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 costed a financial loss of around 2.5 billion dollars.

### Introduction and Motivation

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.

### Data Cleansing in Indoor Moving Objects

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

1. **False negative readings (missing-readings):** RFID tags located 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 preventing 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 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 spent by a bag in sorter if it stays in screening belt for 25 seconds?
- Finding association among the routes visited by bags and passengers.

**Most Relevant Solution:** Probabilistic Workflow [1]

<table>
<thead>
<tr>
<th>Convert RFID Tracking records into Transactional Database</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Clustered Path Database</strong></td>
<td>[1]. Hector Gonzalez, Jiawei Han, Xiaolei Li, “Mining compressed Commodity workflows from massive RFID data sets” CIKM 2006: 162-171</td>
</tr>
<tr>
<td><strong>Probabilistic Workflow</strong></td>
<td></td>
</tr>
</tbody>
</table>

### References

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