

Extracting Location from Contact Traces

Marisa Vasconcelos*
marisav@cs.bu.edu

Mark Crovella
crovella@cs.bu.edu

I. INTRODUCTION

Localization of nodes is essential for many applications. Sensor networks, routing in mobile environments as well as caching can benefit by having this extra information. Applications dealing with emergencies and rescue operations need precise and accurate localization information. Many node localization techniques have been proposed, but most of them lack in some features (non-convex areas, non-symmetric communication, other types of constraints, etc.) that are desirable to have. For example, some of the techniques require special hardware or are only effective when the nodes are close to nodes whose location is known a priori. Sextant [8] is such solution that seeks to incorporate most of the desirable features. It is a constraint-based framework which relies on connectivity information.

The novel aspect of Sextant is that it relies on positive as well as negative spatial constraints. Positive constraints are based on the reception of a beacon from a nearby node. If a node cannot receive direct transmissions from another node, this constitutes a negative constraint. However, Sextant and the other techniques do not give us accurate localization estimates if we incorporate mobility of nodes.

Although mobility brings several issues, we believe that mobility can offer extra information to improve the location estimates. To deal with mobility scenarios, we consider another type of constraint: *temporal constraint*. This constraint takes advantage of properties of time and space that limit the possible locations. For that we use adjacent timestep areas to intersect with areas increased by the maximal rate (speed) of the node's movement.

To quantify the gain obtained using temporal information, we run simulations using synthetic data

and real contact traces obtained from a conference scenario. In both scenarios, to give us relative positioning, we had some landmark, which are nodes we know the location a priori. For the synthetic data, we simulated nodes moving according to the Random Way Point (RWP) model [2]. While RWP is a good starting point, it has been shown not representative of real mobility patterns [11]. We therefore also rely on real contact traces to incorporate more realistic mobility patterns.

The main contributions of this work are as follows:

- Extension of Sextant framework to mobility scenarios incorporating temporal constraints.
- Evaluation with measures, obtained from real scenarios such as conferences.
- Simulation of different topologies and speeds of nodes in a field.

The rest of the paper is organized as follows. Section II discusses related work. Section III describes the Sextant and Sextant with temporal constraints algorithms. Our evaluation and main results are presented in Section IV. Finally, Section V presents conclusions and directions for future work.

II. RELATED WORK

There are several localization algorithms and they can be classified according to: use of landmark nodes (nodes with known location), range estimation method (Time-of-arrival, Angle-of-arrival or connectivity) or computational method (centralized vs. decentralized).

The Convex position method [4] is a centralized algorithm that uses a set of geometric constraints and solves it using convex optimization. All nodes in the network must declare their connectivity to a single node that solves the possible locations as a optimization problem. The accuracy of this method depends on the number of landmark nodes and, since the only information used is positive (nodes

*Supported by CAPES/Brazil

that can hear each other), the resultant areas are only convex.

GPS-Less [1] is a distributed algorithm that positions nodes in the centroid of the landmark nodes they have contact with. However, this approach does not disseminate the nodes position information beyond the first hop. Consequently, nodes that are outside the range of landmarks cannot be located.

APS [6] is a distributed algorithm that provides absolute positioning using hop count, signal strength or Euclidian distance to compute and disseminate transitively the location of the landmark nodes. After receiving this information, each node can compute its position using multilateration techniques. This method performs well if we have a uniform distribution of landmark nodes but it does not guarantee convergence always.

GPS-Free [3] is a distributed algorithm that gives relative position of the nodes with respect to the network topology. Each node calculates and disseminates the distance between its one-hop neighbors, using time of arrival measurements. To use this kind of measurement, each node requires synchronization hardware. Furthermore, if we have a mobile scenario, the relative positions must be recalculated if the reference or intermediate nodes move.

MDS-MAP [10] is a MDS (multidimensional scaling) based method that uses distance or connectivity information to compute relative node positions. If there is a sufficient number of landmark nodes, it can also provide the absolute locations. There is an improved version [9] that deals with the problem of irregularly-shaped networks (presence of holes). In this improved version each node computes a local relative map using MDS and then the algorithm merges all maps to form a big relative map.

Sextant [8] is a distributed algorithm that solves a set of geographic constraints based on connectivity information. It does not require any special hardware, uses a small number of landmark nodes and can model non-convex areas using negative information. In addition, it uses a dissemination algorithm to allow nodes that are more than one-hop distant to compute their locations. The constraints are represented as a set of polygons enclosed in Bezier curves which allows an expressive and compact representation. However, Sextant was

developed mainly for static scenarios, which do not include the use of temporal constraints.

III. NODE LOCALIZATION ALGORITHM

In this section, we describe Sextant algorithm and present the extension using the temporal constraints.

A. Sextant Description

As the original Sextant, we also use three constants named R , r and $FIELD_AREA$ as input of the algorithm. R represents the maximum distance that a node can receive transmissions from. r is the minimal distance that a node has to be away from a node that cannot hear it, and $FIELD_AREA$ is the area of the field. The reason for two separate radii is to handle non-uniform propagation regions.

The algorithm computes three types of areas during its execution. The estimated area E , the maximal coverage area (or might-hear region) M and the assured coverage area (or must-hear region) A . The overview of the algorithm is shown in Figure 1. In our implementation, we assumed that we have global information of all the node areas.

The estimated area is the output of the algorithm, which is the area where a node can be. This area is derived through intersections of the might-hear areas of the nodes that are in contact and subtractions of the must-hear areas of the nodes that are out of range. The might hear area is calculated by doing the union of all circles with radius R centered in the boundary of the node estimated area. The must hear area is calculated by doing the intersection of all circles of radius r centered in the estimated boundary area. Figure 2 shows the estimated area for one node in each timestep after running Sextant.

The might hear areas are the representation of the positive constraints, they are the upper bound where the node can hear, while the must hear areas is the area where we can guarantee that a node can receive transmissions. After computing the estimated area for each node, we also verify all the constraints for that node. This verification is carried out using the following rules:

- For each node that can be heard by this node, verify if their might-hear areas *intersect* with this node estimated area.
- For each node that cannot be heard by this node, verify if their must-hear areas *do not intersect* with this node estimated area.

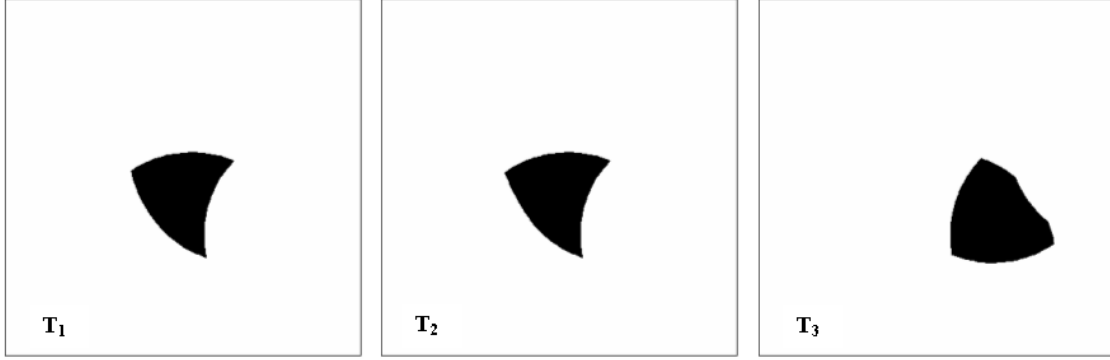


Fig. 2. Sextant estimated areas (E_q) for each timestep for node q

Sextant

Input: connectivity matrix, R , r , and landmark nodes

Output: estimated areas for all nodes in the topology

Initialization:

for all non-landmark nodes $q \in Q$

$$E_q = \text{FIELD_AREA}$$

Computing areas :

for each node $q \in Q$

for each node p that q can hear in the topology

$$E_q = E_q \cap M_p$$

for each node n that q cannot hear in the topology

$$E_q = E_q \setminus A_n$$

Fig. 1. Sextant Algorithm.

B. Sextant with Temporal Constraints

Sextant was originally developed for static scenarios. In this work we studied its behavior in a mobile scenario. For that we include another type of constraint: the temporal constraint. Before computing the estimated areas using this new type of constraint, we have to execute Sextant individually for each timestep (Figure 2).

The next step is to refine the areas intersecting the estimated areas with the motion area from an adjacent timestep. The motion area is computed using the areas calculated in the previous step expanded with the movement rate (speed) in any

Sextant with Temporal Constraints

Input: estimated areas from Sextant algorithm and speed v

Output: estimated areas for all nodes in the topology

Computing areas:

for each T_i

for each node $q \in Q$

$$E_q^{T_i} = E_q^{T_i} \cap S_q^{T_{i-1}} \quad (1)$$

$$E_q^{T_i} = E_q^{T_i} \cap S_q^{T_{i+1}} \quad (2)$$

Fig. 3. Sextant with Temporal Constraints Algorithm.

direction. The computation of this constraint was implemented as shown in Figure 3.

For each node q , we compute two types of constraints: backward and forward. The backward constraint (Figure 4) is computed by intersecting the node estimated area $E_q^{T_i}$ with the expanded area of the same node in the previous timestep $S_q^{T_{i-1}}$. The forward constraint (Figure 5) is calculated by intersecting the node estimated area $E_q^{T_i}$ with the expanded area in the next timestep $S_q^{T_{i+1}}$ of the same node.

These refinements are performed until there is no changing in the nodes' estimated area. While these refinements are made, we perform the same spatial constraints verification made in the first step.

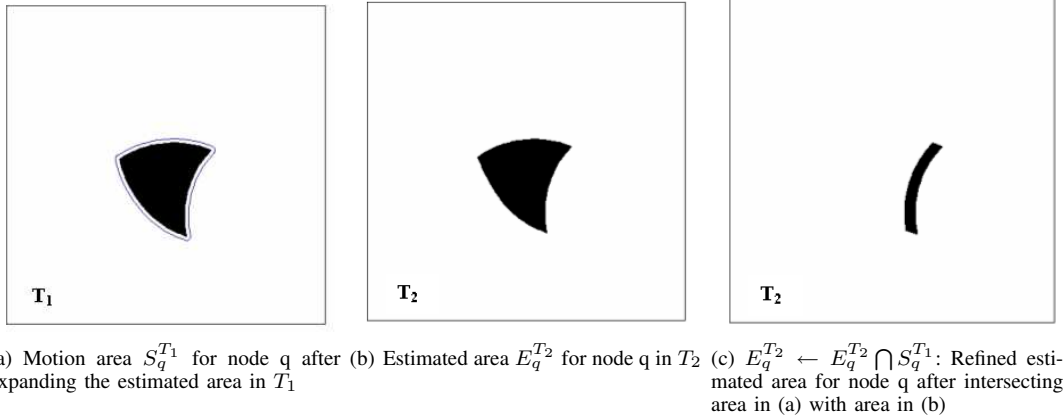


Fig. 4. Sextant with temporal constraints (Backward constraint)

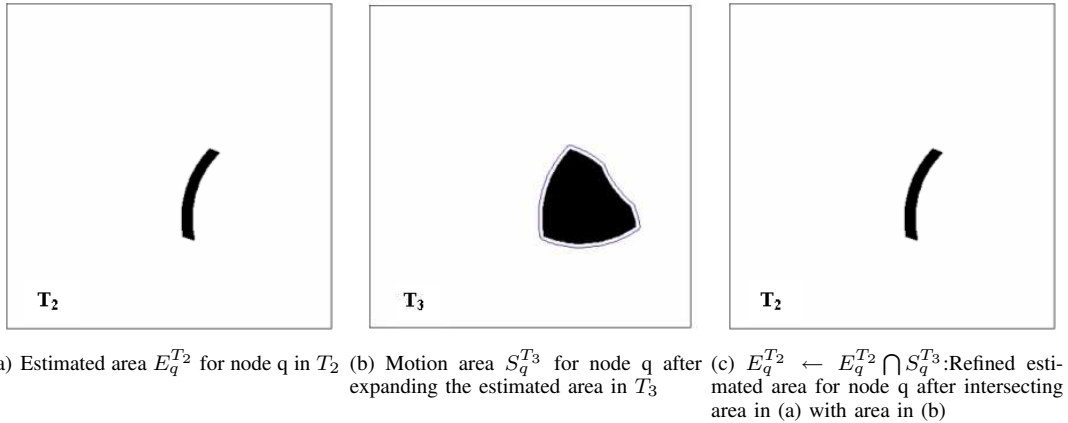


Fig. 5. Sextant with temporal constraints (Forward constraint)

IV. EVALUATION

We validated our algorithm using simulations of synthetic scenarios and using real contact traces.

A. Synthetic scenarios

We have simulated a topology with 30 nodes randomly distributed over an area of $500 \times 500 \text{ m}^2$. The mobility is based on the Random Way Point model [5] with no pause times. We also considered symmetric scenarios: if a node A can hear node B, node B can hear node A. For this data, we also verified for each area if the true location was inside the estimated area.

For each second (timestep) of this simulation, we have extracted the connectivity matrix from

this sub-interval. A node can hear another node if its distance d to the node is $d \leq R$. As in the Sextant paper, for our simulations, we assumed $R = r$. The other parameters used in the simulations are described in Table I.

Parameter	Value
Total number of nodes	30
Number of landmark nodes	3
Grid Size	$500 \times 500 \text{ m}^2$
Speeds considered	$[0.5, 20] \text{ m/s}$
Number of timesteps	15
Duration of each timesteps	1 sec

TABLE I
SIMULATOR PARAMETERS.

B. Contact Traces

We have used a contact trace collected during experiments made during INFOCOM 2006 [7]. The iMotes used have Bluetooth 1.1 radio. The stationary iMotes have ~ 100 meters and the mobile ones have ~ 30 meters radio range. Every 120 seconds plus a random number of seconds, each iMote starts an inquiry scan to contact devices that are in the neighborhood. These active contacts were recorded in a log along with the time of first and last inquiry scan. The other parameters used in the simulations are described in Table II.

For our experiments, we use 30 minutes of this trace. We have considered each timestep being 120 seconds, because of the scan inquiry granularity. We filtered nodes that did not appear in all the timesteps. To estimate the value of r (negative constraints), we ran Sextant for all the nodes in the trace topology until we notice a constraint violation. This methodology gave r equals to 17 meters for this trace.

Parameter	Value
Total number of nodes	62
Number of landmark nodes	5
Landmark ranges	100 m
Mobile ranges	30 m
Grid Size	44.6 x 42.7 m
Average speed	$[8 \times 10^{-3}, 6.6 \times 10^{-3}]$ m/s
Number of timesteps	15
Duration of each timestep	120 sec

TABLE II
CONTACT TRACE PARAMETERS.

C. Results

The results of both algorithms are the areas where a node must be during that timestep. We have varied three parameters for both scenarios: the values of R , r and the nodes' maximum speed. We computed the areas and their centroids using the Monte Carlo method, the same used in Sextant paper [8]. For the synthetic data, all the results have a 95% confidence interval.

1) *Results for Synthetic Data:* For all the plots using the synthetic data, we considered $R = r$, the same assumption made in Sextant paper. In Figure 6, we show the sum of all areas during each timestep varying R and r with all nodes at 20 m/s.

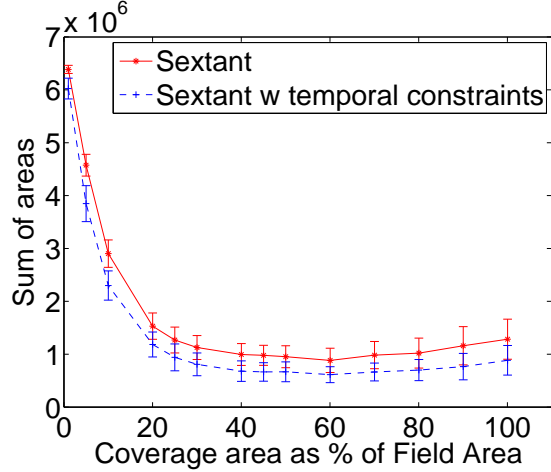


Fig. 6. Synthetic data: Sum of all areas as function of the coverage area with speed 20 m/s

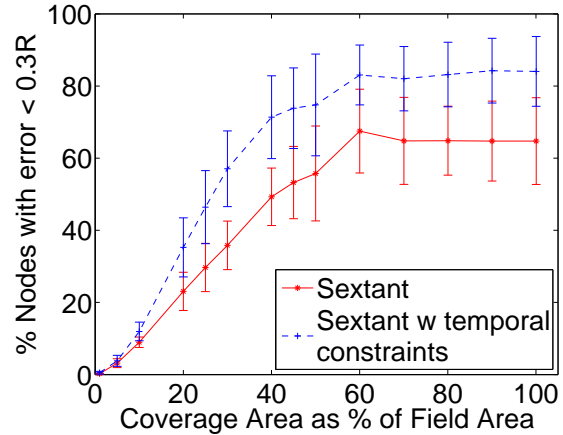


Fig. 7. Synthetic data: Percentage of nodes at 20 m/s located within an error of less than $0.3R$ as function of the coverage area

We can see that both algorithms have their resultant areas reduced when we increase the values of R and r .

This happens because increasing the value of r , we increase the must-hear areas. The effect will be a larger subtraction of must-hear areas from the estimated areas. Another explanation is when we increase R , we increase the connectivity (average number of contacts) with landmarks or nodes that are directly connected to them, making the estima-

tions more refined.

Figure 7 shows the percentage of nodes which had their position estimated accurately as a function of the coverage area. Again, we can see that increasing the coverage area; we observe the number of nodes located grows faster than using only spatial constraints.

Figure 8, 9 and 10 shows the histogram of areas size distribution of both algorithms using the same value of R (50% of coverage area) and different speeds. As we increase speed, we increase the number of nodes with smaller areas, which demonstrates that the temporal constraints make the estimates more accurate.

In Figure 11 and Figure 12, we show the sum of all resultant areas and the percentage of nodes that can localize themselves accurately as a function of the speed using a transmission range of 25% and 50% of the coverage area. We notice that the areas are reduced when the nodes maximum speed is increased.

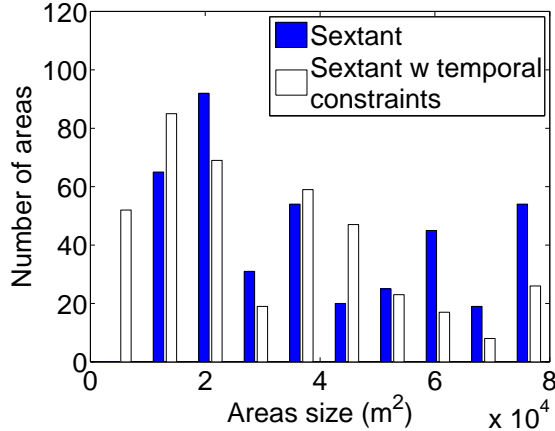


Fig. 8. Synthetic data: Areas size distribution for nodes at speed 5 m/s and $R = 199.52$

Figure 13 shows the real movement of a node and the movement estimated using Sextant and Sextant with temporal constraints. We notice that Sextant gets confused by looking at only local information. Also, for this node, we computed the cumulative error, which gives the sum of all distances between the estimated location and the actual position. We could see that Sextant with temporal constraints has smaller cumulative error (2×10^3) than Sextant ($3 \times$

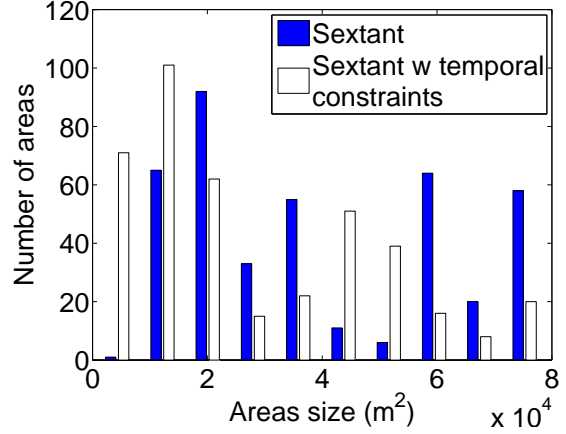


Fig. 9. Synthetic data: Areas size distribution for nodes at speed 10 m/s and $R = 199.52$

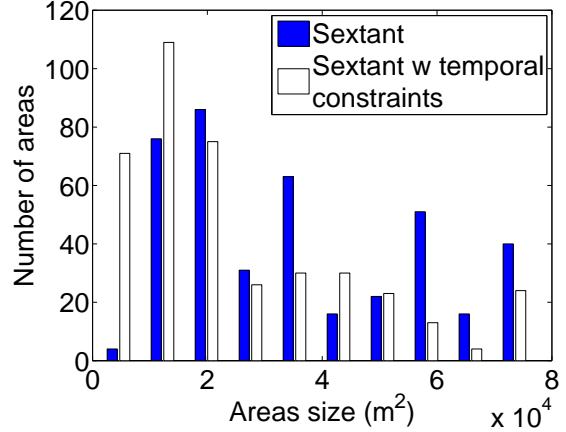


Fig. 10. Synthetic data: Areas size distribution for nodes at speed 20 m/s and $R = 199.52$

10^3).

2) *Results for Contact Traces:* Figure 14 shows the sum of all nodes areas in each timestep. As we can see, we reduced the estimated areas in approximately 10%, when we use temporal constraints. In addition, we observe that Sextant with temporal constraints has better results for lower speeds. This can also be verified also in Figure 15, which shows the distribution of the areas sizes. This can be explained by the fact that the intersections due to temporal constraints are performed using smaller areas. The fact that lower speeds present a better result when compared to the synthetic data may be

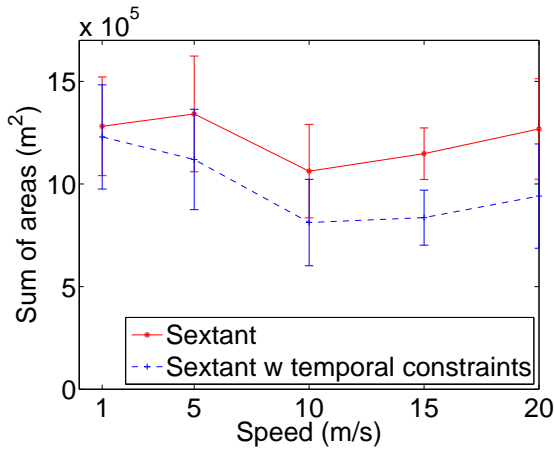
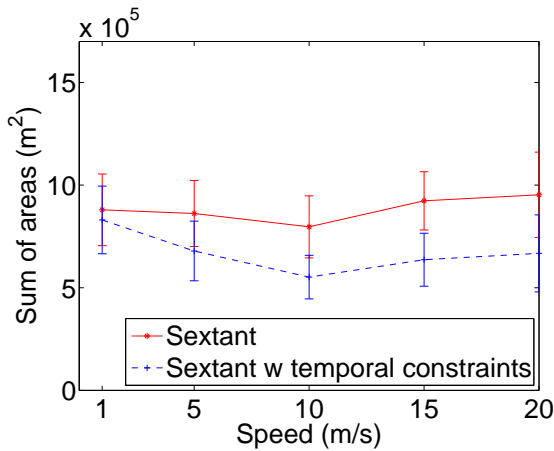
(a) $R = 141.08$ (25 % coverage area)(b) $R = 199.52$ (50 % coverage area)

Fig. 11. Synthetic data: Sum of all areas as function of speed

because the maximum speed of nodes in the trace was lower or include pause times.

It is important to note that using speed equal to zero, the motion areas are nonexistent and the temporal constraints use the same areas computed by the spatial constraints. We did not observe any violations of any constraints during the simulations even with zero speed. This happened because the must-hear areas are nonexistent and the might-hear areas are equal to the whole field. The coverage of the landmark nodes and the field size are responsible for that.

The results for the trace were not accurate enough

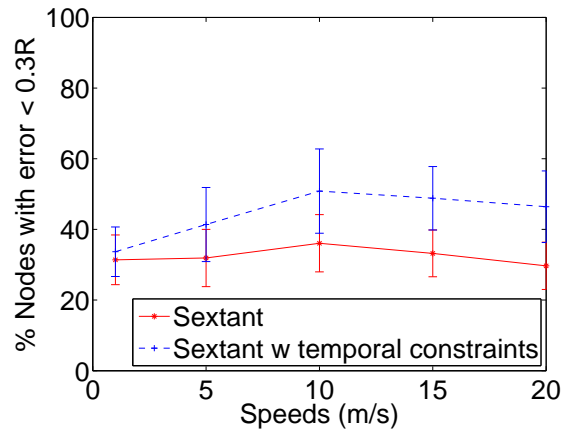
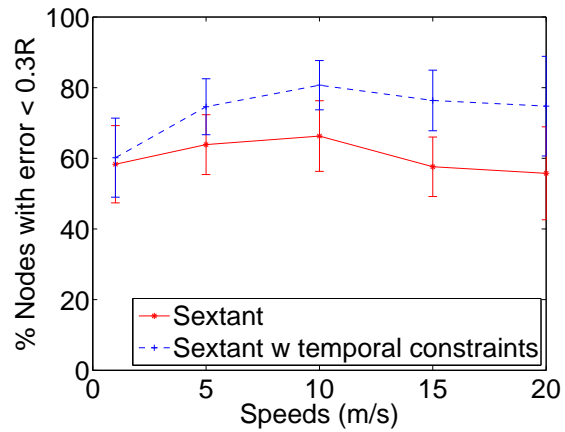
(a) $R = 141.08$ (25 % coverage area)(b) $R = 199.52$ (50 % coverage area)

Fig. 12. Synthetic data: Percentage of nodes accurately determining their position as a function of the node maximum speed

for tracking. We observed that the estimated areas represented a big percentage of the field area and because of that most of must-hear areas were empty and might-hear areas were equal to the field area. Because of that the negative constraints calculated using the must-hear areas are not useful in this case.

We have experienced other difficulties when we used the trace data:

- **Limitation of the number of contacts per device during the inquiry scan.** The maximum number of contacts with the same start time was limited to 8, because of Bluetooth upper bound of number of the responses returned during an inquiry scan. This also explains why

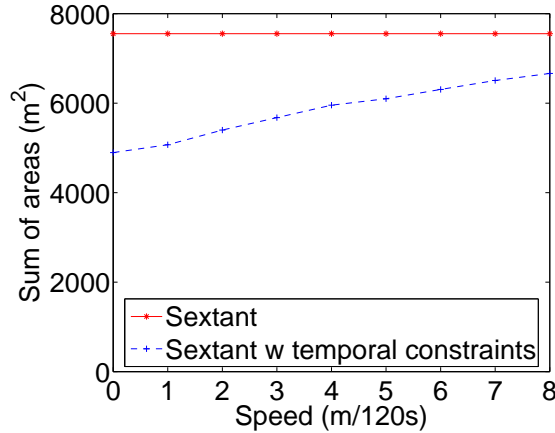


Fig. 14. Contact Traces: Sum of areas as function of speed

the number of contacts in the trace was low.

- **Accurate location of the landmark nodes.** The data did not have the real location of the stationary iMotes.
- **Nodes can be heard in different floors or could be outside of the field.** We could not make sure if a node was in the field during a timestep or it was out of range of all landmark nodes. Some stationary nodes could hear other nodes in other fields located in other rooms. So, even if we have selected from the trace only nodes that had a contact during that interval, we don't know if the node was outside of the field.
- **The granularity of the inquiry scan.** It made difficult to compute accurately the motion constraints. The inquiry scan was set to be 120 seconds during the trace collection. It gives us too much variance and makes harder to recreate the mobility pattern of the trace.

For both scenarios, synthetic and trace, we notice that the more timesteps used the more accurate are the estimations. Unfortunately due to computation limitations, we could not experiment with more timesteps.

V. CONCLUSION AND FUTURE WORK

In this paper we have investigated the problem of localization of wireless nodes in mobile scenarios. Most localization algorithms rely in some kind of special hardware to compute the nodes position.

This has motivated the development of cheaper solutions for example, using connectivity information.

Sextant is one of these solutions that use connectivity information to extract spatial constraints to locate wireless nodes in a field. However, Sextant constraints are not enough to deal with mobile scenarios. We observe that mobility gives extra information about the past and future locations of the nodes. We use this information to refine the nodes location, adding to Sextant, temporal constraints. This type of constraint guarantees an upper bound of the possible location of a node during adjacent timesteps. Using the temporal constraints along with the spatial constraints, we were able to improve the nodes position estimations.

To analyze the performance of this extension, we performed simulations using synthetic and real data varying the nodes transmission range and their maximum speed. The results obtained show that Sextant with temporal constraints was able to refine the location estimates. We showed that our algorithm has a better performance for high speeds (> 10 m/s). For the real contact trace scenarios, we could improve our estimates if we had better information such as real position of the landmarks and small transmission range compared to the field area.

We will further investigate the performance and behavior of this algorithm under different scenarios, bigger fields, such as campus or cities and how the accuracy varies when we increase the number of timesteps.

REFERENCES

- [1] N. Bulusu, J. Heidemann, and D. Estrin. GPS-Less Low Cost Outdoor Localization for Very Small Devices. Technical report, University of Southern California, 2000.
- [2] T. Camp, J. Boleng, and V. Davies. A survey of Mobility Models for Ad Hoc Network Research. volume 2, pages 483–502, 2002.
- [3] S. Capkun, M. Hamdi, and J. Hubaux. GPS-Free Positioning in Mobile Ad-hoc Networks. In *Hawaii Int. Conf. on System Sciences*, 2001.
- [4] L. Doherty, K. Pister, and L. E. Ghaoui. Convex Position Estimation in Wireless Sensor Networks. In *IEEE Conference on Computer Communications (INFOCOM)*, 2001.
- [5] G. Lin, G. Noubir, and R. Rajaraman. Mobility Models for Ad-hoc Network Simulation. In *IEEE Conference on Computer Communications (INFOCOM)*, 2004.
- [6] D. Niculescu and B. Nath. Ad-hoc Positioning System (APS). In *GLOBECOM*, 2001.
- [7] E. Nordström, C. Diot, R. Gass, and P. Gunningberg. Experiences from Measuring Human Mobility using Bluetooth Inquiring Devices. In *Proceedings of the 1st International*

- Workshop on System Evaluation for Mobile Platforms (MobiEval)*, 2007.
- [8] R. M. S. Guha and E. G. Sirer. Sextant: A Unified Node and Event Localization Framework Using Non-Convex Constraints. In *In Proceedings of ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc)*, 2005.
 - [9] Y. Shang and W. Ruml. Improved MDS-based Localization. In *IEEE Conference on Computer Communications (INFOCOM)*, 2004.
 - [10] Y. Shang, W. Ruml, Y. Zhang, and M. Fromherz. Localization From Mere Connectivity. In *MobiHoc*, 2003.
 - [11] L. Song and D. Kotz. Evaluating Opportunistic Routing Protocols with Large Realistic Contact Traces. In *CHANTS '07: Proceedings of the second workshop on Challenged networks CHANTS*, pages 35–42, New York, NY, USA, 2007. ACM.

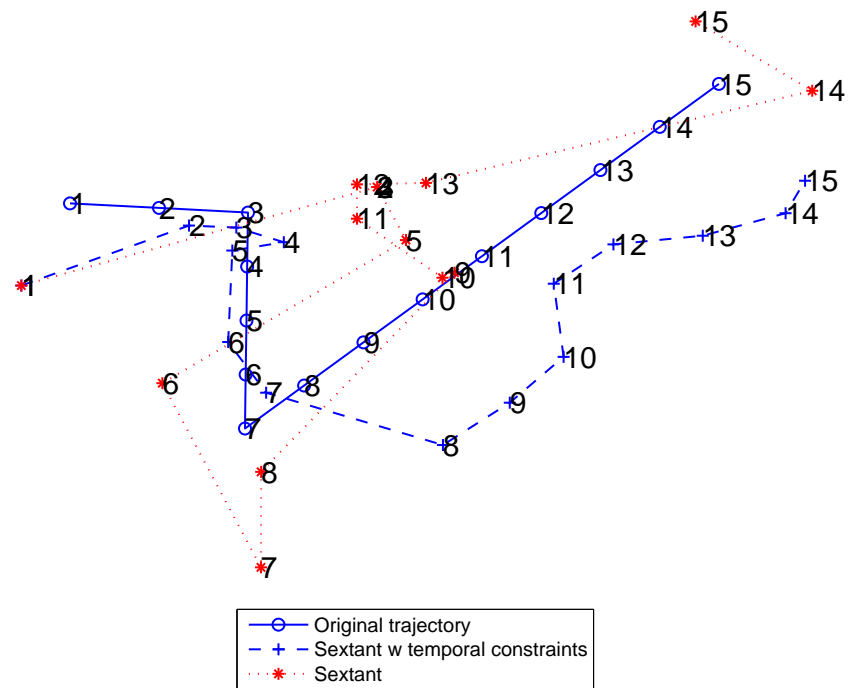


Fig. 13. Synthetic data: Tracking for a node at speed 5 m/s and $R = 45$ (64% coverage area), Sextant cumulative error = 2×10^3 , Sextant with temporal constraints cumulative error = 3×10^3

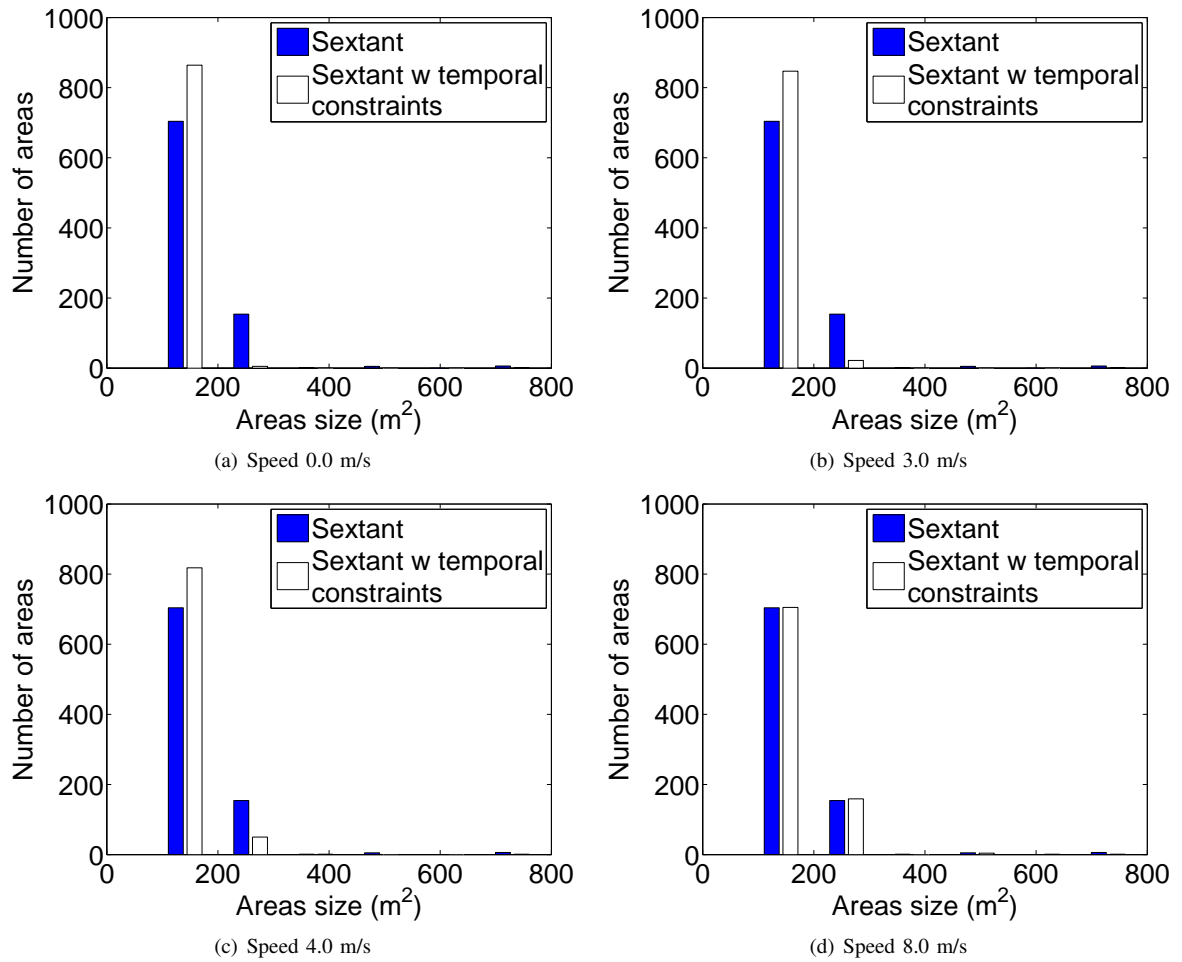


Fig. 15. Contact Trace: Areas size distributions