AALBORG UNIVERSITY

Business Intelligence for RFID-tracked Indoor Moving Objects

Associate Professor Hua Lu

http://people.cs.aau.dk/~luhua

Aalborg University

BagTrack Project

In 2012, more than 34 million bags were mishandled globally; a loss of ~$3.32 billion to the global aviation industry.

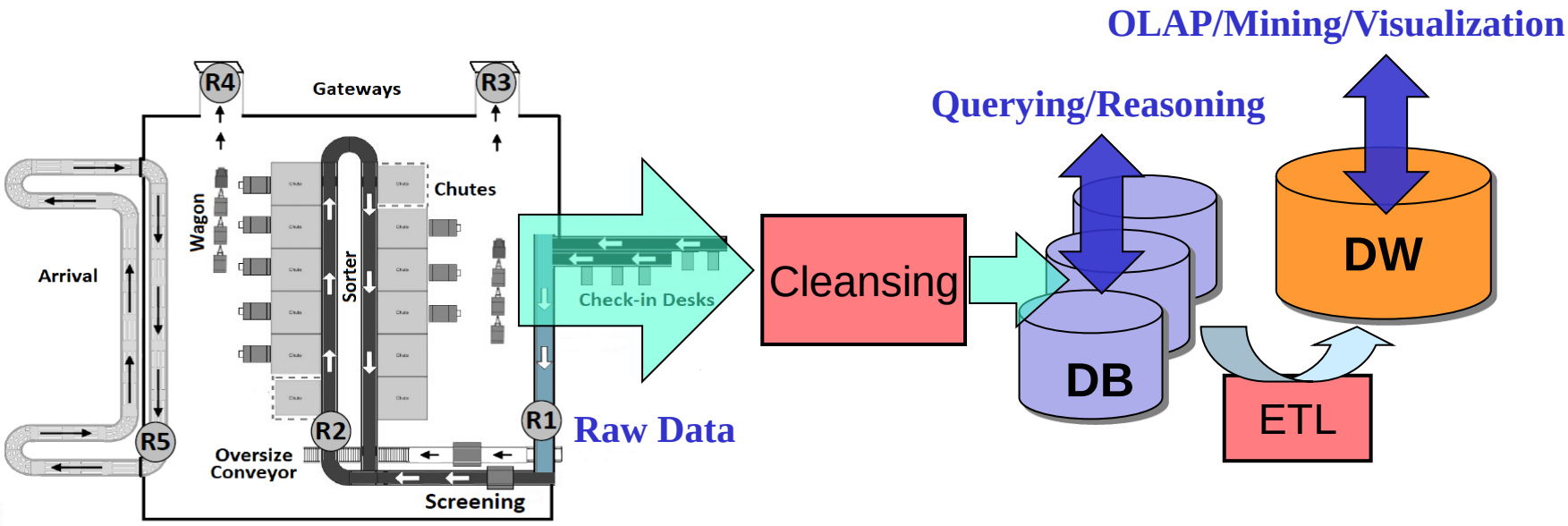
Project goal: to build a global IT solution for baggage tracking and management that will eventually cut the worldwide baggage problems by 50%.

- Aalborg University
- Lyngsøe Systems
- Aalborg Airport
- SAS
- IATA

http://www.youtube.com/watch?v=F16OuQim30g

Daisy's Role in BagTrack Project

- We're responsible for the *data management* work package.
- Specifically, we develop technologies for capturing and analyzing the massive amounts of RFID-based baggage location data that is obtained in airports.
  - Data warehouse (DW) and analysis for RFID data
  - Querying database (DB) of objects tracked by RFID
  - RFID data cleansing



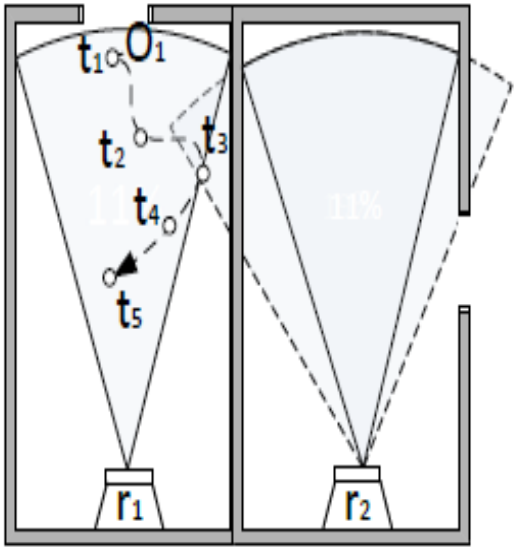
Outline

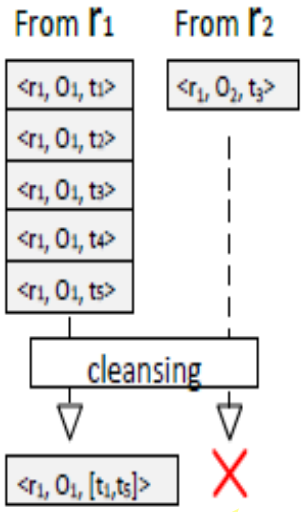
- Background and Introduction
- Cleansing Indoor RFID Tracking Data
- Distance-Aware Queries on RFID-tracked Indoor Moving Objects
- Data Warehouse for Airport RFID-Based Baggage Tracking Data
- Summary

Necessity of Cleansing

- Data errors
  - Temporal redundancy
  - Spatial ambiguities
- RFID reader deployment
  - Detection range
  - Sampling frequency
- Indoor space constraints
  - Distance between readers
  - Moving object speed

Temporal aggregation

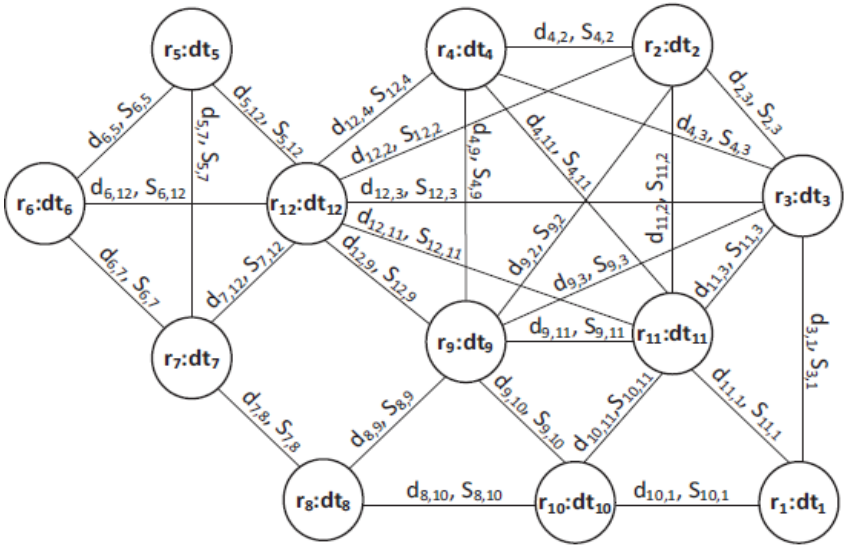


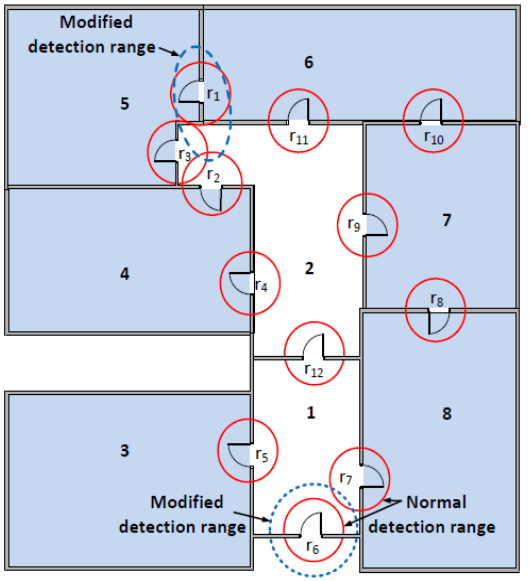


Spatial ambiguity: (a.k.a. cross-readings)
a reader mistakenly reads out a tag that is outside its intended detection range.
We resolve it by *Spatial Cleansing*.

Spatial Cleansing for RFID Data

- Distance-aware RFID deployment graph
  - Each reader is represented by a graph node
  - An edge means an object can move from one reader to another
  - Constraints (e.g., distance and speed) are captured as weights
- Idea of spatial cleansing
  - Scan the raw data w.r.t. the graph to identify spatial ambiguities
  - Resolve the ambiguities according to the connectivity and weights in the graph





Outline

- Background and Introduction
- Cleansing Indoor RFID Tracking Data
- Distance-Aware Queries on RFID-tracked Indoor Moving Objects
- Data Warehouse for Airport RFID-Based Baggage Tracking Data
- Summary

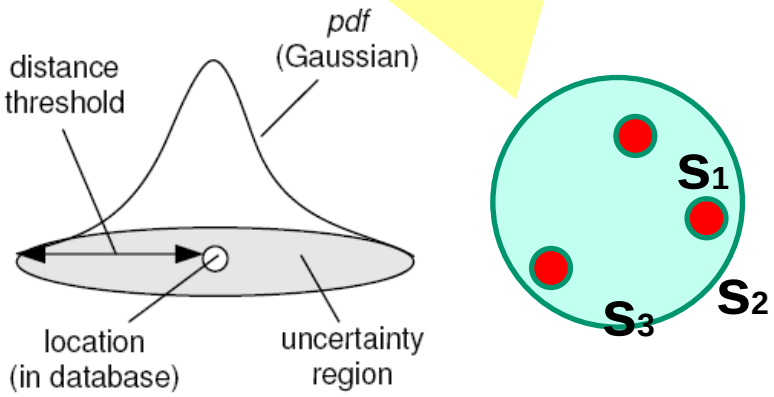
Indoor Distance-Aware Queries

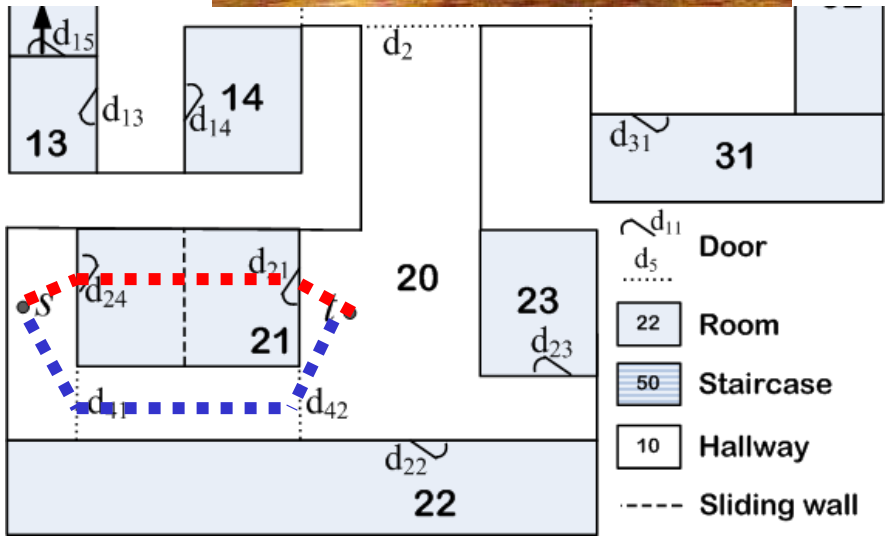
- Query types
  - Indoor range queries
    - Given an indoor location *p* and range value *d*, find all the objects whose *indoor* distance from *p* is not larger than *d*.
  - Indoor *k* nearest neighbor queries:
    - Given an indoor location *p*, find the *k* nearest neighbors for *p* in terms of indoor distance.
- Application scenarios
  - Baggage monitoring in airports
  - Museum guide service
  - Boarding reminder service
  - Indoor navigation

Technical Challenges

1. Temporal indoor topology variation
   - Pre-computed distances invalid!
2. Uncertain indoor object location
   - Continuous or discrete

Expected Indoor Distance is defined and adopted

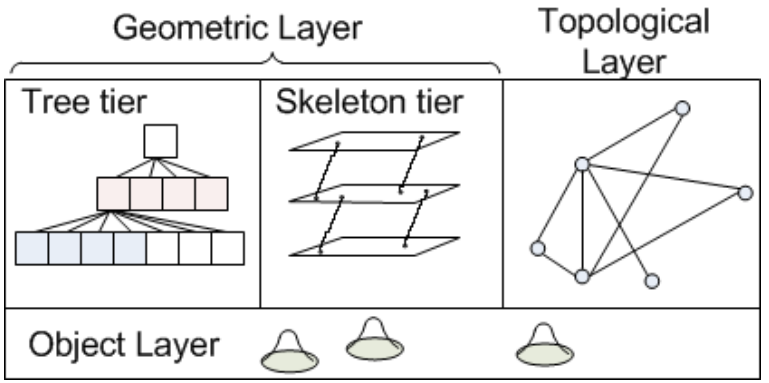


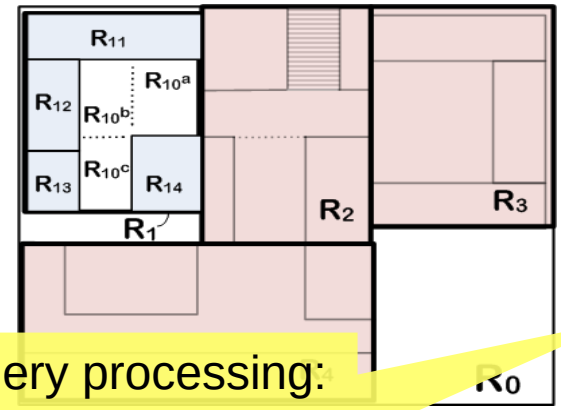


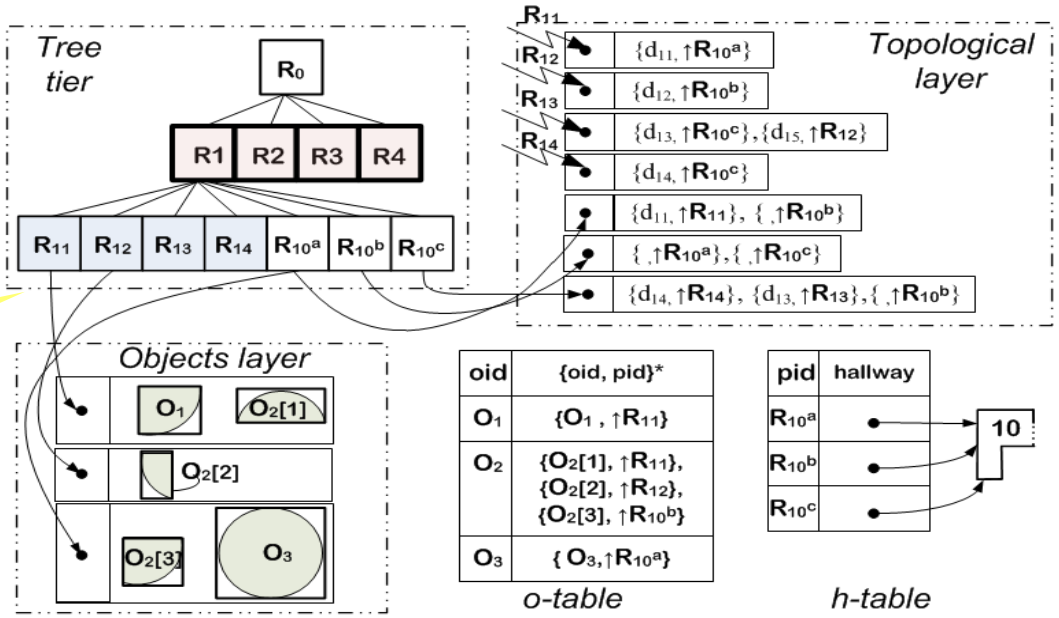
- Efficient distance computation is critical to queries
  - Can accommodate temporal changes and location uncertainty

Our Index Design

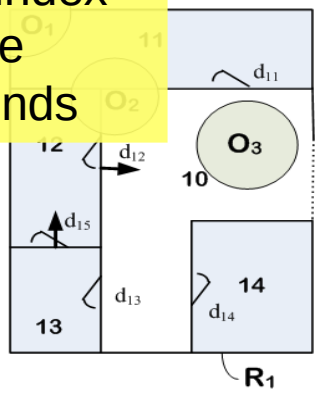
- Layers and Tiers
  - divide the problem and localize changes
- A composite index







Query processing: Search the index with effective pruning bounds



(without the skeleton tier)

Outline

- Background and Introduction
- Cleansing Indoor RFID Tracking Data
- Distance-Aware Queries on RFID-tracked Indoor Moving Objects
- Data Warehouse for Airport RFID-Based Baggage Tracking Data
- Summary

Data Warehouse Needed for Better Analysis

- Example analysis queries
  - How many bags were sent to wrong destination from CPH last week?
  - Find the average number of baggage traveling from CPH on Sunday.
- Problems with raw tracking data
  - The huge volume makes analysis queries very slow
  - Not well structured for powerful analytical queries with multiple dimensions
- From *tracking* records to *stay* records
  - A location is not fully covered by a reader
  - A tracking record does not tell how long an object stay on a particular location
  - A stay record is more informative than tracking record

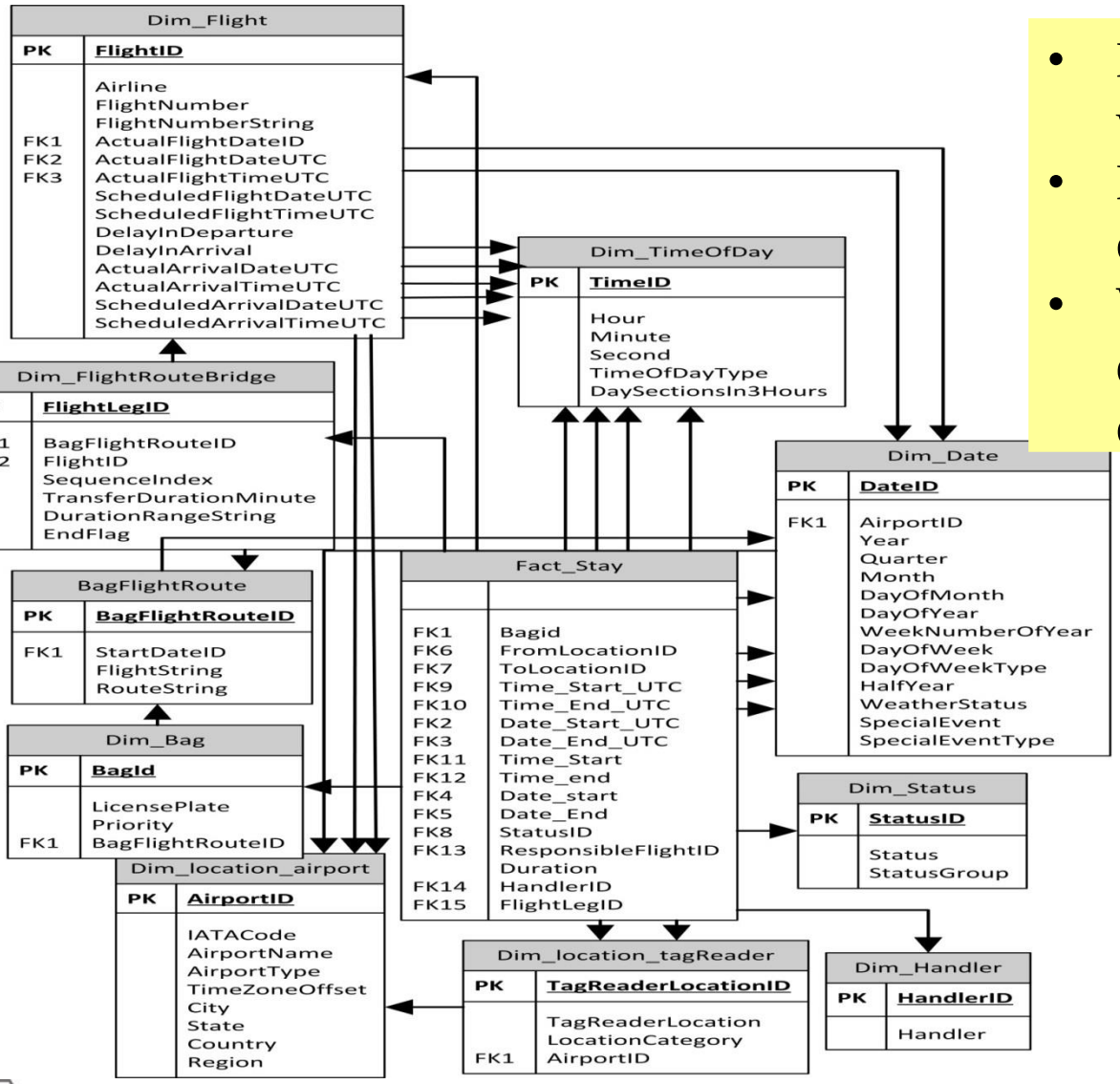
Bagid	Location	Time In	Time Out
B1	L1	1	1
B1	L3	4	5
B1	L4	7	8
B1	L10	19	20
B1	L11	30	31

Tracking records

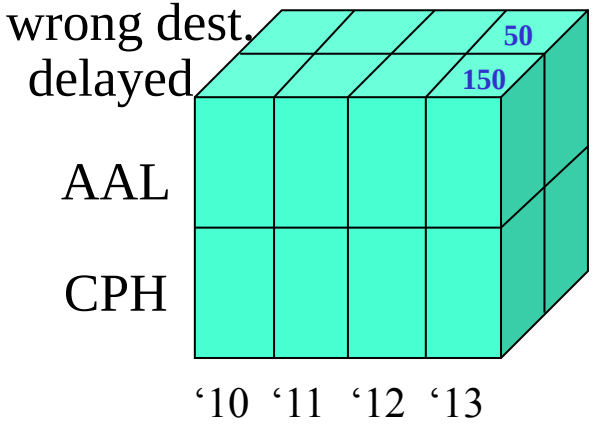
Bag Id	From Location	To Location	Time Start	Time End	Duration
B1	L1	L3	1	4	3
B1	L3	L4	4	7	3
B1	L4	L10	7	19	12
B1	L10	L11	19	30	11
B1	L11	L11	30	31	1

Stay records

A Snow-Flake Schema for DW

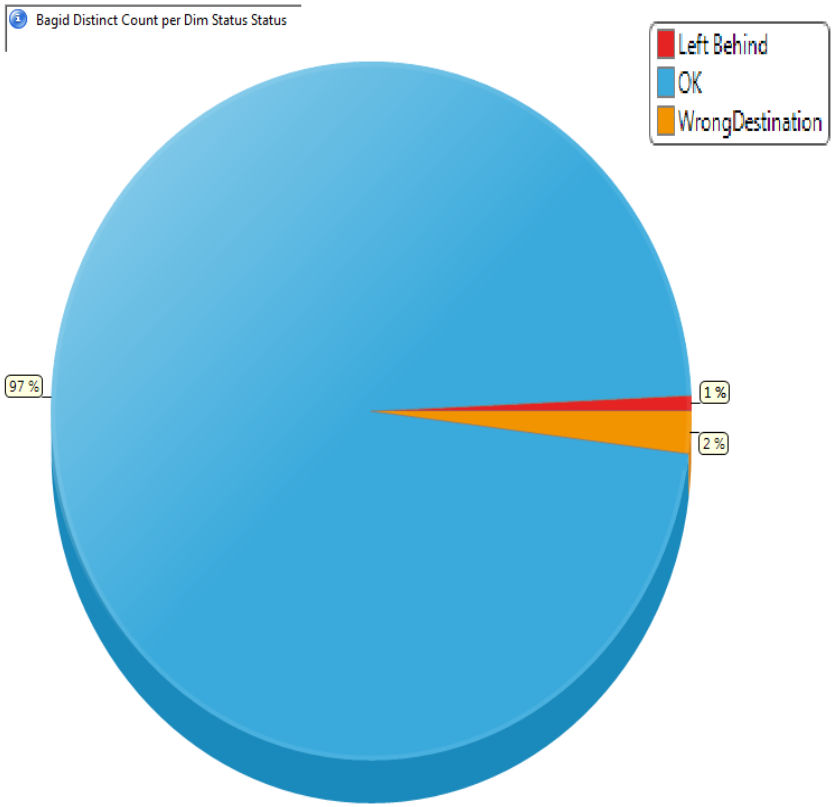


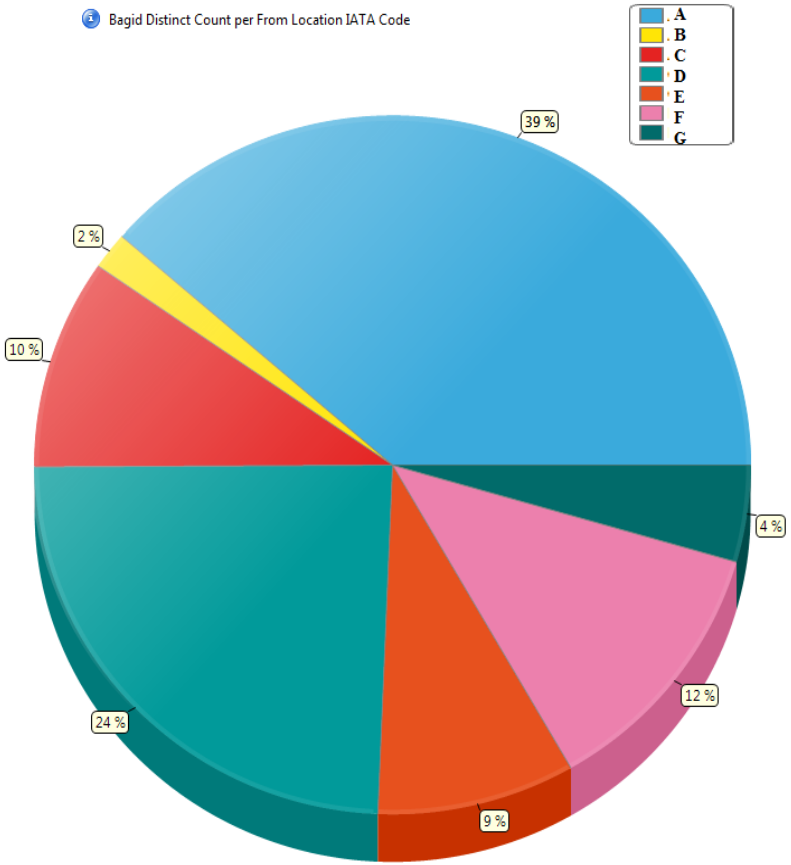
- Data can form a hyper *cube* with multiple *dimensions*.
- Each non-empty cell in the cube tells a *fact*.
- We can do roll-up or drill-down analysis to more rough or detailed results.



Results of Some Queries on the DW

- Find the percent of mishandled bags





Outline

- Background and Introduction
- Cleansing Indoor RFID Tracking Data
- Distance-Aware Queries on RFID-tracked Indoor Moving Objects
- Data Warehouse for Airport RFID-Based Baggage Tracking Data
- Summary

Summary

- Daisy intends to provide a complete set of effective and efficient business intelligence techniques for RFID-tracked indoor moving objects (e.g., airport bags).
  - Cleansing, querying, and warehousing/analyzing
  - Transforming raw RFID tracking data into meaningful and useful information for higher level use
- Daisy's technical proposals covers all the data management issues of the entire process of the IT system envisioned by the BagTrack project.
- Currently, Daisy has two faculty members, two postdocs, and two PhD students working on the BagTrack project.

Selected Research Publications

- Asif I. Baba, Hua Lu, Xike Xie, Torben Bach Pedersen: *Spatiotemporal Cleansing for Indoor RFID Tracking Data*. 14th International Conference on Mobile Data Management (MDM), 2013.
- Xike Xie, Hua Lu, Torben Bach Pedersen: *Efficient Distance-Aware Query Evaluation on Indoor Moving Objects*. 29th IEEE International Conference on Data Engineering (ICDE), 2013.
- Tanvir Ahmed, Torben Bach Pedersen, Hua Lu: *A Data Warehouse Solution for Analyzing RFID-Based Baggage Tracking Data*. 14th International Conference on Mobile Data Management (MDM), 2013.

- Acknowledges
  - Some of the slides used here are adapted from those offered by my coauthors.

Sponsors for Our Research

- The BagTrack project is funded by The Danish National Advanced Technology Foundation

- The work on RFID data cleansing is also supported by the NILTEK (Nordjysk Innovationscenter for Luftfarts Teknologi) project and Daisy Innovation projects

Potential Collaboration with Daisy

- Daisy has world-class researchers on data intensive systems. Our research covers *traditional databases*, *novel data management*, *data warehouse*, *data mining*, etc.
- In addition to relational business data, we also work actively on the following non-traditional data types:
  - GPS data
  - Indoor space data (e.g., RFID, Bluetooth positioning data)
  - Electricity data in smart grids
  - Social media data (e.g., tweets)
- We’re very interested in industrial collaboration that spans
  - student-based projects
  - networking activities
  - funded research projects
- Our website offers more information: http://daisy.aau.dk

daisy

Questions?

Thank you!