Advances in Knowledge-Based and Intelligent Information and Engineering Systems
M. Graña et al. (Eds.)
IOS Press, 2012

doi:10.3233/978-1-61499-105-2-857

Discovering shoppers' journey in retail environment by using RFID

Marin Vukovic*, Ignac Lovrek* and Hrvoje Kraljevic**

*University of Zagreb, Faculty of Electrical Engineering and Computing, Department of Telecommunications, Unska 3, HR-10000 Zagreb, Croatia
{marin.vukovic, ignac.lovrek}@fer.hr
**Agrokor, Trg Drazena Petrovica 3, HR-10000 Zagreb, Croatia
hrvoje.kraljevic@agrokor.hr

Abstract: Understanding the customers' needs is becoming more and more important in retail as it may provide valuable information that can be used to improve customer's shopping experience and product sales. Knowing about customers' behaviour within the store can provide information for improving customer relationship management and including real-time capabilities. This paper presents an experimental system for shopping path tracking and analysis that uses RFID for positioning shopping baskets. The collected data is used for determining customer presence and movement through a store, as well as time spent in specific store segments. The field trial of the proposed system in a real store is presented.
Keywords: Shopping journey, customer tracking, RFID

1. Introduction

Increasing business demands in the retail sector of fast-moving consumer goods places challenges on how to increase the level of consumer satisfaction, especially within large stores. Is it suitable to apply the same, standardized, retail layout across all geographic regions? Do customers really have the same habits regardless of geo-graphical areas they live in? Do we even know exactly how much time the customer spends within the sales area, what are his or her interests in certain categories of goods, and similar? It is difficult to measure the sale potential of certain goods if it is not known how many customers have decided not to purchase a product due to the lack of information, ignorance, unavailability of qualified staff or something else.

For example, consider a customer who spends above average time in one visit to the store, differing from his or her usual time per visit. It may be that the customer noticed a new model of an LCD TV in the technical equipment department and he or she is spending more time there in order to get more information about the product. However, the customer finally leaves the store without actually buying the television. If the store management had the ability to measure and visualize customer movement and observe the time spent in the department, it would enable them to detect that the customer is spending more than the usual amount of time in this department but does not make the purchase in the end. Observing this behaviour on a larger scale, including more customers and more store departments, it may lead to the conclusion that the department in question needs more staff in order to enhance sales. Without it, the increased interest in certain product without the actual sale can and should be treated as a loss.

There are two interrelated issues required to manage relationships with customers in a better way: a) knowledge about customers' behaviour, i.e. shoppers journey in the retail environment that will enable floor layout optimization on a daily basis, and b) understanding the customers' needs which will enable the improvement of shopping quality by providing real-time information about the product the customer is considering, offer alternative products and make it easier for the customer to find the desired product (store plan visualization, customer navigation, etc.).

This paper deals with the experimental system for tracking, analysing and visualizing the shoppers journey in the retail environment based on the wireless technology RFID (*Radio Frequency Identification*). The real-time location of RFID-based shopping baskets is used for collecting data describing the anonymous shopper's journey, while data aggregation, analysis and visualization that enables insight into customer behaviour is provided by a web application.

The rest of the paper is organized as follows: Section 2 deals with related work in the area of RFID in retail business in general and specifically in customer relationship management (CRM) including real-time CRM (RT-CRM). The architecture and functionality of the proposed experimental system is given in Section 3. A brief description of the field-trial follows in Section 4, while Section 5 concludes the paper.

2. Related work

The role of information and communication technology (ICT) in retail business has been widely studied for a long time [1]. The specific role of radio frequency identification is not restricted to the "identification" of objects using electronic product code (EPC) – it also enables individual digital presence and networking of objects as defined by conceptual frameworks of Internet of Things and Machine to Machine Communication (M2M) [2, 3].

Besides logistics, the high impact of ICT, including RFID, is expected in marketing and customer services, especially from and for large companies. Much more research as well as adoption is in the area of supply chain management (SCM) [4], than in customer relationship management. CRM is a critical business strategy in gaining competitive advantages. The ubiquity of the Internet has changed the way businesses are conducted. The term "real time" in retail applications refers to the time during which the customer stays in the retail area. Managing relationships with customers in real time assumes the possibility of interaction with the customer during his or her stay in the exhibition and sales space. CRM in real time is not possible without the collection of data about customer behaviour in real time, and their processing (analytic functions). Only then it may be possible to recommend and offer products in real time(operational functions) and to achieve cooperation between sellers and buyers in real time within a loyalty program (collaborative functions).RT-CRM is becoming increasingly significant to enable the agility of businesses to provide quick, accurate and complete responses to customer needs [5].

Different technologies and solutions can be used for this purpose, such as RFID-tags in loyalty cards, shopping baskets or carts, and even smart screens in shopping carts. Two interesting RFID based approaches to RT-CRM are elaborated in [6, 7]. From the information point of view, RT-CRM in retail collects data about customers and products from different sources in real time and extracts necessary and useful information for managing customers. Besides that, the acquired data stored in a customer database can be analysed which makes it possible to detect behavioural patterns, optimize products distribution in the sales area, customer profiles, personalized offers and more. A low-cost RT-CRM system based on a wireless sensor network equipped with RFID readers, which collects readings from RFID-based loyalty cards, is proposed in [6]. Mechanisms that bring SCM and CRM together by combining inventory management, frequent shopper program, discount pricing and recommender system are discussed in [7].

3. Experimental system for shopping path tracking and analysis

The experimental system is a part of a wider research project dealing with business intelligence and e-commerce aspects of retail business. This involves analysis of the shoppers' journey presented in this work but also the analysis of the customers' purchases [8].The goal of the proposed experimental system is to track, analyse and visualize customer movement within the store. However, customer movement data acquired by path tracking can be used for other purposes besides the one presented in this paper. Previous research in general user movement and its prediction shows that customer movement habits can be extracted from movement [9], which can be used for in-store customer profiling. Useful knowledge regarding store and in-store product layout in real time could be gained by combining movement paths with movement prediction, similar to research presented in [10] which deals with outdoor movement.

The system architecture is divided according to the three specified goals: positioning subsystem, analysis of the subsystem and visualization and administration of the subsystem.

Positioning the subsystem consists of RFID reader(s) and antennas that are deployed at suitable positions throughout the store. Customers are using shopping baskets equipped with RFID tags that are read once the customer moves near the deployed antenna. The tag reading is forwarded to the analysis subsystem.

The analysis subsystem aggregates the raw data acquired by the positioning system and performs analytic functions. The goal of analytic functions is determining whether the reading from the antenna can be verified as a customer presence at that antenna, deriving customer movement paths from raw positioning data, and calculating the amount of time that the customer spent in each store segment during his or her observed visit to the store.

Example store layout is shown on Figure 1, containing two segments S_1 and S_2 that need to be monitored for customer presence. Since each segment in this example has only two possible entrances, two antennas should be deployed in order to successfully monitor customers entering and exiting the segment. In this set-up, antennas $A_1S_1$and A_2S_1 monitor segment S_1 and antennas $A_1S_2$and A_2S_2monitor segment S_2.

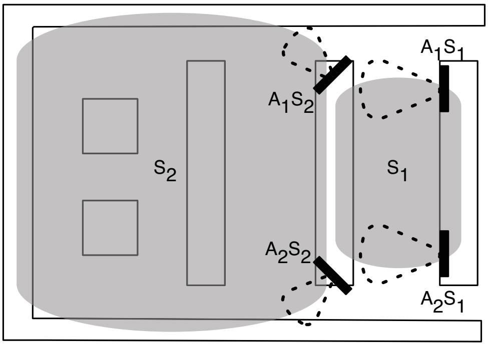

Figure 1 - Store segmentation

In general, a store is divided in segments $S = (S_1, \ldots, S_i, \ldots, S_n)$, where each segments S_i is covered by one or more antennas $AS_i = (A_1S_i, \ldots A_jS_i, \ldots, A_mS_i)$.

Figure 2 depicts the parameters that are derived from raw customer movement. The movement representation used is proposed in [11] for representing movement in general. Each arrow represents an event, i.e. customer presence at antenna (A_iS_j). This is the basis for further analysis, which consecutively determines segment presences $SP[S_j]$ and amounts of time spent in each segment (ΔT_j). In order to visualize customer movement through the store, directions d_k are derived from the recorded readings as well. The process of determining these parameters is explained further in the text.

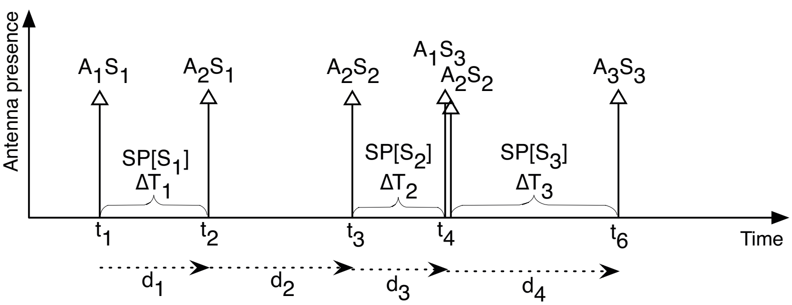

Figure 2 - Customer movement analysis

The visualization and administration of the subsystem developed as a web application enables visualization of the analysed data. It also provides administrative functions over the entire system, such as controlling the readers and antennas, running the analysis process, etc.

3.1 Determining customer presence

The positioning system should detect customer presence in specific store segments such as dairy products, meat products, liquor, etc. and to track the customer's path through the store. For this purpose, there is no need to define the customer's position with great precision. Objects to be tracked are shopping baskets equipped with RFID tags, while positioning is done by RFID readers and antennas deployed in the store. The tag on each basket identifies single customer during one shopping journey around the store.

In order to detect the customer, i.e. shopping basket presence in a certain segment, it is enough to monitor entry and exit points from the segment. These points are called referent points for each segment. Each referent point has an RFID antenna mounted in order to read the tags.

The presence of customer at a referent point is determined by a successful reading of a tag on its shopping basket. The typical store has pathways between shelves up to 5 meters wide so it is necessary to use long-range UHF antennas and appropriate UHF tags for this purpose.

The presence of customer in a segment involves two steps:

(i) Validation of tag reading at antenna(s)which is referred to as antenna presence (*AP*);

(ii) Analysis of the path of the customer journey in order to verify that he or she actually moved in the presumed segment, referred to as segment presence (*SP*).

The customer presence read by an antenna monitoring the pathway to certain segment (e.g. A_1S_2) usually indicates that the customer is entering the segment. However, single reading of a tag may be due to various interferences and errors [12] so it is necessary to differentiate successful tag reading from the verified antenna presence ‚*AP*. In this sense, it is important to observe the total number of readings n_R in a single time frame Δt that corresponds to the average time needed for a customer to pass along the pathway covered by an antenna. Threshold N_R determines the minimal number of readings during observed time frame needed to determine antenna presence. Threshold value is gained experimentally according to the antenna power and the area it covers. The antenna presence $AP[T_k(A_iS_j)]$ refers to the presence of a tag T_k at antenna A_iS_j monitoring the segment S_j and is determined as follows:

$$n_R\left(A_iS_j,\Delta t\right) \geq N_R\left(A_iS_j\right) \rightarrow AP\left[T_k\left(A_iS_j\right)\right]$$

A tag is read a number of times, depending on the antenna power, its sensitivity, the reader implementation, interference, the distance between the tag and antenna and number of tags in the field of an antenna at the time of the reading. In the presented case study, threshold N_R was experimentally set to 30, while the maximum number of reads exceeded 200 when there was a single tag in the area, moving at walking rate along a pathway with an antenna range set up to 5 meters.

The customer's segment presence *SP* is determined by examining consecutive antenna presences. The prerequisite for this approach is that each store segment has limited number of entries/exits that are all covered by antennas. Thus, by examining consecutive antenna presences at times t and t+1 it is possible to determine whether the shopper actually moved in the observed segment. As shown on Figure 1, while examining the shopper presence at segment S_2 the system should first detect the presence at one of the antennas monitoring the pathway to that segment, e.g. antenna A_1S_2 or A_2S_2, as $AP[T_k(A_1S_2)](t)$.Customer presence in segment S_2can be confirmed only if the antenna presence in the next reading at time t+1 corresponds to $AP[T_k(A_2S_2)](t+1)$, indicating the complete trip through the segment, or once again $AP[T_k(A_1S_2)](t+1)$, indicating shopper's loop through the segment.

In general, each segment S_j is defined by n entrances/exits corresponding to antennas A_i, where i=1..n:

$$S_j = \{A_1S_j, A_2S_j...A_nS_j\}$$

The presence of customer identified by tag T_k in segment S_j monitored is determined as follows:

$$\exists AP\left[T_k\left(A_iS_j\right)\right](t) \wedge \exists AP\left[T_k\left(A_mS_j\right)\right](t+1) \wedge \{A_iS_j, A_mS_j\} \in S_j \rightarrow SP\left[T_k\left(S_j\right)\right]$$

The above expression implies that antennas A_iS_j and A_mS_j detected the customer's tag at times t and t+1.

3.2 Customer movement

Movement and direction is estimated from the consecutive antenna presences. The possible set of directions D is determined from in-store antenna layout. Set D consists of n directions d, $D=\{d_1, \ldots, d_k, \ldots, d_d\}$. Each direction d_k, is specified by the source antenna presence AP_s and the destination antenna presence AP_d:

$$d_i = [AP_s, AP_d]$$

The data in the form of consecutive antenna presences AP is transformed to consecutive movements d:

$$AP(t), AP(t+1)...AP(t+m) \rightarrow d_1, d_2, ...d_{(m-1)}$$

In order to visualize customer movement in the web application each defined direction is linked to the store layout map using coordinates.

3.3. Shopping time distribution

The time that the customer spends in each store segment is based on detected segment presences SP. Each segment has two moments in time assigned; time of entrance and time of exit. The time of entrance corresponds to the first antenna presence at an antenna $A_iS_j(t_f)$ monitoring the segment S_j. Accordingly, the time of exit corresponds to the last antenna presence at the same or different antenna $A_mS_j(t_l)$ monitoring the same segment S_j. Therefore, the time spent in segment S_j is noted as ΔT_j and is equal to the difference of exit and entrance times:

$$\Delta T_j = t_l - t_f$$

The average time the customers spend per each segment is visualized as a heat map. Heat map colours are given in gray-scale and indicate the amount of time spent in each segment. The average time is presented with regards to the total time the customer spent in the store. It is calculated from moments of the first and last detected antenna presences during the single shopper trip around the store. The single shopper trip is represented by customer activity, i.e. detected antenna presences. If the customer leaves all observed segments and there is no antenna presence detected for a specified amount of time, the single trip is finished and every new antenna presence triggers a new shopping trip.

4. System deployment and field-trial

The proposed system was deployed in Croatia's largest retail company. In-store evaluation with real shoppers lasted for three months during the summer of 2011. The evaluated system did not cover the whole store due to cabling limitations in the working environment. Since the focus of the system was to evaluate its performance in the real environment, it was necessary to test its main purpose, i.e. detecting customer presence in store segments, but also to examine the influence of possible interferences in such an environment.

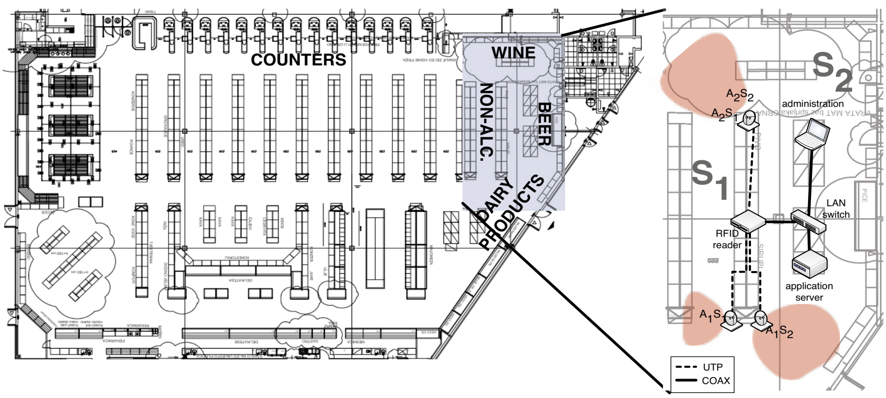

Figure 3 – Field-trial in real store

The evaluation system's position in a real store is shown on Figure 3. This position enables the identification of three segments: the pathway with non-alcoholic beverages (left, S_1), the segment with alcoholic beverages (up and right, S_2) and other (down, "other").

Shopping baskets were tagged with EPCglobal UHF Class 1 Gen 2 RFID tags. The RFID system used operates in a frequency range from 865 MHz to 868 MHz in UHF band. The system consists of three long-range UHF antennas used for monitoring pathways and differentiating segments. It is necessary to note that the upper antenna on Figure 3 monitors both segments S_1 and S_2. Each antenna is connected to a RFID reader hat interprets the signal from antennas and forwards readings to the application server via Ethernet switch. The application server hosts positioning logic, analysis and the web application used for data visualization.

The field trial performed in a real store environment was meant to achieve two goals; to inspect the effect of interferences and to analyse customer behaviour.

Possible interference effects were evaluated by testing various placements of the antennas and tagged shopping baskets with regards to liquids and metals. It was determined that the antennas should be positioned within direct optical sight of the area they are monitoring, covered by non-obstructive material (e.g. plastic, cardboard, etc.). In the final deployment, the antennas were actually placed behind marketing boards made of cardboard. However, significant interference was detected when antennas were placed on the shelves behind or near bottled water and, especially, tin cans containing liquids. Other effects of different origin, such as reflection caused by sheet metal, were not detected at all. The RFID tag placement on shopping baskets is also a matter of evaluation. The baskets used in real store evaluation were made of plastic and did not pose a problem in this manner. However, during experimental evaluation, carts made of wireframe were examined and they showed limited and non-obstructing effects of reflection.

The evaluation, performed for the purpose of customer behaviour analysis, encompassed twenty shopping baskets equipped with RFID tags. Baskets with RFID tags were mixed with ordinary baskets, so customers used any available basket in random manner. During the entire evaluation period, a total of 723 unique shopping trips were tracked and analysed.

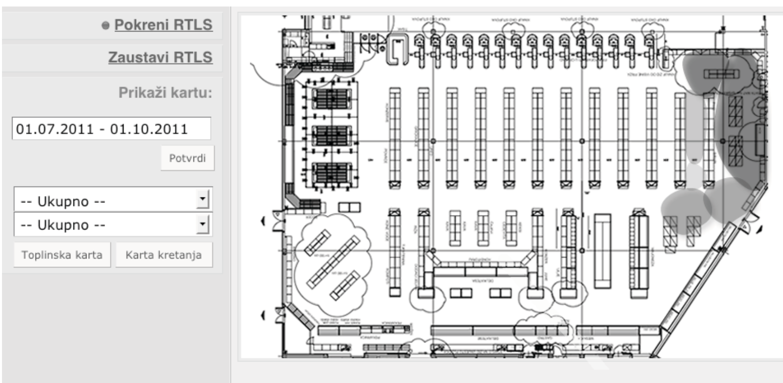

Figure 4 – Heat map of times spent per segment during the summer of 2011

The heat map with average times spent in observed segments for all 723 shopping trips is shown on figure 4. It is clear that the most visited segment is the one on the right, containing alcoholic drinks (S_2 from Figure 3, darkest grey). This is interesting because this segment is not on the expected most common route through the store, which usually leads directly to counters (up). Furthermore, the segment is somewhat secluded because the customers need to do almost a whole loop to get to the counters through the segment.

Figure 5 shows total movement direction aspects during the evaluation period. Most movement, as expected, occurs in the segment named "other" because this is the main hallway through which the customers move more frequently in order to get to any segment of the store. Furthermore, segment S_1 proves to be more popular than S_2

in terms of movement direction, while most customers, whether in S_1 or S_2, move to the counters as expected (to the left on Figure 5).

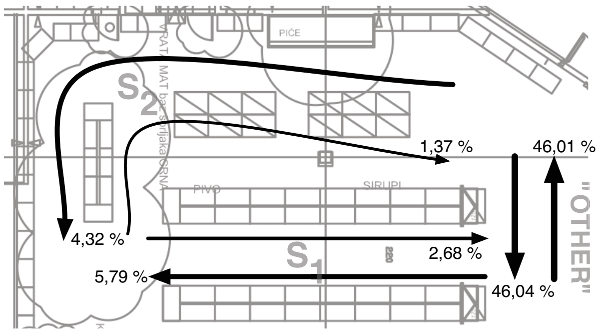

Figure 5 - Total movement direction aspects during the summer of 2011

This observation can actually explain the state of the heat map. Most customers are moving forward to the counters and the shortest path is through the segment S_1, which explains the higher frequency of that segment in terms of movement direction. However, the customers that do enter the segment S_2 are entering it because they are looking for something to buy there. Besides that, segment S_2 contains an empty bottle refund machine, which makes them stay in that segment longer. Furthermore, the upper part of segment S_2 contains an interactive touch screen with wine information, so the search for the desired product can take more time than just passing through the segment S_1 without the actual purchase.

5. Conclusion

The proposed experimental system tracks, analyses and visualizes customer movement in the retail environment by using RFID. The purpose of the performed field trial was to examine the possible effects of interference in the real environment and to see whether it is possible to derive useful customer movement patterns from tracking.

The field trial showed that there are interferences but that they do not significantly affect the performance of the tracking system, which is important for further research.

Tracking customers on just a portion of the store yielded as interesting and useful information about customer paths and objects that attract customers. This information can be used to improve customers' shopping experience by making certain products more accessible in terms of location, availability or support. From the retail point of view, the ability to see where customers are moving makes it possible to position certain products and increase their sales. Observing the time customers spent per segment gives insight about in-store objects that draw more attention and this information can be used to draw more customers to low admittance parts of the store.

The data gathered by the proposed system can be analysed for various purposes besides the one proposed in this paper. Common usage of such data is integration with a loyalty program by associating customer's shopping path with purchases, thus enabling customer profiling. Useful knowledge regarding the store and in-store layout could be gained by combining acquired movement paths with movement prediction. It could be used to determine how to set up an in-store product layout depending on the time, season, holidays and various other behaviour-related parameters. The integration of the proposed system and movement prediction will be investigated within future work.

Acknowledgement. The authors acknowledge the support of the research project "Business Intelligence and Electronic Commerce" funded by the company Agrokor, as well as of the research projects "Content Delivery and Mobility of Users and Services in New Generation Networks" (036-0362027-1639) and "Knowledge-based network and service management" (036-0362027-1640), funded by the Ministry of Science, Education and Sports of the Republic of Croatia.

References

1. ICT and e-Business Impact in the Retail Industry, Impact Study No. 04/2008 European Commission, DG Enterprise & Industry, empirica GmbH (2008).
2. Radio frequency identification (RFID) in Europe: steps towards a policy framework, Communication COM(2007) 96, European Commission, Brussels (2006).
3. X. Huang, S. Le, D. Sharma: A Taxonomy for RFID Systems. Proceedings IEEE International Conference on Signal Processing and Communication Systems, Gold Coast (2007).
4. A. Sabbaghi, G. Vaidyanathan: Effectiveness and Efficiency of RFID technology in Supply Chain Management: Strategic values and Challenges, Journal of Theoretical and Applied Electronic Communication Research, Vol. 3, No. 2, pp. 71-81 (2008).
5. J. O. Chan: The Anatomy of Real-Time CRM, Communications of International Information Management Association, 6(1), 115-124 (2006).
6. P. Schloter , H. Aghajan: Wireless RFID Networks for Real-Time Customer Relationship Management, Proceedings First International Workshop on RFID and Ubiquitous Sensor Networks (USN), Nagasaki, Japan (2005).
7. T. Inaba: Realization of SCM and CRM by Using RFID-captured Consumer Behavior Information, Journal of Networks,Vol. 4, No. 2 (2009).
8. M. Vranic, D. Pintar, Z.Skocir: Integrating quantitative attributes in hierarchical clustering of transactional data, 6th International KES Conference on Agents and Multi-agent Systems, KES AMSTA 2012(accepted for publication)
9. Vukovic, M.: "Adaptive location information management for value-added telecommunication services", Doctoral thesis, University of Zagreb, Faculty of Electrical Engineering and Computing, 2011. (in Croatian)
10. Vukovic, M.; Jevtic, D.; Lovrek, I.: User Movement Prediction Based on Traffic Topology for Value Added Services. // Lecture Notes in Artificial Intelligence, Subseries of Lecture Notes in Computer Science. 6881 (2011) , 1; 357-366
11. Lovrek, I. V. Sinkovic, "Knowledge-based mobility management for personal agents in the all-mobile network", *Multiagent and Grid Systems*, 4(2008), 4; 387-400.
12. Chonggang Wang, Bo Li, Mahmoud Daneshmand, KazemSohraby, and Rittwik Jana. 2011. On Object Identification Reliability Using RFID. *Mob. Netw. Appl.* 16, 1, 71-80. (2011).