Same-Day Delivery: Tactical Design

Alejandro Toriello

Stewart School of Industrial and Systems Engineering
Georgia Institute of Technology

joint with Alex Stroh, Alan Erera

IEOR-DRO Seminar
Columbia University
September 8, 2020

E-Commerce

- Pre-COVID, e-commerce was already a large and growing sector of retail and overall economy.
  - About or above 10% of all US retail since 2013 (Forrester Research).
  - Average annual online spending to reach $2,000 per buyer in 2018 (Forrester Research).
  - Amazon alone accounts for almost half of US e-retail (eMarketer).
  - Amazon now second to Walmart in terms of global employment numbers (566K vs. 2.3M); both very active in e-retail (Fortune).
- COVID has only accelerated these trends.

E-Commerce


Amazon Retail Ecommerce Sales
US, 2016-2019
$149.79
$199.85
$258.22
$325.04
38.3%
44.1%
49.1%
53.7%
2016
2017
2018
2019
billions
% of total retail ecommerce sales
Source: eMarketer, June 2018
www.eMarketer.com

Same-Day Delivery

- Intense competition, constant need for innovation – the customer wants it NOW.
- Same-day delivery (SDD) further erodes brick-and-mortar advantage. But...

Same-Day Delivery

- Intense competition, constant need for innovation – the customer wants it NOW.
- Same-day delivery (SDD) further erodes brick-and-mortar advantage. But...
  - Extremely costly "last mile".
  - Lower order numbers, fewer economies of scale.
  - Fewer than 1/4 of customers willing to pay, and then only small amount (McKinsey).
  - Flat fees (e.g. Amazon Prime) may help amortize costs.

Same-Day Delivery

What's new?

- Traditional delivery: order acceptance, picking and packing *before* last-mile distribution.

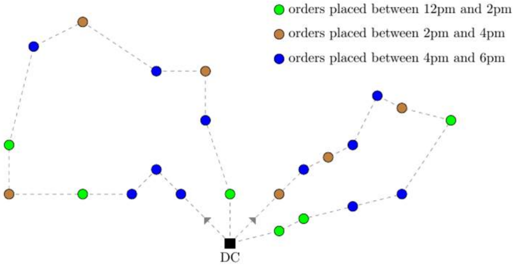
Same-Day Delivery

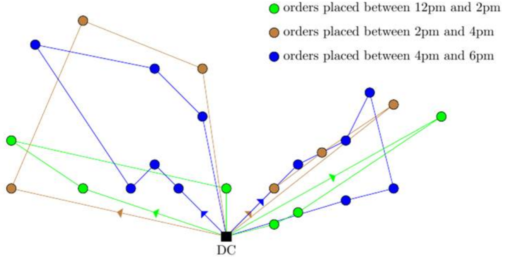
What's new?

- Traditional delivery: order acceptance, picking and packing *before* last-mile distribution.
- Same-day delivery: *simultaneous* order acceptance, picking, packing and last-mile distribution.
  - This talk: Delivery by end of day/common order deadline.
  - Food/grocery delivery: order-specific delivery times, 30 minutes to two hours (Amazon Restaurants, GrubHub, Uber Eats, pizza delivery).

Same-Day Delivery

What's new?





Source: A. Erera

Same-Day Delivery

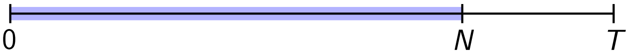
- Operational Models
  - Azi/Gendreau/Potvin (12,14), Campbell/Savelsbergh (05), Klapp/Erera/T. (18a,b,20), Ulmer (17a,b), Ulmer/Thomas (18), Ulmer/Thomas/Mattfeld (19), Voccia/Campbell/Thomas (17), ...
  - Can be used for tactical analysis, but complex and not transparent.

Same-Day Delivery

- Operational Models
  - Azi/Gendreau/Potvin (12,14), Campbell/Savelsbergh (05), Klapp/Erera/T. (18a,b,20), Ulmer (17a,b), Ulmer/Thomas (18), Ulmer/Thomas/Mattfeld (19), Voccia/Campbell/Thomas (17), ...
  - Can be used for tactical analysis, but complex and not transparent.
- Our Goal: Simple, “higher-level” model capturing typical system behavior.
  - What does the “average” SDD operating day look like?

Outline

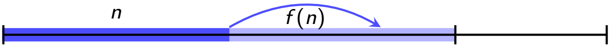
Tactical Dispatching Model



- Single depot with vehicle fleet serving fixed region.
- Orders appear at constant unit rate from 0 to N.
- All orders must be served, dispatches complete by $T > N$.
- Objective: Minimize total dispatching time.

Tactical Dispatching Model

Dispatch time

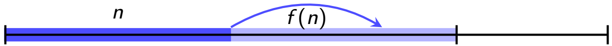


- A dispatch to serve n orders takes $f(n)$ time, where

$$f(0) = 0, \quad f \text{ is increasing}, \quad \text{concave}, \quad \text{can ``keep up''}.$$

Tactical Dispatching Model

Dispatch time



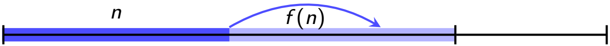
- A dispatch to serve n orders takes $f(n)$ time, where

$$f(0) = 0, \quad f \text{ is increasing}, \quad \text{concave}, \quad \text{can ``keep up''}.$$

- Motivation: $f(n) = a + bn + c\sqrt{n}$ for $n > 0$, where
  - $c\sqrt{n}$ is a BHH (59) routing time approximation,
  - assuming order locations are randomly distributed.

Tactical Dispatching Model

Dispatch time



- A dispatch to serve n orders takes $f(n)$ time, where

$$f(0) = 0, \quad f \text{ is increasing}, \quad \text{concave}, \quad \text{can "keep up"}.$$

- Motivation: $f(n) = a + bn + c\sqrt{n}$ for $n > 0$, where
  - $c\sqrt{n}$ is a BHH (59) routing time approximation,
  - assuming order locations are randomly distributed.

- *Continuous approximations* widely used in logistics (Franceschetti/Jabali/Laporte 17), including urban logistics (Carlsson/Song 18, Figliozzi 07, van Heeswijk/Mes/Schutten 17).

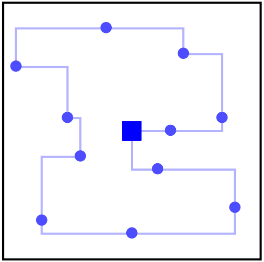
Tactical Dispatching Model

Dispatch time

- For example, for
  1. unit square service region, center depot,
  2. Manhattan distances,
  3. roughly 30 locations sampled uniformly,

  we estimate TSP length as $1.04\sqrt{n}$.

$\mathbb{E}[\text{TSP}_n] \approx 1.04\sqrt{n}$



- Asymptotic constant in this case estimated at ≈ 0.89 (Johnson/McGeoch/Rothberg 96).

Tactical Dispatching Model

Dispatch time

- Realistic situation:
  1. 8 mile by 8 mile service region (center depot)
  2. 25 mph average vehicle speed, Manhattan distances
  3. an order every 6 minutes
  4. 5-minute dispatch setup, 2-minute delivery per order
- We convert this to

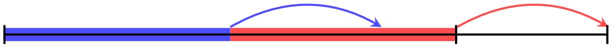
$$f(n) = 5/6 + 1/3n + 3.3\sqrt{n} \quad (\times\ 6 \text{ minutes}).$$

Optimal Structure

Concavity abhors balance

Dispatches should be as unbalanced as possible:

- This looks nice,

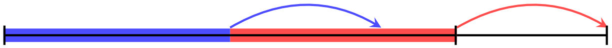


Optimal Structure

Concavity abhors balance

Dispatches should be as unbalanced as possible:

- This looks nice,



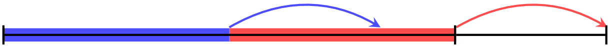
- but this is better,

Optimal Structure

Concavity abhors balance

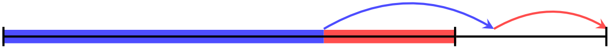
Dispatches should be as unbalanced as possible:

- This looks nice,



- but this is better,

- and so is this!



Consequences and Intuition

1. Decreasing dispatch lengths as day progresses.
   - Matches empirical observations in operational models (KET 18a,b).

Consequences and Intuition

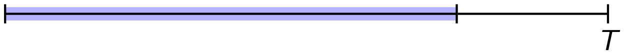
1. Decreasing dispatch lengths as day progresses.
   - Matches empirical observations in operational models (KET 18a,b).
2. Dispatching (and each vehicle) start inactive, then become active and remain so for rest of day.
   - Useful for shift scheduling.

Consequences and Intuition

1. Decreasing dispatch lengths as day progresses.
   - Matches empirical observations in operational models (KET 18a,b).
2. Dispatching (and each vehicle) start inactive, then become active and remain so for rest of day.
   - Useful for shift scheduling.
3. A dispatch takes all currently unserved orders.
   - Vehicles can be "pre-loaded".
   - Not necessarily true with geographic order discrimination.

Many Vehicles

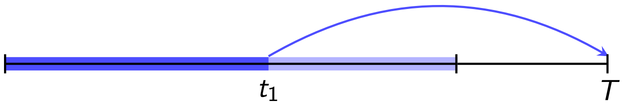
Optimal policy



- Each vehicle
  1. takes all available orders,
  2. leaves such that its dispatch ends at T.
- Compute by solving equations of the form

Many Vehicles

Optimal policy

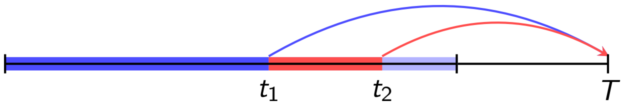


- Each vehicle
  1. takes all available orders,
  2. leaves such that its dispatch ends at T.
- Compute by solving equations of the form

$$t_1 + f(t_1) = T,$$

Many Vehicles

Optimal policy

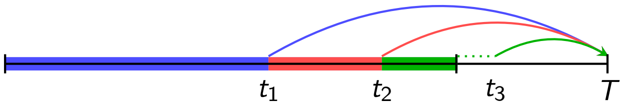


- Each vehicle
  1. takes all available orders,
  2. leaves such that its dispatch ends at T.
- Compute by solving equations of the form

$$t_1 + f(t_1) = T, \qquad t_2 + f(t_2 - t_1) = T,$$

Many Vehicles

Optimal policy



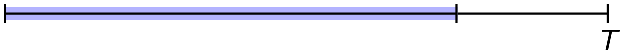
- Each vehicle
  1. takes all available orders,
  2. leaves such that its dispatch ends at T.

- Compute by solving equations of the form

$$t_1 + f(t_1) = T, \qquad t_2 + f(t_2 - t_1) = T,$$
$$t_3 + f(N - t_2) = T, \ldots$$

One Vehicle

Optimal policy

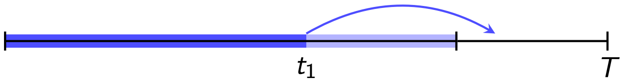


1. Each dispatch takes all available orders.
2. No waiting between dispatches.
3. Last dispatch returns at T.

* Minimum dispatch quantity for all dispatches except possibly last one.

One Vehicle

Optimal policy

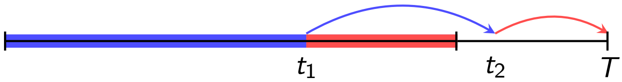


1. Each dispatch takes all available orders.
2. No waiting between dispatches.
3. Last dispatch returns at T.

* Minimum dispatch quantity for all dispatches except possibly last one.

One Vehicle

Optimal policy

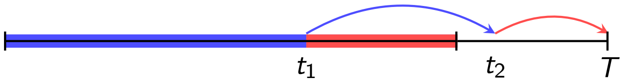


1. Each dispatch takes all available orders.
2. No waiting between dispatches.
3. Last dispatch returns at T.

* Minimum dispatch quantity for all dispatches except possibly last one.

One Vehicle

Optimal policy



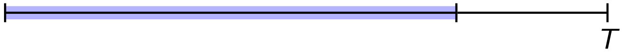
1. Each dispatch takes all available orders.
2. No waiting between dispatches.
3. Last dispatch returns at T.

* Minimum dispatch quantity for all dispatches except possibly last one.

- Try solving progressively higher-order equations:

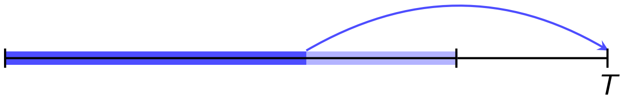
$$t_1 + f(N) = T, \qquad \text{(one dispatch)}$$
$$t_1 + f(t_1) + f(N - t_1) = T, \qquad \text{(two)}$$
$$t_1 + f(t_1) + f(f(t_1)) + f(N - t_1 - f(t_1)) = T, \ldots \qquad \text{(three)}$$

Finite Fleet



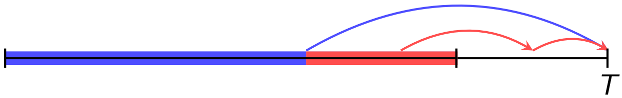
- Optimality depends on parameters; no general structure.
- Hybrid heuristic: For m vehicles,
    1. first $m-1$ follow many-vehicle policy,
    2. last one serves remainder with one-vehicle policy.

Finite Fleet



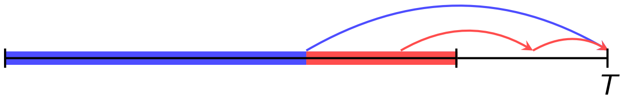
- Optimality depends on parameters; no general structure.
- Hybrid heuristic: For m vehicles,
    1. first $m-1$ follow many-vehicle policy,
    2. last one serves remainder with one-vehicle policy.

Finite Fleet



- Optimality depends on parameters; no general structure.
- Hybrid heuristic: For m vehicles,
  1. first $m - 1$ follow many-vehicle policy,
  2. last one serves remainder with one-vehicle policy.

Finite Fleet



- Optimality depends on parameters; no general structure.
- Hybrid heuristic: For m vehicles,
  1. first $m-1$ follow many-vehicle policy,
  2. last one serves remainder with one-vehicle policy.
- For $f(n) = bn + c\sqrt{n}$, heuristic has approximation guarantee

$$\frac{m - 1 + D_m\sqrt{D_m}}{m - 1 + D_m},$$

D_m is number of dispatches for m-th vehicle.

Tactical Design

Fleet sizing

1. 8×8 mile region, uniformly random locations.
2. An order every 8 minutes for 10 hours, 12-hour day.
3. Manhattan norm, 25 mph, 1 minute service per order.

Tactical Design

Fleet sizing

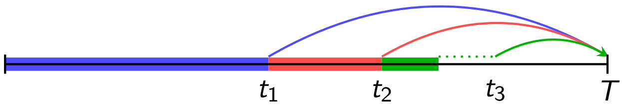
1. 8×8 mile region, uniformly random locations.
2. An order every 8 minutes for 10 hours, 12-hour day.
3. Manhattan norm, 25 mph, 1 minute service per order.

- Many Vehicles: Two dispatches, 64 and 11 orders.
- Single Vehicle: Two dispatches, 55 and 20 orders.
  - Dispatch time increase of only 4%!

Tactical Design

Choosing order cutoff N

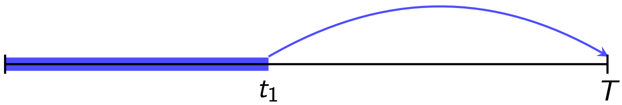
- If revenue is linear in orders served, how long do we accept orders?
  - Assume fleet can be as large as necessary.



Tactical Design

Choosing order cutoff N

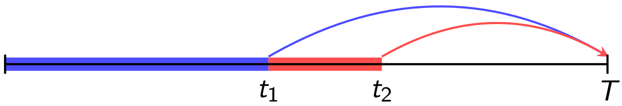
- If revenue is linear in orders served,
  how long do we accept orders?
  - Assume fleet can be as large as necessary.
- Optimal to maximally utilize dispatched vehicles:



Tactical Design

Choosing order cutoff N

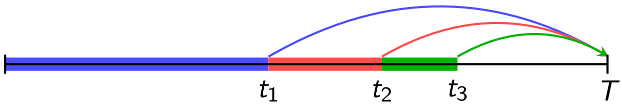
- If revenue is linear in orders served, how long do we accept orders?
  - Assume fleet can be as large as necessary.
- Optimal to maximally utilize dispatched vehicles:



Tactical Design

Choosing order cutoff N

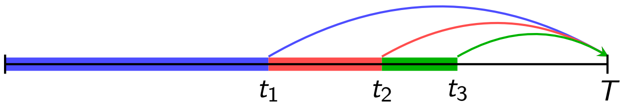
- If revenue is linear in orders served, how long do we accept orders?
  - Assume fleet can be as large as necessary.
- Optimal to maximally utilize dispatched vehicles:



Tactical Design

Choosing order cutoff N

- If revenue is linear in orders served, how long do we accept orders?
  - Assume fleet can be as large as necessary.
- Optimal to maximally utilize dispatched vehicles:



One vehicle: Can prove similar result for one, two dispatches.

Tactical Design

Other potential applications:

1. Service region partitioning.
   - Small areas served by single vehicle, or large area served by many?

Tactical Design

Other potential applications:

1. Service region partitioning.
    - Small areas served by single vehicle, or large area served by many?
2. Combining SDD and overnight deliveries.
    - Starting the day with orders accumulated.

Tactical Design

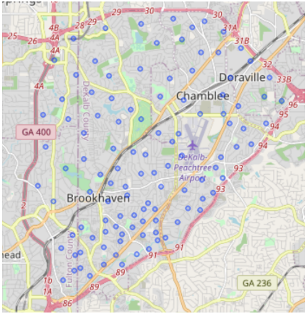
Other potential applications:

1. Service region partitioning.
    - Small areas served by single vehicle, or large area served by many?
2. Combining SDD and overnight deliveries.
    - Starting the day with orders accumulated.
3. Length of work day, size of service region, ...

Computational Validation

Case study in Northeastern Atlanta

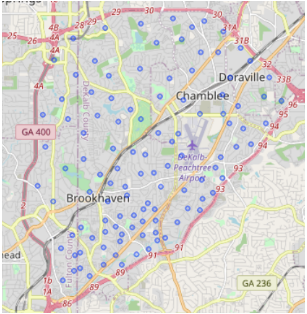
- 22 census tracts, about 92,000 people.
  - Five addresses per tract, 110 total.
  - Depot in northeast border.



Computational Validation

Case study in Northeastern Atlanta

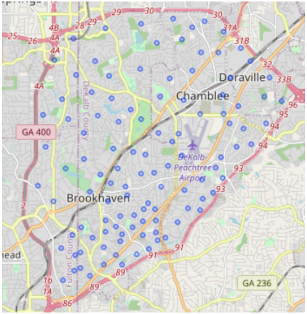
- 22 census tracts, about 92,000 people.
  - Five addresses per tract, 110 total.
  - Depot in northeast border.
- Service day: 9AM - 6PM.
  - Orders every six minutes.
  - Location chosen proportional to tract's population times median income.



Computational Validation

Case study in Northeastern Atlanta

- 22 census tracts, about 92,000 people.
  - Five addresses per tract, 110 total.
  - Depot in northeast border.
- Service day: 9AM - 6PM.
  - Orders every six minutes.
  - Location chosen proportional to tract's population times median income.
- Driving times given by Google API.
  - Driving time calibrated to $24\sqrt{n}$ minutes.
  - 10-min setup per dispatch, 1.5-min service per order.



Computational Validation

Benchmarks

- Two-vehicle fleet:
  - Order cutoff at 3:40 ($N = 66.7$) for full utilization.
  - Model predicts 389 minutes of dispatch time.

Computational Validation

Benchmarks

- Two-vehicle fleet:
  - Order cutoff at 3:40 ($N = 66.7$) for full utilization.
  - Model predicts 389 minutes of dispatch time.
- Operational benchmark:
  - Poisson arrivals (6-min. rate).
  - Compute TSP for all accumulated orders, dispatch when

$$\text{setup} + \text{service time} + \text{TSP} = \text{remaining time.}$$

Computational Validation

Benchmarks

- Two-vehicle fleet:
  - Order cutoff at 3:40 ($N = 66.7$) for full utilization.
  - Model predicts 389 minutes of dispatch time.
- Operational benchmark:
  - Poisson arrivals (6-min. rate).
  - Compute TSP for all accumulated orders, dispatch when

$$\text{setup} + \text{service time} + \text{TSP} = \text{remaining time.}$$

- Hindsight-optimal benchmark:
  - Dispatch with full knowledge of each order's time and location.
  - Lower bound for any operational policy.

Computational Validation

Results

	Tactical	Operational	
Dispatch 1	48.40 units	48.20 units	
	249.58 min.	249.69 min.	
Dispatch 2	18.26 units	18.45 units	
	139.95 min.	139.16 min.	
Total	66.66 units	66.65 units	
	389.53 min.	388.85 min.	

- Benchmark metrics computed over 300 simulations.
- Tactical predictions vs. operational observations within 1%.
- Similar results for one-vehicle case, different cutoff.

Computational Validation

Results

	Tactical	Operational	HSO
Dispatch 1	48.40 units	48.20 units	43.90 units
	249.58 min.	249.69 min.	228.07 min.
Dispatch 2	18.26 units	18.45 units	22.75 units
	139.95 min.	139.16 min.	144.88 min.
Total	66.66 units	66.65 units	66.65 units
	389.53 min.	388.85 min.	372.95 min.

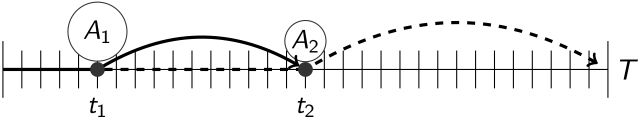
- Benchmark metrics computed over 300 simulations.
- Tactical predictions vs. operational observations within 1%.
- Similar results for one-vehicle case, different cutoff.

Conclusions

- Expect unbalanced dispatches in SDD.
  - Decreasing dispatch lengths.
  - Divide day into inactive/active parts.
- Use policy structure for tactical design.
  - Fleet sizing, cutoff time, partitioning, ...
  - Accurate operational predictions (within 1% or less).

Ongoing Work

- Choosing service region(s) and cutoff time(s).
  - Should we serve different customers differently?
  - In-town versus suburban, near versus far...



- Region partitioning and fleet sizing in tandem.
  - How many vehicles do we need assuming they serve different regions differently?

atoriello@isye.gatech.edu
http://www.isye.gatech.edu/~atoriello3/