

# Supply chain Duplicate transportation RFID data stream Filtering

Nahla Mezzi<sup>1</sup> and Jalel Akaichi<sup>2</sup>

<sup>1&2</sup>Computer science department, University of Tunis  
41, rue de la Liberté, 2000 Le Bardo Tunisia

## ABSTRACT

*The supply chain is a system characterized by the mobility between the various processes of the chain as well as within the process themselves. Data capturing, gathering and management technologies are always needed by companies to support their decision-making and plans, develop their strategies and improve their mobility in supply chain. One of the technologies that could help companies handle data related to mobile object is RFID (Radio Frequency Identification), this later define a new and a rich information environment that opens up new horizons for efficient management of supply chain processes and decision support. However, due to the multiple readings in many cases in detecting an RFID tag and the deployment of multiple readers, RFID data contains many duplicates. Since RFID data is generated in a streaming fashion, it is difficult to remove these duplicates in real time. We propose in this paper an algorithm to detect and remove duplicated data streams in order to store trajectories data related to mobile object into trajectories data warehouse to enhance decision making for supply chain.*

**Keywords:** Supply chain, Mobile objects, Trajectory data warehouse, RFID technology

## 1. INTRODUCTION

Supply chain management is important for modern businesses to coordinate and synchronize activities of partner businesses, giving higher efficiency and improving productivity. However supply chain is a system characterized by the mobility between the various processes of the chain as well as within the process themselves. This mobility present an important challenge to deal with, because a huge amount of data was generated by the function of transport which is characterized by mobile objects such as trucks, trolley, and cars .Mobile objects have an influence on the functioning of the supply chain because they provide correct information about the location and the time in order to reduce the uncertainty, increase the visibility of products and increase the global efficiency of the supply chain. Therefore, using new technologies such RFID is necessary to capture information in real time coming from different data sources, internal and external in order to increase the logistic performance and reduce costs. Thus it is an important technology that serves to change the way business conducts and improves decision making in supply chain [9].

RFID is a wireless communication technology that uses radio-frequency waves to transfer information between tagged objects and readers without line of sight [10]. However, this technology generates a large volume of data streams which include among them an important number of duplicates. These duplicates are most of the time useless and may alter the decision making and decrease computing resources such as processing and storage resources. Removing these duplicates from the RFID data stream is essential as they do not provide an instant information to the system and as they waste system resources. Real-time information capturing and decision support present complex technical challenges, related to managing huge streams of data coming from multiple data sources and converting them into meaningful information so that to support decisions.

Our contributions are described as follows:

- The management of data streams generated by mobile objects, we propose an algorithm that efficiently detects and removes duplicated readings from RFID data streams.
- The modeling of the Trajectory Data Warehouse (TDW) related to supply chain activities.
- The analysis of TDW in order to extract useful knowledge to enhance decision making in the supply chain field.

The paper is organized as follows: Section 2 reviews existing approaches that pushes us to propose an algorithm to remove duplicated data, whereas section 3 presents the mobility in the supply chain using RFID technology. Section 4 deals with the trajectory data warehouse modeling to support trajectory data generated by RFID, while section 5 presents the algorithm DUP\_Filtering. Section 6 shows the analysis and reporting data extracted in order to enhance the decision making in supply chain and finally the conclusion.

## 2. RELATED WORKS

A RFID system consists of a tag which is attached to the object to be identified, a reader for detecting radio waves, the Middleware and a database system for supporting the big number of information related to the mobile objects, and other associated information [4]. The middleware gathers and processes the information captured from readers that will be used later on by company.

The processing, filtering and aggregation aims to transform raw data into meaningful information for the application [4]. So the middleware is very necessary to pre-process data. This paper discusses problems related to filtering RFID data and reviews the existing duplicate RFID data filtering approaches that are related to this work. However, the volume of RFID data is generally very big and produced in a stream which requires a method to eliminate duplication that should process and analyze data instantly.

Therefore, redundancy can happen at two different levels, redundancy at reader level and redundancy at data level.

Duplicate reading at the reader level happens when the tagged object is being read by more than one reader. This can happen when there are overlaps in the reading vicinity of several readers. As opposed, duplicate readings at the data level happens when an RFID reader continues reading the same object repeatedly. The problem of duplicate readings in RFID required an efficient approach to solve it.

### 2.1. Duplicate reading at the data level

The article focuses on the redundant readings at the data level thus we will present some algorithms that treat this problem. This problem occurs when multiple RFID readers are deployed to control and track tagged objects. However, the fast capturing of data by RFID readers provide challenges for data management such as errors, data duplication, and miss-reads. There are many approaches in the literature review to filter duplicate readings [1]–[3]. Algorithms proposed in these two works focus on data level filtering, as they focus on a single RFID reader system.

However, these algorithms proposed do not only deal with the deletion operation which removes unnecessary information; but they also consider just the fixed objects. Moreover, the tags' movement is not suitable in RFID scenarios because they cannot easily be predicted [3]. The major weakness of the sliding window approach is that the size can be not large enough to perform filtering correctly.

Approaches that explore the Bloom filter for filtering duplicate readings in RFID have recently emerged in the late researches [5]–[6]. The main idea of the Bloom Filter [4] is to define an element that represents a positive counter in a bit array of a predefined size and with number of hash functions. All bits in BF are initially set to 0 and will be changed to 1 if it is hashed by the element. To check the existence of element in a set, the element will be executed by the same hash functions used to insert the elements in the array. That is the element is the member of the set if all the bits in which the element was mapped to is positive. For each new element, the corresponding bits in the array take the value 1 [4]. But, Bloom Filter has some weaknesses such as: the impossibility of deletion which can result in the generation of false positive errors. Thus the Bloom Filter tests the membership of any element with errors [4]. So it can provide wrong information which can result in a wrong decision making. That's why Bloom filter requires some improvement.

Moreover, in another work [3] authors present an approach that filters the detected data including the duplicate elimination. They discuss the need to filter RFID stream efficiently and they proposed a filtering method to eliminate noise from RFID readings. The method uses the number of reading occurrence per time as the basis of the filtering process. The same authors in [2]–[4] discussed the problem of detecting duplicated data. The second work is an extension of the previous one. Authors enhance their works by reducing the number of filters from two to only one unlike the previous algorithm that make use of two filters. Thus, eliminating one of the filters reduces the time required to filter duplicate readings. Nevertheless, this approach has a low false positive rate which illustrates the improved correctness of the filtering process. It is more efficient in terms of time and memory usage.

In addition, in the same context, the objective of this work [8] is to propose one pass approximate methods based on Bloom Filters using a small amount of memory. Thus, researchers propose Time Bloom Filter to detect duplicate in RFID data streams, which extends original Bloom Filter to support sliding windows. They propose also Time Interval Bloom Filters to reduce errors which need more space than Time Bloom Filters. So, this method can reduce space for Time Interval Bloom Filters to effectively remove duplicates in RFID data streams. But the major weakness of Time Bloom Filter is to detect duplicates in the server side. However, since all the data will flow through this module in the server, it will soon become the bottleneck.

Another article discusses the same idea but with a slight difference which is the following: First, in order to reduce the network bandwidth requirement, authors proposed using a two layer approach to eliminate the local duplicates in the "network edge", then detecting the global duplicates in the server side. Second, since all the RFID data will be sent to the server, the latter will be submerged. The method proposed is distributing the Time Bloom Filters to several computing nodes to enhance the scalability [7].

In the following, the paper is going to discuss the effective and efficient solutions for the previously mentioned problem of the elimination redundancy at data level.

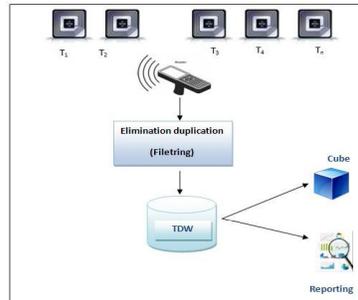
### 3. SCENARIO OF COLLECTING RFID DATA STREAM

In the context of our work we are interested in external mobility that contains the distribution process and transport. Tags are existed in the mobile object; readers detect these tags and transfer them to middleware in order to filter duplicated data. The proposed model in our work consists of an implementation of RFID technology in the supply chain with passive tags. This is achieved by an installation of readers in many places in order to track the mobile objects. These tags will be placed in different steps to capture information in real time because it is very important for control of deviation in production.

We focus just on the external mobility. The tags are integrated in the trucks in order to monitor truck drivers and assist them if there are problems in the trajectory path or in the case of a wrong destination.

Therefore, in our case, we propose that readers are installed in the supply chain, during the process of production and also in the proximity of each destination that trucks must deliver the command in a well-defined time.

Data collected from these mobile objects without contact that is if a RFID tag is within a proper range from an RFID reader, the RFID tag will be detected. This data serves to know the location of mobile object, the trajectories of trucks, as it monitors them to reduce the risks that may occur in order to improve the decision making and the productivity which is based on the analysis in the real time of this data. To makes things clear, the following figure is a schema that describes the mechanism of work starting from RFID data stream collection till reaching the analysis and decision making.



**Figure 1**System architecture

Hence the gathered data will be sent to the middleware. Our algorithm proposed in the next section presents a middleware between collected data and the trajectory data warehouse. This means that gathered data must be filtered and should not contain duplicated and redundancy data before sending them into the data warehouse. It is necessary to send clean data to the data warehouse because, duplicated data can adversely affect the supply chain performance and thus affect decision making.

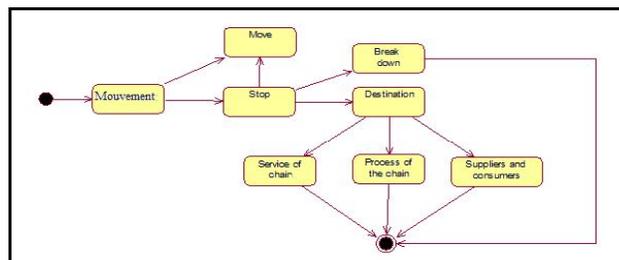
### 4. MODELING TRAJECTORY DATA WAREHOUSE FOR SUPPLY CHAIN

In this work, we present the use of RFID technology in order to gather and control in real time data related to these mobile objects. This will help us to know the location of the mobile objects and the time of departure and arriving.

#### 4.1. TRAJECTORIES OF MOBILE OBJECTS IN THE CHAIN

The mobile objects represent the different means of transport in the supply chain. Mobile objects follow a path which is the path from the starting point to the arriving point; the information on this trajectory represents an important basis for decision-making since they are comprehensive and contain all the data about the object in its environment and travel conditions.

A path consists of a set of movements and a set of stops. Even a mobile object in the supply chain can take either the status, movement (court movement) or stop for different reasons such as the breakdown, climate, or arriving to the destination. The diagram below shows the different states that can occur to an object moving in a path.



**Figure 2**The states of the mobile objects in the supply chain

A mobile object moving in its trajectory can be in movement or in stop; the movement is characterized by a start time and an end time. The start of movement indicates the end of a stop and the end of a movement indicates the beginning of a stop, so the time interval of movement is delimited by two consecutive stops.

Stops are important parts of a path or a non-moving. It can be caused by different reasons; we can distinguish the stop in three types:

- A strategic stop: for the delivery of the transported products, when arrived in the destination.
- A private stop: to have a break ...
- An unforeseen stop: when there is a breakdown or a bad climate.

#### 4.2. Trajectory data modeling in the supply chain

The successive movements of mobile objects in the supply chain are represented by trajectories. The analysis of the path traveled by the mobile objects can help to optimize supply chain and improve decisions. In order to perform mobility analysis in a supply chain, it is necessary to record the data paths, as well as the characteristics of mobile objects.

Data streams must be stored in a multidimensional model to be analyzed. The model of the classic data warehouse is not adapted to the storage of the information related to the mobile objects. For this reason, these applications need to join spatial and temporal types of data, which requires a spatiotemporal data warehouse, but this data warehouse is not sufficient for the treatment of the trajectories of mobile objects.

Besides, the TDW must be able to work in an environment, where the observations of trajectory data are received in a continuous way. The analysis of trajectories will help the decision-makers to observe the movement of trajectories of the mobile objects in the supply chain. The modeling of our TDW requires the precision of contents of supply chain and to organize according to the expected results and the statues of the mobile object. In our work, TDW is an effective way to model mobility data and its characteristics. It allows the monitoring of mobiles objects, movements throughout a path at any moment. It also gives the possibility of analysis and extracts knowledge from historical data, which ensures better decision making.

The following figure is the conceptual model for the resulting data trajectories related to mobile objects in the supply chain represented by a UML class diagram.

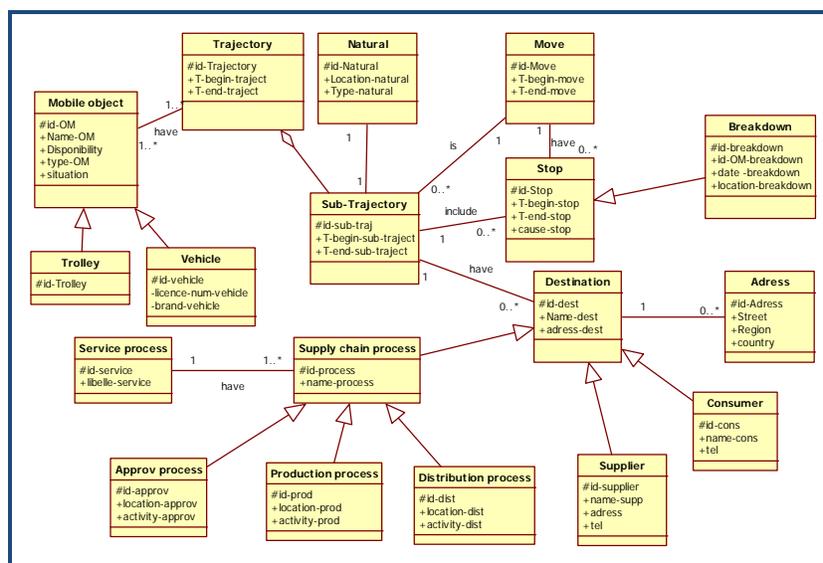


Figure 3Supply chain’s mobile objects ‘Trajectory data Conceptual Modeling

We will describe in what follows the main classes and relationships of our diagram:

**Mobile object:** the mobile objects insure the movement of the goods between the various services of the chain; these objects also transfer the data to the decision-making. The type of mobile object can be either a vehicle for the external transport, or a wagon for the internal transport.

**Trajectory:** the trajectory is the road crossed by a mobile object. Every trajectory is characterized by a time of begins and of the end. The end of the trajectory presents the arrived to the destination. A trajectory can consist of one or several sub-trajectories .Trajectory =  $\sum$ sub- trajectory

**Sub-trajectory:** every sub-trajectory consists of a set of movement. The sub-trajectory limited by two successive stops.

$$\text{Sub-trajectory} = \sum \text{Movement} = \sum_{i=1}^{n-1} S_i + M_i$$

**Movement:** every movement of mobile object in a trajectory has an identifier and characterized by time and location of movement.

**Stop:** stops in a trajectory are identified and characterized by time and location of the stop. A stop can be for several reasons either for reason of breakdown, or in case when the mobile object arrived at its destination.

**Destination:** the destination of the trajectory can be the consumers, the suppliers or the processes of the chain, or the services of the supply chain.

**Natural:** the natural present the natural characteristic for each trajectory (urban, mountain, desert...)

### 4.3. Modeling trajectory data warehouse

We used the multidimensional modeling of the TDW, which is the star schema. It is a method of designing a data warehouse where a central fact table is connected to one or more dimension tables. The fact table is a table of reference to get to the center allowing events or activities archived. In our case, we chose the trajectory as fact table and main subject of analysis since it includes all the information and data relating to the moving object, the analysis will be based on several axes or dimensions, such as time, movement, the mobile object, destination etc. The fact table contains the 'measures', which are numerical fields on which we will do our analyses, we can find the number of stops, distance travelled, travel time, the average speed etc..as shown in fig.4.

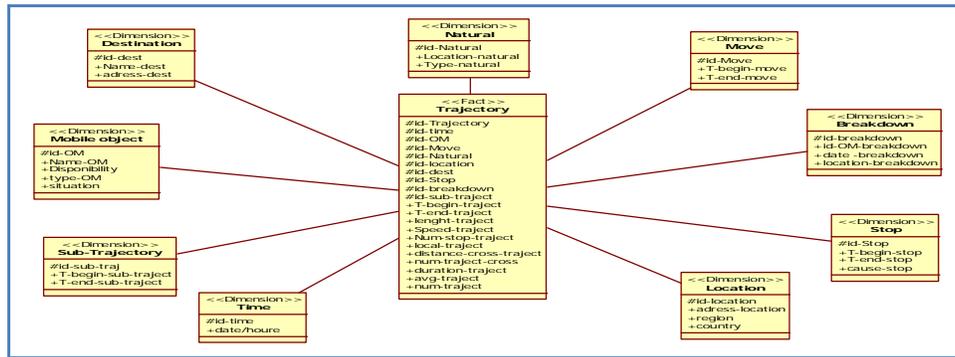


Figure 4 Star schema

In the next section we present our algorithm to eliminate duplicated data generated by RFID in order to load into trajectory data warehouse only data cleaned that serves subsequently to analysis .

### 5. DUP\_FILETERING ALGORITHM

In this section, we present our effective and efficient solutions for the problem of the redundancy elimination at data level. Before begging, we provide two definitions quoted from this work [8] and on which our algorithm is based:

**Definition 1.** An RFID data stream S is a set of {s<sub>1</sub>,..., s<sub>n</sub>}, where s<sub>i</sub> consists of a triple (TagID,Loc,Time) such that "TagID" is the unique identification number for tag , "Loc" is the location of the reader which detects the tag and "Time" is the time of detecting the tag.

**Definition 2.** (Duplicate). In a RFID data stream, item x is considered as a duplicate if there exists item y, such that x.TagID = y.TagID and x.Timestamp - y.Timestamp < τ, where τ is application specific positive value.

Therefore, we develop an algorithm for the detection of the RFID tags distributed with proper range in a given area. If an RFID tag is detected, the RFID data with the same tag\_id is detected repeatedly over a period. Thus, the structure of the Bloom filter optimized is shown by the time of the flow detection and "tempsprec" means the time of the previous stream detection of the same tag\_id . To see if there is a probable duplication or not we will make the difference between both time and controlling the variable T .This later is set to  $\frac{3}{4} \tau$  with τ is in milliseconds.

To better understand this parameter, we give this definition:

**Definition 3.** For detecting mobile objects which are not in working, we created a parameter T, which is initialized to :

$$T = \frac{3}{4}t \times \frac{1}{f} \quad (1)$$

We define a moving object which is not working if T reaches 0. T is decremented each time we remove a duplicated data in the same Tag\_id.

We chose  $T = \frac{3}{4}t \times \frac{1}{f}$  since the probability to detect a mobile object in work is  $p = [(1 - \frac{3}{4}t)] = 0,25$ . We note that t must be in millisecond since the detection of frequency generated by readers is very fast. We choose to multiple  $\frac{3}{4}t$  by  $\frac{1}{f}$  because  $\frac{1}{f}$  is the period of reading reader.

For more explanation, we give a simple example that is  $\frac{1}{f} = 2$  and  $t = 8s$  then  $T = 12$  because the reader read the stream each 0,5 second thus if we do not multiply  $\frac{3}{4}t$  by  $\frac{1}{f}$  T will equal to 6 that is to say if it detects a duplicate T reaches 0 after 3s (since Frequency: 2 readings per second) rather than after 6s that is why we must multiply by the period  $\frac{1}{f}$ .

Figure 5 presents The Dup\_Filtering algorithm to eliminate duplicated data.

```

DUP_FILTERING Algorithm

INPUT : Integer tagid ,x, T;
OUTPUT : Stream

1. Begin
2. x ← Researchtag(tagid) // Looking in the table tag the tagid
3. IF (x == false)
4.   Begin
5.     Call AddInTag(tagid); // Adding tag_id in the tag table
6.     Call AddInStream(tagid) // Add in stream(tag_id, loc, time )
7.     T ←  $\frac{3}{4}t \times \frac{1}{f}$ 
8.     Call AddInVerif(tagid, T); // Add the value of T=6 , tag_id and tempspec in verif table
9.     Call LoadDataInDataWarehouse(getStructureTableStream())
10.  End;
11. ELSE
12.  Begin
13.    diff ← VerifDiffTime(tagid) // Calculate the difference between detection times
14.    IF (diff <= Threshold) THEN // Data is duplicated , check the T value
15.      Begin
16.        T ← VerificationInTableVerif_T(tagid) // Check the T value
17.        IF (T ≠ 0)
18.          UpDateVerif_t(tagid, T) // Decrimente T
19.        ELSE
20.          Display message ("Check your mobile object")
21.        End
22.      Else // data not duplicated
23.        Begin
24.          tp ← UpDateTableStream(tagid) // Changing the time of detection time
25.          UpDateTableVerifTime(tagid, tp) // Changing T (becomes again T= 6)
26.          Call LoadDataInDataWarehouse (getStructureTableStream())
27.        End;
28.      End;
29.    End.

```

Figure5Dup\_Filtering algorithm

The algorithm functions as follows:

**Add\_tag(tagid)**

"INSERT INTO tag (tagid) VALUES "tagid""

**Insert\_Stream(tagid)**

"INSERT INTO stream(tag\_tagid , loc, time)VALUES ("tagid", "Loc", "time")"

**Add\_verif**

"INSERT INTO verif(tag\_tagid,t,tempspec) VALUES(" tagid ", " 6 ", "time")"

In case we find that the data is duplicated, we have to check the value of T (the variable that will help us in decision-making) by verification\_t function (tagid)

**Verification\_t(tagid)**

```
"SELECT t from verif where tag_tagid=" tagid ""
```

if  $T \neq 0$  then we decremented the value of T with the function Up\_Date\_verif\_t(tagid, T)

**Up\_Date\_verif\_t(tagid, T)**

T=T-1

```
"UPDATE verif SET t=" T " where tag_tagid=" tagid ""
```

Whereas, we conclude data is not duplicated according to the update\_stream function (tagid), we should update the stream table Up\_Date\_verif\_time and "verif" table by changing the time of detection and the value of T which becomes again equal to 6, then we should send the new stream to the data warehouse.

**Update\_flux(tagid)**

```
"UPDATE stream SET time =" time " where tag_tagid=" tagid "" with time is the new time detection of data stream.
```

**Up\_date\_verif\_temps**

```
"UPDATE verif SET T=" 6 " , tempsprec=" time " where tag_tagid="tagid "" with time is the new time detection of data stream.
```

As a result, the data warehouse contains only clean data, so we can analyse these data to extract useful knowledge in order to improve decision making in the supply chain.

## **6. DECISION MAKING: ANALYSIS AND REPORTING**

In this section we present analysis and reporting data in order to improve decision making in the supply chain and reduce uncertainty.

### **6.1 Data warehouse analysis**

In this phase, we have filtered RFID data streams and loaded them into the previously mentioned trajectory data warehouse. Thus the data analysis related to mobile objects in the supply chain stored in the TDW, is an important step to derive knowledge necessary to the decision. In our context, the analysis concerns first then data of mobile objects and second their trajectories:

- Analysis of mobile object data: analysis is used to display data relating to mobile objects (the location of the mobile object at a definite time, details regarding the use of mobile objects in a time period, mobile objects available for use, the places visited by the object ...). These data allow decision makers to better manage all types of mobile objects in the business.
- Analysis of mobile objects' trajectories: the analysis concerns the trajectories of the mobile objects and their characteristics (the downtime of a trajectory, time, path length...) these data are used to improve the mobility management of the mobile objects which characterize the trajectories travelled, and subsequently improving the transport function.

Data analysis allows operating with maximum efficiency and developing new ideas to better understand and optimize the act of decision making. In our context, we aim to improve decision making in supply chain. In other words we aim to extract useful knowledge related to trajectories data that can represent a support for taking the best decision and reducing costs and uncertainty. In this work, we have supported the analysis of data from the data warehouse using Pentaho Schema Workbench. PSW can create Mondrian OLAP cubes from a model of star data warehouse.

```

<?xml version="1.0"?>
- <Schema name="Trajectory">
  - <Cube name="trajectory_cube" enabled="true" cache="true" visible="true">
    <Table name="trajectory" alias=""> </Table>
    <Dimension name="mobile object" visible="true" type="StandardDimension"> </Dimension>
    <Dimension name="location" visible="true" type="StandardDimension"> </Dimension>
    <Dimension name="Time" visible="true" type="StandardDimension"> </Dimension>
    <Measure name="nbr_mobile-object" visible="true" aggregator="distinct-count"
      column="Location_idLocation"> </Measure>
  </Cube>
  - <Cube name="cube_trajectory" enabled="true" cache="true" visible="true">
    <Table name="trajectory"> </Table>
    <Dimension name="Destination" visible="true" type="StandardDimension"> </Dimension>
    <Dimension name="Time" visible="true" type="StandardDimension"> </Dimension>
    <Dimension name="mobile object" visible="true" type="StandardDimension"> </Dimension>
    <Dimension name="breakdown" visible="true" type="StandardDimension"> </Dimension>
    <Measure name="nbr-breakdown" visible="true" aggregator="distinct-count"
      column="Stop_idStop" datatype="Numeric"> </Measure>
  </Cube>
</Schema>

```

Figure6 XML cube schema

Constructing OLAP cube allows to answer different questions. A cube consists of one or more fact tables with their dimension tables. It aims to organize the data to be analyzed by area / topic and to highlight relevant results for the decision maker.

In our work, asking information about the mobile objects and their trajectories allows us to guide the trucks drivers in their trajectories, to improve mobility ...

There are some questions (queries) that a decision maker can ask:

- What is the number of stops caused by breakdowns depending on the location by truck throughout the trajectory in a month?
- What are the places where there is frequently breakdown?
- What is the number of moving objects by location in a well-defined period?
- What is the nature of the places visited by the trajectory?
- What is the number of the retarding mobile objects?
- Etc.

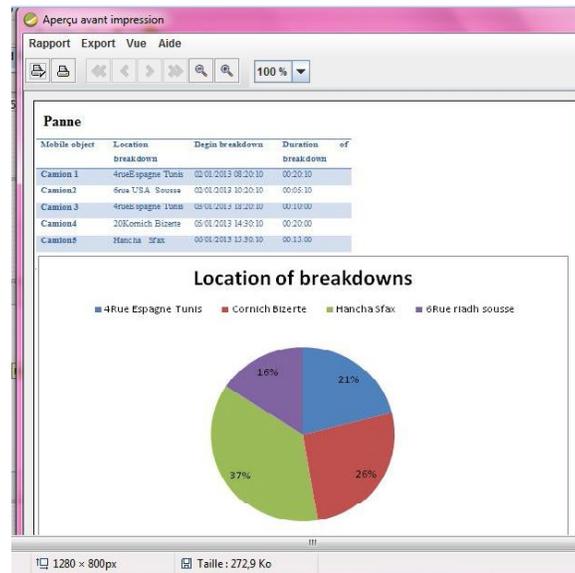
## 6.2 Data warehouse Reporting

We present in this section, the analysis reports issued through queries using the report designer tool and decisions that may be proposed by the decision makers of the company. Our final goal along this work is to discover hidden relationships between variables including space and time in order to find explanations of some problems related to trajectories of mobile objects and consequently make best decision to enhance mobility in the supply chain.

We use interrogations based on several axes of analysis which allow us to extract important knowledge that can help in the decision making. First, the interrogation concerns the mobile objects in the stop for a given moment, allows making instant status tracking of a moving object, we can know whether an object is in stop and why, thus allowing effective and instant intervention.

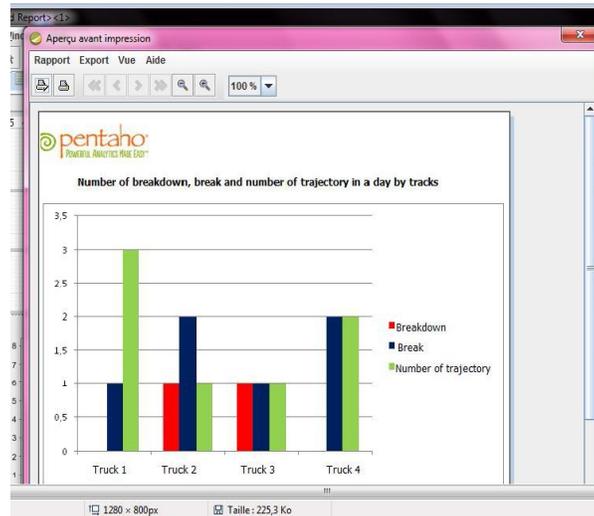
Interrogations are used to know the number and duration of stops on a trajectory to assess the effectiveness of this trajectory by calculating the lost time, which allows a strategic intervention to understand the reasons for this loss and try to reduce it. This can also be a good way to compare the trajectories, which allowing better future choices.

To better understand the reasons for delay and stops, interrogation about the nature of the trajectory can be useful, thanks to RFID technology we are able to have an idea about the climate conditions of a trajectory, we note that both trajectory 1 and 2 are formed of deserted areas and highways. This will allow decision makers to better anticipate failures and potential problems related to climate and take into account the necessary precautions to mountainous areas and desert regions as shown in Figure 7.



**Figure7** Location breakdowns

Moreover, thanks to interrogation about the type of nature of the trajectory, this allows the decision maker to know the performance of trucks according to nature and thereby they can follow the truck which has better performance depending on the nature in order to assign the suitable trajectory to each truck to save time. Therefore information on the nature of the places visited helps decision makers to take precautions depending on the climate, and therefore reduce the risk of breakdowns and accidents on the road. Therefore, the decision maker can decide to buy moving objects the most adapted to the nature frequently crossed.



**Figure8** Number of breakdowns, pause and number of trajectory travelled in a month by trucks

The decision-making system which is interested in mobility in the supply chain based on a multi-dimensional trajectory data warehouse enables online monitoring of the transport function in the supply chain, it facilitates the direct intervention and provides an excellent model for the analysis of data related to the path on different axes. So a focused and specialized analysis, allows to extract knowledge and provide decision makers with relevant elements for better decision making can reduce costs and time to strengthen the supply chain and performance

In our case the scheme of the data warehouse and the questions proposed give decision makers an overview of the mobile object in the supply chain by specifying the behavior of the moving object and the characteristics of its trajectories such as the number of stop by trajectory, the location of the paths, distance, time trajectories...

This information allows decision makers to make decisions taking into account the mobility in the chain, which can reduce the costs related to transportation, improve product traceability transported, find a better adaptation to unforeseen events, control the interactions and reduce the uncertainty of information which leads subsequently to optimize different functionalities of the supply chain in particular and business in general.

## 7. CONCLUSION AND FUTURE WORKS

In order to improve decision making in the supply chain based on data analysis, we presented in this paper an algorithm designed to eliminate duplicated data generated by RFID readers from the supply chain activities. We presented also the model of trajectory data warehouse that supports trajectories data related to mobile objects.

We have not considered in our analysis the case of several readers that detect the same tag which that may be the aim of our future researches

## REFERENCES

- [1.] Y.Bai ,F. Wang, P.Liu .“Efficiently filtering RFID data streams”, VLDB Workshop on Clean Databases, 2006,pp.1-8.
- [2.] H.Mahdin, ,J.Abawajy.” An approach to filtering duplicate RFID data streams”, Communications in Computer and Information Science 124, 2010, pp.125–133.
- [3.] H.Mahdin, J.Abawajy. “ An approach to filtering RFID data streams”. In Proceedings of the 10th International Symposium on Pervasive Systems, Algorithms, and Networks, Kaoshiung, Taiwan, 14–16 December 2009; pp. 742-746
- [4.] H.Mahdin, J.Abawajy. “ An Approach for Removing Redundant Data from RFID Data Streams “, ISSN 1424-8220 [www.mdpi.com/journal/sensors](http://www.mdpi.com/journal/sensors) , 2011,pp.9863-9877
- [5.] X.Wang, Q. Zhang, Y. Jia. ”Efficiently filtering duplicates over distributed data streams”. In Proceedings of the International Conference on Computer Science and Software Engineering, Wuhan, Hubei, China, 12–14 December 2008; pp. 631-634.
- [6.] H.Shen, Y. Zhang. “Improved approximate detection of duplicates for data streams over sliding windows”. J. Comput. Sci. Technol. 2008, 23,pp. 973-987
- [7.] W. Jiang, W.Yongli, G. Zhang .”A Two-Layer Duplicate Filtering Approach for RFID Data Streams”.Communications in Computer and Information Science Springer ,Volume 312, 2012, pp 226-233
- [8.] C.H.Lee, C.W. Chung. “An Approximate Duplicate Elimination in RFID Data Streams”. Data & Knowledge Engineering 70(12), 2011, pp.1070–1087.
- [9.] G.Matičević, M. Čičak , T. Lovrić .” RFID and Supply Chain Management for Manufacturing Digital Enterprise”. Engineering “ Industrial Engineering and Management “ "Supply Chain Management - New Perspectives", book edited by SandaRenko, ISBN 978-953-307-633-1, Published: August 29, 2011 under CC BY-NC-SA 3.0 license.
- [10.]J. P.T. Mo, Xue Li. “Event Management of RFID Data Streams for Fast Moving Consumer Goods Supply Chains”, in Unique Radio Innovation for the 21st Century: Building Scalable and Global RFID Networks, Eds.DamithRanasinghe, Michael Sheng, SHERALI Zeadally, pub. Springer-Verlag. ISBN: 978-3-642-03461-9,2011, pp.89-110.

## AUTHOR



**Nahla Mezzia** master student at the Higher Institute of Management of Tunis, will receive the Master degree in Business Intelligence from the Higher institute of Management of Tunis in June 2013 and has received the fundamental license of computing science of management in 2010 from the Higher Institute of Management.



**Jalel Akaichi** received his PhD in Computer Science from the University of Sciences and Technologies of Lille (France) and then his Habilitation degree from the University of Tunis (Tunisia) where he is currently an Associate Professor in the Computer Science Department. Jalel Akaichi has published in international journals and conferences, and has served on the program committees of several international conferences and journals. He is currently the Chair of the Master Science in Business Intelligence. Jalel Akaichi visited and taught in many institutions such as the State University of New York, Worcester Polytechnic Institute, INSA-Lyon, University of Blaise Pascal, University of Lille 1, etc.