# BUSINESS INTELLIGENCE BASED ON A WI-FI REAL TIME POSITIONING ENGINE
## *A Practical Application in a Major Retail Company*

Vasco Vinhas, Pedro Mendes and Pedro Abreu
*FEUP - Faculdade de Engenharia da Universidade do Porto, Rua Dr. Roberto Frias s/n, Porto, Portugal*
*DEI - Departamento de Engenharia Informática, Rua Dr. Roberto Frias s/n, Porto, Portugal*
*LIACC - Laboratório de Inteligência Artificial e Ciência de Computadores, Rua do Campo Alegre 823, Porto, Portugal*

Keywords: Wi-Fi, Real-Time Tracking, Business Intelligence, Retail.

Abstract: Collecting relevant data to perform business intelligence on a real time basis has always been a crucial objective for managers responsible for economic activities on large spaces. Following this emergent need, the authors propose a platform to perform data gathering and analysis on the location of people and assets by automatic means. The developed system is retail business oriented and has a fairly distributed architecture. It couples the core elements of a real-time Wi-Fi based location system with a set of developed functional views so to better explicit the information that one can observe for each tracked entity, the undertaken path on the space, demographic concentration patterns. Tests were conducted on a real production environment as a partnership outcome with a major player in the retail sector and the obtained results were completely satisfactory having the managers confirmed the provided knowledge relevance.

# 1 INTRODUCTION

People and asset's tracking and positioning data collection – even if with the sole purpose of locating those entities – on a given space, has always been a need for those with management responsibilities, regardless their economic sector. Up until thirty years ago, the mechanisms to pursue such objective were manifestly out of reach of the masses mainly due to financial reasons. At that time, most systems were part of military projects and made use of radar technology to passively detect air or marine traffic (LaFollette et al., 1991), although in this last case the positioning was achieved through an active procedure – the traceable water vehicle had to intentionally transmit a radio signal with its latitude and longitude encoded (Yu, 2005).

However the recent advances in several domains such as computer technology, video capturing and positioning sensors changed the previously mentioned premises. Nowadays video surveillance systems, with mid-end camera resolution are cost-effective even in the context of small and medium enterprises – SMEs – with the non minor advantage of complete digital storage and processing. In spite of a video feed does not provide data on the positioning of entities, in the corresponding camera's scope, in way that such information can be retrieved automatically. With the latest wireless networks proliferation, other approaches emerged, with the obvious advantage of automatically providing clean data at a semantic level, displayable and storable in a straightforward way. Among the technologies that fit in this situation one shall refer to: Bluetooth; Wi-Fi; RFID; GPS; amongst others. Naturally, each of them encloses its strengths and flaws. Some require the traceable items to have an active behavior while others process them as passive items. The coverable area and the error involved are also important factors to be considered as well as the required resources, such as power and equipment density. Some of them remain unwavering with layout changes while others require a full system recalibration.

The research work proposed in this document materializes itself on a system that takes advantage of typical redundant Wi-Fi networks and is based on a positioning engine built on top of these. The system provides a visualization platform of such data on a real time basis. The information can be displayed through several perspectives, including fully scalable concentration grids, clean positioning



Vinhas V., Abreu P. and Mendes P. (2009).
BUSINESS INTELLIGENCE BASED ON A WI-FI REAL TIME POSITIONING ENGINE - A Practical Application in a Major Retail Company.
In *Proceedings of the 11th International Conference on Enterprise Information Systems - Databases and Information Systems Integration*, pages 11-16
DOI: 10.5220/0001858000110016

of the elements at hand or even a vision inference assuming that the items to be tracked are associated with people. This system was originally developed for a corporate entity operating in the global retail market, whose institutional designation is obfuscated by commercial reasons. Following the same line of reasoning all data provided; application's screenshots and shop floor layouts were based on real production data and were obfuscated, without compromising functionality and intelligibility.

# 2 STATE OF THE ART

Despite the advances in this field, leading to noteworthy breakthroughs, some issues still remain to be addressed. From these, one ought to point out those concerning occlusion, which are classifiable in three distinct categories: self-occlusion, where part of the object to trace, typically articulated ones, occludes another part; inter-object occlusion where two traceable objects occlude each other; and occlusion by the background scene where the physical space's properties propitiate a camouflage of the object to track. For the inter-object occlusion research works like (MacCormick and Blake, 2000) (Elgammal et al., 2000) exploit the a priori knowledge of the object's position and attempt to predict possible occlusions and solve them smoothly.

Considering other domain issues, the hurdles that arise when tracking entities with non-linear movements must be addressed as this point constitutes one of the major problems in tracking persons in non controllable environments. A pragmatic approach to this situation could point out to time resolution diminishment in which tracking is achieved or the loosening of the real time requirements. When the solution is based on video feed processing, there are some additional problems directly related to the inherent technology. The scene's illumination should be adequate so to facilitate the image binarization processes and the network shall be enhanced and optimized in order not to become a system's bottleneck.

## 2.1 Non-Image based Systems

GPS is commonly used to perform real time detection of different types of vehicles and as a base tool to analyze their motion (Yu, 2005) (Nejikovsky et al., 2005). Yet in this scenario, the technology is applicable in three distinct ways: Cellular Based Tracking is a solution based in a conventional mobile phone with a GPS receiver that emits the vehicles position every five minutes. Wireless Passive Tracking has core advantage in using GPS, because once it is set up, there is no monthly fee associated, and with it is possible to collect information like for instance, how many pit stops are made by a vehicle in a given route and how fast is it moving. Its worldwide coverable area constitutes an ideal solution for transporting companies.

The radio frequency identification (RFID) is a non-standardized wireless tag location method. This technology requires a RFID receiver and a set of tags which can be divided in two different groups: Passive - only detectable on a 13-meter radius from the receiver; Active Group - have their own internal power source, offer both reliable detection on a larger scale, and more resilience to occlusion problems caused for possible obstacles in the environment. The two major issues about this technology are the receiver's cost and the active tag's average unit price, mainly due to the need of an independent power supply (C. et al., 2007).

Wi-Fi 802.11 technology allows establishing connectivity between a set of devices allowing an easy setup of wireless data networks in academic campus, industrial facilities, public buildings, etc. The technology behind these networks can also be used for designing a tracking system. By reusing commonly existing data networks and a low level protocol it is possible to create a tracking system on top of this infrastructure. Another advantage of this technology is the possibility of tracking an object by using a single access point, though the precision will weaken due to the lack of signal triangulation. Because of its technical details, the impact of issues such as occlusion and signals loss is reduced to a residual level especially in environments, which do not have high concentration of metallic materials (Mingkhwan, 2006).

Bluetooth is a wireless protocol available on any modern mobile equipment, allowing data exchange between multiple devices. It is exclusively used for short-range communications, which is the cause for its poor applicability on tracking systems. The battery consumption is also remarkably high (Jappinen and J., 2007).

## 2.2 Image based Systems

Thermal signature systems are one of the most expensive technologies for locating an object on a scene. The main purpose of these solutions is the reconnaissance and processing of thermal images. These systems attempt to recognize specific thermal signatures of the entities being tracked although some items might not have them.

Multi-camera video surveillance is a technique that uses a set of cameras to track entities in a given environment. By accurately crossing the information coming from cameras which have intersecting frustums one can enhance the precision of the system, despite the possible processing overhead (Mittal and Davis, 2003). Camera calibration and its positioning on a tridimensional space may be performed/inputted manually (Collins et al., 2001) (Cai and Aggarwal, 1999) or even automatically despite the obvious errors that may occur if the last method is undertaken on an unsupervised way (Lee et al., 2000) (Khan and Shah, 2003). Although these systems are actually used in some scenarios, some issues still persist. The need to have a dedicated network for the system, the expenditure required for high-resolution equipment and the computational demands, are still major concerns. In the literature, some research work tries to optimize the performance of these systems by minimizing the need of brute force computation (Mittal and Davis, 2003) and by using overlapping camera views (Huang and Russel, 1997) (Javed et al., 2003) (Kettnaker and Zabih, 1999).

# 3 PROJECT DESCRITPION

In this section, the undertaken project is described in detail in what regards its several components and analysis perspectives.

## 3.1 Project Description

As revealed through Figure 1, the elaborated technical design contemplates several independent modules that communicate in order to achieve a systematic unit. The first action, in offline mode, consists in conducting a complete map creation\edition. The user shall specify, amongst other details, depicted in subsection 3.2, the image file representing the shop floor layout, the used scale and identify, by using a draw-like tool, what items are to be visible by visitors as well as spawn areas. This information is compiled in a XML file for both the position engine and real-time monitoring tool and submitted to the mentioned database for the historical BI application.

The Wi-Fi tag consists in a miniaturized active 802.11 a/b/g board with a couple of power batteries attached. These are configured to connect to a specific Wi-Fi network – security, DHCP and other network configurations are also possible – and to directly communicate with the position engine. By using this kind of wireless technology, it is possible to reuse partially or totally the client's network infrastructure, having only, for special requirement, a high density of access points as the accuracy naturally increases with this factor.

The used position engine periodically collects data from the tags and updates their position against a pre-loaded localization model. This model is very similar to the produced from the map editor differing only in the available information regarding visible objects. This model also requires a previous offline site survey for measuring Wi-Fi signal strength and for network items – routers and access points – precise localization. The engine is also web-enabled and supports a HTTP/XML API so that third-party applications can interact with it, therefore accessing localization and status information regarding each individual registered tag.

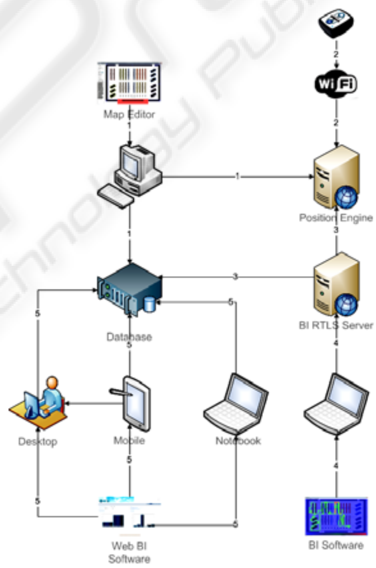


Figure 1: System's Global Architecture.

Using this communication protocol, the developed real-time monitoring server is responsible for gathering, at a specific periodicity – typically equals to the position engine frequency – every tag's valid location data. With this information, this module is directly responsible for updating the database and caching the session's data for the real-time monitoring application.

Having the continuous up-to-date database as a solid information reference, it was possible to enable both real-time and historical business intelligence applications. For real-time knowledge extraction, it was only used data referring to active sessions, for historical analysis, and delegating all the process

effort to the database engine, specific and dynamic time windows were used to filter data.
The versatility of such application must be referred as it congregates both web-enabled features and zero data process – it is all delegated to the database engine and allocated in a dedicated server – enabling its usage in a wide range of devices, including PDAs and mobile phones.

## 3.2 Real-Time Monitoring Application

This unit is responsible for accessing location data from the online position engine and, simultaneously, using a multithread sliding-window approach, commit new data to the database and compute data into visual information following distinct approaches. Each of these tool's facets is mapped into a distinct GUI tab enabling independent analysis.

Before BI extraction, there are two mandatory configuration requirements that must be met: the first consists in loading the shop floor layout; the second consists in establishing a HTTP connection with the position engine. The third, optional, requirement resides in opening a database connection for online data insertion. If this is not met, there are a virtual infinite number of application's instances that can be run at the same time, enabling simultaneous BI extraction for numerous organization's members.

### 3.2.1 Real-Time Tracking

This feature enables complete item tracking by overlapping current item's position directly over the loaded map information. This capability is independent of the GUI's windows size and/or shape as the coordinate systems are always synchronized.
As depicted is Figure 2, it is also possible to enable session's path history directly over shop floor image, therefore enhancing visual perception of both current positions and session routes.

In the presented screenshot, it is possible to see that at the time it was taken, there is only a single client in the shop, whose location is near the layout center and that his visit concentrated mainly in direct routes in the north corridors.

If there would be more clients present, it would be possible to perceive their current location by their blue dots representation as well as, optionally visualize their session routes.

This feature tries to emulate a bird's eye view of the all shop floor, with the possibility of recalling all the client's routes as if they left a visible trail while they are touring the facility.

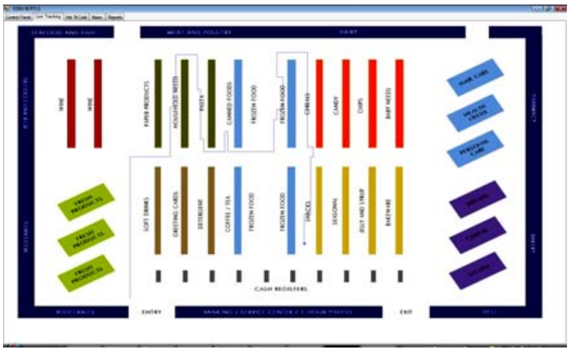

Figure 2: Real-Time Tracking with Visible Path.

### 3.2.2 Demographic Concentration

The demographic concentration feature enables space division using a completely dynamic, in terms of cell size, matrix grid that is colored according to demographic concentration at given time. Once again, this computation can be performed using only strictly real-time positions or by recalling all current session's route positions.

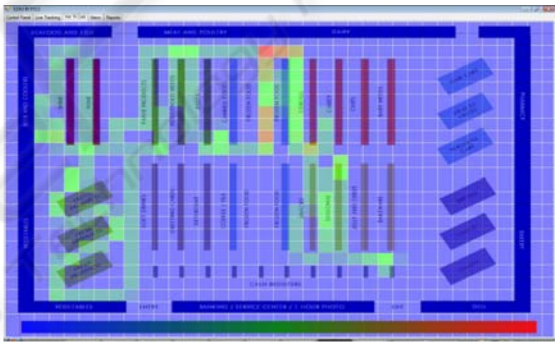

Figure 3: Real-Time Tracking with Visible Path.

Each cell is then colored following a three-color gradient where blue stands for lower levels of concentration, green for intermediate levels and red for high levels. A special remark is due to the fact that the entire colour spectrum is used for the mentioned purposes.
As illustrated in Figure 3, the concentration matrix is drawn, with a partially transparency, over the shop floor plant. In this practical example, the illustration corresponds to a situation where there were two clients present in the shop, and it was selected, for concentration calculus concerns, not only current position but all routes' positions.

This tool's feature enables swift, yet effective, hot and cold zones analysis, current bottlenecks, unvisited versus most visited areas and online queue alerts and management. Also, as reported in the previous subsection, all the graphical information is independent of the GUI's window size or shape.

### 3.3 Real-Time & Historical BI Application

In order to extract significant business intelligence knowledge based on historical data and not only real-time information, the authors decided to make an immediate use of the raw position data stored in the database – in conjunction with visible structures pre-processed information – so that significant retail BI facts could be easily extracted and presented to end-users – both shop managers and organization's top executives. Taking advantage of using Oracle as equally laboratory and production database, Oracle's Application Express was used to generate a web application responsible for processing data and, most important, aggregate information in an understandable way.

Apex's engine is directly embedded to the database, thus directly dealing with client's web requests. With this architecture, several systems can easily access BI application as all heavy processing is database server's responsibility, leaving the client with only chart rendering computation.

Despite extensive tool's analysis being object of discussion in the next section, by using the described architecture and technology, several practical measures were considered for extraction, namely: hot and cold zones, average visit duration and distance and number of visits. All these benchmarks can be targeted to use only real-time data or recall specific historical time windows.

## 4 RESULTS

One is able to state that the system performs swiftly as a whole as its constituting parts are able to exchange data harmoniously both in laboratory and production environments. As for the core tool, the real-time monitoring application, it was observed that the displayed data is quickly interpretable, thus making its purpose: enabling real time business intelligence. In fact, just by looking, for instance, to the grid tab, managers are able to see what the current hot and cold zones are and make decisions based on such information. According to the retail store managers, by using the application, one is able to, real time, assess the effectiveness of several decisions, namely: the store's layout, number of open cash registers or the work being carried out by employees in the zones where the customer service is made face-to-face.

Figure 4: Web Application – Zone Distribution.

Figure 4 represents the retail store grid concentration which has the layout as shown on Figure 2. The data comprehends one week of activity where each chart entry corresponds to the number of accounted presences, on a specific zone, every two seconds. The zones are numbered from one to the number of rows times the number of columns. Zone 1 is located on the top left corner.

By analyzing the exposed charts, some interesting aspects on the customer's behaviour when visiting the shop, can be assessed. Although zones 13 and 14 correspond to the store's entry, it is observable that very little time is spent by the customers. Another interesting point resulting from this analysis concerns the zones where more presences were accounted. All of them include central corridors, which are naturally part of the paths that clients have to cross to reach the products they are searching for.

## 5 CONCLUSIONS & FUTURE WORK

One ought to affirm that all the most important goals were fully accomplished. First, a fully functional item real-time location and tracking system were pursued. The Wi-Fi based solution, not only complied but did it reusing most of the client's network infrastructure, thus reducing negative impact in both financial and logistic terms. Secondly, the designed system's architecture proved to be reliable, efficient and flexible enough to contemplate vast and diverse application scenarios. These features are more visible in what concerns dynamic and user-friendly shop floor layout definition. Also within this scope the distributed communication architecture performed as predicted enabling computation across distinct machines, therefore improving overall performance and reliability. This feature also enabled simultaneous multi-terminal access, both to the real-time analysis tool and the historical statistical software.

Taking into consideration the project's tools, both were classified, by the retail company's end-users – mainly shop managers, marketing directors and board administrators – as extremely useful and allowed swift knowledge extraction, preventing them the excruciating, and not often useless process of getting through massive indirect location data. The immediate visual information provided by the system proved to be effective in direct applications such as queue management and hot and cold zones identification, and most significant, in what concerns to visit's pattern extraction across different time dimensions, thus enhancing marketing and logistic decisions' impact. One must refer to Oracle's APEX technology adoption. It has demonstrated to be able to allow multiple simultaneous accesses and, consequently, dramatically enhancing analysis empowerment, while, at the same time, eliminated heavy data computation from end-users terminals, concentrating it in controlled and expansible clusters.

Regarding future work areas, there has been identified a set of potential project enhancements that would be able to suppress some hurdles and, somehow, wide potential new applications. Considering business intelligence extraction, it would be useful to build or reuse an inference engine capable of determining the odds of a given customer turn right or left in the next decision point, taking for that, into account his past actions and comparing them to other customers' action that are classified in the same cluster. This aspect should be also applied to historical data so that efficient customer clusters would be defined and maintained.

There have been identified several application domains that go beyond the retail segment. Amongst these, one shall mention the possible system's adoption by large warehouse management where traffic jams are not unusual.

As a summary, it is fair to state that the project's initial ambitions were fully met and that the close cooperation with an important stakeholder in the global retail market was extremely important for better measuring the system's positive impact and potential firstly unseen applications. The technology transparency, allied with the future work areas, is believed to greatly improve potential applications in several domains, thus significantly widening the project's initial horizons.

## REFERENCES


C, C., J, Y., and W, J. *Determining technology trends and forecasts of RFID by a historical review and bibliometric analysis from 1991 to 2005*. In *Technovation27, Elvisier Ltd*, page 268-279, 2007.

Cai, Q. and Aggarwal, J. *Tracking human motion in structured environments using a distributed camera system*. *In IEEE Trans. Patt. Analy. Mach. Intell*.,pages 1241–1247, 1999.

Collins, R., Lipton, A., Fujiyoshi, H. and Kanade, T. *Algorithms for cooperative multisensory surveillance*. In *Proceedings of IEEE*, pages 1456–1477, 2001.

Elgammal, A., Duraiswami, R., Harwood, D. and Davis, L. *Background and foreground modeling using nonparametric kernel density estimation for visual surveillance*. *In Proceedings of IEEE*, pages 1151–1163,2000.

Huang, T. and Russel, S. *Object identification in a Bayesian context*. *In Proceedings of International Joint Conference on Artificial Intelligence*, pages 1276–1283, 1997.

Jappinen, P. and J, P. *Preference-aware ubiquitous advertisement screen*. *In e-commerce 2007 Proceedings*, IADIS Press, pages 99–105, 2007.

Javed, O., Rasheed, Z., Shafique, K. and Shah, M. *Tracking across multiple cameras with disjoint views*. *In IEEE International Conference on Computer Vision (ICCV)*, pages 952–957, 2003.

Kettnaker, V. and Zabih, R. *Bayesian multi-camera surveillance*. *In IEEE Conference on Computer Vision and Pattern Recognition*, pages 117–123, 1999.

Khan, S. and Shah, M. *Consistent labeling of tracked objects in multiple cameras with overlapping fields of view*. *In IEEE Trans. Patt. Analy. Mach.Intell*., pages 1355–1360, 2003.

Lee, L., Romano, R. and Stein, G. *Monitoring activities from multiple video strams:establishing a common coordinate frame*. *In IEEE Trans. Patt.Recogn. Mach.Intell*., pages 758–768, August 2000.

MacCormick, J. and Blake, A. *Probabilistic exclusion and partitioned sampling for multiple object tracking*. *In Int. J. Comput. Vision*, pages 57–71, 2000.

Mingkhwan, A. *Wi-fi tracker: An organization wi- fi tracking system*. *In IEEE CCECE/CCGEI*, Ottawa,page 1353-1356, 2006.

Mittal, A. and Davis, L. *M2 tracker: A multiview approach to segmenting and tracking people in a cluttered scene*. *In Int.J. Comput. Vision*, pages 189–203, 2003.

Nejikovsky, B., Kesler, K. and Stevens, J. *Real time monitoring of vehicle/track interaction*. In *Rail Transportation*, pages 25–31, 2005.

Yu, Z. *GPS train location and error analysis which based on the track fitting of the railway's geometric locus*. In *ICEMI 2005: Conference Proceedings of the Seventh International Conference on Electronic Measurement*.