Towards Real-time Analysis of Smart City Data:
A Case Study on City Facility Utilizations

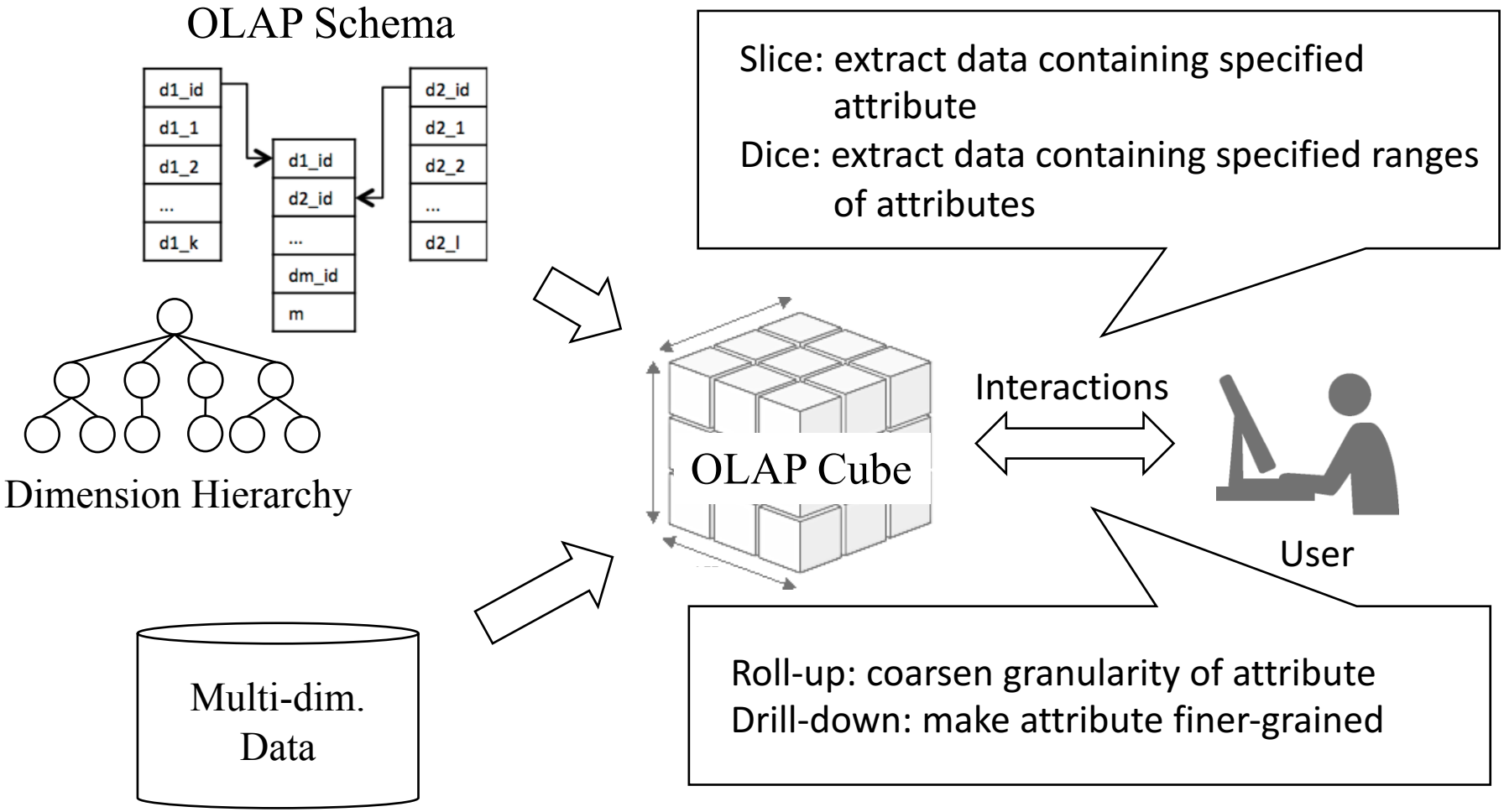
Takahiro Komamizu, Toshiyuki Amagasa, Salman Ahmed Shaikh, Hiroaki Shiokawa, Hiroyuki Kitagawa

University of Tsukuba, Japan

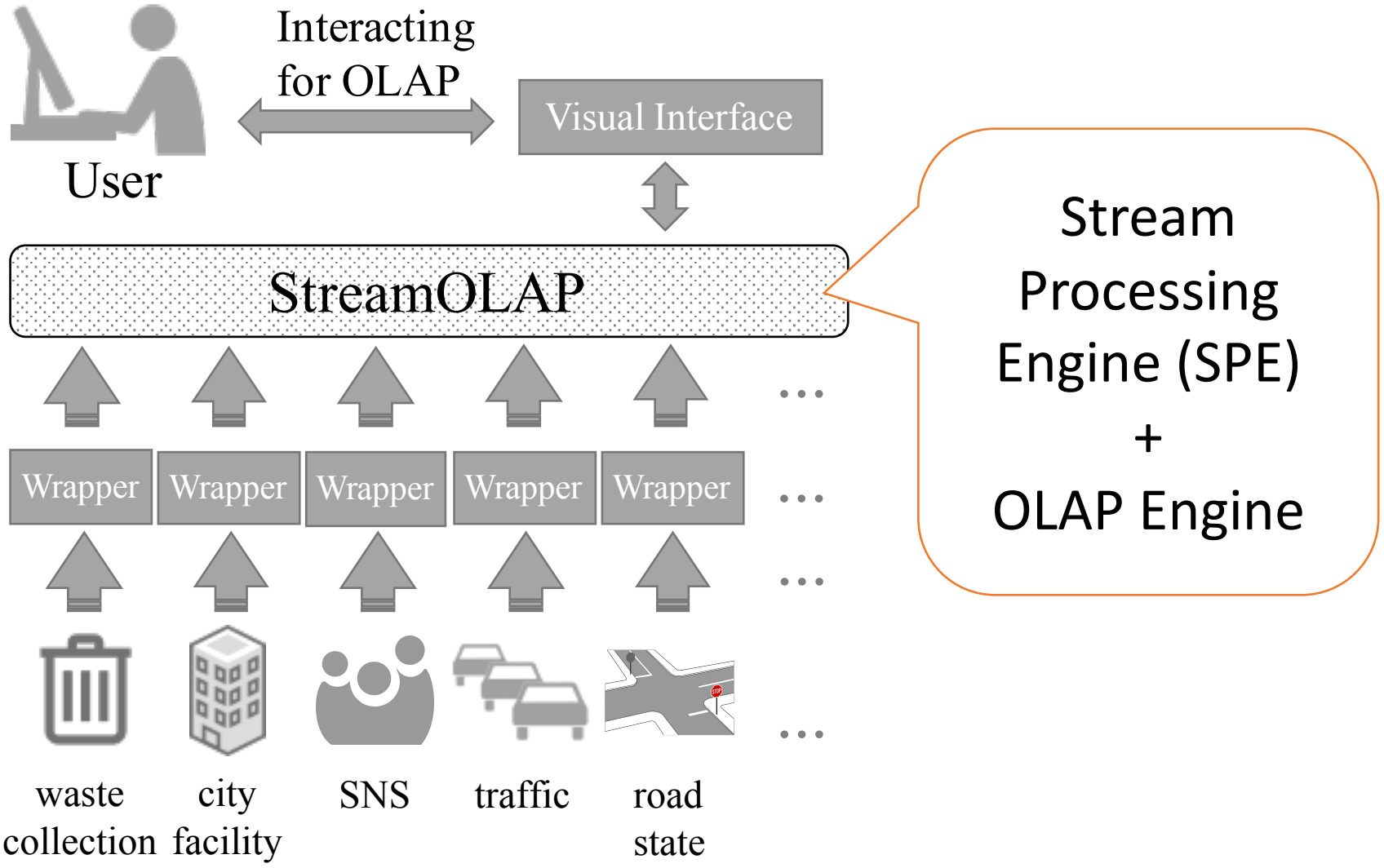
Data from Smart City

- Increasing demands for Smart City applications
  - Monitoring real-world
    - environmental conditions (e.g., air and water)
    - artificial objects (e.g., roads and buildings)
    - citizens opinions (e.g., social networks)
  - Analyzing for smart lives
    - decision making (e.g., air pollution prevention policy)
    - facility management (e.g., roadway management)
- Data from smart city tend to be multi-dimensional.
- Real-time analysis of smart city data has been becoming an increasing demand.

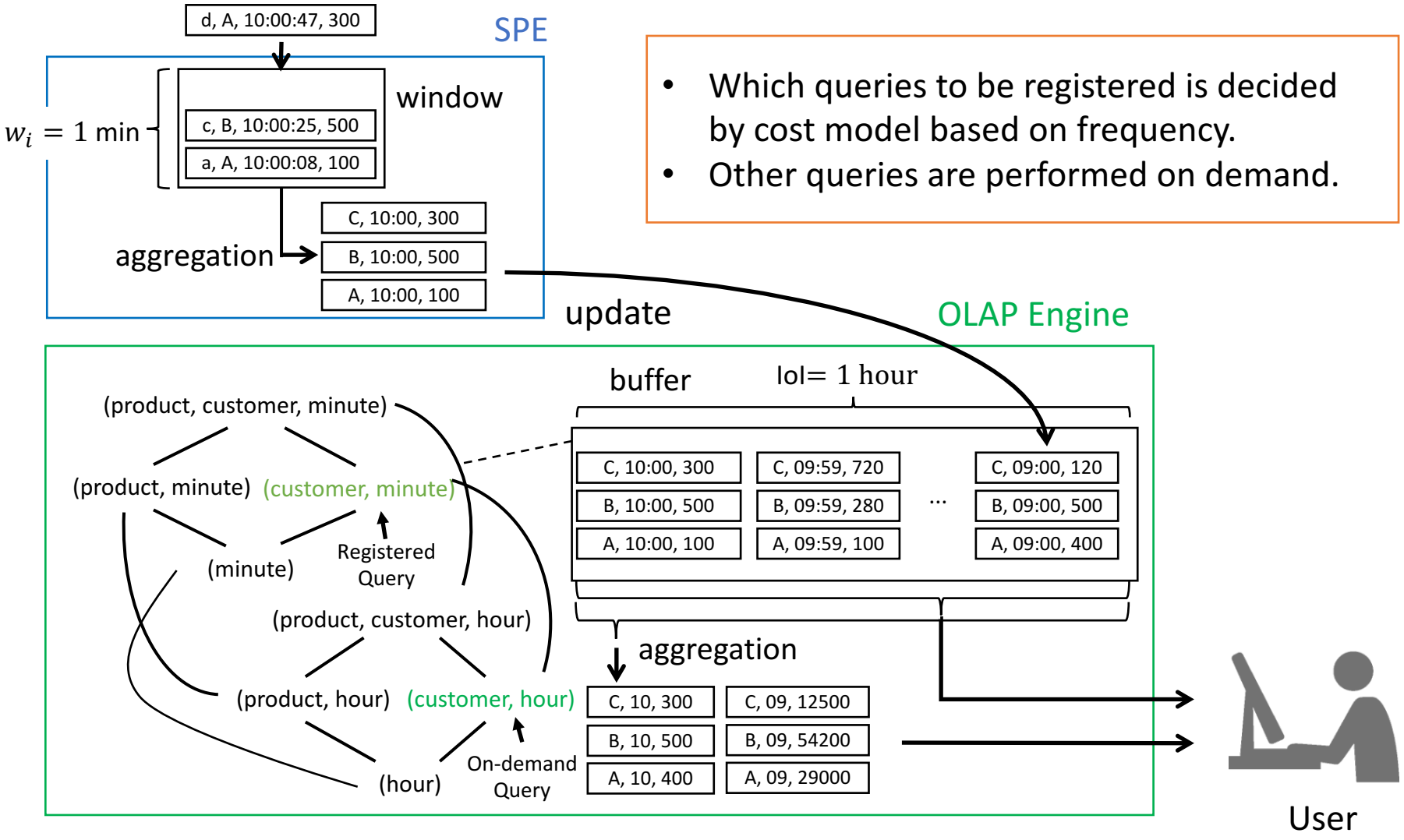
OLAP: multi-dimensional analysis



OLAP + Real-time analysis



StreamOLAP [Nakabasami et al. Big Data 2015]



- Which queries to be registered is decided by cost model based on frequency.
- Other queries are performed on demand.

Purpose of this paper

- Realize that StreamOLAP system is applicable for real-world (esp. smart city) data.
- Share experiences for StreamOLAP application using *city facility utilization* logs.
  - OLAP operation case study
  - Visualization concerns

City Facility Utilization logs

- Field: Tsukuba city, Japan
- Log: who, when, and where, a user used
- User info.: Attributes of users
  - For privacy prevention, we cannot obtain raw data.
    - e.g., address is anonymized to postcode.

(a) City facility utilization log.

Attribute	Description
Log ID	Record identifier.
Facility	Name/ID of utilized facility.
Room	Name/ID of utilized room.
Date	Date of utilization.
Time	Starting and ending time of utilization.
Purpose	Purpose of utilization.
User	Name/ID of utilizing users.
Fee	Amount of payment for utilization.
Num. of People	Number of utilizing users.
Location	Latitude and longitude of facility.

(b) User information

Attribute	Description
User	Name/ID of user (group in reality).
Purpose	Purpose of user for his/her utilizations.
Num. of Members	Number of users.
Postcode	Anonymized address of representative user.
Ward	Ward of user. This is higher abstraction of Postcode.
User Type	Type (e.g., *Kids*, *elders*, etc.) of user.

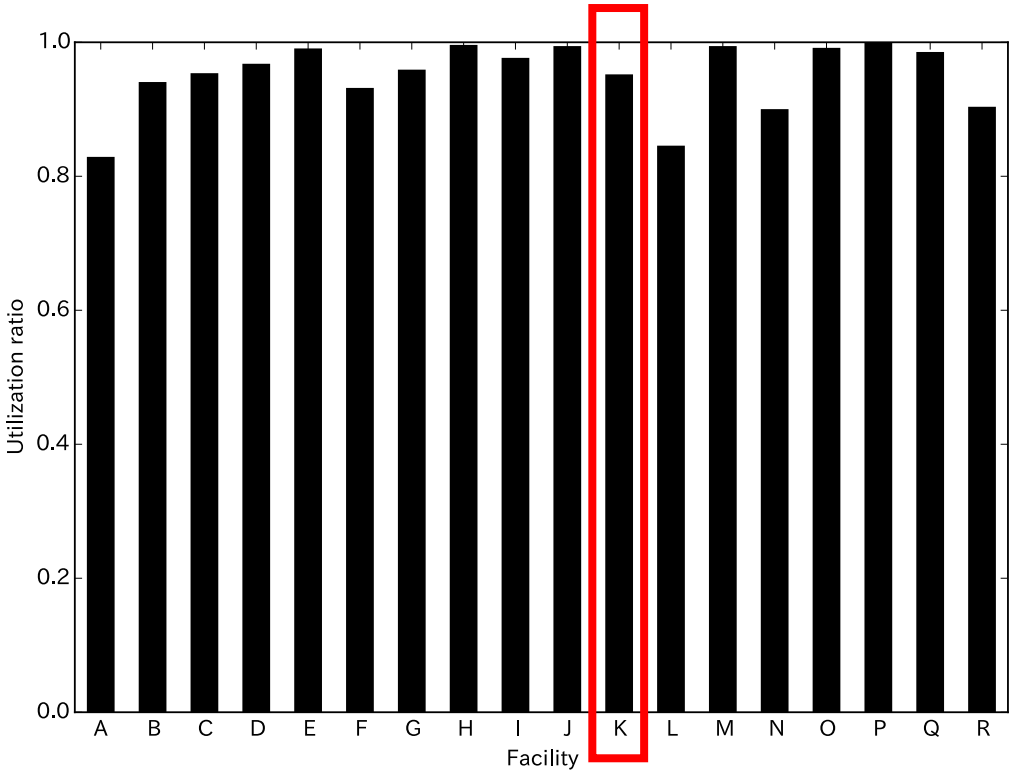
Analytical Scenarios

We discuss practical scenarios with practitioners in Tsukuba city and the following are three of them.

1. Analysis with drill-down operation
   - Showcasing capabilities of drill-down operations.
2. User location-based analysis
   - Investigation w.r.t. locations of users and facilities
3. Map visualization
   - Visual results of the location-based analysis

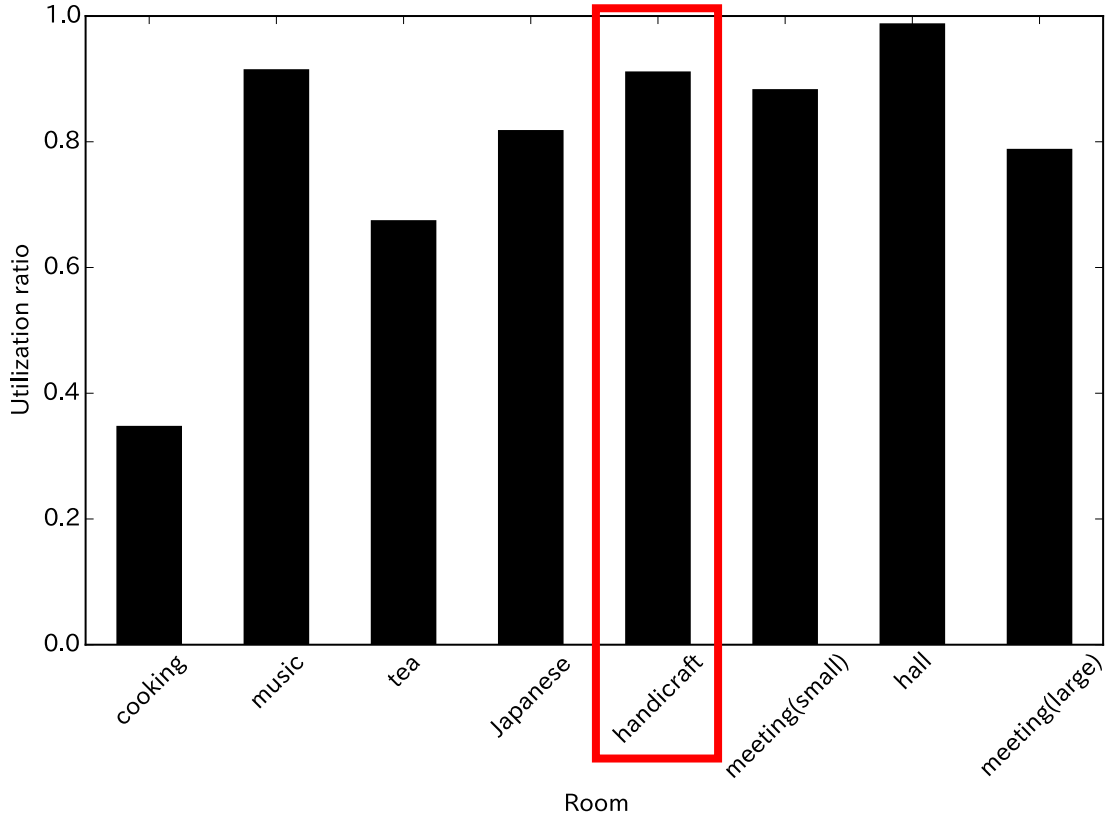
Analysis with drill-down operation

- Suppose a user is interested in utilization ratio of facilities, then the system displays overall utilization ratio for each facility.



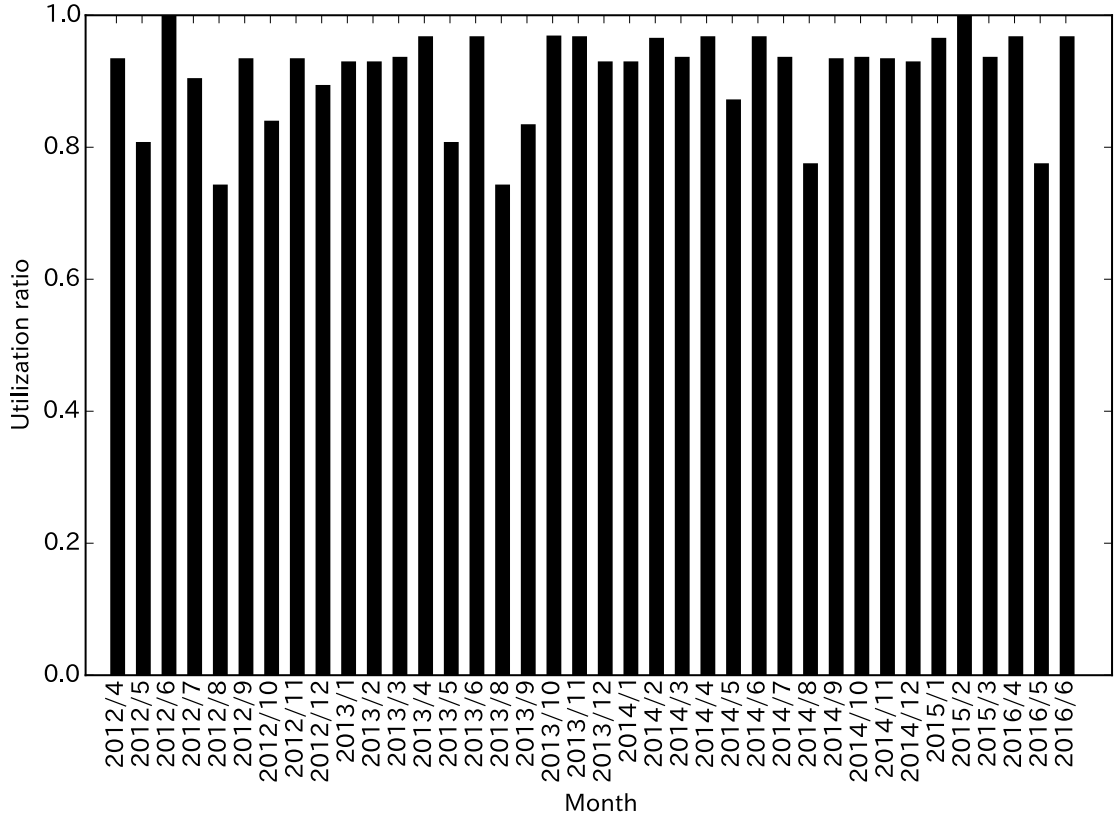
Analysis with drill-down operation

- As she has a concern about facility E, she selects to slice and drill-down for rooms in E.



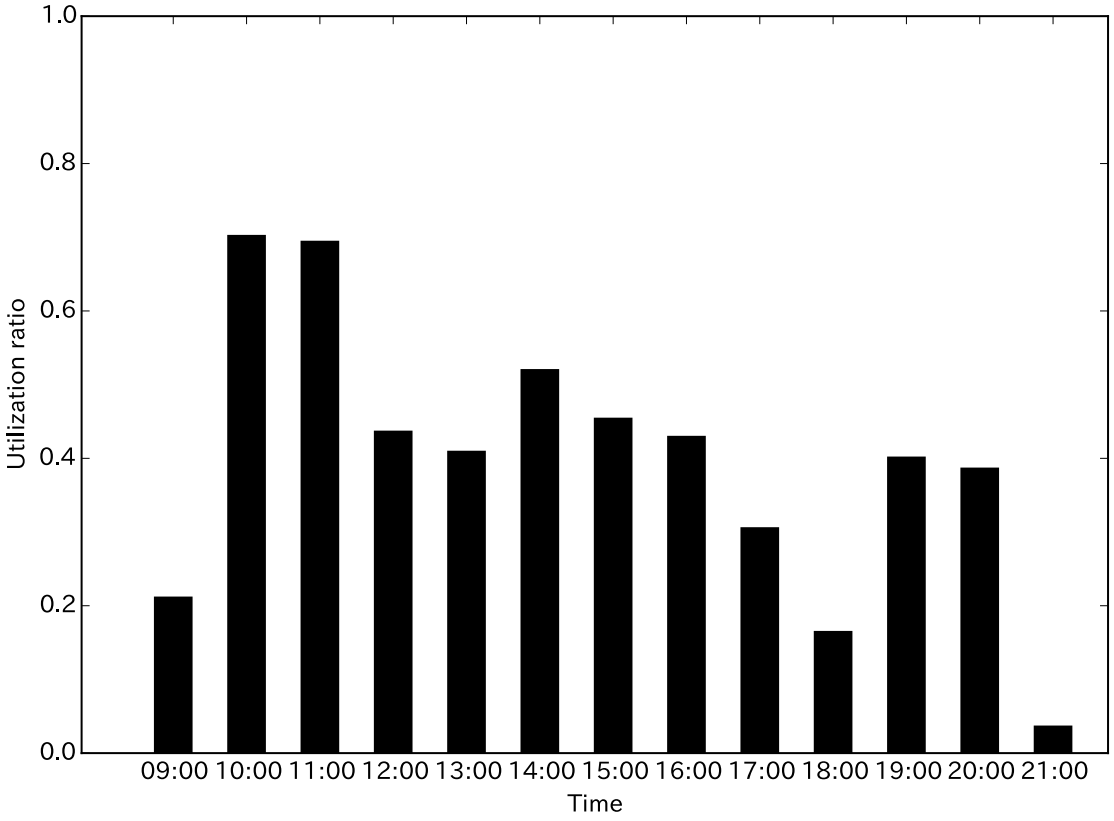
Analysis with drill-down operation

- Then, she gets interested in handicraft room of E, so she drill-down into months.



Analysis with drill-down operation

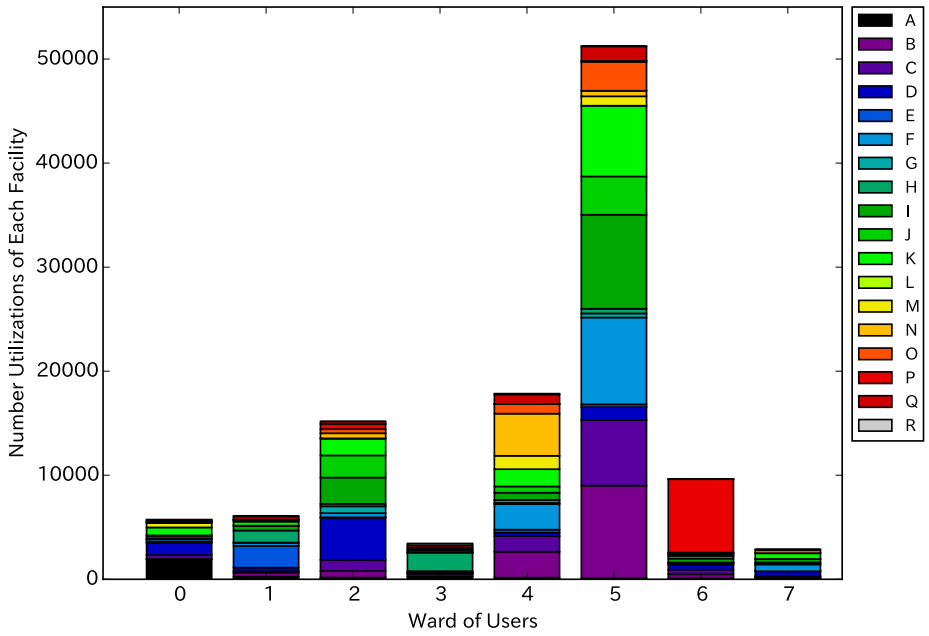
- She drills down into hour-level in order to observe what are the popular time in a day.



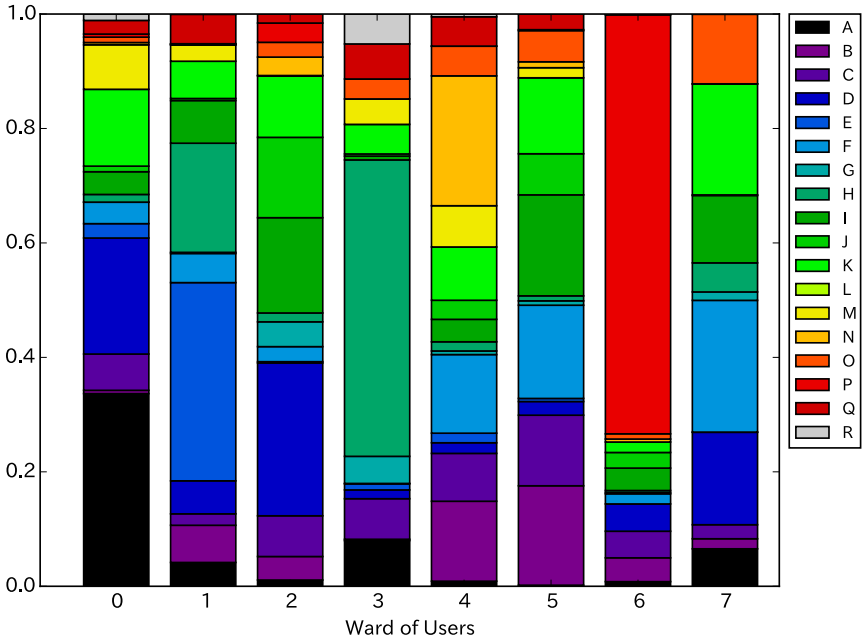
User Location-based Analysis

- City officers assume there are relationships between user locations and facilities.

Count-based view



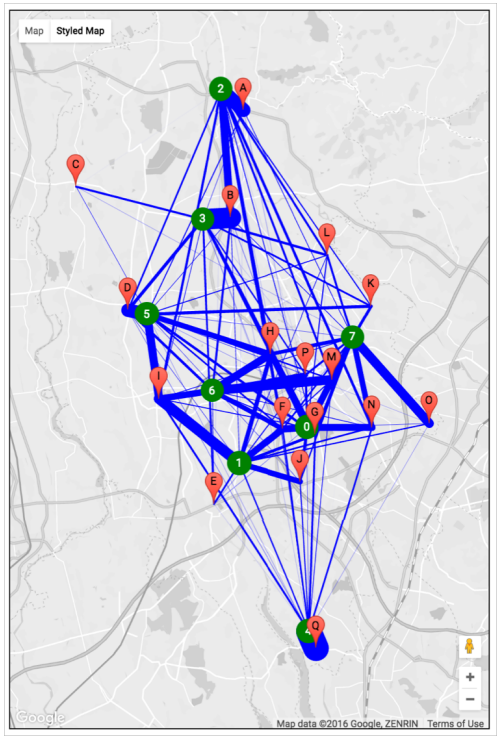
Ratio-based view



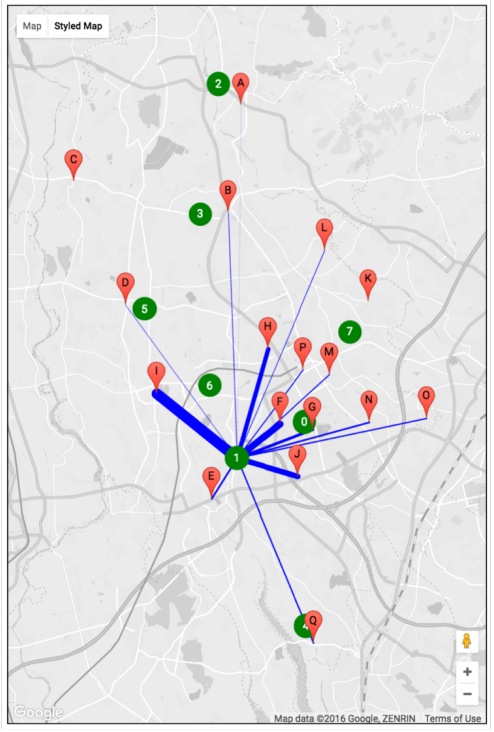
Map Visualization

- Visualize the relationships on maps in order to understand geographical relationships.

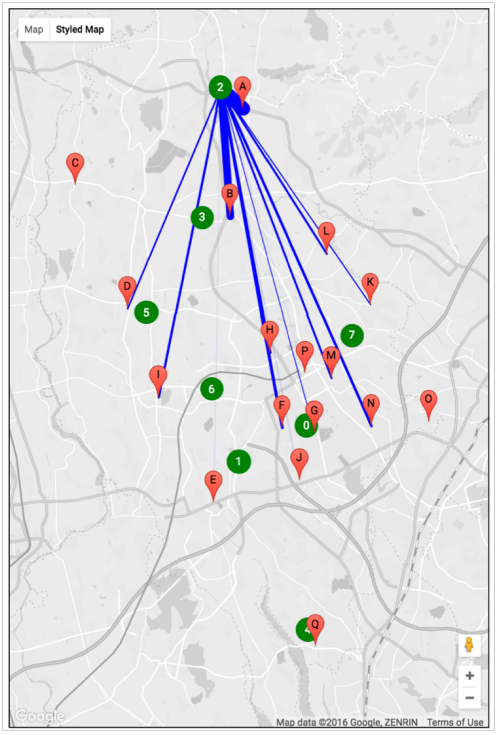
Overall



Ward 1



Ward 2



Conclusion

- Real-time analytic framework for smart city
- Case study on real-world smart city data based on practitioners' advices
  - City facility utilization logs from Tsukuba city
  - Basic analytical scenario
  - Location-based visualization
- Future work
  - Extend the framework for more heterogeneous data.

Thank you for
your kind attentions.