# Transactional Data Analysis and Visualization Lab

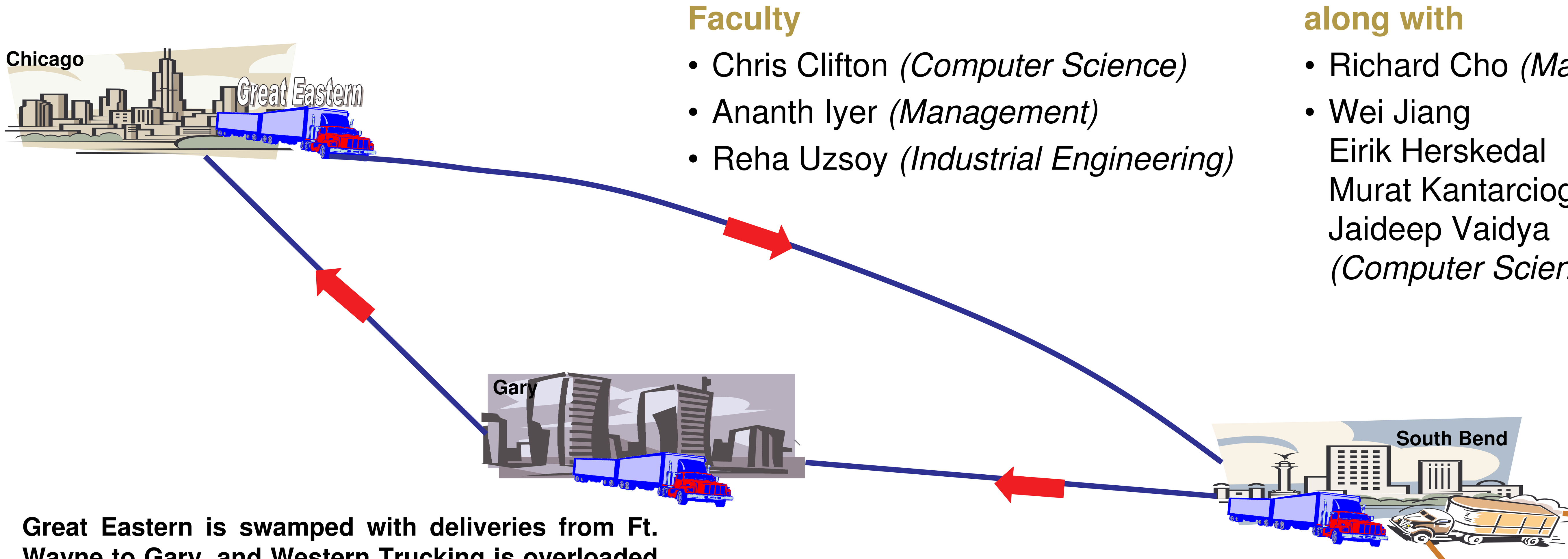

**Faculty**
- Chris Clifton *(Computer Science)*
- Ananth Iyer *(Management)*
- Reha Uzsoy *(Industrial Engineering)*

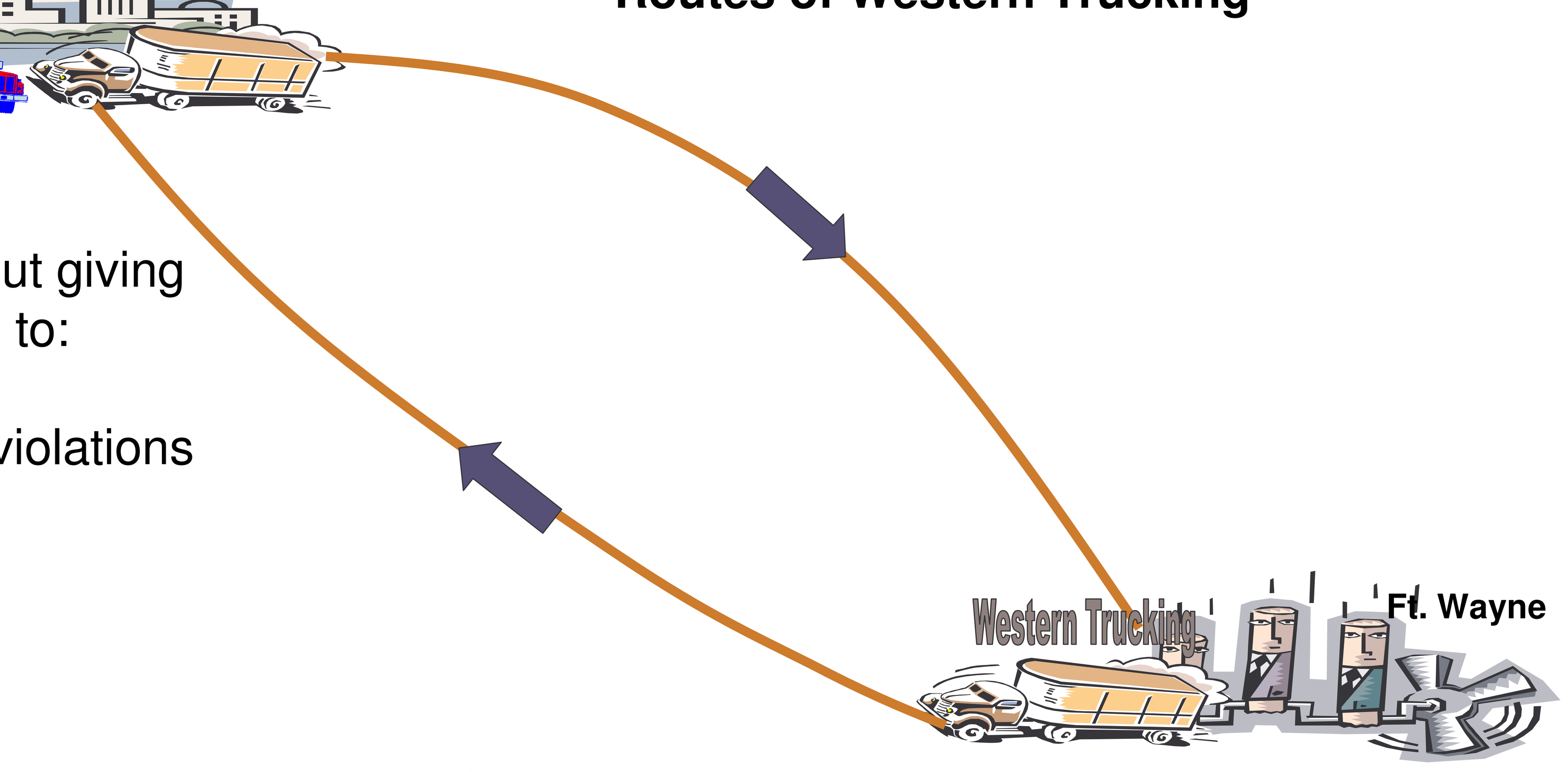
**along with**
- Richard Cho *(Management)*
- Wei Jiang
  Eirik Herskedal
  Murat Kantarcioglu
  Jaideep Vaidya
  *(Computer Science)*



**Funded by**
e-Enterprise Center

**In cooperation with**
LEEAP
Indiana Center for Database Systems

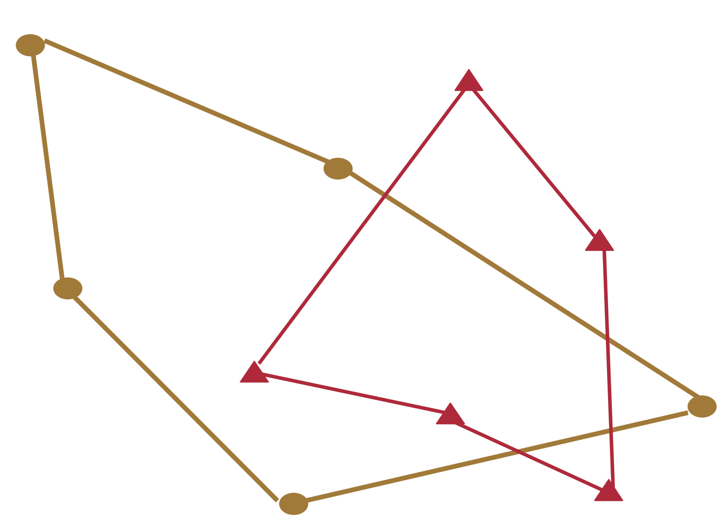

**Great Eastern is swamped with deliveries from Ft. Wayne to Gary, and Western Trucking is overloaded with business from Chicago to South Bend. Collaborations between Great Eastern and Western Trucking are possible at either Gary and south Bend.**

**The above diagram represents a collaborative route between the two companies, and the collaboration takes place as South Bend.**

Problem: How to identify this opportunity without giving away information that would enable companies to:
- Steal customers
- Give the appearance of price-fixing / antitrust violations

## Example: Two Traveling Salesman

The optimal rout of salesman A
The optimal rout of salesman B
Customers of salesman A
Customers of salesman B

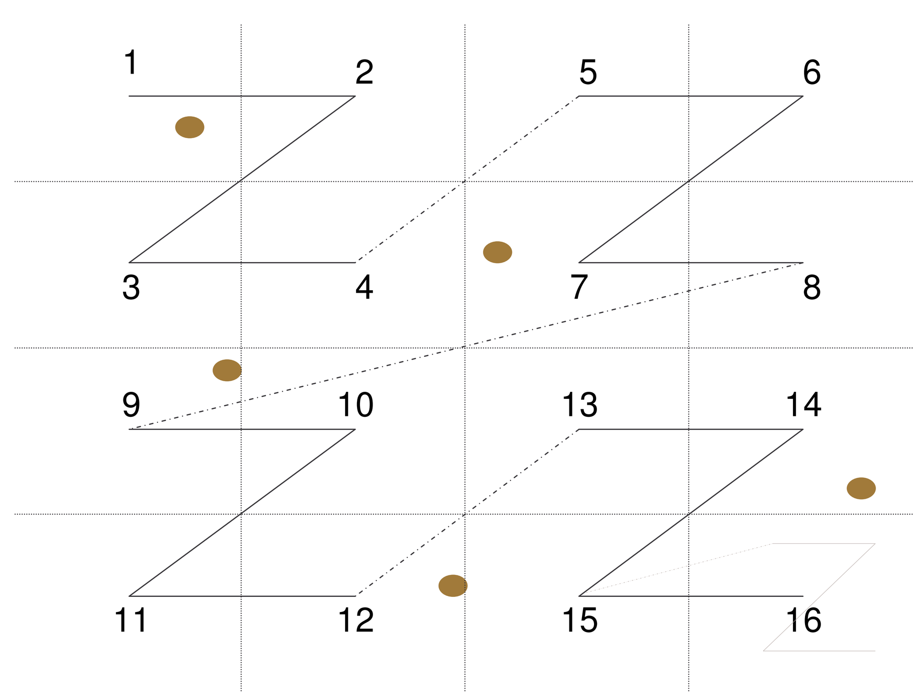


Swapping some customers would give both a shorter route

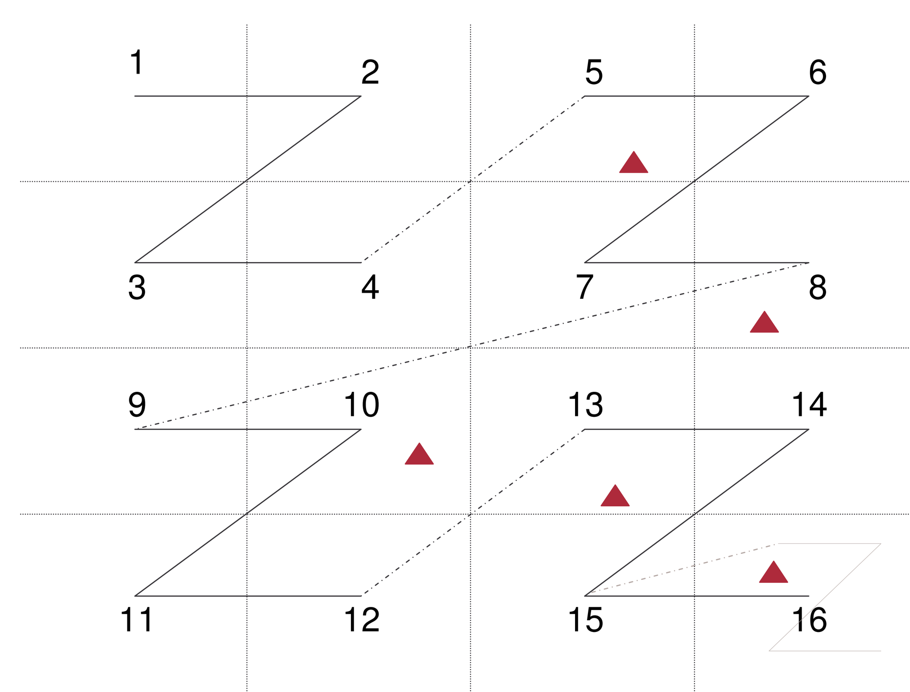
*But they don't want to share information on all customers!*

### Private solution using space-filling curves

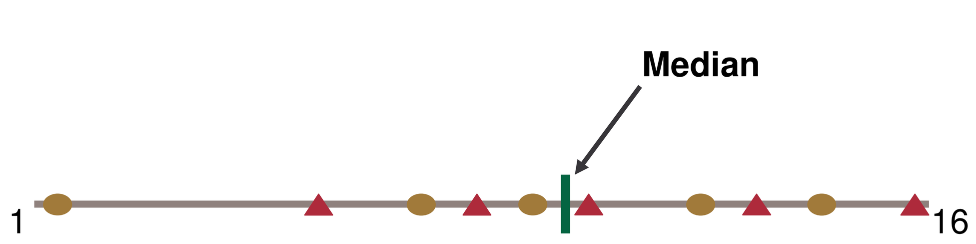
For each customer of salesman A, calculate a corresponding position on an interval.



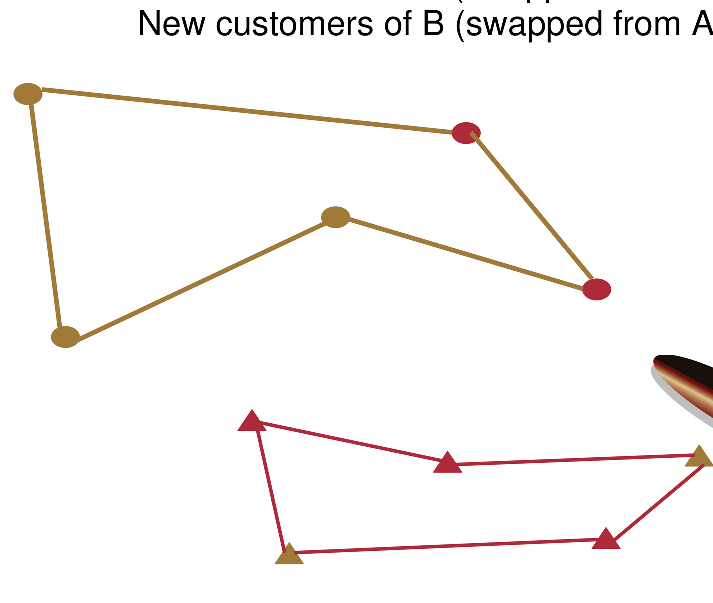

For each customer of salesman B, calculate a corresponding position on an interval.




Now find the median and swap items on the "wrong" side




The new rout of A
The new rout of B
Original customers of A
Original customers of B
New customers of A (swapped from B)
New customers of B (swapped from A)



## Bigger Problem: Linear Programming

Start with *The Decomposition Algorithm for Linear Programs* (Dantzig and Wolfe '61)

**Block-structured Linear Programming Problem**
- Master: All variables, some constraints, function to be minimized
- Subproblems: Additional (separable) constraints

| | $X_1$ | $X_2$ | ••• | $X_n$ | |
|---|---|---|---|---|---|
| Min | $c_1$ | $c_2$ | ••• | $c_n$ | |
| s.t | $A_1$ | $A_2$ | ••• | $A_n$ | b |
| | $B_1$ | | | | $b_1$ |
| | | $B_2$ | | | $b_2$ |
| | | | ••• | | |
| | | | | $B_n$ | $b_n$ |

- **Master Problem**
  - For given solution sets for $SP_i$, say $\{X_i^1, X_i^2, \bullet\bullet\bullet, X_i^K\}$, the feasible set à $X_i^1 = \Sigma_k s_{ik} X_i^k$, and $\Sigma_k s_{ik} = 1$.
  - Master Problem is changes to use variable $S_i = \{s_{ik}\}$, for each $X_i$,

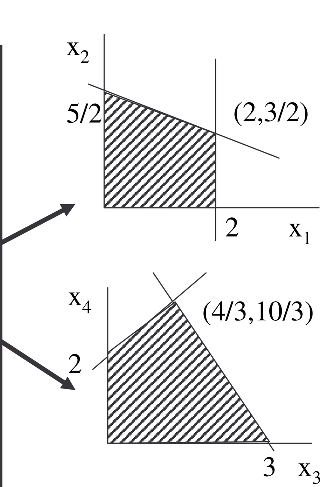
| | $s_{i1}$ | $s_{i2}$ | ••• | $s_{iK}$ | |
|---|---|---|---|---|---|
| Min | $c_1X_i^1$ | $c_2X_i^2$ | ••• | $c_nX_i^n$ | |
| s.t | $A_1X_i^1$ | $A_2X_i^2$ | ••• | $A_nX_i^n$ | |
| | 1 | 1 | ••• | 1 | 1 |

  - From the dual optimal of this LP, say $\{\pi_1^*, \pi_n^*, \bullet\bullet\bullet, \pi_n^*\}$ generate new objective function for $SP_i$
    Ł $\text{Min } \pi_1^* S_i = \text{Min } (\pi_1^* A_i) X_i$
- **Subproblem $SP_i$**
  - Solve the following LP,
    $\text{Min } (\pi_1^* A_i) X_i$
    s.t $B_i X_i = b_i$
- **Iterations**
  - $SP_i$ gives the optimal solution set $X_i^*$ to *MP*
  - If $X_i^* \in \{X_i^1, X_i^2, \bullet\bullet\bullet, X_i^K\}$, then ignore it. If all new solutions are turned down, the current solution is optimal.
  - Otherwise, calculate new objective function for *MP*.
  - Repeat until obtaining the optimal solution.

### *Example*

| | | | | | | |
|---|---|---|---|---|---|---|
| Min | $-2x_1$ | $-x_2$ | $-x_3$ | $+x_4$ | | |
| s.t | $x_1$ | | $+x_3$ | | ≤ | 2 |
| $SP_1$ | $x_1$ | $+x_2$ | | $+2x_4$ | ≤ | 3 |
| $SP_2$ | $x_1$ | | | | ≤ | 2 |
| | $x_1$ | $+2x_2$ | | | ≤ | 5 |
| | | | $-x_3$ | $+x_4$ | ≤ | 2 |
| | | | $2x_3$ | $+x_4$ | ≤ | 6 |

(All variables are non-negative)




(Iteration 1) Initial solution = (0, 0, 0, 0)
Ł Min $-2x_1 - x_2$ and Min $-x_3 + x_4$

(Iteration 2) New solutions (2, 3/2) & (3, 0) from $SP_1$ & $SP_2$
à *MP* : $\pi_1^* = (1, 1)$ and Opt. Value = -5
à new obj. functions are
Min $-2x_1 - x_2$ and Min $-x_3 - 2x_4$

(Iteration 3) No new sol'n from $SP_1$ and (4/3, 10/3) from $SP_2$
à *MP* : No more improvement from Opt. Value = -5.
à Thus, this is an optimal and
$X_1 = (x_1, x_2) = 1/7*(0, 0) + 6/7*(2,3/2) = (12/7, 9/7)$
$X_2 = (x_3, x_4) = 19/21*(0, 0) + 2/21*(3,0) = (2/7, 0)$

**Goal – Private Computation**
- Subproblem constraints not shared
- Master constraints not shared
- Everybody only learns the part of the result that matters to them
  - *And nothing else!*
  - Not trivial to accomplish this efficiently – *but we're working on it*


PURDUE UNIVERSITY
CERIAS®
e-Enterprise Center