Performances of geographical routing protocols combined with a position estimation process in wireless heterogenous networks

Erwan Ermel, Anne Fladenmuller, and Guy Pujolle

Laboratoire d'Informatique de Paris 6
Université Pierre et Marie Curie
Paris, France
{`erwan.ermel@, anne.fladenmuller@, guy.pujolle@`}`lip6.fr`

Abstract. This paper addresses the performance of geographical routing protocol in wireless networks, where only few nodes possess self-locating capability such as GPS. To be able to apply end-to-end geographical routing protocols, it is necessary every node know their position coordinates. We propose a method to infer such positioning information to any node, based only on connectivity and localization information obtained from the neighborhood. Three metrics are used to evaluate the performance of such a scheme: the density of useful nodes for geographical routing protocol, the reachability and the path length.

***Index Terms*— ad hoc network, localization, positioning, wireless network, self-locating.**

1 Introduction

Wireless systems are widely spread all around the world, but many problems remain to offer ubiquitous access. In internet networks, the routing process is based on the IP address of a destination, which uniquely identifies a node and contains a network prefix used to locate the node in the network. In case of mobile user, the node keeps its IP address but the network prefix looses its significance. This problem becomes even more crucial in Ad Hoc or Sensor Networks, as no dedicated routing infrastructure exists, leaving each node in charge of its auto-configuration and of routing issues. As the number of wireless terminals grows, the complexity of the routing process also increases and the main difficulty for such systems is to remain scalable.

One proposition to deal with this scalability issues consists of using a geographical routing protocol. Such a protocol bases the routing decision only on the geographical position - coordinates - of the destination node, thus suppressing the need for routing tables.

For example, the nearest node to the destination is chosen as the next hop node. These routing alternatives raise their own problems such as how to store and retrieve the mapping between an IP address - still required by applications - and the localization coordinates. Many research work is currently undertaken to overtake this problem, a survey proposed in [1] references some of the papers dealing with localization issues for ubiquitous computing. Another issue raised by geographical routing is that it requires that every node knows its position, therefore they can not be integrated by the geographical routing process.

Most of the papers in this research field deals with geographical routing in an homogenous networks, where all nodes have got self-location capability - like GPS - [2, 3]. We argue that this assumption is somehow restrictive as a wireless network is more likely to be composed of heterogeneous than homogenous devices.

For these reasons, we focus our work on evaluating the performances of geographical routing protocols in heterogeneous wireless networks. *Simple nodes* - nodes with no self-locating capability - could not be used in such scheme until they can estimate their own position. Some approaches try to overcome this problem by using IP routing solutions when nodes don't know their position [4]. This nonetheless leads to mix routing protocols which may not be adapted to low resources Sensor networks. Other proposals get the nodes to estimate their position based on physical measurements such as an angle of arrival (AOA) [5], signal strength [6], time of arrival (TOA) [7] or based on connectivity based approach [8]. The latter proposal has for main advantage that no specific devices is requested to infer a position.

We argue in this paper that although the position precision obtained when nodes get their position from their neighborhood is not always great, it is sufficient to operate routing operations.

Thus we combine geographical routing protocols with a simple position estimation process, based on a connectivity approach.

This paper is organized as follows : Section 2 will present our approach followed in Section 3 by our results. Section 4 will conclude the paper.

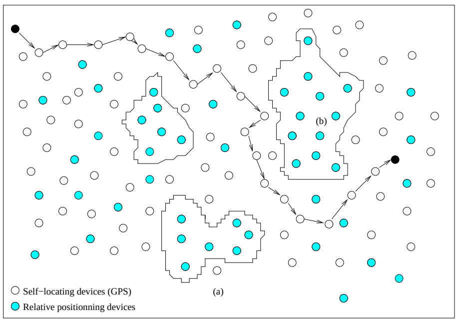
2 Context

We propose in this article to evaluate the performances of geographical routing protocol combined with our localization process for *simple nodes* (Fig. 1).

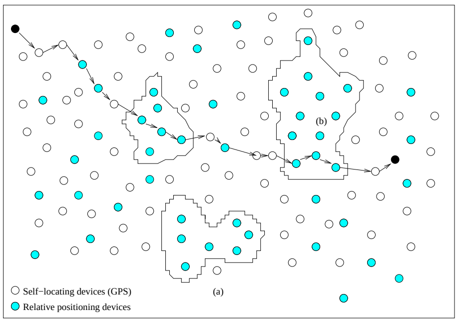
A greedy approach has been chosen for the geographical routing protocol, like [2] due to its simplicity: the next forwarding node is the nearest in distance to the destination.

The position estimation process is based on the connectivity approach [9]. With a convex hull selection method among the neighbor nodes, only useful nodes remain for the position estimation process. The estimated position is then a simple average of the position of selected nodes.

The main contribution of this position estimation proposal comes from the fact that we do not require any specific equipment to get a node to estimate its position. To run our solution we require that only a few nodes in the topology



(a) Simple nodes with no position estimation process



(b) Simple nodes with a position estimation process

Fig. 1. Geographical routing in heterogeneous wireless networks.

have self-locating capability. The position information of a node (coordinates and the estimated accuracy of the position) is then simply exchanged in “Hello Messages” adding no particular overhead in terms of the information propagation. Each time a node receives a “Hello Message”, it decides or not to update its position and it forwards it to its own neighborhood. This process is done continuously to take into account the modifications of the environment. It allows step-by-step to infer a position to a node distant of x-hops from self-located nodes - the self-located nodes being the only ones with a very precise position information. The position estimation process is independent of the underlaying network technologies and protocols. Thus no further considerations will be made on the underlaying routing protocols in this paper.

In some previous work [10], we have shown that our position estimation proposal gives optimal results when selecting some neighbors. Obviously, our resulting estimated position is not always precise, as it depends on the network topology and the distribution and number of self-located nodes in the network.

The aim of this paper is first to evaluate the convergence of position information based on the number of time “Hello Messages” are exchanged. First we want to confirm that the estimation process will tend to become more and more precise as the number of "Hello Messages" propagated in the network increases. This point is not obvious, as divergences could appear if estimation errors propagate in the network. Second, we want to evaluate the impact of the topology density on the node reachability depending on the number of self-located nodes in the network.

2.1 Convergence time of the position estimation

We first consider a simple static case to evaluate the convergence time of the position estimation process (Fig. 2). We define five nodes $N1, N2, N3, N4, N5$, where only $N1$ and $N5$ are self-located. The real positions of the nodes are $N1_x(0)$, $N2_x(45)$, $N3_x(70)$, $N4_x(100)$ and $N5_x(130)$. R_{max} , the theoretical transmission range is set to 50m. $N5_x(130)$ emits its first "Hello message" at iteration 3.

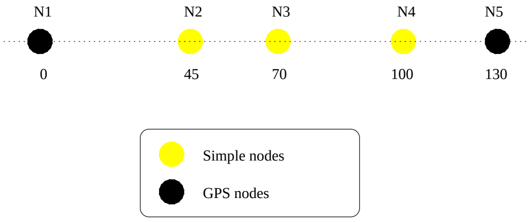


Fig. 2. Simple topology case.

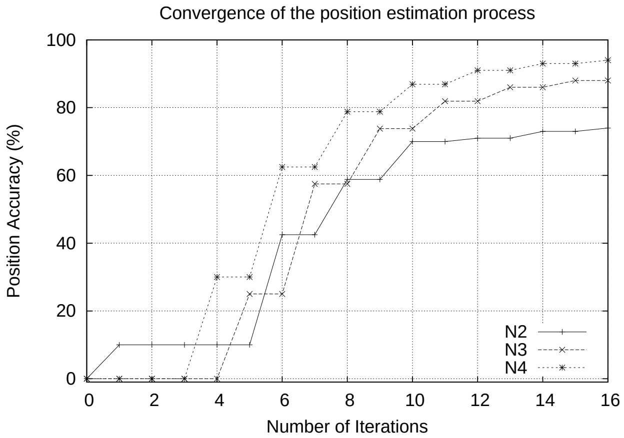


Fig. 3. Convergence time of the position estimation process

Fig. 3 shows the convergence time of the position estimation process combined with a geographical routing protocol. From 0 iteration to 3which also correspond to the number of position messages exchanged or broadcasted by the neighbor nodes - only $N1$ broadcasts its “Hello Message”. Simple nodes $N2, N3, N4$ can’t estimate their position until $N5$ broadcast its ‘Hello Message” at iteration 3. in this example a simple node has a good estimation of its positions after 4-5 iterations - and the maximum precision of position is reached after 10 steps.

Before reaching the stationary state, the precision of position increases by steps. For example, between iteration 3 and 4, nodes $N2$ and $N4$ get position messages from $N1$ and $N5$ while $N3$ receives nothing. $N2$ and $N4$ estimate their position and then broadcast it. Then $N3$ estimates its position between steps 4 and 5, and broadcasts the result. And so on...

We showed that with a simple position estimation process, simple nodes can estimate their position with a simple connectivity approach.

Even if this example is basic, we observe that the convergence time of the precision of position is linked to the number of self-located nodes present in the neighborhood: as the density of SLN nodes increases, the convergence time decreases.

2.2 Definitions

We will now define the metrics we use to evaluate the performance of a geographical routing protocol combined with a position estimation process.

Useful density We define the useful density $\mathcal{D}_{useful}$ as the density useful for a geographical routing protocol i.e. only nodes with a position are taken into account. Thus $\mathcal{D}_{useful}$ of a node represents the number of SLN nodes and the number of SN nodes with a position in its neighborhood.

We also define $\mathcal{D}_{max}$ the maximum density of useful nodes. $\mathcal{D}_{max}$ is obtained when all the nodes are self-located i.e. 100% of the nodes are GPS nodes.

Accessibility As simple nodes gets a position, new paths can be found and used by the geographical routing protocol. Thus, a node can “discuss” with more nodes. We call *accessibility* or *reachability* the percentage of nodes that can be reached by another one with a geographical routing.

To estimate $\mathcal{R}$, we divide the number of real paths found by the theoretical number of paths. By construction $\mathcal{R} \in [0; 1]$ with $\mathcal{R} = 0$ when no path has been found and $\mathcal{R} = 1$ when the network is a connected graph. The higher $\mathcal{R}$ is, the higher the connectivity of the network is.

Path length We define by *average path length* $\mathcal{L}_{mean}$ the average number of hops necessary to establish a path between two nodes. If no path is found, it will not have a length since it does not exist and it will be discarded for the estimation of $\mathcal{L}_{mean}$.

Homogeneous and Heterogeneous routing We define two geographical routing protocol cases : the first one consists of a geographical routing protocol without a position estimation process. In such case, only SLN nodes are useful in the routing process, and SN nodes are purely ignored. We called such an approach a *homogeneous geographical routing protocol.* Inversely, when a geographical routing protocol is combined with a position estimation process, all nodes of the network can be used to achieve the routing and we call this approach a *heterogeneous geographical routing protocol.*

3 Results

As our position estimation process is independent from the underlayer technologies, we choose not to use NS2 or Glomosim and we have developed our own Java code simulator for its simplicity. 50 nodes are randomly placed in a 1000m x 1000m square. Self-locating nodes and simple nodes are also randomly elected. The maximum theoretical transmission range R_{max} is set to 170m. We consider only static cases. Mobility will be considered in future works.

3.1 Useful density and accessibility

We first study the impact of the percentage of SLN nodes from 0 to 100% for a 50 nodes topology (Fig. 4).

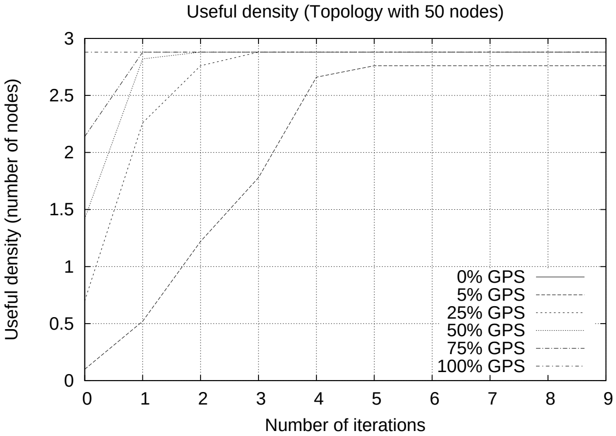


Fig. 4. Performance of $\mathcal{D}_{useful}$ for different percentage of GPS nodes and as a function of the number of iterations of the position estimation process

With 100% of SLN nodes, the useful density reaches the maximum density with $\mathcal{D}_{max} = 2.8$ (every node has got 2.8 neighbor nodes). Fig. 4 shows that as the percentage of SLN nodes increases, the convergence time of $\mathcal{D}_{useful}$ decreases.

With 5% of GPS nodes (2 SLN), $\mathcal{D}_{useful}$ never reaches $\mathcal{D}_{max}$ mostly due to the sparse topology. Several nodes are alone and won't be able to participate to the network and then won't be useful at all.

Focus now on the worst possible case : low density (50 nodes) with a low percentage of SLN nodes (5%) (Fig. 5).

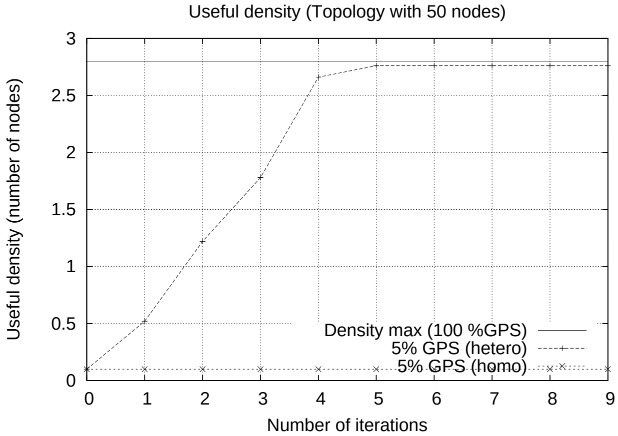


Fig. 5. Performance of $\mathcal{D}_{useful}$ for a topology of 50 nodes with 5% GPS

For a homogeneous routing, the number of useful nodes is equal to 0.1 and doesn't vary in time (the number of SLN is constant), whereas in the case of a heterogeneous routing, $\mathcal{D}_{useful}$ increases in time to reach a maximum of 2.8. The $\mathcal{D}_{useful}$ of a heterogeneous routing converges toward the maximum density $\mathcal{D}_{max}$ but never equals it due to the pathology of the network topology. Nevertheless heterogeneous routing performs better than homogeneous routing.

With a position estimation process, a geographical routing protocol can use more nodes to achieve its goals than in the case of using only SLN nodes.

We now study the reachability of a node into a heterogeneous network. This is an important performance factor because it determines if a path exists between every node of the network. $\mathcal{R}$ is a good indicator of the connectivity of the network.

Fig. 6 shows the impact on the reachability according to the number of iterations of the position estimation process. We varied the percentage of GPS nodes from 0 to 100%.

The study of the Fig. 6 is very similar to that which we have just done for the density Fig. 4. As the number of iterations of the position estimation process increases, $\mathcal{R}$ grows and converges towards a maximum value depending

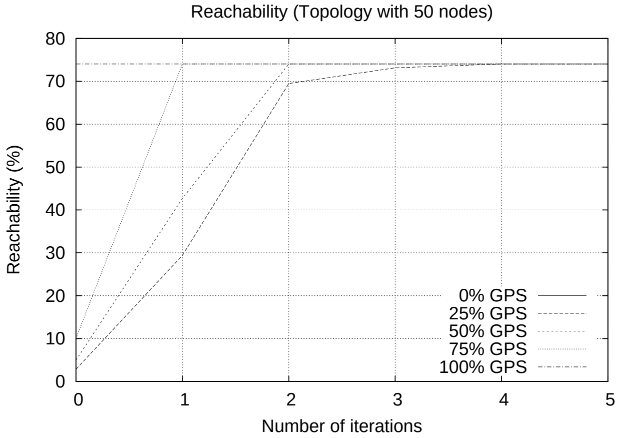


Fig. 6. Reachability for a given percentage of GPS nodes and as a function of the number of iterations of the position estimation process.

on the density of the topology. As the simple nodes estimate their position, the useful density grows, and as the useful density of nodes increases, the feasibility of roads increases as well.

As the simple nodes acquire a position, the connectivity of the network increases to reach a maximum. Here $\mathcal{R}_{max}(50) = 73\%$. We also see that the more the percentage of self-locating nodes in the network is, the faster the convergence is. $\mathcal{R}(50) < 100$ implies that the graph representative of the network is not complete but composed of under complete graphs.

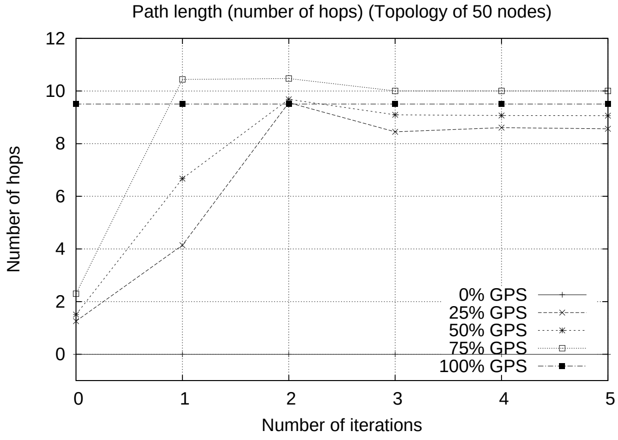
As the useful density, the reachability (accessibility of the nodes) benefits from the packing of a position estimation process to a geographical routing protocol.

3.2 Path length

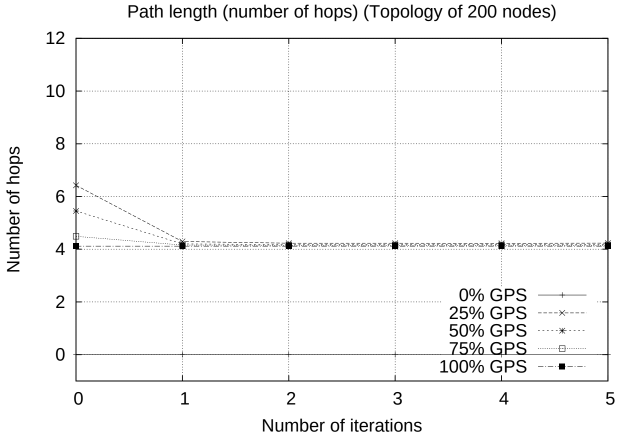
The figure Fig. 7 shows the average path length according to the number of iteration of the position estimation process. We looked at $\mathcal{L}_{mean}$ in topologies of 50 and 200 nodes. For each particular topology, we varied the percentage of SLN from 0 to 100%.

These curves comprise two phases: one *transitory* and the other one *stationary*:

- The stationary phase is reached when $\mathcal{L}_{mean}$ becomes constant. It is observed in every case.
- The transitional phase is the stage before the stationary phase. A difference of *pathology* of the results can be noticed according to the total number of



(a)



(b)

Fig. 7. Average path length in function of the percentage of SLN nodes and in function of the number of iteration od the positions estimation process.

nodes of topologies. Two types of transitional stages are defined: the *progressive transitory phase* and the *regressive transitory phase.*

The progressive transitional stage is observed when the theoretical density and the percentage of SLN nodes are low. In this case, the average path length increases quickly as the density and/or the percentage of SLN nodes is low. Increasing connectivity increases the probability of existence of a way between two nodes. In a spare topology case (50 nodes), the increase of connectivity is the result of the union of under complete graphs, as the simple nodes estimate their position. Thus the average path length can only be longer than before. But that also implies that news paths can be found to reach new destinations.

The regressive transitional stage is observed when the theoretical density and/or the percentage of SLN nodes are significant enough. In this case, $\mathcal{L}_{mean}$ decreases gradually towards the optimum path length. This decrease is all the more significant as the percentage of nodes GPS is weak. These cases of figures are the concrete example of a *Swiss-Cheese topology* as shown in Fig. 1(b).

At the beginning, when the simple nodes do not have a position, the geographical routing can not use them: the routing path has to circumvent all the empty zones of position information, which implies a longer path. As soon as simple nodes acquire their position, geographical routing protocol uses these nodes. Thus the *hole* grow blurred. Fig. 1(a) and Fig. 1(b) illustrate these remarks.

Now let us compare the average path length in the case of a homogeneous and/or heterogeneous routing.

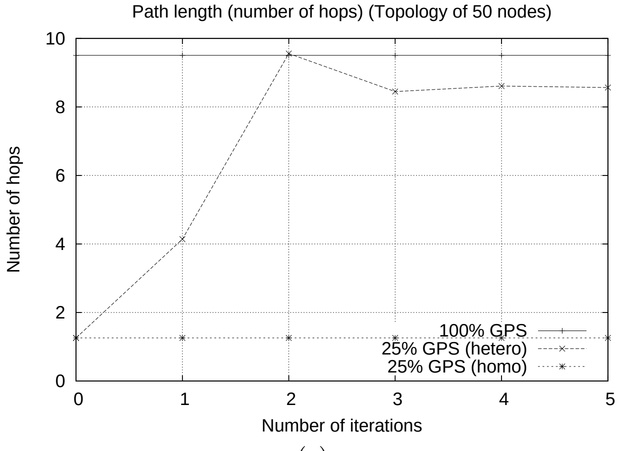
The Fig. 8 compares $\mathcal{L}_{mean}$ obtained by a homogeneous and heterogeneous geographical routing. Moreover we have to study $\mathcal{L}_{mean}$ in the cases of progressive transitional stage Fig. 8(a) and regressive transitional stage Fig. 8(b).

In the case of a progressive transitional stage - Fig. 8(a) - the average path length in a homogeneous routing is constant. Within a topology of 50 nodes, $\mathcal{L}_{mean} \simeq 1.2$. Thus the average path length is no more than a hop. Routing in such topology is not very useful since a node can discuss only with its immediate neighbors and no further.

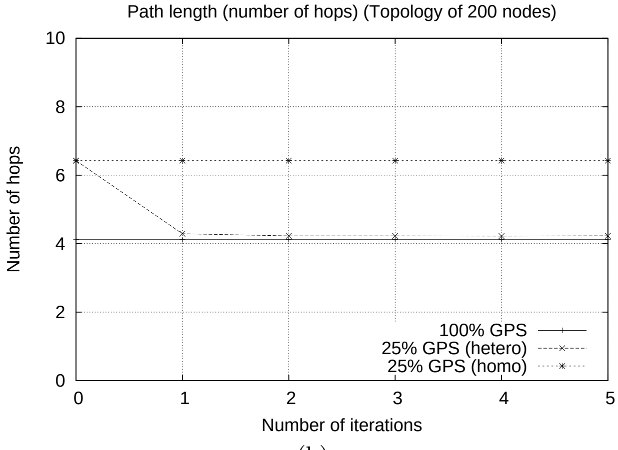
When simple nodes have a position, an important growth of the path length is noticed; $\mathcal{L}_{mean} \simeq 7$ is almost 6 times greater than for the homogeneous routing. Indeed this increase average path length is only a translation of the increase in probability of finding a road between two nodes. Thus when two under complete graphs are linked - by the means of simple nodes - the average path length increases.

In the case of a regressive transitional stage - Fig. 8(b) - the average path length in a homogeneous routing is constant, $\mathcal{L}_{mean} = 6.5$, whereas for a heterogeneous geographical routing, $\mathcal{L}_{mean}$ decreases quickly before being stabilized towards 4.3. These results confirm our idea to use simple nodes with position into geographical routing in order to largely minimize the path length.

In these two figures - Fig. 8(b) and Fig. 8(a) - the *theoretical curve* represents the theoretical average path length between different nodes. In the case of a progressive transitional phase, the theoretical average length is the upper limit, whereas in the decreasing transitional phase, it undervaluates the lengths obtained by the mean of the homogeneous and heterogeneous geographical routings. In all cases, the heterogeneous routing converges towards this theoretical



(a)



(b)

Fig. 8. comparison between average Length of roads in a number of jumps of a heterogeneous routing and a homogeneous routing for two topologies of 50 and 200 nodes including 25% of which autolocalized nodes.

length, without however reaching it. The path length, when using our approach becomes closer to the theoretical path length when the density of the nodes increases.

4 Conclusion

The addition of a position estimation process improves the total performances of a geographical routing in a heterogeneous wireless network.

We showed that despite the fact our position estimation process may not be precise, it allows ti increase nodes connectivity and reachability. The average path length reflect the benefices of such a routing for a heterogeneous network. This work underlines the possibility to use a simple geographical routing protocol, in an environment with a very little constraints, as it only requires that some nodes in the topology possess a precise position information.

We will focus our future works on the position estimation process and specially simple techniques to enhance the precision of estimated position to study the impact of the position precision on the performance of geographical routing for a heterogeneous network.

References

1. J. Hightower and G. Borriello, "Location systems for ubiquitous computing," *IEEE Computer*, vol. 34, pp. 57–66, August 2001.
2. B. Karp and H. T. Kung, "Gpsr : Greedy perimeter stateless routing for wireless networks," in *Proceedings of ACM/IEEE MOBICOM'00*, Aug. 2000.
3. Y. Ko and N. Vaidya, "Locaton-aided routing (lar) in mobile ad hoc networks," in *Proceedings of ACM/IEEE MOBICOM'98*, pp. 66–75, Aug. 1998.
4. S. Capkun, M. Hamdi, and J. P. Hubaux, "Gps-free positioning in mobile ad-hoc networks," *Cluster Computing*, vol. 5, April 2002.
5. D. Niculescu and B. Nath, "Ad hoc positioning system (aps) using aoa," in *Proceedings of IEEE INFOCOM'2003*, April 2003.
6. C. Savarese, J. M. Rabaey, and J. Beutel, "Localization in distributed ad-hoc wireless sensor networks," *Proceedings of the ICASSP*, May 2001.
7. A. Savvides, C.-C. Han, and M. B. Strivastava, "Dynamic fine-grained localization in ad-hoc networks of sensors," July 2001.
8. L. Doherty, K. S. J. Pister, and L. E. Ghaoui, "Convex optimization methods for sensor node position estimation," in *Proceedings of IEEE INFOCOM'2001*, (Anchorage), April 2001.
9. E. Ermel, *Localisation et Routage géographique dans les réseaux sans fil hétérogènes.* PhD thesis, Laboratoire Informatique de Paris VI, Université Pierre et Marie Curie (Paris VI), France, 2004.
10. E. Ermel, A. Fladenmuller, G. Pujolle, and A. Cotton, "Improved position estimation in wireless heterogeneous networks," in *Networking 2004*, May 2004.