1. INTRODUCTION

Since the development of the Bluetooth Low Energy (BLE) standard, this technology has been constantly gaining attention in various fields, such as Health-care, Home Automation, Internet of Things, etc. However, yet another BLE application of great importance would be in indoor positioning systems. Such a system typically suggests the deployment of a network of broadcasting nodes and a receiver that is able to “listen” to the transmitted signals. Then, depending on the positioning technique being used (triangulation, trilateration, fingerprinting, etc.), these signals are processed to finally produce an estimation of the receiver’s position. As it has been noted (Faragher, Harle, 2015), Received Signal Strength (RSS) fingerprinting is the de facto localization technique for indoor positioning on consumer devices today. This specific approach involves matching sensed patterns with already known ones that have been georeferenced and stored within a database. It received this specific name, since these patterns are as unique as a fingerprint can be.

Every mesh/network deployment could be evaluated based on some performance metric. On the contrary, knowing this metric beforehand could enable the deployment of the network in a way that its performance becomes optimal. An example of such an optimization can be the distribution of cellular antennas in a way that maximizes the total coverage or the deployment of Wi-Fi APs in a university so that the disconnections of walking users are minimized. In the case of a network of 2.4GHz nodes (e.g. Wi-Fi APs or BLE Beacons) used for fingerprint-based positioning, this performance is directly related to the accuracy of the positioning estimations that the system can offer.

A considerable amount of research has already been done for finding the optimal placement that would maximize the performance of an indoor positioning system like the aforementioned one. Although in each case, the optimization objective varies, two general approaches can be distinguished. In the first one, the goal is to minimize the expected error of the positioning estimation (Baala et al., 2009; Sharma et al., 2010; He et al., 2011; Picco et al., 2013; Li et al., 2015; Laitinen; Lohan, 2016; Voronov, 2017). However, it can be argued that making an accurate prediction of the error of a positioning estimation is quite challenging due to the complexity of properly modelling the error’s sources themselves. The second approach is based on the perception that in fingerprint-based positioning applications the more discrete the estimations are the better, since statistical uncertainty cannot be avoided. According to that, the optimization process tries to maximize the vector distance of RSS fingerprints in the area of interest. Although this idea has been favoured by the latest papers on this field (Meng et al., 2012; Chen et al., 2013; Chen et al., 2014; Du; Yang, 2017; Alsmady; Awad, 2017; Eldeeb et al., 2018), it may still have some points of criticism, mainly related to the way this distance is measured. It is worth mentioning that in many cases, the objective of the studies was to also optimize the number of nodes needed to be deployed (and thus the installation cost) (He et al., 2011; Picco et al., 2013; Chen et al., 2014; Li et al., 2015; Laitinen; Lohan, 2016).

Thus far, the leading interest of the associated research has been essentially the improvement of the general accuracy of the positioning estimations in terms of numerical coordinates, or in other words, the minimization of the difference between the coordinates of the estimated position and the actual position.
However, in practice, a fingerprint-based indoor positioning system can often offer only broader area/proximity estimations, instead of specific positioning estimations of high accuracy. In such systems, where basically area or zone approximations are being offered, the benefits of the so far proposed optimizations are not maximal, since the optimization process will highly be consumed at enhancing aspects that have limited effect. Therefore, in this paper, a new optimization methodology targeting those exact systems is proposed. More specifically, a methodology to adjust the placement of positioning nodes used in fingerprint-based indoor positioning systems, with the goal to increase the location prediction among different area sections (e.g. room, building, etc.). Besides the improvement in both the localization accuracy and navigation functionality, such an optimization also increases the cost-effectiveness ratio. That, can ultimately lead to lower deployment costs, since less transmitting nodes might be needed to reach sufficient levels of performance.

The paper continues having the following structure: Section 2 discusses the notion of zone partitioning in an indoor positioning system, along with how to suitably model the indoor space for such an optimization. Section 3 presents the performance metric that will be used to evaluate each different deployment scenario. Section 4 describes how this process may be executed for ultimately finding the optimal node placement. Finally, Section 5 will present our conclusions and also suggestions for future work.

2. MODELLING THE INDOOR ZONES

Often, when the subject of study is to identify where an entity is situated within space, the leading interest is essentially the numerical coordinates of its physical position in some reference system. Although this notion is highly applicable to various scientific fields (e.g. Surveying Engineering, Radio Navigation, GNSS Tracking, etc.), there are still cases such as Indoor Positioning and Navigation, where a numeric position may need some spatio-symbolic enrichment before it becomes valuable. This importance and generally the difference between localization and positioning aspects, has been acknowledged even from plainly technical sources (Karl, Willig, 2005).

Even for indoor positioning systems offering highly accurate positioning coverage (e.g. sub-meter), it is easy to depict the value of grouping different points in space, into distinct spatial sets (or zones) of specific semantic properties. For example, a university student searching for the "Lecture Hall B3", would prefer making a lookup based on the room’s name in a hypothetically provided indoor positioning system App, instead of some specific coordinates. In a similar way, other users having mobility impairments would recognize the worth of an indoor positioning system that supports space semantics as described by Liu et al., 2019, to be able to search for navigation routes via zones that are accessible by them.

In a zone-aware indoor positioning system, the more correctly the system can estimate the zone within which an entity is located, the better this is for its performance. This, requires the radio identity of each zone to be as distinct as possible and so, proposing an optimization method to achieve that is, essentially, the scope of this paper. Such an optimization mechanism has to begin with properly modelling the indoor environment, along with the zones of interest; a space subdivision process, which is a known problem in literature and has comprehensively been discussed (Worboys, 2011, Zlatanova et al., 2014, Diakité, Zlatanova, 2018). With respect to that, there are two aspects needed to be considered. The geometry part, which is required for the signal propagation simulation and the semantics part, which will define the zones.

Starting with the first aspect, a major decision needs to be taken regarding the dimensions of the model. As mentioned, to maximize the radio distinctiveness among different zones by adjusting the node placement, one needs to be able to simulate the radio propagation within the indoor environment. Therefore, since an accurate radio propagation model requires the utilization of an accurate representation of the propagation space, the more detailed this indoor model is, the better. In theory, a point cloud based 3D model that would include even furniture surfaces, would perform the best. However, since the model its complexity affects highly the speed of the optimization, a more efficient approach is needed.

In a similar work of node placement (Dalla’Rosa et al., 2011), where both 2D and 3D indoor models where examined, it was shown that the results between the two cases were similar. However, the 3D case took (for a small model) 500% more time, while additionally, this percentage gets exponentially higher as the model gets enlarged. Therefore, we opt to use a 2D indoor model. Nevertheless, this also means that we neglect the exact geometry of any windows, doors, or half wall openings, and that, should be noticed.

The need to decompose the geometry of the 2D model defined above, into different zones, introduces the second aspect; the semantics. Dividing a small indoor-space (e.g. a house) into distinct zones, might sound intuitively straightforward. For example, one could distinguish a living room, a kitchen, a bathroom, etc., being separated by walls. However, what happens if walls were not present (e.g. a kitchen being connected with the living room, with no walls in between)? A problem becoming even more evident as the area increases (e.g. airports, train stations). At the same time, quite often we might be interested in merging sections that are physically separated. For example, a museum might want to cluster different rooms into thematic zones (e.g. Paleolithic, Mesolithic, etc.). On the other hand, the zone subdivision process may also include some constraints. Therefore, considering that the zoning process may not always be straightforward, it needs to be defined.

With respect to the aforementioned, we apply the following 3 rules:

- **Zones must not overlap:** A physical position or area should not belong to different zones, since that would contradict with the zone distinctiveness notion.

- **The interior of each zone must be continuous:** Having zones that are discontinuous introduces impracticality to their utilization and thus, it should be avoided.

- **Zone’s borders must be perpendicular to the reference grid:** During the optimization process, the indoor model (i.e. the obstructions along with the zones) need to be spatially indexed into a reference grid. Ensuring that the borders of the zones are perpendicular (or parallel) to the axes of this grid (Figure 1) is crucial to the optimization’s speed, due to the reduction of the geometrical calculations needed to be done during the radio propagation modelling.
3. DEFINING THE PERFORMANCE METRIC

To improve the accuracy of the zone predictions, one needs to consider the way this prediction is made. A positioning algorithm based on RSS-fingerprints is typically a distance check between a new (unclassified) set of RSS measurements and a number of other (classified) sets of RSS measurements, which have been gathered during a training phase. Each one of these sets can be considered as a vector of RSS values and thus, this distance is essentially a vector distance. In an indoor positioning system that offers awareness of the location (or zone) to which a physical position belongs (e.g. kitchen, corridor, etc), all vectors corresponding to the same zone, form a single distinct class. This notion can be illustrated through the following figures, where an example of a 2-node setup within an indoor environment has been used to illustrate how these vectors of RSS values (in symbolic units), form the different classes. More specifically, the left part of Figure 2 shows 4 different zones being divided by a grid of sub-space cells, having a total resolution of 10x10. At the corners, 2 transmitting nodes (blue & red) have been installed and their radio coverages have been simulated (based on a simplistic radio propagation model) and presented on the right parts.

At each sub-space cell, the combination of the signals produces a distinct vector of 2 RSS values. Plotting these vectors results in Figure 3, where each dimension corresponds to a single node. Therefore, illustrating a 3-node setup would result in a 3-dimensional graph, while a bigger setup would require a hyper-dimensional representation. Grouping all RSS vectors by their zones can help us determine the different class regions. In Figure 3 these can be found colored respectively using an approximated alpha shape. Every point within a shape belongs to the corresponding class (i.e. zone), however, all points outside these shapes (separation space) belong to no class.

To accurately model a radiomap is quite difficult due to random noise and that, has direct impact to the accuracy of the class borders. With that said, let us assume that the cell grid became continuous (which is the reality) and that we measured the RSS vector at a new physical position within a specific zone. Then, the probability that this vector would lie within the corresponding class region (in RSS vector space), would be higher if the physical position was in the center of the zone, and not its borders. On the contrary, if its physical position was close to the zone’s borders, there is even a chance that this vector would now lie closer to a neighbour class, which fundamentally characterizes the difficulty in classification. Ultimately, the more separated these class regions are (Separation Area in Figure 3), the less this problem exists and so, this metric can be used to measure the distinctiveness among different zones.

This contribution has been peer-reviewed.

An indoor positioning system may often involve hundreds of reference nodes and so, determining the n-dimensional separation area, would require a series of highly demanding hyper-shape calculations. Consequently, an alternative approach which would still respect the class-separation notion, is also proposed. Namely, the minimum (to ensure sufficient levels of accuracy per zone) and the combined separation distances, which are presented in Figure 4. These distances can be considered as the class interconnections, having direct correspondence to the borders of the zones in the physical propagation environment (Figure 5).

4. OPTIMIZING THE NODE PLACEMENT

Having defined the performance metric for measuring the zone distinctiveness, the next step is to execute the optimization process, during which, various node placement scenarios will iteratively be evaluated to ultimately select the one offering the maximum performance. This process requires the utilization of a) a radio propagation model for simulating the RSS values within the indoor environment and b) an optimization function to control which node placement scenarios to evaluate. In each case, several options exist, offering different trade-offs between computational complexity and accuracy. Nevertheless, this section presents our recommended approaches.

4.1 Simulating the Signal Propagation

One of the most crucial parts of the optimization mechanism is the radio propagation model, since its accuracy is directly affecting the final optimization’s worth. Self-evidently, a generic radio propagation model (e.g. free-space path loss) that does not consider the physical obstructions in the propagation environment, would not suffice for accurately simulating the RSS values within that. The radio energy is highly affected by phenomena such as reflection and absorption (caused by obstructions like walls and doors) and so, a deterministic radio propagation model is required, taking these into account. Therefore, we propose the utilization of a ray launching model (Luo, 2013), since it offers the best trade-off between computational complexity and accuracy for this demanding (due to its iterative nature) process.

Based on the Geometrical Optics phenomena of reflection and absorption, ray launching considers analytically both the electromagnetic properties and the propagation environment. More specifically, assuming that all nodes transmit omnidirectional, a sufficient number of rays is evenly (in terms of angle) generated and traced (Figure 6), for estimating the associated power fields at every sampled cell. The term sufficient is used to denote the importance of delivering ultimately (even after many reflections and absorptions) the generated ray to every cell, which in reality would indeed receive the corresponding signal. In the end, the attenuation of each ray will be the result of a) the distance path-loss during its propagation in free space, b) the attenuation due to reflections and c) the attenuation due to absorptions.

During the simulation, the attenuation coefficient of each obstruction type needs to be used. Although generic estimations can be found in literature, an even better approach would be to compute the optimal ones, based on the specific characteristics of the propagation environment. To achieve that, one could deploy first several nodes across the area at known positions \( P_0 \) and then, sample at known positions \( P_1 \) their signal strengths...
After finding the optimal coefficients for the radio propagation model, the next challenge is to utilize it efficiently. An optimization scenario could well exceed trillions of ray intersection checks in total. Therefore, the more these checks are reduced without sacrificing the accuracy, the better. For that, several enhancing approaches exist, such as Spatial Indexing (e.g. this obstruction lies within those cells or, vice versa, this cell includes those obstructions), visibility pre-calculations, hierarchical clustering, reduction in the resolution of the grid, parallelization of the process, etc..

### 4.2 Selection of the Optimization Function

Having defined a) the metric to evaluate the performance of each different node placement scenario and b) the radio propagation model needed for the simulation of the signal propagation within the indoor environment, the last step is to choose a function to control and drive the whole optimization process by selecting which specific placement scenarios to examine. Undoubtedly, simulating in a brute-force approach, all possible cases, would certainly return the best scenario. In practice, however, such a computational load renders this approach highly impractical. For example, if we had 100 different nodes, there are countless combinations of how these could be deployed within a building. Therefore, decreasing the number of the evaluation cases is crucial for this optimization problem.

While various optimization algorithms exist that are suitable for solving the problem of an optimal node placement (e.g. Ant Colony Optimization, Particle Swarm Optimization, Simulated Annealing, etc), the ones that have been recognized and applied the most, are the Genetic Algorithms (GAs). In our case, they remain as the suggested approach. Although GAs have been comprehensively discussed in literature when utilized for node placement optimizations (Yoon, Kim, 2013), it can be said that their general goal is to translate the principles of Charles Darwin’s natural selection, into an iterative procedure for solving the optimization problem.

As shown in Figure 7, this procedure is mainly the repetition of 3-steps: the selection, crossover and mutation steps. Initially, a population of individual chromosomes (or optimization solutions) is generated. Then, the strongest chromosomes (or best solutions) are selected in order to be preserved or mixed in pairs, producing the next generation of chromosomes. Some of these new chromosomes are then randomly mutated (producing again a slightly different solution) to ensure that the vast search-space is explored better. This 3-step process is then repeated, until some threshold is reached.

### 5. CONCLUSIONS

A new performance metric has been proposed that can be used to evaluate and increase the performance of an indoor position system, where Received Signal Strength fingerprinting is used as the localization technique. Since the scope of this paper was primarily to suggest a general methodology on how to use this metric to perform such an optimization, its evaluation is still needed. For that, we could compare the localization accuracy between 2 different node deployment scenarios. The first one would be based on the optimal solution that the proposed optimization methodology would produce, while the second one would be based on an unbiased regular node deployment, or even based on the intuition of an already experienced installer.

This new metric has been formulated according to the usual way the signals are utilized during the localization process. Therefore, the key factors affecting the optimization performance the most, are expected to be a) the accuracy of the used indoor model and radio propagation model and b) the exact implementation of the optimization function. These are subject to individual research and stand as ideas for supplementary future work. Nevertheless, besides the evaluation and the improvement of the optimization performance in terms of speed and accuracy, other suggestions for future work include the introduction of weights during the zoning process and the reduction of the number of nodes needed, until sufficient levels of performance have been achieved.
REFERENCES


Chen, Guofeng, Zhang, Yan, Luan, Fengyu, Xiao, Limin, Gosier, Guadeloupe, 12–17.


OGC, 2016. OGC IndoorGML Corrigendum.


Voronov, Roman Vladimirovich, 2017. The problem of optimal placement of access points for the indoor positioning system. 13, 61–73.

