Bicycle-Sharing System Expansion: Station Re-Deployment through Crowd Planning

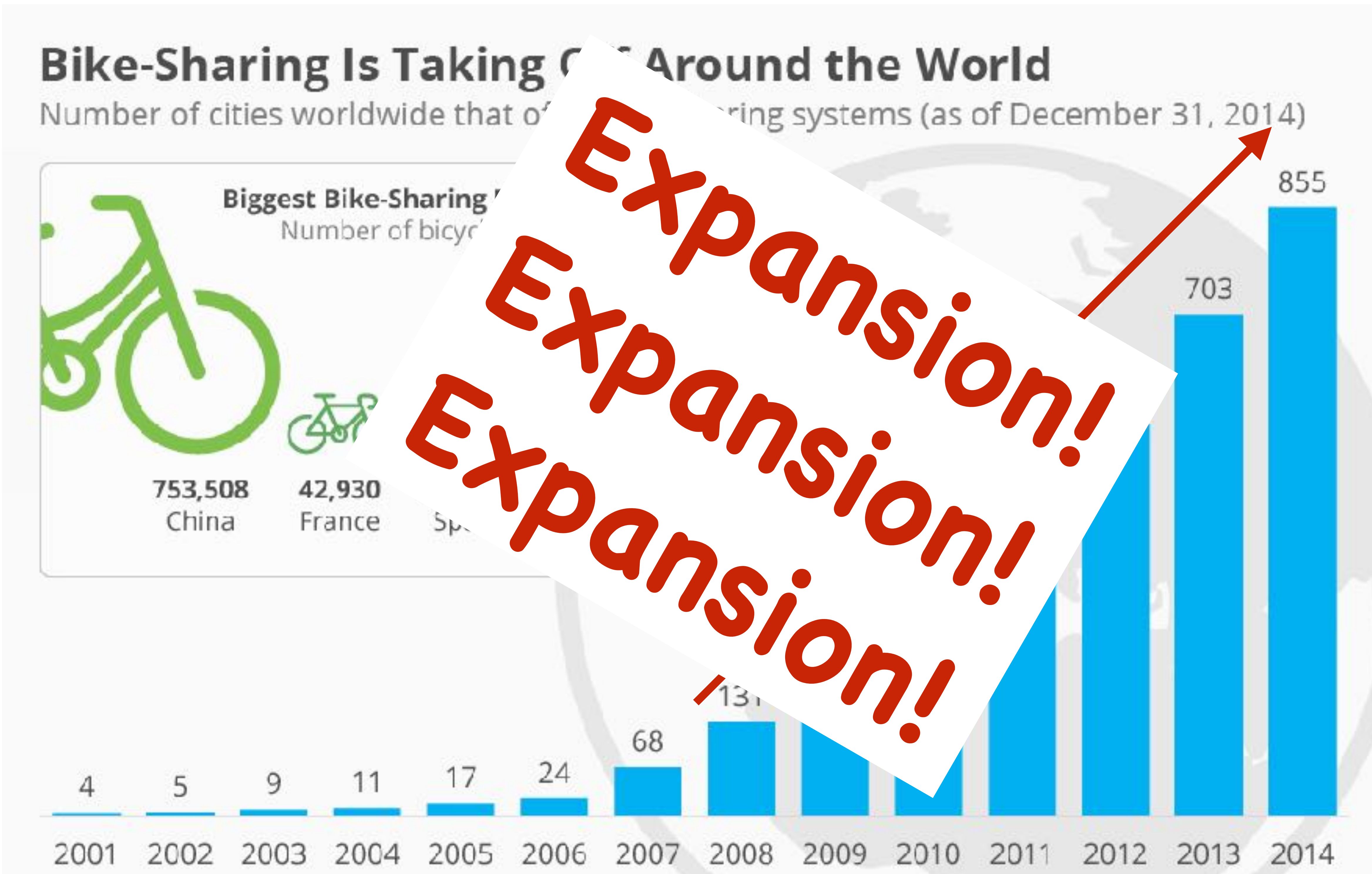
Jiawei Zhang[1], Xiao Pan[2], Moyin Li[1], Philip S. Yu[1,3]

[1]University of Illinois at Chicago, Chicago, IL, USA
[2]Shijiazhuang Tiedao University, Shijiazhuang, Hebei, China
[3]Institute for Data Science, Tsinghua University, Beijing, China

Big cities need public bicycle-sharing systems to alleviate traffic congestion

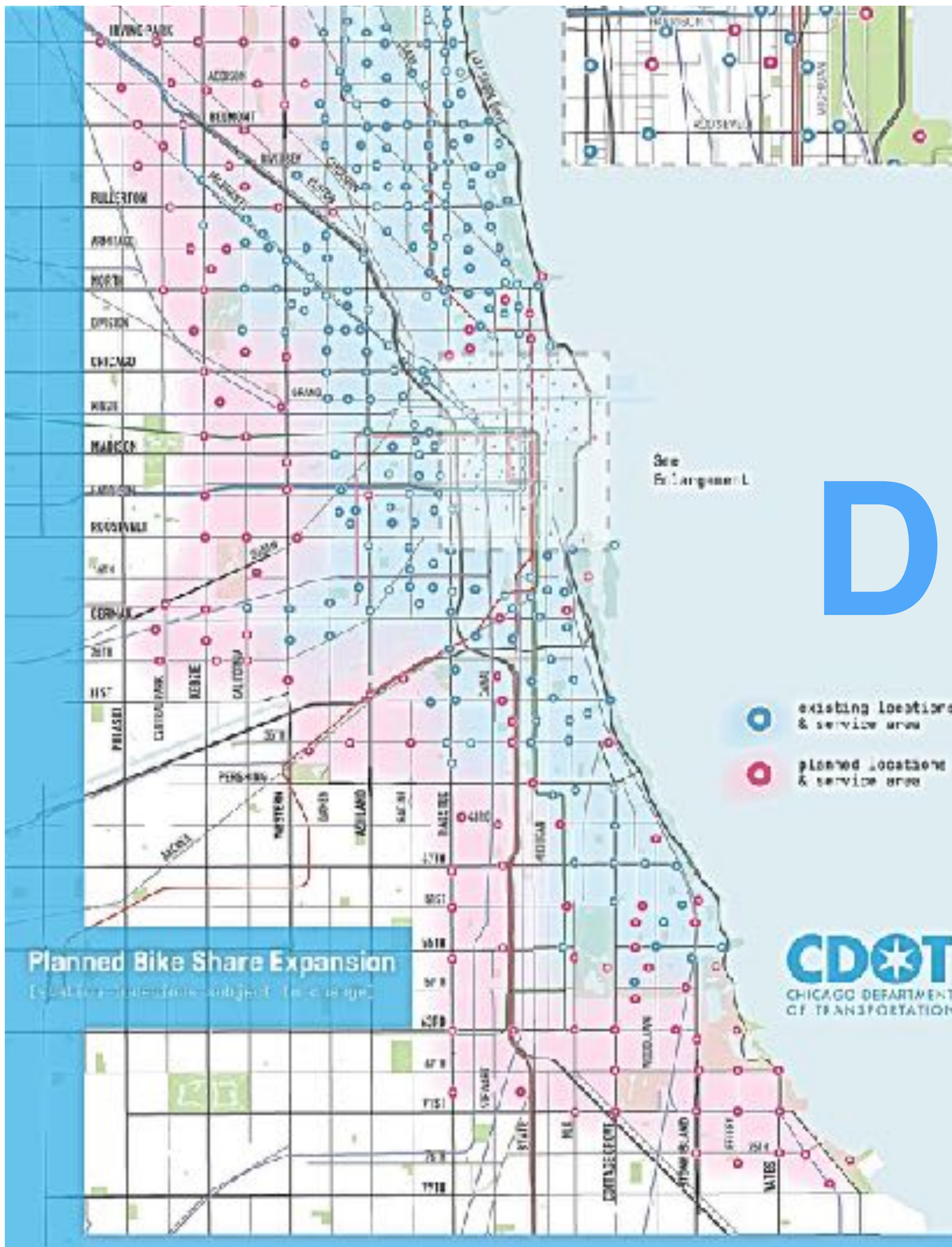
Amount of space required to transport the same number of passengers by car, bus, or bicycle.
Event info at www.facebook.com/Urban.Ambassadors - Photos by www.tobinbennett.com
(Des Moines, Iowa - August 2010)

Bicycle-sharing systems are launched in many big cities and keep expanding



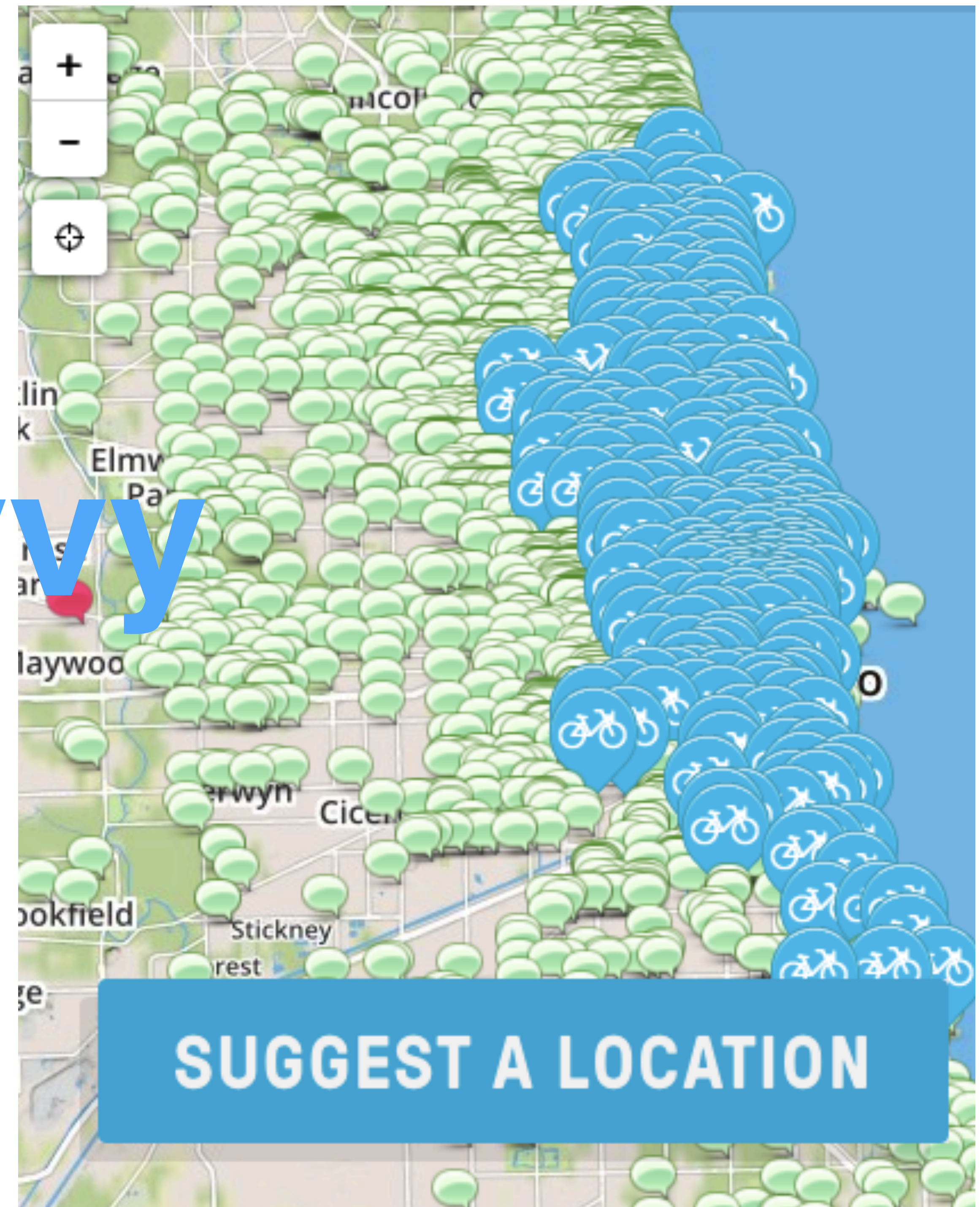
Problem Studied: Bicycle-sharing system expansion via crowdsourcing

Expansion Plan



Divvy

Crowd Suggestions



Bicycle-sharing system expansion actions

- Actions performed in expansion
  - Stations
    - *add* new stations
    - *remove* existing stations,
    - *move* existing stations to a new place (remove it first, and add it again)
  - Bikes
    - *add* new bike docks
    - *remove* existing bike docks

Bicycle-sharing system expansion objective

- Prior to expansion: service provider determines the target expansion size (**# stations K, # bikes C**)
- Expansion objective
  - **maximize** the **usage convenience** for customers
  - **minimize** the **cost** involved in the expansion

Bicycle-sharing system expansion objective

- Prior to expansion: service provider determines the target expansion size (**# stations K, # bikes C**)
- Expansion objective
  - **maximize** the **usage convenience** for customers
  - **minimize** the **cost** involved in the expansion
- Objective Optimization Function

$$\mathcal{S}_F^* = \arg\max_{\mathcal{S}_F} convenience(\mathcal{S}_F) - \beta \cdot cost(\mathcal{S}_F)$$

$$s.t.\ |\mathcal{S}_F| = K, \sum_{s \in \mathcal{S}_F} capacity(s) = C,$$

optimal stations and bikes after expansion

objective station number

objective bike number

Proposed Method: CrowdPlanning

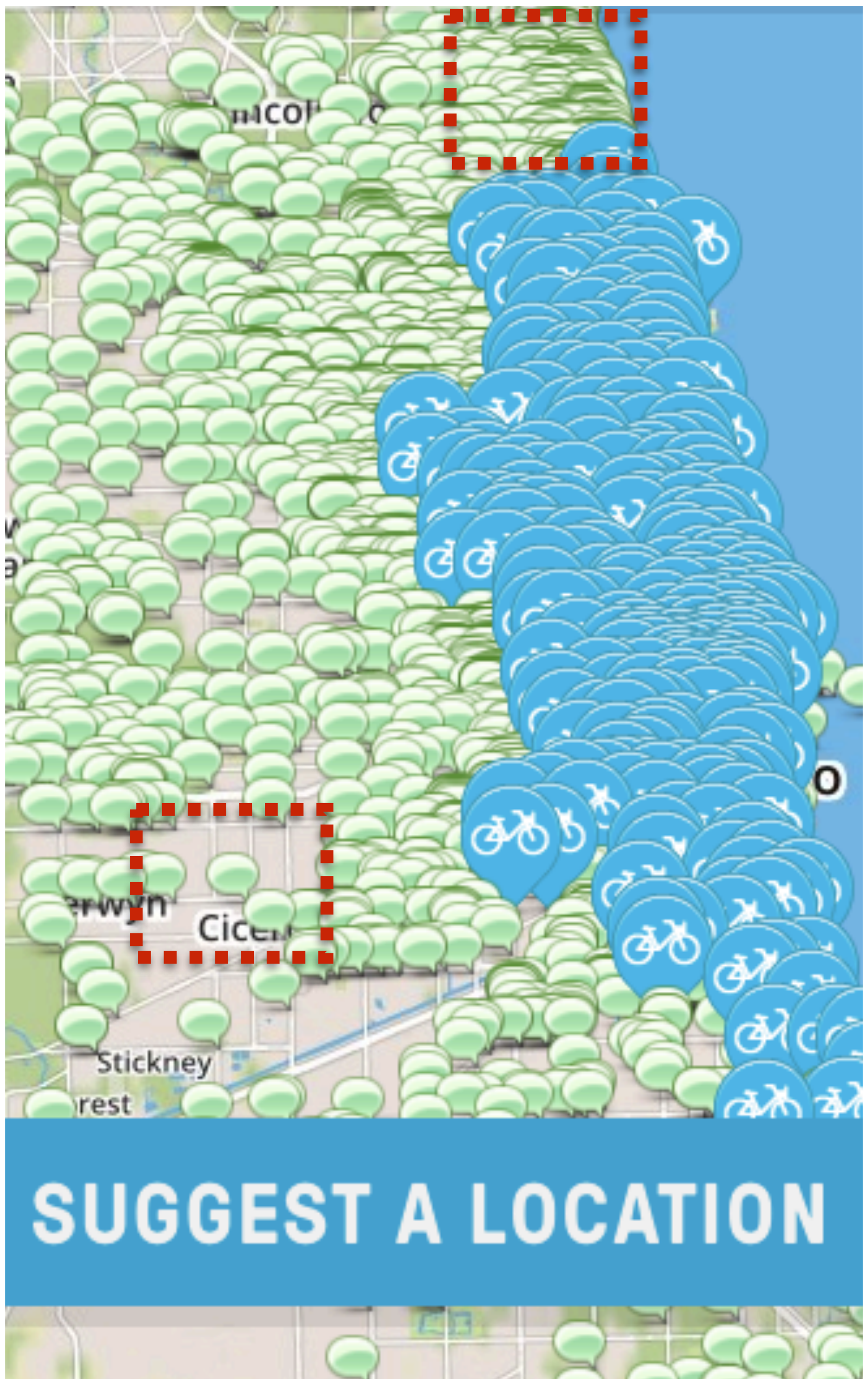
- CrowdPlaning: two-phase planning method
  - Step 1: Station Planning (i.e., station location determination)
    - usage convenience
      - current station convenience: historical trips
      - future station convenience: crowd suggestions
    - cost in adding/removing/adjusting bike stations

Proposed Method: CrowdPlanning

- CrowdPlaning: two-phase planning method
  - Step 1: Station Planning (i.e., station location determination)
    - usage convenience
      - current station convenience: historical trips
      - future station convenience: crowd suggestions
    - cost in adding/removing/adjusting bike stations
  - Step 2: Station Capacity Planning (i.e., bike assignment)
    - usage convenience
      - current bike number: historical trips
      - future bike number: crowd suggestions
    - cost in adding/removing bikes from existing stations

Step 1: Station Planning

- Usage convenience based on **crowd suggestions**



$$\text{convenience}_c(\mathcal{S}_F) = \sum_{g_i \in \mathcal{G}} y(g_i)$$

$$s.t. \;\; y(g_i) \geq y(g_j), \text{ if } |\Gamma_H(g_i)| \geq |\Gamma_H(g_j)|, \forall g_i, g_j \in \mathcal{G}.$$

$y(g_j)$: {0,1} has stations in the cell

$|\Gamma_H(g_j)|$: crowd suggestions at the cell

$\mathcal{G}$: service region grid cells

Step 1: Station Planning

- Usage convenience based on **historical trip records**
  - **frequently** used existing stations are **more likely** to be preserved in the expansion

Step 1: Station Planning

- Usage convenience based on **historical trip records**
  - frequently used existing stations are more likely to be preserved in the expansion
  - existing station preserving probability

P(gi) is monotone

$$P(g_i) = \frac{e^{k|\Gamma_T(g_i)|}}{1 + e^{k|\Gamma_T(g_i)|}}$$

trips start/end at the cell

Step 1: Station Planning

- Usage convenience based on **historical trip records**
  - frequently used existing stations are more likely to be preserved in the expansion
  - existing station preserving probability

$$P(g_i) = \frac{e^{k|\Gamma_T(g_i)|}}{1 + e^{k|\Gamma_T(g_i)|}}$$

trips starting/end at the cell

  - historical trip usage based convenience

$$\text{convenience}_u(\mathcal{S}_F) = \sum_{g_i \in \mathcal{G}} y(g_i) \cdot P(g_i)$$

Step 1: Station Planning

- Station deployment costs

$$\text{cost}(g_i) = \begin{cases} \text{cost}_s^+, & \text{if } \bar{y}(g_i) = 0, y(g_i) = 1; \\ \text{cost}_s^-, & \text{if } \bar{y}(g_i) = 1, y(g_i) = 0; \\ 0, & \text{otherwise.} \end{cases}$$

station adding cost

no station before expansion

one station after expansion

no changes, no cost

Step 1: Station Planning

- Station deployment costs

station adding cost

no station before deployment

one station after deployment

$$\mathrm{cost}(g_i) = \begin{cases} \mathrm{cost}_s^+, & \text{if } \bar{y}(g_i) = 0, y(g_i) = 1; \\ \mathrm{cost}_s^-, & \text{if } \bar{y}(g_i) = 1, y(g_i) = 0; \\ 0, & \text{otherwise.} \end{cases}$$

no changes, no cost

- Overall station deployment cost

$$\mathrm{cost}(\mathcal{S}_F) = \mathrm{cost}(\mathcal{G}) = \sum_{g_i \in \mathcal{G}} \mathrm{cost}(g_i)$$

$$= \sum_{g_i \in \mathcal{G}} \max\{y(g_i) - \bar{y}(g_i), 0\} \cdot \mathrm{cost}_s^+ + \max\{\bar{y}(g_i) - y(g_i), 0\} \cdot \mathrm{cost}_s^- .$$

Step 1: Station Planning

- Station deployment objective function

convenience terms

$$\mathcal{Y}^* = \arg\max_{\mathcal{Y}} \text{convenience}(\mathcal{S}_F) - \beta \cdot \text{cost}(\mathcal{S}_F)$$

$$= \arg\max_{\mathcal{Y}} \sum_{g_i \in \mathcal{G}} \Big(y(g_i) + \alpha \cdot y(g_i) \cdot P(g_i) \Big) - \beta \cdot$$

cost term

$$\Big(\max\{y(g_i) - \bar{y}(g_i), 0\} \cdot \text{cost}_s^+ + \max\{\bar{y}(g_i) - y(g_i), 0\} \cdot \text{cost}_s^- \Big)\Big)$$

$$s.t.\ \underline{D} \le \text{density}(g_i, n) \le \overline{D}, \forall g_i \in \mathcal{G},$$

$$y(g_i) \ge y(g_j), \text{ if } |\Gamma_H(g_i)| \ge |\Gamma_H(g_j)|, \forall g_i, g_j \in \mathcal{G},$$

$$\sum_{g_k \in \mathcal{G}} y(g_k) = K; y(g_k) \in \{0, 1\}, \forall g_k \in \mathcal{G}.$$

quantity constraint

Step 2: Bike Planning

number of bikes assigned to station g_j

- Bike assignment planning
  - more bikes assigned to stations with more **suggestions**

$$c(g_i) \geq c(g_j), \text{ if } |\Gamma_H(g_i)| \geq |\Gamma_H(g_j)|, \forall g_i, g_j \in \tilde{\mathcal{G}},$$

  - more bikes assigned to stations with more **historical usages**

$$c(g_i) \geq c(g_j), \text{ if } |\Gamma_T(g_i)| \geq |\Gamma_T(g_j)|, \forall g_i, g_j \in \tilde{\mathcal{G}}.$$

  - **construction cost** should be as low as possible

$$\mathrm{cost}(g_i) = \begin{cases} \mathrm{cost}_d^+ \cdot (\mathrm{c}(g_i) - \bar{\mathrm{c}}(g_i)), & \text{if } \bar{y}(g_i) = y(g_i) = 1, \mathrm{c}(g_i) \geq \bar{\mathrm{c}}(g_i); \\ \mathrm{cost}_d^- \cdot (\bar{\mathrm{c}}(g_i) - \mathrm{c}(g_i)), & \text{if } \bar{y}(g_i) = y(g_i) = 1, \bar{\mathrm{c}}(g_i) \geq \mathrm{c}(g_i); \\ 0, & \text{otherwise.} \end{cases}$$

$$\mathrm{cost}(\tilde{\mathcal{G}}) = \sum_{g_i \in \tilde{\mathcal{G}}} \mathrm{cost}(g_i)$$

$$= \sum_{g_i \in \tilde{\mathcal{G}}} \bar{y}(g_i) \cdot y(g_i) \cdot (\mathrm{cost}_d^+ \cdot \max\{\mathrm{c}(g_i) - \bar{\mathrm{c}}(g_i), 0\} + \mathrm{cost}_d^- \cdot \max\{\bar{\mathrm{c}}(g_i) - \mathrm{c}(g_i), 0\}).$$

Step 2: Bike Planning

derivation in the paper

- Bike assignment planning
  - objective function

$$\min_{\{c_{g_i}\}_{g_i \in \tilde{\mathcal{G}}}} \sum_{g_i \in \tilde{\mathcal{G}}} \bar{y}(g_i) \cdot y(g_i) \cdot \Big(\mathrm{cost}_d^{+} \cdot \max\{\mathrm{c}(g_i) - \bar{\mathrm{c}}(g_i), 0\}$$

$$+ \mathrm{cost}_d^{-} \cdot \max\{\bar{\mathrm{c}}(g_i) - \mathrm{c}(g_i), 0\}\Big)$$

$$s.t.\ \ \mathrm{c}(g_i) \geq \mathrm{c}(g_j),\ \text{if } |\Gamma_H(g_i)| \geq |\Gamma_H(g_j)|, \forall g_i, g_j \in \tilde{\mathcal{G}},$$

$$\mathrm{c}(g_i) \geq \mathrm{c}(g_j),\ \text{if } |\Gamma_T(g_i)| \geq |\Gamma_T(g_j)|, \forall g_i, g_j \in \tilde{\mathcal{G}},$$

$$\sum_{g_i \in \tilde{\mathcal{G}}} \mathrm{c}(g_i) = C; \mathrm{c}(g_i) \in \mathbb{N}^{+}, \forall g_i \in \tilde{\mathcal{G}}.$$

Chicago Divvy bicycle sharing system Dataset

- Divvy and crowd suggestion Datasets

Table 1: The Divvy Datasets

datasets	trip	station	bike
2013 Q3-Q4	759,788	300	5,040
2014 Q1-Q2	905,699	300	5,209
2014 Q3-Q4	1,548,935	300	5,040
2015 Q1-Q2	1,096,239	474	8,274
2015 Q3-Q4	2,087,204	474	8,274

- 179,610 bike trips per month in the past two years
- Station number increases to 474 due to the system expansion at early 2015

Table 2: Crowd Suggestion Dataset

datasets	suggestions	comments
Crowd Suggestion	1,098	775

Chicago Divvy bicycle sharing system Dataset

- Settings:
  - stations after the expansion as the station redeployment ground truth.
  - Trips and suggestions received before the end of 2014 Q4 as the known information

Chicago Divvy bicycle sharing system Dataset

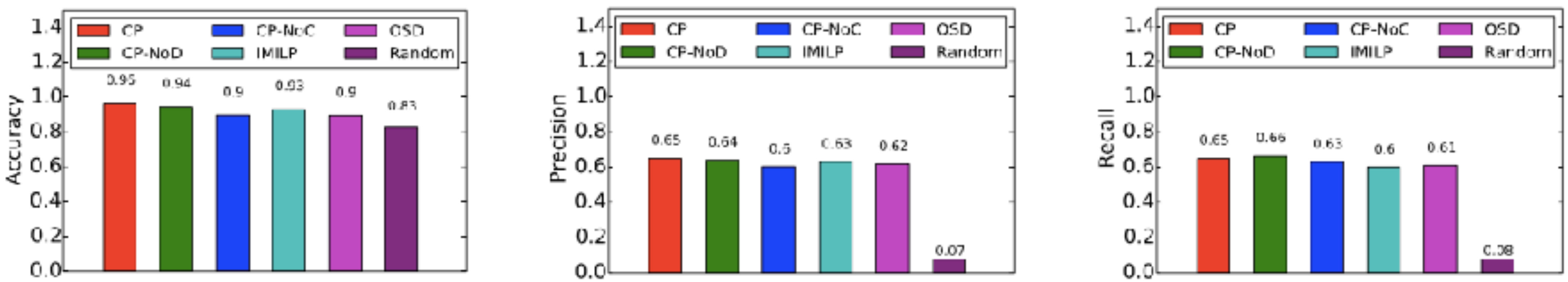
- Settings:
  - stations after the expansion as the station redeployment ground truth.
  - Trips and suggestions received before the end of 2014 Q4 as the known information
- Comparison Methods
  - *CrowdPlanning*: method proposed in this paper
  - *CP-NoDens*: no density constraints is considered
  - *CP-NoCost*: no construction cost is considered
  - *IMILP*: existing method for station deployment only, no capacity assignment
  - *OSD*: extension of existing method with construction costs
  - *Random*: random station and bike planning

Chicago Divvy bicycle sharing system Dataset

- Settings:
  - stations after the expansion as the station redeployment ground truth.
  - Trips and suggestions received before the end of 2014 Q4 as the known information
- Comparison Methods
  - *CrowdPlanning*: method proposed in this paper
  - *CP-NoDens*: no density constraints is considered
  - *CP-NoCost*: no construction cost is considered
  - *IMILP*: existing method for station deployment only, no capacity assignment
  - *OSD*: extension of existing method without construction costs
  - *Random*: random station and bike planning
- Evaluation Metrics:
  - Accuracy, Precision, Recall for station deployment
  - MSE, MAE, R2 for capacity assignment

Experiment Results

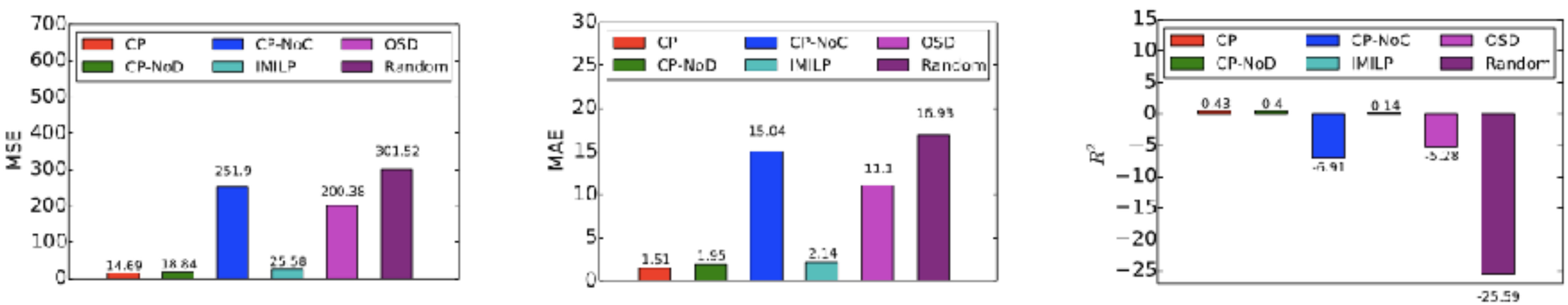
- Station Deployment Result



(a) Accuracy (b) Precision (c) Recall

Figure 4: Station deployment result evaluated by Accuracy, Precision and Recall.

- Bike Capacity Assignment Result



(a) MSE (b) MAE (c) R^2

Figure 5: Local station capacity assignment result evaluated by MSE, MAE and R^2.

Summary

- **Problem Studied**: bicycle-sharing system expansion with crowd planning
  - *station redeployment*: add new stations and remove/adjust existing stations
  - *station capacity assignment*: add/remove bikes from existing bike stations
- **Proposed Method**:
  - **convenience maximization**
    - convenience of existing stations/bikes based on historical trip records
    - convenience of new stations/bikes based on crowd suggestions
  - **cost minimization**
    - cost introduced in add/removing stations and bikes

Related Works: Bicycle Sharing Systems

Bicycle-Sharing System Analysis and Trip Prediction
MDM' 16

Bicycle-Sharing System Expansion
SIGSPATIAL' 16

Bicycle-Sharing System Trip Route Planning
IEEE CIC' 16

……
More Opportunities

Bicycle-Sharing System Expansion: Station Re-Deployment through Crowd Planning

Q & A

Jiawei Zhang[1], Xiao Pan[2], Moyin Li[1], and Philip S. Yu[1]
jzhan9@uic.edu, smallpx@gmail.com, mli60@uic.edu, psyu@cs.uic.edu