

Real world implementation of belief function theory to Detect Dislocation of Materials in Construction

Saiedeh N. Razavi

Civil Engineering Department
University of Waterloo
Waterloo, ON, Canada
snavabza@engmail.uwaterloo.ca

Carl T. Haas

Civil Engineering Department
University of Waterloo
Waterloo, ON, Canada
chaas@civmail.uwaterloo.ca

Philippe Vanheeghe

LAGIS UMR 8146, Ecole Centrale de Lille,
Cite Scientifique, BP 48,
59651 Villeneuve d'Ascq Cedex, FRANCE
philippe.vanheeghe@ec-lille.fr

Emmanuel Duflos

LAGIS UMR 8146, Ecole Centrale de Lille,
Cite Scientifique, BP 48,
59651 Villeneuve d'Ascq Cedex, FRANCE
emmanuel.duflos@ec-lille.fr

Abstract – *Dislocations of construction materials on large sites represent critical state changes. The ability to detect dislocations automatically for tens of thousands of items can ultimately improve project performance significantly. A Belief_function -based data fusion algorithm was developed to estimate materials locations and detect dislocations. Dislocation is defined as the change between discrete sequential locations of critical materials such as special valves or fabricated items, on a large construction project. Detecting these dislocations in a noisy information environment where low cost Radio Frequency Identification tags are attached to each piece of material, and the material is moved sometimes only a few meters, is the main focus of this study. This work is a continuation of previous research, in which we tackled the location estimation problem by fusing the data from a simulation model. The results indicate the potential of the Belief_function based algorithm to detect object dislocation.*

Keywords: Locating, dislocation detection, belief function theory, construction materials.

1 Introduction

Material tracking is a key element in a construction materials management system. The unavailability of construction materials at the right place and at the right time has been recognized as having a major negative impact on construction productivity. Moreover, poor site materials management potentially delays construction activities, and thus threatens project completion dates and stands to raise total installed costs [9].

While automated controls are often established for engineered and other critical materials during the design

and procurement stages of large industrial projects, on-site control practices are still based on necessarily fallible direct human observation, manual data entry, and adherence to processes. These are inadequate for overcoming the dynamic and unpredictable nature of construction sites.

Node location approaches using signal strength and based on triangulation or relaxation algorithms [3, 4, 8, 10] are limited because of the cost of required node electronics (no current high volume demand exists), and because the anisotropic, dynamic transmission space on a construction site, for example, can not feasibly be mapped at the temporal or spatial resolution required. In addition, even sophisticated and expensive solutions experience multipath, dead space, and environmentally-related interference to some extent. For example, the Wi-Fi RTLS (real time location systems), such as commercial solutions from AeroScout, Ubisense, Ekahau, and the PanGo Network, require extensive calibration to map the Wi-Fi signals to locations throughout a building while the existence of 802.11 access points is not guaranteed for any facility being built. Thus we have selected a more cost-effective approach that is applicable to construction job site specifications. However, developing a method for location estimation that is robust to measurement noise but still has a reasonable implementation cost is a challenge.

Wireless sensor network-based data collection technologies such as GPS and RFID (Radio Frequency Identification) are being developed for a wide spectrum of applications. Specifically, more recent research is demonstrating that, coupled with mobile computers, data collection technologies and sensors can provide a cost-effective, scalable, and easy-to-implement materials location sensing system in real world construction sites [1, 5, 9, 11, 15, and 16]. The evident drawback of the current

cost-effective and scalable systems is lack of accuracy, precision, and robustness.

This study presents an improved formulation for robustly processing uncertainty and imprecision in proximity methods and is based on the Belief function theory [7, 12, 14]. By proximity we mean a binary spatial-constraint-based method [16]. The approach presented here gracefully manages the issue of dislocated tags, and results are presented graphically in an intuitive format.

For the purpose of this research, an integrated solution for automated identification and localization of construction materials was incorporated for a large industrial construction project. The main focus of the field trial was to develop a data fusion method for location estimation that is robust to measurement noise and has a reasonable implementation cost. The field trial involved a continuous site presence over 16 months by three graduate and three undergraduate students.

For the subset of data used in this paper, the tags' location data were logged by GPS-enabled readers for 109 tags, three times per day, for four consequent days. RFID read rates were sporadic, ranging from ten reads of a tag per minute to periods of hours without reads.

This paper is organized into the following sections. A brief introduction provides background to occupancy cell framework and proximity localization methods. Then, a practical elaboration on formulating belief function theory for locating materials and detecting dislocated items is presented. A brief description of the construction field experiment and the acquired data set follows. The results indicating the potential of the Belief function theory to detect materials dislocation make up the next section. Finally, the conclusion summarizes the findings of this research study, and suggests additional further works.

2 Background

In general, there are two approaches to localization. One is fine grained localization using detailed information, and the other is coarse-grained localization using minimal information. The tradeoff between the two approaches is obvious: minimal techniques are easier to implement and more likely to consume fewer resources and incur lower equipment costs, but they provide less accuracy than detailed information techniques.

Fine-grained node localization methods are based on specific detailed information and can be categorized into these measurement techniques:

- Time of Flight
- Received Signal Strength (RSS)
- Lateration and Angulation
- Distance-estimation using time difference (TDoA)
- Pattern Matching (RADAR)
- RF sequence decoding.

Coarse-grained node localization or connectivity-based localization algorithms are those which do not use any of the measurement techniques described above. In this

category, some sensors called anchors have a priori information about their location. The locations of other sensors are estimated based on connectivity information, such as who is within communication range of whom.

Proximity is the basis of another localization model that does not attempt to actually measure an object's distance from reference points, but rather determines whether an object is near one or more known locations. The presence of an object within a certain range is usually determined by monitoring of physical phenomena with a limited range, e.g., physical contact to a magnetic scanner, or communication connectivity to access points in a wireless cellular network. Some of the proximity-based methods introduced in this section make up a part of the proposed solution for this research.

In proximity models, for reduction of computational complexity, a discrete representation in 2D is employed instead of a more realistic continuous model. In the discrete view, a rover (any reader carrier) moves around in a square region, Q , with sides of length s ; Q is partitioned into n^2 congruent squares called "cells" of area $(\frac{s}{n})^2$. The RF communication region of a read is modeled as a square centered at the read and containing $(2\rho + 1)^2$ cells, instead of a disk of radius r . Thus, the position of reads as well as tags is represented by a cell with grid coordinates, rather than a point with Cartesian coordinates, and one is only interested in finding the cell(s) that contains each RFID tag (Figure 1). This paradigm is applied in the proximity approaches in particular.

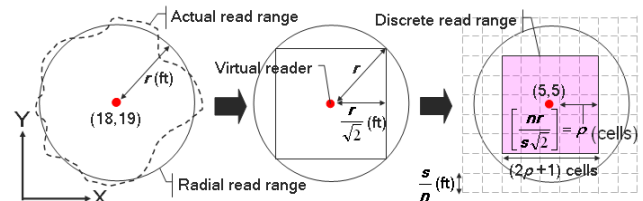


Figure 1: Modeling the RF communication region under the occupancy cell framework [17]

Simic and Sastry [13] presented a distributed algorithm for locating nodes in a discrete model of a random ad hoc communication network and introduced a bounding model for algorithm complexity. Song et al [17] adapted this discrete framework, based on the concept that a field supervisor or piece of materials handling equipment is equipped with an RFID reader and a GPS receiver, and serves as a "rover" (a platform for effortless reading). The position of the reader at any time is known since the rover is equipped with a GPS receiver, and many reads can be generated by temporal sampling of a single rover moving around the site. If the reader reads an RFID tag fixed at an unknown location, then RF communications connectivity exists between the reader and the tag, contributing exactly one proximity constraint to the problem of estimating the tag location. As the rover comes into communication range with the tag time and time again, more reads form such proximity constraints for the tag. Combining these

proximity constraints restricts the feasible region for the unknown position of the tag to the region in which the squares centered at the reads intersect with one another (Figure 2).

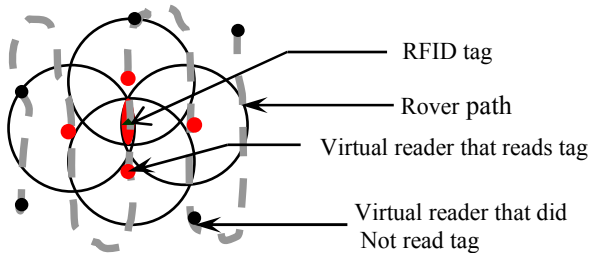


Figure 2: An illustration on how proximity methods work

Song et al [16] also implemented Simic and Sastry's algorithm in large-scale field experiments, including as parameters (1) RF power transmitted from an RFID reader, (2) the number of tags placed, (3) patterns of tag placement, and (4) the number of reads generated based on random reader paths. Analyzing data collected shows that in 51% of the total 4,200 instances, the true location of a tag was expected to be within ± 3 cells from the center of the region estimated to contain that tag. Although this approach was proven adequate (3-4 m accuracy) for static distributions of tags, it is not easily extended to tracking moving or moved tags.

3 Belief function theory for locating materials

The theory of belief function is a generalization of Bayesian theory. This theory has been initially developed by Dempster [7] and mathematically formalized by Shafer [12]. Belief function theory is a popular method to deal with uncertainty and imprecision with a theoretically attractive evidential reasoning framework [2]. Caron et al [6] showed that it can also manage the issue of moving tags while it is scalable and the granularity of the frame of evidence can shift in real-time.

In this theory the source of information is called evidence and the possible basic hypotheses are called the frame of discernment (E). The frame of discernment is our problem world that we are trying to observe and understand.

For the present application, the hypotheses for each tag will be of the form "It is in location C_{ij} " [Formula (1)]. A set of mutually exclusive propositions based on these hypotheses should be defined to build the frame of discernment. So, for each tag, the frame of discernment is the set of non overlapping square cells of the region (Figure 3). A circular area based on the ideal reading range of an RFID reader is the closest shape to the antenna's reading zone, but they can not cover the whole area of interest without overlapping and adding a lot of computational complexity to the problem. In addition, the actual range of any read position in the field is so highly

dependent on antenna orientation, multipath interference, and other factors such as RFID tag signal strength that it is highly ill-formed and is thus not more unreasonably modeled as a square than any other shape. Moreover, the square is computationally attractive. If the construction site is virtually partitioned into $n \times m$ square cells of C_{ij} , $i=1, \dots, n, j=1, \dots, m$, then the frame of discernment for each tag is

$$E = \{C_{ij} \mid i = 1, \dots, n, j = 1, \dots, m\} \quad (1)$$

For the model proposed here, the RF communication region of a read is modeled as a square, centered at the read and containing $(2R+1)^2$ cells while R is the ideal read range for the defined framework. Thus, the position of reads as well as tags is represented by a cell with grid coordinates, rather than a point with Cartesian coordinates, and only the cell(s) that contains each RFID tag are of our interests. To lower the error rate of the solution, cells are defined as 1×1 squares. The idea is adopted from the discrete framework presented in section 2.

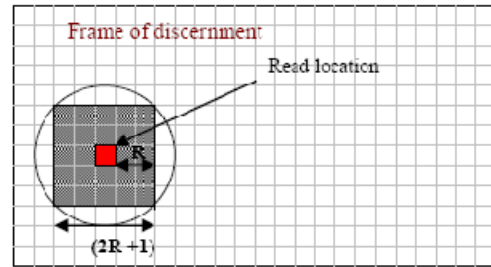


Figure 3: Discrete frame of discernment and modeling the RF communication region

Any observation of the GPS/RFID sensor GR_i , will be presented by assigning its beliefs over E. This is called the "mass function" of the sensors GR_i , and is denoted by m_i . Thus according to the sensors' observation of location L_i , the probability that "the specific tag is in location L_i " can not be lower than the belief confidence $Belief_i(L_i)$.

$Belief_i(L_i)$ is the belief confidence on all the evidences E_k that supports the given proposition. To take into consideration the whole belief that is in the given proposition, we need to include all the subsets of E which imply L_i . The belief is calculated as:

$$Belief_i(L_i) = \sum_{E_k \subseteq L_i} m_i(E_k) \quad (2)$$

For the current purpose, when an RFID reader reads a tag, the combination of GPS/RFID data gives information about the location of the tag, which is a hypothesis. This information can be modeled by a basic belief assignment because of the uncertainty in RFID read range due to the surrounding environment. To deal with this uncertain read range we assign different beliefs to different subsets of cells centered on the GPS/RFID sensor set of GR_i such that the sum of the all beliefs are equal to one. Let j be the index for p nested square areas centered on the RFID reader such that $E_1 \subset E_2 \subset E_3 \subset E_4$ [6].

$$\sum_{j=1}^p m_i(E_j) = 1 \quad (3)$$

For the solution proposed here, we assume for the area with half a read range, and for the rest of the area within read range (Figure 4). The two values in belief assignment are set up based on location error distribution of the sample data (Figure 10). The distribution shows that there could be a higher belief for the area closer to the tag.

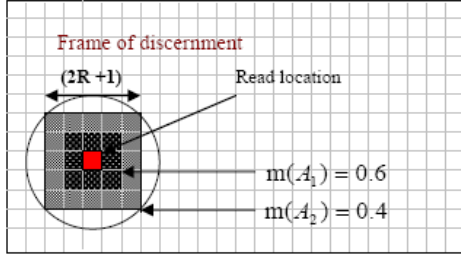


Figure 4: The proposed solution mass allocation

For fusing the data of more than one sensor, Belief function theory provides a means to combine these observations. In this sensor fusion we suppose that the sensors are independent. For the application proposed here, we can have multiple RFID readers coupled with GPS or multiple GPS chips with RFID interrogators or any other upcoming related technology.

In the simplest scenario, due to the environmental and other factors, GPS and RFID are having different reliabilities for each event of “read”. Therefore different read events can be considered as the independent observations that can be fused by the Belief function theory.

For each possible proposition (e.g., It is at location L_i) Belief function theory gives a rule for combining the sensors S_k 's observation E_k with the sensors $S_{k'}$'s observation $E_{k'}$:

$$(m_i \oplus m_j)(L_i) = \frac{\sum_{E_k \cap E_{k'} = L_i} m_i(E_k) m_j(E_{k'})}{1 - \sum_{E_k \cap E_{k'} = \phi} m_i(E_k) m_j(E_{k'})} \quad (4)$$

The Dempster's rule of combination (Formula 4) reallocates the conflict. This rule is commutative and associative and can be used in a sequential mode, to combine more than two acquired reads.

The system needs to decide at which cell the tag is located. Therefore it needs to compare the masses allocated to each cell after the fusion process. The center of gravity of the cells that have the maximum mass will be chosen as the location of the tag.

4 Detecting dislocation

For different essential causes conflict may exist: the sources may not be reliable, or the basic belief assignments may be modeled incorrectly.

When a tag is dislocated, a new reading may be made whose associated basic belief assignments contradict the past measurements. Conflict value is used here to detect this contradiction and so the movements of tags in the field.

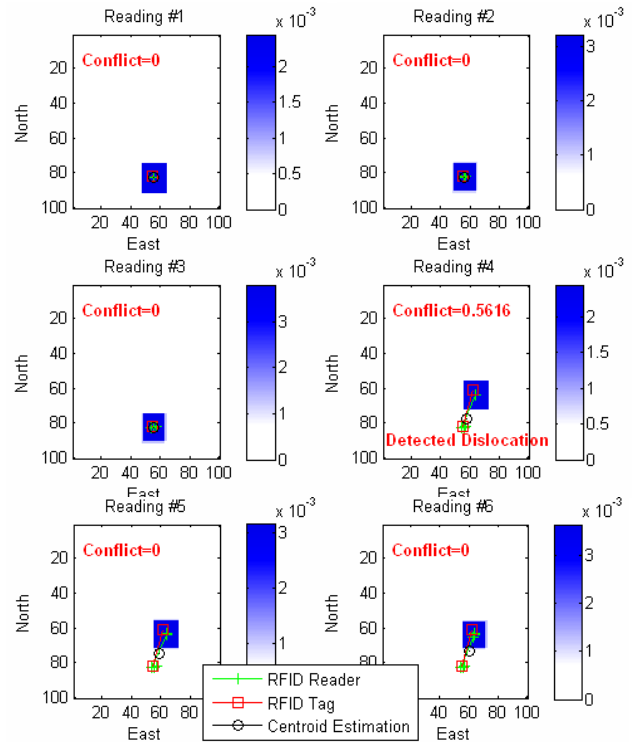


Figure 5: Detecting dislocation with conflict ratio

The conflict management can be handled in four different ways:

- Discounting of previous basic belief assignments at each time step to favor the latest
- Reject the old basic belief assignments when the conflict is over the threshold
- Decrease conflict by enlarging the focal sets
- Develop a new method in the frame of the Dezert-Smarandache Theory (DSmT).
- Use other combination rules such as Dubois-Prade's rule (DP), Dezert-Smarandache Theory (DSmT), or some others.

To deal with conflict in this study, old basic belief assignments are rejected when the conflict is beyond a pre-defined threshold. This approach was chosen due to the simplicity of computation and the scalability of the method.

Figure 5 illustrates a sample case study in which a dislocated tag is detected. In this case study, 6 read events of an RFID tag have been introduced sequentially to the fusion algorithm. The final estimated location is equal to the center of the darkest blue area that corresponds to the highest allocated mass.

5 Field experiments

Field trials were conducted to obtain experimental data to validate the data fusion model and to demonstrate the feasibility of employing the components, methods, and technologies developed. A large industrial construction project in Toronto hosted one field trial.

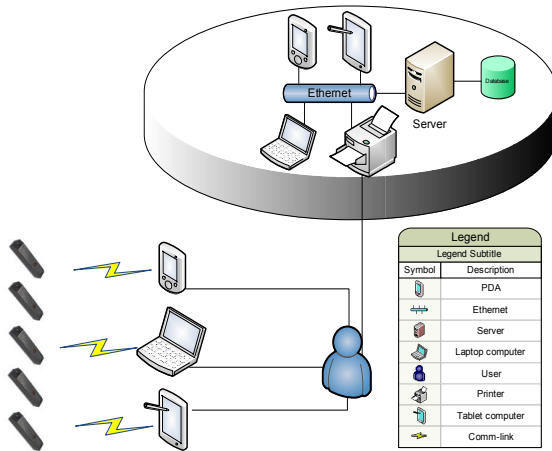


Figure 6: System network diagram

An RFID-GPS-based location estimation prototype was used in a comprehensive series of experiments. The technology for this research combines Global Positioning System (GPS) and Radio Frequency Identification (RFID) technology to automatically locate materials on the job site. Each RFID tag is assigned to a critical component. Then, a person traveling around the site with GPS-enabled readers detects the presence of tags around his/her own position (Figure 6).

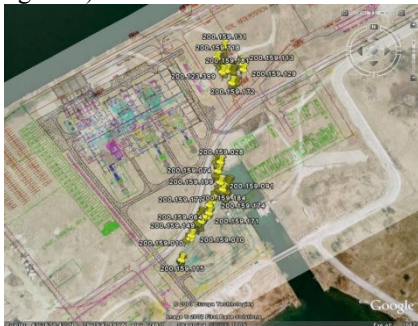


Figure 7: Sample map, including several RFID tag locations



Figure 8: Sample tagged items

During the whole course of the experiment, 375 tags were used to test the feasibility of tracking and locating certain critical components on a construction site and its supply chain. The data for testing the model are the coordinates of each tag ID on the lay down yards, which have been logged on a daily basis for five months. The estimated size of the data set is 94 days of data logging multiplied by, on average, 110 tags on the site per day multiplied by, typically a dozen reads per tag per day [11].

The daily location data is saved in the kml format so that it can be opened in the Google Earth map environment and the location information visualized. An AutoCAD drawing of the site overlaid on the Google Earth aerial photo provided more landmark reference details about the locations on the site. Maps were created in different granularity and various scales to allow proper visualization by field workers. The figures 7 and 8 present a tagged item and a sample map.

6 Results

Outputs of the used RFID-GPS-based prototype were used as the inputs for the developed Belief-function-based algorithm. The prototype outputs were estimated locations of the tags, based on the observed read events for each tag. These estimated locations were calculated in the centroid model. The figure 9 illustrates the hierarchical relationship among tag read events, estimated locations of the prototype, and the Belief-function-based fused estimation.

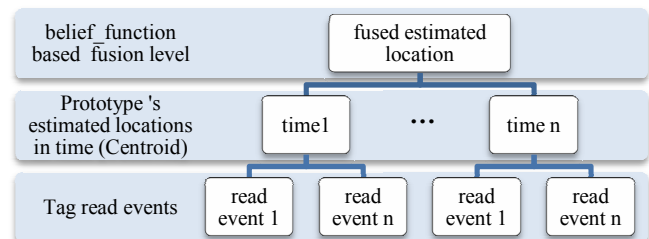


Figure 9: Hierarchical relationship representation among read events, estimated locations, and the estimations of the belief_function_based fusion level

A case study of a sample data subset is presented in this section. This subset was captured at the site warehouse area over four sequential days and is based on three daily cycles of reading and estimating locations. Each reading cycle might result in a typical number of ten to fifteen reads per tag. The following table represents the data specifications with respect to the distance between the estimated location – of the prototype based on the reads – and the real location of a tag.

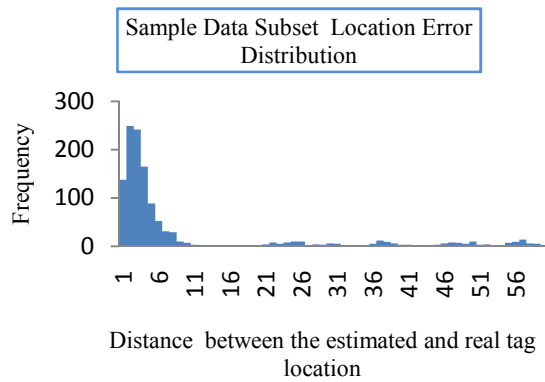


Figure 10: Location error distribution in the data subset

A total number of 57 dislocations were created in the sample data subset, among which more than one dislocation may be assigned to some of the tags. Benchmark measurements in the sample data subset – with sub-foot accuracy GPS – shows only a few dislocations for the period of observation in the sample data subset. To provide a greater number of dislocated samples in a controlled manner, 57 samples were created by transposing a sequence of estimates and their associated benchmarks for two stationary tags. The figure 11 illustrates this method of creating dislocation samples.

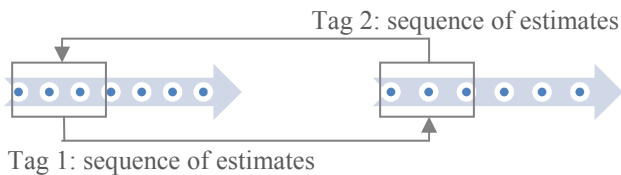


Figure 11: Creating dislocated samples by transposing a subset of estimates sequence and the corresponding benchmarks for two tags

The algorithm was run through the whole sample data subset to allow the observation of the dislocation detection rate with respect to the conflict threshold and the assumed read range in the frame of discernment. The results show a real potential for using the Belief function theory to detect the dislocated materials.

Dislocation detection rates with respect to the conflict threshold are shown in Table 1. Figure 12 presents the Receiver Operating Characteristic curve (ROC). Results indicate that a low conflict threshold causes high sensitivity and may result in false dislocation-detections. Conversely, with a high conflict threshold, some dislocations may not be detected.

Table 1: Dislocation detection rate with respect to conflict threshold

Conflict threshold	True-positive (TP)	False-positive (FP)	True-negative (TN)	False-negative (FN)
0	57	1160	0	0
0.2	39	39	1121	18
0.4	37	34	1126	20
0.6	35	33	1127	22
0.8	35	34	1126	20
1	22	29	1131	35
20	0	0	1160	57

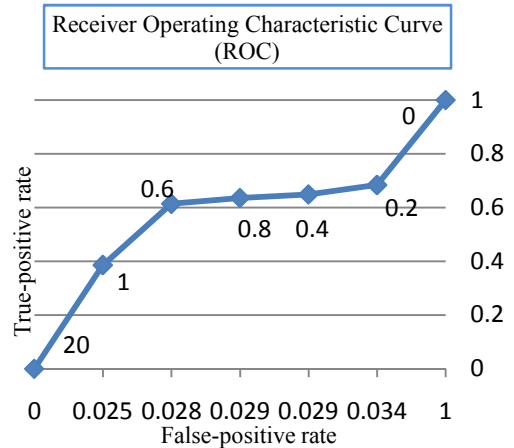


Figure 12: Receiver Operating Characteristic Curve (ROC) for different conflict thresholds

Table 2 presents the dislocation detection rates for different hypothetical read ranges in the frame of discernment (Figure 4). Figure 13 also shows the ROC diagram for the data shown in table 2. The results are obtained for the conflict threshold of 0.4 and indicate that a hypothetical range similar to that experienced on the site works best.

Table 2: Dislocation detection rate with respect to the hypothetical read range (Conflict threshold = 0.4)

Hypothetical read range (m)	True-positive (TP)	False-positive (FP)	True-negative (TN)	False-negative (FN)
0	56	981	187	1
2	54	347	813	3
6	46	87	1091	14
8	40	40	1120	17
10	37	34	1126	20
14	35	33	1127	22
16	31	30	1130	26
18	27	32	1127	30

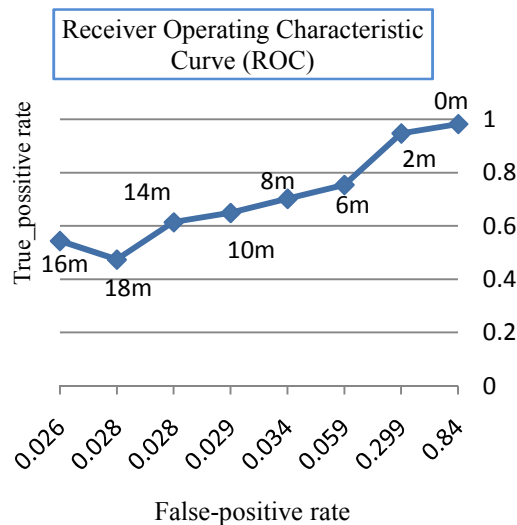


Figure 13: Receiver Operating Characteristic Curve (ROC) for different hypothetical read ranges

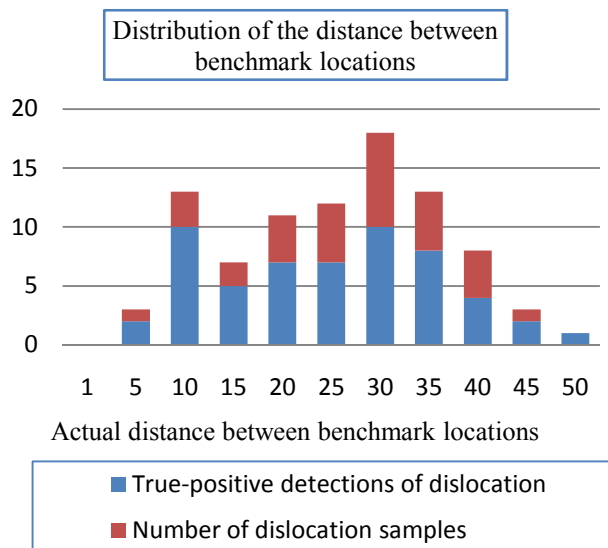


Figure 14: Distribution of the distance between benchmark locations of all dislocated sample as opposed to true-positive detections, for conflict-threshold = 0.8 and hypothetical read range=16m

Figure 14 presents a histogram for the distance between benchmark locations of the true-positive detections comparable to the same distribution for the entire dislocated sample’s population. The true-positive detection rates shown in this figure are based on the conflict-threshold of 0.8 and the hypothetical read range of 16m. A comparison of these two histograms shows that the probability of detection increases when dislocation distances are higher.

7 Conclusion

The targeted application described in this paper is to detect dislocation of materials on a construction site. Each material to detect is equipped with a RFID tag. A rover equipped with a RFID receiver and a GPS receiver is moving on the construction. The RFID receiver allows detections and the GPS localizations. Imprecision and uncertainty are the two main characteristics of this process. Belief function theory is therefore well adapted to propose a solution to improve the detection especially in case of dislocation. The main contribution of this work has been to show that a large and real world application based on the belief theory framework has been technically possible and has allowed saving money.

The detection process is based on a discretization of the search space. Each element of few nested subsets are then characterized by a basic belief assignment corresponding to the belief that an RFID tag belongs to this subset and therefore that a material is present within the border of the subset. The location of the subset is determined using a GPS receiver. When several readings are made, a fusion is made to obtain a new basic belief assignment. A dislocation is detected when the conflict resulting of the fusion process is greater than a known given threshold. This method has been implemented on a real large construction site in Toronto. The estimated size of the data set is 94 days of data logging multiplied by, on average, 110 tags per days. A dozen reads per tag per day were logged. The work focused on dislocations detections. The results obtained showed quite good detection rates and ROC curves have been plotted. As awaited the false alarm rate increases as the threshold used to detect the conflict and therefore the dislocation.

The conflict management is the heart of this method and the ROC curves show a rather high sensitivity to the value of the threshold. A deeper study must therefore now be carried out on this point. This may be done in four different ways:

- Discounting of previous basic belief assignments at each time step to favor the latest
- Reject the old basic belief assignments when the conflict is over the threshold
- Decrease conflict by enlarging the focal sets
- Develop a new method in the frame of the Dezert-Smarandache Theory.
- Use other combination rules such as Dubois-Prade’s rule (DP), Dezert-Smarandache Theory (DSmT), or some others.

Finally, we can notice that the results presented here can be easily reproduced in the frame of a large sensors network where the aim is to detect the moving of sensors.

Acknowledgement

The authors gratefully acknowledge the collaboration and helpful advice of Dr. Francois Caron.

References

- [1] Akinci, B., Patton, M., and Ergen, E. . "Utilizing Radio Frequency Identification on Precast Concrete Components - Supplier's Perspective." *the Nineteenth International Symposium on Automation and Robotics in Construction (ISARC 2002)*, Washington, DC USA, September 2002, pp. 381-386
- [2] Basir O, Karray F, Zhu H W., "Connectionist-based Dempster-Shafer evidential reasoning for data fusion" *IEEE Transactions on Neural Networks*, 16(6), pp. 1513_1530. 2005
- [3] Boyd, S., and Vandenberghe, L. *Convex Optimization*, Cambridge University Press, Cambridge, UK., 2004.
- [4] Bulusu, N., Heidemann, J., and Estrin, D. "GPS-less low-cost outdoor localization for very small devices". *IEEE Personal Communications*, 7 (5), 28-34. 2000.
- [5] Caldas, C., Grau, D., and Haas, C., "Using Global Positioning Systems to Improve Materials Locating Processes on Industrial Projects." *Journal of Construction Engineering and Management*, 132 (7), 741-749, 2006
- [6] Caron, F., Duflos E., Haas, C., Vanheeghe, P., « Modélisation de mesures de proximité par la théorie des fonctions de croyance ». *Revue Traitement du Signal*, vol. 4(2). 2007
- [7] Dempster, A.P. "A generalization of Bayesian inference", *Journal of the Royal Statistical Society*, 30, 1968.
- [8] Doherty, L., Pister, K., and Ghaoui, L., "Convex position estimation in wireless sensor networks". *Proc., INFOCOM, IEEE*, 2001, pp.1655-1663.
- [9] Grau Torrent, D. and Caldas, C.H., "Field Experiments of an Automated Materials Identification and Localization Model" *Proceedings of the 2007 ASCE International Workshop on Computing in Civil Engineering*, Pittsburgh, PA, 2007, pp. 689-696.
- [10] Hightower, J.; Boriello, G.; Want, R. "SpotON: An Indoor 3D Location Sensing Technology based on RF Signal Strength", *Technical Report #2000-02-02, University of Washington*, 2000
- [11] Razavi, S.N., Young, D., Nasir, H., Haas, C., Caldas, C., Goodrum, P., "Field Trial of Automated Material Tracking in Construction", *CSCE 2008 conference*, Quebec, Canada, June 2008
- [12] Shafer, G. "A Mathematical Theory of Evidence", *Princeton University Press*, 1976, Princeton.
- [13] Simic, S.N., and Sastry, S. "Distributed localization in wireless ad hoc networks". *Technical Report UCB/ERL M02/26, Department of Electrical Engineering and Computer Sciences*, 2002, University of California, Berkeley, CA.
- [14] Smets, P. and Kennes, R. "The Transferable Belief Model". *Artificial Intelligence*, 66:191-234. 1994
- [15] Song, J., Haas, C., and Caldas, C. "Tracking the Location of Materials on Construction Job Sites" *Journal of Construction Engineering and Management*, Vol.132, No.9. pp 911-918. 2006
- [16] Song, J., Haas, C.T., and Caldas, C.H. "A Proximity-based Method for Locating RFID Tagged Objects". *Special Issue of Journal of Advanced Engineering Informatics on RFID Applications in Engineering*. 2006.
- [17] Song, J., Haas, C., Caldas, C., and Liapi, K. "Locating Materials on Construction Sites Using Proximity Techniques", *ASCE Construction Research Congress*, 2005, San Diego, CA