves 1 to $N-1$. Note that $Q_{1k}, \dots, Q_{N-2,k}$ are known.

Finding $Q_{N-1,k}$ is the goal. Based on the definition of slack we want the following to hold for every site that receives a delivery in wave N :

$$T_1 + \frac{Q_{1k} + \dots + Q_{N-2,k} + Q_{N-1,k}}{L_k} - X_{Nk} = K_N \quad (12)$$

From this, we can determine $Q_{N-1,k}$ for each site k :

$$Q_{N-1,k} = (K_N + X_{Nk} - T_1)L_k - (Q_{1k} + \dots + Q_{N-2,k}) \quad (13)$$

The best solution will occur by picking K_N as large as possible so that inventory and vehicle capacity constraints are satisfied. Let I_{N-1} be the total amount shipped to the depot in the first $N-1$ waves. Let R_N be the set of sites that will receive a delivery in wave N . Then, we have the following constraint on K_N based on the inventory available:

$$K_N \leq \left(I_{N-1} - \sum_{w=1}^{N-2} \sum_{k \in R_N} Q_{wk} - \sum_{k \in R_N} (X_{Nk} - T_1)L_k \right) / \sum_{k \in R_N} L_k \quad (14)$$

If vehicle v performs route j in wave $N-1$, then we have a similar (but approximate) constraint based on the vehicle’s capacity:

$$K_N \leq \left(C_v P + \sum_{w=1}^{N-2} \sum_{k \in \sigma_{vj}} Q_{wk} - \sum_{k \in \sigma_{vj}} (X_{Nk} - T_1)L_k \right) / \sum_{k \in \sigma_{vj}} L_k \quad (15)$$

To illustrate this approach, let us consider the five-site example presented in Section III. In this case, all of the sites are visited in every wave, so $R_N = \{1, 2, 3, 4, 5\}$ for $N = 2$ and $N = 3$. We start with $N = 2$ to maximize the slack in the wave 2 deliveries by setting appropriate delivery quantities for the first wave. $I_{N-1} = 200,000$ regimens, and $X_{Nk} = 240 + w_{vjk}$ minutes. The term $-\sum_{k \in R_N} (X_{Nk} - T_1)L_k = 336,624$ regimens, and $\sum_{k \in R_N} L_k = 1,126.98$ regimens per minute. Thus, K_N must be no greater than $(200,000 + 336,624)/1,126.98 = 476.16$ minutes. (If it were larger, then there would not be enough material available in the first wave to satisfy the resulting delivery quantities.) (The vehicle capacity constraints are not tight for this wave.) From this value, we can determine the delivery quantities for Wave 1 using (13). These are shown in Table III.

For $N = 3$, $I_{N-1} = 440,000$, and $X_{Nk} = 480 + w_{ijk}$. Thus, K_N must be no greater than 449.12 minutes. From this value, we can determine the delivery quantities for Wave 2 using (13). These are shown in Table III. (The vehicle capacity constraints are again not tight for this wave.)

Wave 3 must deliver enough to complete each POD's requirements for the 10-hour dispensing campaign. The delivery quantities are shown in Table III, and the slacks in Table IV.

After delivery volume improvement, PODs 1 and 4 received significantly more regimens in the first wave, about the same in the second wave, and fewer regimens in the third wave. The other three PODs received significantly fewer regimens in the first wave, about the same in the second wave, and more regimens in the third wave. In the second wave, every POD had a slack of 476 minutes; in the third wave, every POD had a slack of 449 minutes. (The first wave slacks remained the same.) Thus, delivery volume improvement increased the minimum slack from 429 to 449 minutes.

TABLE III. DELIVERY QUANTITIES FOR EXAMPLE AFTER DELIVERY VOLUME IMPROVEMENT. ALL QUANTITIES IN REGIMENS.

POD	Wave 1 quantity	Wave 2 quantity	Wave 3 quantity
1	36,067	38,989	34,764
2	33,910	42,439	43,221
3	41,333	50,833	51,054
4	44,796	51,522	48,842
5	43,863	56,217	58,309

TABLE IV. SLACK CALCULATIONS FOR EXAMPLE AFTER DELIVERY VOLUME IMPROVEMENT. ALL TIMES IN MINUTES.

POD	Wave 1 slack	Wave 2 slack	Wave 3 slack
1	519	476	449
2	546	476	449
3	543	476	449
4	531	476	449
5	550	476	449

V. RESULTS

For testing the approaches, we considered two scenarios that we will briefly describe here.

The first scenario used 50 PODs from the TourSolver example [15]. We assumed that each POD will open 12 hours after medication distribution begins and will operate for 12 hours. The total number of regimens dispensed is given by the example. The PODs dispense between 10,000 and 53,000 regimens. The total number of regimens is 950,389. Medication will arrive at the depot in five waves every three hours. Each one of the first four waves will supply 200,000 regimens, and the fifth will supply 150,389 regimens. The nine

trucks provided in the example have different capacities from 5 to 40 pallets, and each pallet can hold at most 11,200 regimens.

First, we created a CVRP instance in which just over 21% of each POD's total quantity is delivered. (The total number of regimens was 199,982.) We first ran TourSolver using a 3 hour run time for each truck. This yielded a solution that used all nine trucks, but some trucks were busy for all of the three hours, and others had much shorter routes. Using routes with such duration variability will yield a poor solution. Because the average time that a truck was busy was approximately 2:20, we ran TourSolver again using a 2:20 run time for each truck. In the resulting solution, six trucks were busy for nearly the entire run time. The three other trucks required some overtime (5, 10, and 40 minutes). Thus, the routes were more nearly equal in duration.

We constructed schedule in which each route starts after each wave. The delivery quantities were the same for the first four waves and smaller for the fifth wave. The minimum slack was 354 minutes, and this occurred in the fifth wave. The slack decreased with each wave because the depot receives only 2.4 hours' worth of medication every 3 hours.

We then used delivery volume improvement to adjust the delivery quantities. The minimum slack in every wave after the first increased, and the overall minimum slack was 418 minutes (an 18% improvement).

We also considered a simple dispatching scheme in which the entire quantity for a POD is delivered in two shipments. We reused the same routes and prioritized them by the travel time to the POD so that the routes with the longest travel times were done first. The minimum slack was 181 minutes, which occurs in the fifth wave.

The second scenario used realistic data from three counties in the state of Maryland. Medication arrives at the state RSS (depot) in seven waves, one every two hours, with roughly the same amount of medication in each wave. A total of 189 PODs will dispense medication from $T_1 = 24$ hours to $T_2 = 48$ hours. The total number of trucks available is 71, and each has a capacity of 268,800 regimens.

First, we ran TourSolver to create routes. Then, we created a schedule in which each vehicle started its route after each wave. The minimum slack was 360 minutes, which occurred during the seventh wave. The slack of some vehicles' routes decreased with each wave because their route durations exceeded the time value of the deliveries (in the single wave case, each delivery supplied about 3.4 hours worth of medication).

To improve the slack of these solutions, we adjusted the delivery quantities so that PODs that were visited later in a route received more material in the first delivery. This increased the time at which the POD would run out. The delivery volume improvement technique described above set the delivery quantities of one "wave" so that the slacks at every POD during the next wave were the same. Using this technique dramatically increased the minimum slack to 552 minutes.

However, the potential improvement of delivery volume improvement was limited because the minimum slack often occurred on the first delivery to a POD with the latest delivery time, which was determined by the routes generated by TourSolver.

VI. SUMMARY AND CONCLUSIONS

This paper introduced the medication distribution problem, an important part of planning the response to a bioterrorism attack, and presented a delivery volume improvement technique that can be used with a two-stage routing and scheduling approach for constructing solutions. Because a robust plan is desirable, our objective was to maximize the minimum slack of the solution. Instead of attempting to solve the problem as a large integer program, we adopted a two-stage solution approach that separates the problem into two subproblems. This practical separation reduces the solution effort, though it is not guaranteed to find an optimal solution.

To demonstrate the approach, we applied it to two scenarios, including one for three counties in the state of Maryland.

Delivery volume improvement increased slack, dramatically in some cases. The extra hours of slack could be critical in an emergency. They show that a careful analysis of the scenario is necessary to construct an effective medication distribution plan.

Future work is needed to automate the routing and scheduling approach to enable a decision support tool for public health emergency preparedness planners, to develop optimization techniques for finding even better solutions, and to test these approaches on other scenarios.

ACKNOWLEDGMENT

The work of Sara Lu and Kristen Schalliol occurred during their participation in the NSF REU site “Introducing the Systems Engineering Paradigm to Young Researchers and Future Leaders” (NSF grant EEC 0243803). Kay Aaby, Rachel Abbey, and Kathy Wood at the Montgomery County, Maryland, Public Health Services provided excellent guidance and assistance. Cooperative Agreement Number U50/CCU302718 from the CDC to NACCHO supported this publication. Its contents are solely the responsibility of the University of Maryland and do not necessarily represent the

official views of CDC or NACCHO. The discussion of related work relies in part on material prepared by Zhi-Long Chen.

REFERENCES

- [1] Centers for Disease Control and Prevention (CDC), “Key Facts about the Cities Readiness Initiative (CRI),” <http://www.bt.cdc.gov/CRI/facts.asp>, accessed September 16, 2008.
- [2] M.S. Daskin and E.H. Stern, “Hierarchical objective set covering model for emergency medical service vehicle deployment,” *Transportation Science*, vol. 15, pp. 137 – 152, 1981.
- [3] M.O. Ball and F.L. Lin, “A reliability model applied to emergency service vehicle location,” *Operations Research*, vol. 41, pp. 18 – 36, 1993.
- [4] A. Ceyhun, H. Selim, and I. Ozkarahan, “A fuzzy multi-objective covering-based vehicle location model for emergency services,” *Computers and Operations Research*, vol. 34, pp. 705 – 726, 2007.
- [5] R. Sivanandan, A.G. Hobeika, S.A. Ardekani, and P.B. Lockwood, “Heuristic shortest-path method for emergency vehicle assignment - a study on the Mexico City network,” *Transportation Research Record*, no. 1168, 1988, pp. 86 – 91, 1988.
- [6] A. Weintraub, J. Aboud, C. Fernandez, G. Laporte, and E. Ramirez, “An emergency vehicle dispatching system for an electric utility in Chile,” *Journal of the Operational Research Society*, vol. 50, pp. 690-696, 1999.
- [7] A. Haghani, Q. Tian, and H. Hu, “Simulation model for real-time emergency vehicle dispatching and routing,” *Transportation Research Record*, no. 1882, pp. 176 – 183, 2004.
- [8] M. Dror, M. Ball, and B. Golden, “Computational comparisons of algorithms for the inventory routing problem,” *Annals of Operations Research*, vol. 4, pp. 3 – 23, 1985.
- [9] A. Campbell, L. Clarke, A.J. Kleywegt, and M.W.P. Savelsbergh, “The inventory routing problem,” in *Fleet Management and Logistics*, T.G. Crainic and G. Laporte, eds. Dordrecht: Kluwer Academic Publishers, pp. 95-114, 1998.
- [10] F. Baita, W. Ukovich, R. Pesenti, and D. Favaretto, “Dynamic routing-and-inventory problems: A review,” *Transportation Research*, Part A, vol. 32, pp. 585 – 598, 1998.
- [11] N.H. Moin and S. Salhi, “Inventory routing problems: A logistical overview,” *Journal of the Operational Research Society*, vol. 58, pp. 1185 – 1194, 2007.
- [12] Z.-L. Chen, “Integrated production and outbound distribution scheduling: Review and extensions,” *Operations Research*, in press.
- [13] J.W. Herrmann, S. Lu, and K. Schalliol, “A Routing and Scheduling Approach for Planning Medication Distribution,” Proceedings of the 2009 Industrial Engineering Research Conference, Miami, Florida, May 30 – June 3, 2009.
- [14] P. Toth and D. Vigo, “Exact Solution of the Vehicle Routing Problem,” in *Fleet Management and Logistics*, T.G. Crainic and G. Laporte, eds. Dordrecht: Kluwer Academic Publishers, pp. 1-31, 1998.
- [15] C2Logix, “SNSTourSolver,” <http://cdcestockpilerouting.c2logix.com/Citrix/AccessPlatform/auth/login.aspx>, accessed July 25, 2008.