MAPPING MOVING OBJECT EVENTS INTO A NETWORK OF OBJECT FLOWS TO SUPPORT DECISIONS

Anastasia Griva  
*Athens University of Economics and Business, an.griva@aueb.gr*

Cleopatra Bardaki  
*Athens University of Economics and Business, cleobar@aueb.gr*

Katerina Pramatari  
*Athens University of Economics & Business, k.pramatari@aueb.gr*

Doukidis Georgios  
*Athens University of Economics and Business, gjd@aueb.gr*

Follow this and additional works at: [http://aisel.aisnet.org/ecis2016_rp](http://aisel.aisnet.org/ecis2016_rp)

Recommended Citation  
Griva, Anastasia; Bardaki, Cleopatra; Pramatari, Katerina; and Georgios, Doukidis, "MAPPING MOVING OBJECT EVENTS INTO A NETWORK OF OBJECT FLOWS TO SUPPORT DECISIONS" (2016). *Research Papers*. 112.  
[http://aisel.aisnet.org/ecis2016_rp/112](http://aisel.aisnet.org/ecis2016_rp/112)
MAPPING MOVING OBJECT EVENTS INTO A NETWORK OF OBJECT FLOWS TO SUPPORT DECISIONS

Research

Anastasia, Griva, ELTRUN E-Business Research Center, Department of Management Science and Technology, Athens University of Economics and Business, Athens Greece, an.griva@aueb.gr

Cleopatra, Bardaki, ELTRUN E-Business Research Center, Department of Management Science and Technology, Athens University of Economics and Business, Athens Greece, cleobar@aueb.gr

Katerina, Pramatari, ELTRUN E-Business Research Center, Department of Management Science and Technology, Athens University of Economics and Business, Athens Greece, k.pramatari@aueb.gr

Georgios, Doukidis, ELTRUN E-Business Research Center, Department of Management Science and Technology, Athens University of Economics and Business, Athens Greece, gid@aueb.gr

Abstract

Recent technological developments have facilitated the continuous identification and tracking of individual objects/things moving in space. Business analytics tools can handle the resulting vast amount of object tracking data. Thus, tracking technologies could be viewed as information facilitators that can directly improve decision-making. This research suggests a data-driven approach that transforms the simple object movement events captured by tracking devices in a monitored area into objects’ flows composing a network. In addition, we devise two new metrics, the volume and the mobility of the moving objects per flow to characterize the objects movement patterns. The proposed approach offers a structured way to transform massive tracking data into valuable, new knowledge of the moving behavior of objects that can support a wealth of business decisions. We demonstrate the utility of the proposed approach with real Radio Frequency Identification (RFID) data representing garments’ movements in a retail store of a fashion retailer.

Keywords: Object tracking, network flows, business analytics, decision-making
1 Introduction

The popularity of object tracking devices, such as Global Positioning System (GPS) receivers, RFID and Bluetooth Low Energy (BLE), have facilitated the identification of individual objects moving into space. Thus, a vast amount of data depicting the spatiotemporal events generated by objects while they are moving, have been gathered and remain unutilized. At the same time, business analytic techniques have been developed to connect large datasets and enable broader and deeper analysis than previously possible (Phan and Vogel, 2010; Provost and Fawcett, 2013). Therefore, there is a growing enthusiasm for the notion of Big Data and data-driven decision-making. Big Data analytics now drive near every aspect of our modern society, including retail industry, financial services etc. (Bertino et al., 2011; McKinsey Global Institute, 2011).

Thus, there is a growing research interest on how we can exploit different kinds of data and extract knowledge to support decision-making in different domains (Abbasi et al., 2012; Boone and Roehm, 2002). However, only few papers discuss how to analyze and model data captured by tracking devices to support decisions related to the context the data derived from. On the one hand, some studies focus solely on what decisions such data can facilitate and how (e.g. Chatziantoniou et al., 2011; Lee et al., 2011; Ngai et al., 2007; Oztekin et al. 2010). On the other hand, some studies are limited to the modeling part of the tracking data, overlooking the potential impact on decision making (e.g. Gonzalez et al., 2010; Kang and Yong 2010; Zhong et al. 2015). Further, other relevant studies focus on privacy, technical, and data quality issues of tracking systems (Bardaki et al., 2011; Lee and Kwon, 2015; Xu et al., 2010). Hence, in the literature, the modeling of object tracking data and decision-making are isolated and examined separately. Therefore, there is a lack of research examining how we can structure and present the tracking data in such ways that reveal the movement behavior of the objects in the tracked area, with a view to facilitate decision-making.

This research utilizes advanced data management tools for handling immense volume of data to propose a data-driven approach that transforms spatiotemporal object events recorded by tracking devices into a set of object flows composing a network. These object flows reflect actual tracked object movements between locations according to the business processes of the monitored environment. We devise two measures/characteristics of object flows: the volume and the mobility that reveal the quantity and the motion frequency of objects per identified flow. These two metrics combined with the network of object flows provide valuable knowledge that can support a wealth of business decisions. We demonstrate the utility of our approach analysing real RFID data representing captured garments’ movements in a RFID-enhanced fashion retail store.

The remainder of the paper is organized as follows. Section 2 analyses the research area, surveys the literature, and points out the research gap. The proposed approach and its evaluation are given in sections 3 and 4, respectively. Finally, section 5 summarizes the main outcomes of the paper, presents the theoretical contribution and the practical implications of our approach; and highlights further research.

2 Background

The advent of the location based technologies and the explosive growth of positioning equipment, has facilitated users to track and observe several events with the marks of time and locations, named as spatiotemporal events, which are generated by different “things” while moving in space, termed as moving objects (Civili et al., 2005; Jin et al., 2013). These moving objects could be shoppers moving in a store, visitors moving in a museum, medical patients in hospitals, products moving in a fashion store, delivery vehicles and public transit buses, and so on (Chang et al., 2014). At the same time, due to this popularization of the tracking equipment, such as GPS receivers, RFID and BLE tracking devices, a vast amount of moving object data originating in supply chain operations, road network monitoring, geo-positioning, and other RFID applications, has been gathered (Civili et al., 2005; Giannotti et al., 2007; Gonzalez et al., 2010). These data gathered by sensor-based devices are opening up exciting new steams of innovative applications (Chen and Storey, 2012). Thus, the major challenges are:
(A) How we can exploit the common characteristics the spatiotemporal events have, and model them in an organized and generalized way, (B) How the representation of these events into a model could be utilized to extract knowledge, (C) What knowledge we can extract; and (D) how we can use this knowledge to support decision-making.

In the meantime, business analytic techniques have been developed that can connect large datasets to enable broader and deeper analysis than previously possible (Phan and Vogel, 2010; Provost and Fawcett, 2013), and there is growing enthusiasm for the notion of Big Data. It is widely acknowledged that we have entered in the era of Big Data (Chang et al., 2014; Chen and Storey, 2012). Big Data research looks at how to analyze data in different domains in a way that generates deeper knowledge and adds value to the decision-making process in businesses (Sharda et al., 2013). Big data analytics now drive nearly every aspect of our modern society, including retail industry, financial services etc., and affect global economy; these unprecedented amounts of data have an impact on the way we work, we live, and we run businesses (Bertino et al., 2011; Breur, 2015; Constantiou and Kallinikos, 2014; Goes, 2014; McKinsey Global Institute, 2011). Since, big data analytics are now everywhere, the key to success for any firm is to support data-driven decision-making. Businesses that leveraging their data and facilitate decisions based on business analytics, have a higher performance. Nowadays, it is difficult to find a successful business that does not utilize business analytics (Chaudhuri et al., 2011). As Chang et al. (2014) said, “The key to competitive advantage is to accelerate decision-making by providing managers with guidelines for the application of analytics in their business processes”. There are plenty of papers exploiting data-driven approaches to support decision-making in different domains, such as retailing, health, finance etc. For instance, Abbasi et al. (2012) utilize business analytics to detect financial frauds, Boone and Roehm (2002) use neural networks to identify customer segments and support decisions in the retail context.

Major companies, like Metro and Wal-Mart, have recognized the need of data-driven decision-making; thus they took advantage of tracking technologies and mostly RFID to support decisions in the supply chain, such as out of stock prevention (Xu et al., 2010). However, in the literature there is limited research looking into ways of exploiting tracking data of moving objects and modelling them in a structured way that assists decision-making. Since RFID is the most well-known tracking technology, the majority of studies focus on how we can utilize RFID data to support specific decisions. For example, Guo et al. (2015) propose an RFID-based intelligent decision support system (DSS) architecture to handle production monitoring and scheduling in a distributed manufacturing environment. Chatziantoniou et al. (2011) present a DSS that integrates RFID streams with product information to support real-time supply chain decisions. Lee et al. (2011) develop a logistics workflow system to achieve effective demand management. In the same context, Larson et al. (2005) identify shoppers paths generated by RFID-enabled shopping charts. Apart from the retail domain, Ngai et al. (2007) develop an RFID-based system to facilitate the traceability of specific items in an aircraft company. Furthermore, Oztetik et al. (2010) provide a DSS to maximize the sensor field coverage via an optimum placement of a fixed number of RFID readers in a medical center. Beyond the RFID technologies, Civlis et al. (2005) utilize GPS data corresponding to real road network to identify the moving patterns, in the form of route, with an aim to predict each moving object’s location.

Although, there are papers that use tracking technologies to facilitate decisions in the literature, there is a lack of papers, which show how we can exploit the common characteristics of moving object data in order to model them with a view to support decision-making. Concerning the modeling issues, there is a group of papers that utilize GPS data and model the objects trajectories into sequences/vectors in order to mine the spatiotemporal patterns (Giannotti et al., 2007; Kang and Yong 2010). Nonetheless, they purely focus on the algorithmic part in order to extract these common patterns, setting aside how this modeling could be utilized for decision-making purposes. The most relevant work is that of Gonzalez et al. (2010) that proposes a graph-based object movement cube, and an algorithm that performs simultaneous aggregation of both spatiotemporal and item dimensions on a movement graph. However, apart from the modeling of the RFID events into graph-based cubes, this paper focuses mainly on technical issues and overlooks the decision-making part. Furthermore, another relevant
work is that of Zhong et al. (2015), where the authors propose a big data approach to discover frequent trajectory patterns from RFID-enabled manufacturing data, and they model these data into RFID-cuboids. Nevertheless, regarding the modeling part, they solely focus on the database modeling and structure of the RFID data. Moreover, once again, the decision-making part is absent. Apart from the aforementioned studies that discuss solely the decision-making or the modeling part of the objects tracking data, other papers focus on privacy, architectural design, and data quality issues derived from the tracking systems. Indicatively, Xu et al., (2010) and Lee and Kwon, (2015) investigate privacy issues derived from the location-based services and the tracking technologies, Bardaki et al., (2011) evaluate the information completeness of object tracking systems.

Nowadays, the increasing pervasiveness of the location-based technologies with the vast amount of data they collect necessitates novel data-driven business models and decision-making mechanisms to support this new environment (Zhou et al., 2015). In the literature, few scholars discuss how we can model the moving object data with a view to facilitate decisions. Furthermore, the issues of modeling of object tracking data and decision-making are isolated and examined separately. Hence, taking advantage of big data capabilities, we propose an approach, which shows how we can handle and model this kind of data to support decisions. To the best of our knowledge, there is no other approach that examines the modeling of object tracking data and the decision-making part as an inseparable entity. We provide specific steps and guidelines on how to exploit the vast amount of the object moving data generated by tracking devices, and model them into networks to mine knowledge and facilitate decision-making. Additionally, we introduce two new terms, flows volume and mobility, to give extra info on the moving behaviour of objects.

3 Mapping moving objects into object network flows: Approach

Following the "Design Science" research paradigm (Hevner et al., 2004), according to which, research achieves knowledge and interprets a problem via developing and evaluating artifacts; we suggest an artifact which is a data-driven approach, and then we evaluate it by a real case. This artifact transforms the object movement events captured by object tracking devices, into objects’ flows composing a network. It consists of three main phases. The first one concerns the data understanding and identification of model’s components, the second the creation of the object flow network and the alternative network views; and the last one the decision-making. Figure 1 illustrates the 3-phase proposed approach. The volume of the moving objects and their mobility characterize each flow. These two characteristics combined with the identified flow patterns of the objects can be enriched and viewed under different analysis lens to facilitate a wealth of decisions. The following sections detail our approach and present the way we can model the common data recorded from tracked moving objects in order to get new views of the object movements that can support decision-making.

**Figure 1. Proposed approach**
3.1 Data understanding and modeling

First, we need to understand the available data, as well as the sensor-enabled place and the related processes they were generated in. The original dataset comprises of simple spatiotemporal events reflecting objects that move in an area monitored by tracking equipment. Each event depicts the object’s location (event’s spatial dimension) at a specific time moment (event’s temporal dimension), while it was recorded by the tracking infrastructure. For example, events of RFID-tagged products moving through an RFID-gate in a retail store, or visitors in a museum identified by their mobiles and whose movements are captured by beacons. Each row in the available dataset should absolutely reveal where and when the event happened; and, certainly, which object was involved. In other words, we need the exact location of the object at the specific date and time of the event, and the unique identifier of the object. Optionally, the data may include the tracking device that captured each event, the process during which the event took place, e.g. inventory replenishment process in a retail store; and more characteristics of the examined object, such as the expired date of a food product (Table 1).

<table>
<thead>
<tr>
<th>Object’s unique identification</th>
<th>Tracking device identification</th>
<th>Place of tracking device</th>
<th>Related process</th>
<th>Date-Time</th>
<th>Additional attributes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1234</td>
<td>Reader 1</td>
<td>Location X</td>
<td>ABC</td>
<td>04-09-15 12:21:00</td>
<td>...</td>
</tr>
</tbody>
</table>

Table 1. Row of Dataset example

This kind of datasets have a noisy nature, thus we have to de-noise them by eliminating duplicate values, data inconsistencies, and corrupted values (Keller et al., 2014). Hence, we select the appropriate data suitable for the forthcoming analysis and clean/ de-noise them. Then, we propose to model the recorded object movement events as object flows composing a network. To transform the data into a network of object flows, we need to know and understand more on the movements of objects in the monitored area. For example, which are the common routes of objects, who/ what process initiates each object’s movement, where the tracking devices are installed, what kind of objects we deal with, at which locations the objects stand or just pass through etc.

The role of the modeling phase is to identify: (a) the objects (e.g. products, shopping carts etc.) we are interested in, (b) the actors, namely the actual triggers of the objects’ movements, for example staff or shoppers moving into a store holding RFID-tagged items, or carrying a smartphone that broadcasts Bluetooth signals; and (c) the tracking devices, as well as the locations (referred to as capture points) they are installed. These are the structural entities of the network of object flows. A mock-up of the monitored area could be useful to place the model’s components in it (Left figure of Figure 2).

3.2 Network creation and viewing

We continue with the formulation of the network of objects’ flows. First, the pre-identified locations equipped with tracking devices (i.e. capture points) are the nodes of the network. Then, according to the available cleaned data and the relevant processes, we identify the objects’ movements between
locations; and we map them as objects flows between the network’s nodes. Essentially, each object flow between network nodes reflects spatiotemporal events captured by the tracking device on the respective locations/ nodes. For example, the nodes of objects flow network at the right of Figure 2 and the links/flows (Left of Figure 2) between nodes reflect captured events of objects movements between the capture points. We devise two measures that characterize each objects flow. The measures of the network flows are volume (the number at the right of each flow) and mobility (the number at the left of each flow, calculated in hours).

- **Flow volume**: it represents the number of objects that passed through a network node at the same time moment and are directed to another node. For example, in Figure 2 we can figure out that 5000 objects passed from the “capture point 2” to the “capture point 3”, and so on.
- **Flow mobility**: it measures the time an object spends at a network node before it moves to another (residence time), e.g. how much time an RFID-tagged blood product stays in a hospital’s blood bank prior its usage. For example, in Figure 2 “capture point 2” is the location that objects spend the most time, as they stayed on average 3 hours before they moved to capture point 3.

The two proposed flow characteristics carry new knowledge valuable for decision-making. Essentially, using the two metrics we can generate meta-data that describe the original tracking data and, respectively, the generated object flows. We depict these meta-data in an object flow network as a dashed meta-flow. In essence, the meta-flows are flows, which do not depict the actual/sequential flows of objects, but meta-data derived from the sequential flows. Assuming that the right part of Figure 2 represents a museum’s monitored area, and visitors’ movements are tracked, the dashed flow (meta-flow) from museum’s entrance (capture point 1) to exit (capture point X) reveals 4.300 visitors with average spending time in it 4 hours. Thus, an objects flow network may include several meta-flows, according to the knowledge we desire to extract. For example, the average time visitors spend from the exhibit in “capture point 2” until the exit of the museum (“capture point X”).

The transformation of events into object flows between network nodes also provides unprecedented inside knowledge of the actual flows of objects between monitored locations. According to the standard processes in the monitored area, we have identified the expected, standard flows of objects between locations. For example, in a monitored fashion store, a standard flow of a garment begins from the backroom, passes the backroom’s exit to sales floor, and then into the fitting room. However, exceptional movements may happen in a store, which result to exceptional tracked events and, consequently, exceptional flows. For instance, an RFID-tagged garment has been tracked for the first time in the backroom and the next captured event comes from the fitting room. This could mean that there are missing intermediate reads in the way of the garment from the backroom to the sales-floor, perhaps due to RFID-reader inconsistencies. Additionally, in the network we may have some self-join flows. The self-join flows (node named “capture point 4” of Figure 2) are more than one contiguous in time events captured by the same reader, for example, a shopper who holds a smartphone that transmits Bluetooth signals may be captured by a beacon more than one times when standing in front of a shelf. These flows are also divided into standard and exceptional. In general, the exceptional flows can affect the validity of the network and consequently the validity of the derived knowledge that will be used to support decisions. For this reason, the relevant data should be eliminated from the dataset, and reconsider the network creation phase.

**Network views** are different network structures caused by the selection of different measures, as described above, nodes e.g. process, and analysis lens e.g. dimensions, with a purpose to support different decisions. Apart from the representation of the locations into nodes, the network could be enriched and formed in different views, by adding the relevant processes e.g. business process. Thus, we could notice how the objects are moving and react passing through the different relevant processes (middle of Figure 3). Additionally, more complex network views are created by depicting a concatenation of the above two (left of Figure 3). The last one could happen in case a tracking device serves more than one processes, for example a tracking device that captures both products’ store floor replenishment made by clerks, and shoppers’ movement. Alternatively, in cases the same process is taking place in
more than one location, for example the inventory count process that happens in the sales-floor and in the backroom by the same handheld tracking device. Different nodes could be chosen according to the case, and to the business value of the derived network. The dimensions represent the network structural attributes we want to examine and analyze from a decision-making perspective. Different analysis lens, are served by different dimensions, and construct alternative views of the network, with a prospect to support business decisions. For instance, we could have product dimension (category, SKU), and/or the time dimension with the whole hierarchy it could contain (month, days, hours etc.); the last one may generate different views of the network according to the selected time window, and in different granularity levels based on the selected hierarchy.

**Figure 3.** Network enrichment. Left: Simple network view, using locations as network’s nodes. Middle: using relevant processes as nodes. Right: left’s network enrichment by adding the relevant processes of the middle figure (complex view).

### 3.3 Decision-making

This phase points out how the mapping of the moving object events into object network flows could be used to extract valuable knowledge with a view to support a wealth of decisions, and what these decisions could be. We provide some examples to show how a decision maker could elicit knowledge utilizing the network with its flow metrics (mobility and volume). In general, the network, at its simplest view could offer full visibility of the objects. More specifically, from the network at Figure 2 apart from the identification of possible objects’ flows, we can observe that “capture point 2” is that, in which objects spend most of the time; thus if the network depicts the movements in a retail store, this is the point that a promotion action should take place. In the same spirit, exploiting the network that depicts the flows volume, we could identify the most voluminous flows (i.e. the flows that most of the objects pass), or the less frequently visited flows; hence we could use this view of the network as a heat map. For example, if the nodes are capture points posed in supermarket aisles, we could use the generated knowledge to track the number of shoppers passing from each capture point, and change the products displayed in the aisles. In the same context, we could also use the flow mobility, and extract the residence time of the shoppers in an aisle. Alternatively, by categorizing the flows according to their mobility into high, medium and slow-moving could help managers support many decisions, varying for instance from replenishment strategies, promotions and dynamic pricing, to category relocation in the retail context. Additionally, via exploiting both flows mobility and volume we could identify delays in the network, and normalize slow-moving flows, or even the voluminous ones. For example, if the objects are cars and the nodes are specific roads, we could normalize the flow of cars in the road networks. Moreover, we could predict objects flows, the next node an object will visit, and the residence time at each node. For instance, we could predict the time a product lying in a specific aisle, will be purchased, this way a manager could eliminate out of stock situations, or we could predict the flow path a customer will follow in a supermarket. As well, via identifying the wrong network flows and detect the missing reads, we could diagnose readers’ inconsistencies, and improve the tracking system; thus we can also improve the quality of the derived data.
4  Case: RFID-captured moving objects

Here, we testify the approach’s credibility and effectiveness by applying it in a RFID-enabled retail fashion store.

4.1  Data understanding and modeling

Our dataset involves captured events of RFID-tagged garments moving in a retail fashion store during a 6-month period, from May to November 2014. More specifically, garments’ attached RFID tags generate these movements; because of the in-store business processes and RFID-enabled processes performed either by clerks, or by shoppers, such as replenishments, purchasing etc. The dataset included the exact places/locations of the objects/garments, the RFID processes that are connected with these places, and the RFID readers that are located in each place (Table 2). For example, we knew that a garment, which is being uniquely identified via an electronic product code (EPC), has just been received in a specific date time, and this receiving process has been captured by an RFID cage, which is located in the backroom. Then this product has been transferred from backroom to store floor and this event captured via the RFID-enabled replenishment gate, and so on. Moreover, apart from the location dimension, garment dimension (product category, stock keeping units’ (SKUs), garment name, color, size) has been also selected as an input for our analysis. After the data cleansing operation, the 99.8% of the initial dataset (307.243 records) had been chosen to be used for the forthcoming analysis.

<table>
<thead>
<tr>
<th>EPC</th>
<th>Location</th>
<th>Process</th>
<th>Readpoint</th>
<th>Time Stamp</th>
<th>SKU</th>
<th>Category</th>
<th>Product Name</th>
<th>Size</th>
<th>Color</th>
</tr>
</thead>
<tbody>
<tr>
<td>urn:epc:tag:gid-96:31161137:20037:100004</td>
<td>storefloor</td>
<td>Replenishment</td>
<td>replenishment gate</td>
<td>2014-09-17 16:26:09.000</td>
<td>31161137020037</td>
<td>shirt</td>
<td>MOGOL</td>
<td>48</td>
<td>3</td>
</tr>
<tr>
<td>urn:epc:tag:gid-96:31161137:20037:100004</td>
<td>storefloor</td>
<td>Inventory</td>
<td>handheld</td>
<td>2014-10-06 12:01:59.000</td>
<td>31161137020037</td>
<td>shirt</td>
<td>MOGOL</td>
<td>48</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 2  Sample of selected RFID data

Model’s structural Entities are: (A) store locations and (B) RFID readers. Moreover, we will enrich the model and thus, the network by adding also the relevant (C) RFID processes. Concerning the store locations, we have the backroom, and its sub-locations, which are the backroom entrance, from where the garments enter the store, and the stockroom where the backroom inventory is kept. There is also the replenishment gate, which is the intermediate location between backroom and store floor. Regarding the store floor, it also has two other sub-locations, the first one is the store aisles in which the garments are displayed, and the second is checkout desk, where the garments are moved when they are purchased. Each store location has RFID readers, these readers are the tracking devices; we have two kinds of readers, the stationary and the non-stationary. The non-stationary readers (or else handhelds) could be used either in backroom or in store floor during the inventory process. This process may happen at any time by clerks either to count the amount of the garments that are stored in the stockroom, or to enumerate the garments that remain unsold in the store aisles. Furthermore, we have three other RFID processes; each process is connected with a reader. Hence, we have the receiving process, which takes place when the garments enter the store and are captured in the RFID cage. The replenishment process that happens when the garments are moved from the backroom to the store floor via the replenishment gate and vice versa. In addition, there is the checkout process during which the checkout reader recognizes the garments, and thus they are sold. In Figure 4 there is a list of the model’s entities, in which we have also add the relevant processes, and in Figure 5 are illustrated these entities in the store layout. At the left side, the locations and readers are depicted, and at the right, there is the complex form of the previous one, in which the related processes have been also added. As it is legit, model’s moving objects are the garments; model’s actors are clerks that perform all the RFID processes, as described above, and shoppers that are associated only with the checkout process.
A. Store locations:
1. Backroom
   i. Backroom Entrance \(\ni\) Backroom
   ii. Stockroom \(\ni\) Backroom
2. Replenishment Gate
3. Store floor
   i. Checkout Desk \(\ni\) Store Floor
   ii. Store Aisles \(\ni\) Store Floor

B. RFID Readers:
1. Stationary:
   i. RFID Cage (is placed in location A1i)
   ii. Replenishment Gate (A2)
   iii. Checkout reader (A3i)
2. Non-stationary:
   i. Handhelds: (A1ii, A3ii)

C. RFID Processes:
1. Receiving (B1i)
2. Replenishment (B1ii)
3. Inventory (B2i)
4. Checkout (B1iii)

Figure 4. List of Model's Structural Entities and of relevant processes

Figure 5. Model entities in the store layout. Left: placing tracking devices and capture points in store layout. Right: left’s model’s enrichment by adding the relevant processes.

4.2 Network creation and viewing

According to the model’s entities, we identified the RFID-captured garments movements, and we created the garments’ flow network. The different views of this network assist us to transform the simple RFID events into flows and deliver them in a visual manner. At Figure 6 it is illustrated the network at its simplest view. In this view, the nodes are the different locations that are shown at the left of Figure 5. Moreover, this network depicts flows mobility for all the garments, in an aggregated view, throughout the available period of time (6 months). In addition, two meta-flows (backroom entrance \(\rightarrow\) checkout desk, and backroom entrance \(\rightarrow\) store aisles) are also shown. Hence, from Figure 6, we could figure out that it takes on average less than half an hour (0.01 days) for the garments to be transferred from the backroom entrance to the stockroom. Moreover, garments stay at about 34 days in the stockroom before they pass the replenishment gate and move to the store aisles, then they spend 15.98 days in the store aisles before they been transferred back in the stockroom. Furthermore, it takes 17.37 days on average for an object to be sold when it has been already moved to the stores aisles. Also, when a garment is delivered into the store it stays 43.12 days in the backroom, before it is displayed for the first time at the store aisles. This number is different from that at the flow stockroom \(\rightarrow\) replenishment gate \(\rightarrow\) store aisles, as a garment may follow this flow more than one times. Thus, we can calculate that garments stayed on average two months (43.12 days: backroom entrance \(\rightarrow\) store aisles, and 17.37 days: store aisles \(\rightarrow\) checkout desk) in the whole store area until they have been sold.

At this point we have to mention that network’s dimensions could be store, time (month, week, days, hours etc.), and garments (garment’s category, name, SKU, color, size). Moreover, in the given examples all the available 6 months (time dimension) dataset is utilized, and the networks represent the movements of all the available garments in this time window. Now, we enrich the network to point out the flows volume, via utilizing the more complex network view (right figure of Figure 5). Thus, Figure 7 illustrates garments’ standard flows.
As described before, each network node represents a concatenation of store’s location and of the process that takes place in this location; the number in each node depicts the garments passed from this node. The flows from the start node to another one represent the first time the garments have been captured. The flows from a node to the end node represent that this node was the last known node of these objects, so after this event we cannot detect them in any other node during the selected time window. The numbers attached in each flow depict flows volume (i.e. the garments passed from this flow). For instance, at Figure 7 we can notice that the majority of the garments (70,211) had been captured for the first time in the backroom entrance during the receiving process via the RFID cage. At about 35% of them (24,125) stayed at the backroom, and had never been displayed at the store floor. We have only eleven captures in the backroom during the inventory count process, via the handheld, as during the selected time window the backroom inventory process was in a trial period. At about 42,852 garments passed from the backroom to the store floor through the replenishment gate, and 27,720 garments had been sold. In the same context as above, Figure 8 depicts all the remaining exceptional RFID flows. For example, this network depicts that 1,755 object instances have been captured in the receiving, and
the next captured event happens in the checkout node, hence, there are missing reads in the replenishment gate. As shown in Figures 7 and 8, there are some self-join flows in the network, that are also divided into standard and exceptional. The exceptional self-join flows are events, which happen in greater contiguous timestamps, and lead to missing reads/events. For instance, a garment that is captured two times when passing from backroom to store sales-floor, with a time difference of two days, is not a right self-join read, since there are other missing events that intervene. Hence, we have to identify the adequate threshold between the time these events happened, such that $\Delta T (e2-e1) \leq$ threshold. According to retailer’s guidance the threshold that distinguishes the standard and the exceptional self-joins is thirty seconds. Thus, the standard self-joins are depicted in Figure 7, whereas the exceptional self-joins are illustrated in Figure 8. The exceptional flows (Figure 8), revealed us data inconsistencies, caused by RFID readers’ inconsistencies, hence we eliminated and cleaned these flows from the dataset.

![Figure 8. Garments' exceptional flows volume → Missing captured events/reads](image)

### 4.3 Decision-making

In this section, we will show how the network flows could be transformed into knowledge in order to support business decisions. The business value of the proposed approach will be presented via some examples. More specifically, via enriching Figure 6 with flows mobility, we could generate a report as shown in Figure 9. Moreover, this report could be more meaningful and could be also generated per product category, or per color, size and SKU. Such a visualization of both mobility and volume of network flows could offer to a manager full garments visibility, could be also utilized as input to predict the residence time at each location, and detect any in-store delays.

![Figure 9. Transformation of network flows to support decisions](image)
Further, by isolating specific nodes of the network, for example the nodes “store aisles” and “checkout desk”, and also by using the hierarchy product categories of the product dimension, and utilizing the mobility flows, we could extract knowledge about each garment category. Then, we could use this knowledge to classify the categories into fast, medium, slow, and no moving, according to the average time spent in the store floor until they are sold. According to five-number summary (box plot) the garments that stayed from 4.33 to 10.88 days in the store floor before they have been sold are considered as fast-moving (green histogram bars of Figure 10). Those which stayed from 10.88 days to 23.44 days are medium-moving (orange bars), and slow-moving (red bars) are those that stayed in the store aisles till 58.08 days, before they have been sold. Then utilizing the box plot we generated Figure 10 that depicts the product categories in classes according to how fast they are purchased by shoppers.

![Figure 10. Classification of garments categories into slow, medium and fast-moving according to flows mobility](image)

The above could be helpful for the store manager to determine the in store space allotted per each category. Moreover, in the cases that there is not enough space in the store aisles, he could move back in the backroom the slow-moving categories or even those that are unsold for a long time and are labeled as “no-moving”; in our case these categories are “leather skirt” and “short trousers”. In addition, the above knowledge could be also utilized to design dynamic pricing strategies and promotions for instance to get rid of the garments that no or slow-moving. Furthermore, this could be also useful for the characteristics of the product dimension, such as colors, to examine for example, the items per color that should be displayed in the store floor area. At this point, we have to mention that, the main difference when utilizing RFID instead of point-of-sales (POS) data is that the last of them could be used solely to characterize garments in terms of sales. Thus, we could classify for instance, the garments into high, low selling. However, via utilizing RFID data we can also take into consideration the time (flows mobility) they stayed in the different store locations, and how fast they purchased, hence we could classify them into slow or fast moving etc., as described above. Furthermore, via utilizing the flows volume, in the SKU level we could provide to the store’s manager a daily replenishment report. This report could exploit the flow volume of the checkout node to identify the high selling SKUs. Then, using the volumes of the other flows it could generate a list with: (A) the garments that are not present on the store aisles but are available in the stockroom (red color in Table 3). (B) The garments that should be refilled, as there are few pieces remaining on the store aisles (orange color), and (C) the garments that are not available in the whole store area (grey color). In the same spirit, using solely POS data the knowledge of Table 3 is not possible to be extracted, as this way we could only be aware of the column “times purchased”, omitting the remaining pieces in both stockroom and store aisles.
Mapping Moving Objects into Network Flows

<table>
<thead>
<tr>
<th>SKU</th>
<th>Category</th>
<th>Product name</th>
<th>Size</th>
<th>Color</th>
<th>Times purchased</th>
<th>All remaining store pieces</th>
<th>Store aisles’ pieces</th>
</tr>
</thead>
<tbody>
<tr>
<td>31161137020037</td>
<td>shirt</td>
<td>MOGOL</td>
<td>48</td>
<td>3</td>
<td>10</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>65140643050601</td>
<td>bag</td>
<td>KUBO</td>
<td>36</td>
<td>60</td>
<td>3</td>
<td>11</td>
<td>1</td>
</tr>
<tr>
<td>80140440050042</td>
<td>cloak</td>
<td>NADAR</td>
<td>38</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>10861423060012</td>
<td>coat</td>
<td>SORTE</td>
<td>38</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 3. SKU daily replenishment list – based on the high sellers of a specific day

5 Conclusions and discussion

The advent of tracking technologies has resulted in a massive volume of data generated by tracking equipment capturing the location of moving objects in a continuous manner. It is important to study how such data, in combination with the tracked context that generates them, can reveal new unprecedented knowledge essential for decision makers. For example, we don’t provide to managers just massive datasets but we capacitiate them to see products moving in the facility, production times at different stages, errors in the production process etc. Namely, we propose a transformation of object tracking data to object flows formulating a network that depicts the moving behavior of objects in the tracked environment. We offer to the decision makers these movement patterns of objects along with the volume and the mobility of the objects on each flow.

To the best of our knowledge, there is not an approach that gives specific steps and guidelines on how to exploit the object moving data generated by tracking devices, and model them into networks of object flows that provide new knowledge and facilitate decision-making. Additionally, this is the first effort that introduces the terms flows volume and mobility to characterize the behaviour of moving objects. From a practical point of view, the proposed modeling and visualization of object tracking offers an effortless and rapid way to transform the complex events into knowledge and support decision-making. The network of object flows at its simplest view allows us to identify the most voluminous flows and the less frequently visited flows (flows volume), and the slowest and fastest moving flows (flows mobility); this knowledge is the keystone of decision support. More specifically, this visualization and mapping of the knowledge into networks could facilitate a wealth of decisions varying from rearrangement of the network nodes to prediction of the objects’ flow.

Moreover, the sequential representation of object movement events into network flows could be used as the base of a business analytics object tracking system. The alternative network views derived from the selection of measures (flows mobility and volume) and dimensions (e.g. product, time etc.), which is similar to the concept of OLAP (online analytical processing) cubes, could offer a new network-based technique of analyzing events generated by different objects/things. Thus, it could facilitate technological issues arising of the evolution of Internet of Things (IoT). This effortless representation of complex events could also support real time analytics, which is an emerging topic in the new environment of IoT commands. In addition, this mapping of the complex events into network flows could also inspire a decision support system (DSS). This system could exploit the different views of the network in different aggregation levels combined with the flows volume and mobility to facilitate decision-making.

Further research can examine the implementation of the aforementioned emerging issues from a technical point of view. Moreover, we intend to assess the approach’s effectiveness utilizing tracking data derived from different domains. For instance, data generated by the mobile devices of visitors moving between exhibits in a museum and captured by BLE tracking devices, such as beacons.
Acknowledgements

The authors acknowledge Sensor-Enabled Real-World Awareness for Management Information Systems (SERAMIS) project partners (http://seramis-project.eu) for providing the necessary data for this study.

References


