The BagTrack Project - An Overview

Asif Iqbal Baba, Sari Haj Hussein, Tanvir Ahmed

Supervisors: Torben Bach Pedersen and Hua Lu

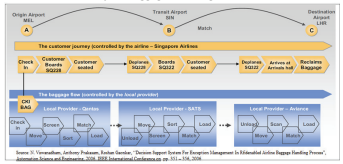
Introduction and Motivation

According to the Baggage Report for year 2011 compiled by SITA, the specialists in providing ICT solutions to the Air Transport Industry (ATI), the number of mishandled bags has been increasing over the past few years, thus accompanying the increase in passenger traffic. The direct consequence of this mishandling is reduction in the net profit of global airlines. To put facts in figures for the year 2010, around 29.5 million bags were mishandled which begot a financial loss of around 2.5 billion dollars.

Project Objective

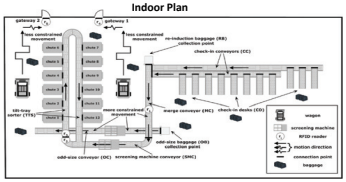
The BagTrack project which relies on the RFID technology in order to develop a computerized baggage handling solution that ensures better processing time and reduction in the amount of mishandled baggage in the air transport industry. For it to meet its goals, the BagTrack project explores innovative proposals in the areas of indoor tracking, cleansing of RFID data, and intelligent mining of that data, in addition to the study of statistical inference models for estimating the baggage handling quality at each site in the handling process.

Airport Baggage Handling

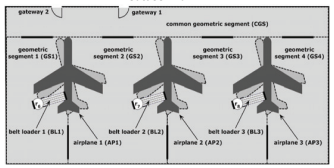


Topological Model and Continuous Query Processing

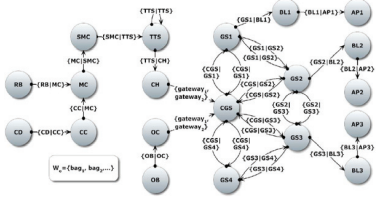
Indoor Plan



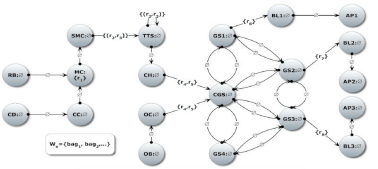
outdoor Plan



Graph Based Model



RFID Deployment Graph



i	$\kappa_i(t)$	i	$\kappa_i(t)$
CO, CC, CC, RB, OB, AP1, AP2, AP3	0	GS1	$\{r_y -0.1\}$
MC	$\{r_y -1\}$	GS2	$\{r_y -0.1\}$
SMC	$\{r_y -0.8, r_y -0.2\}$	GS3	$\{r_y -0.1\}$
TTS	$\{r_y -0.2, r_y -0.8\}$	BL1	$\{r_y -0.9\}$
OC	$\{r_y -0.95, r_y -0.95\}$	BL2	$\{r_y -0.9\}$
CH	$\{r_y -0.95, r_y -0.95\}$	BL3	$\{r_y -0.9\}$
CGS	$\{r_y -0.09, r_y -0.09\}$		

Data Cleansing in Indoor moving Objects

Research aims at reducing the errors that are inherent in RFID-based indoor applications. Errors came from different sources including hardware limits and device deployment in a complicated indoor environments like airports. Data errors need to be minimized before any meaningful query/analysis can be made subsequently.

RFID raw data are inherently unreliable. Three typical undesired scenarios:

1. *False negative readings(missing-readings)*: RFID tags locates in the range of a reader, may not be read by the reader at all.

- When multiple tags are simultaneously detected, RF collisions occur and signals interfere with each other prevents the reader from identifying any tags.
- A tag is not detected due to water or metal shielding or RF interference.

2. *False positive readings (cross-readings)*: RFID tag located outside the range of a reader are captured by this reader.

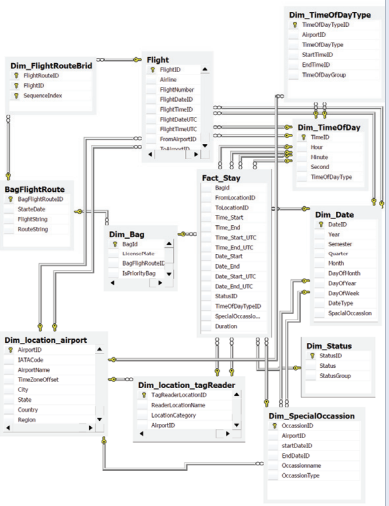
- may arise by the reflection of metal items, the abrupt strengthen of RF, and the change of antenna directions.

3. *Redundant Readings (Duplicate readings)*: Readings generated by a reader about a same tag continuously.

- Tags in the scope of a reader for a long time (in multiple reading frames) are read by the reader multiple times.
- Multiple readers are installed to cover larger space or distance, and tags in the overlapped areas are read by multiple readers.
- To enhance reading accuracy, multiple tags with same EPCs are attached to the same object, thus generate duplicate readings.

Data warehouse Design

- How many bags were lost at the sorter of Aalborg airport in the Easter-Holiday of 2012 that came from Copenhagen?
- How many baggage was handled by the Aalborg airport in the evening of 12th of January?
- What is the average time taken by bags to go from check-in to sorter in Aalborg airport during rush hours?
- Average number of bags that traveled the route AAL-CPH-BRU in the weekend of June.
- How many bags were lost in rush hours of Aalborg airport that travelled by flight SK1202?



Data Mining

- Finding interesting pattern
- Sequential Pattern mining
- Where should be the next location of the bag if the current location is sorter?
- How much time should be spend by a bag in sorter if it stays in screening belt for 25 second?
- Finding association among the routes visited by bags and passengers.
- Most Relevant Solution: Probabilistic Workflow [1]

Convert RFID Tracking records into Transactional Database

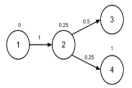
ObjectID	Path with Time
C1	L1, L2, L4
C2	L1, L2, L3
C3	L1, L2, L3
C4	L1, L2

ObjectID	Path with Duration
C1	(L1,5),(L2,145),(L4,0)
C2	(L1,3),(L2,62),(L3,0)
C3	(L1,3),(L2,62),(L3,0)
C4	(L1,5),(L2,86)

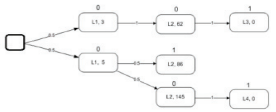
Clustered Path Database

Path with Time	Count
(L1,5),(L2,86)	1
(L1,5),(L2,145),(L4,0)	1
(L1,3),(L2,62),(L3,0)	2

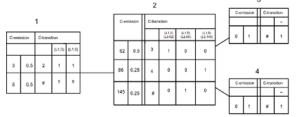
Probabilistic Workflow



Complete Workflow



Compressed Workflow



References

[1]. Hector Gonzalez, Jiawei Han, Xiaolei Li, "Mining compressed commodity workflows from massive RFID data sets" CIKM 2006: 162-171