Statistical Significance Analysis of the *R-DPSO*

Towards an Understanding of the Relationship Between the Population of Robots and the Maximum Communication Distance

Micael S. Couceiro, Nuno M. F. Ferreira
RoboCorp, Department of Electrical Engineering (DEE)
Engineering Institute of Coimbra (ISEC)
Coimbra, Portugal
e-mail: {micael, nunomig}@isec.pt

Fernando M. L. Martins
RoboCorp, Coimbra College of Education (ESEC)
Institute of Telecommunications (IT)
Coimbra, Portugal
e-mail: fmlmartins@esec.pt

Rui P. Rocha
Institute of Systems and Robotics (ISR)
University of Coimbra (FCTUC)
Coimbra, Portugal
e-mail: rprocha@isr.uc.pt

Filipe Clemente
RoboCorp, Faculty of Sport Sciences and Physical Education, University of Coimbra (FCDEF)
Coimbra, Portugal
e-mail: fclemente@esec.pt

***Abstract*— This paper presents a statistical significance analysis of a modified version of the Particle Swarm Optimization (*PSO*) on groups of simulated robots performing a distributed exploration task, denoted as *R-DPSO* (Robotic *DPSO*). This work aims to evaluate this novel exploration strategy studying the performance of the algorithm under communication constraints while increasing the population of robots. Experimental results show that there is no linear relationship between the number of robots and the maximum communication range. In general, the decreased performance by the developed algorithm under communication constraints can be overcome by slightly increasing the number of robots as the maximum communication range is decreased.**

Keywords: multi-robot systems; statistical analysis; swarm robotics; analysis of variance.

I. INTRODUCTION

The biological world abounds in collective phenomena that have important adaptive functions, ranging from coordinated movement to nest building and all the way to communication [1]. The principles of self-organization are appealing for explaining biological collective phenomena where the resulting structures and functionalities greatly exceed in complexity the perceptual, physical, and cognitive abilities of the participating organisms. Examples of biological self-organization include the construction of beehives, the foraging strategies of ants, and the regulation of colony life in social insects. In all these cases, the resulting structure emerges from the collective work of individual organisms that execute simple behaviors based on local information and do not possess a global plan of the end result or a central coordinator.

The examples of behavior-based collective approaches described above inspired the design of novel machine-learning techniques and swarm robotics [2] [3]. This area of research, also known as swarm intelligence [4] [5], studies large collections of relatively simple agents that can collectively solve problems that are too complex for a single robot or that can display the robustness and adaptability to environmental variation displayed by biological agents.

One of the most well-known swarm algorithms is the Particle Swarm Optimization (*PSO*) developed by Kennedy and Eberhart [6]. This optimization technique models a set of potential problem solutions as a swarm of particles moving around in a virtual search space. However, a general problem with the *PSO* and other optimization algorithms is that of becoming trapped in a local optimum, such that it may work in some problems but may fail on others. In search of a better model of natural selection using the *PSO* algorithm, the Darwinian Particle Swarm Optimization (*DPSO*) was formulated by Tillet *et al*. [7] enhancing the ability to escape from local optima.

Just like in Multi-Robot Systems (*MRS*), where groups of robots interact to accomplish their goals [8], both *PSO* and *DPSO* use groups of interacting virtual agents (*aka*, particles) in order to achieve their optimization. However, contrarily to virtual agents, robots are designed to act in the real world where communication constraints and obstacles need to be taken into account.

In our previous work [9], an extension of the *DPSO* to *MRS* was proposed. Each robot is then responsible for each virtual agent, which it needs to evaluate at each iteration. After each set of evaluations, the robots communicate to share the objective information (cost or fitness) needed to progress to the next iteration of the algorithm.

However, the design of new swarm robotics systems cannot provide us with quantitative prediction of the collective performance. Real robot experiments and simulations are the most direct way to observe the behavior of the system under different conditions (*i.e.*, population size and communication constraints). However, trials with real or simulated robots do not scale well as the size of the system grows.

Therefore, it is hard to predict the ideal number of robots necessary for a given task.

Bearing these ideas in mind, this paper carries out a statistical analysis of the previously proposed algorithm in order to evaluate the relationship between the population of robots in the *R-DPSO* and the maximum the communication range between robots. To that end a Multivariate Analysis of Variance technique (*MANOVA*) is used to evaluate the performance of the algorithm based on the number of robots and the communication distance.

The paper is organized as it follows. Section II presents an overview of the *R-DPSO* algorithm. A brief description of the *MANOVA* is given in Section III. Experimental results analyzing the performance of the algorithm are demonstrated Section IV while Section V outlines the main conclusions.

II. ROBOTIC DARWINIAN PARTICLE SWARM OPTIMIZATION

This section briefly presents the *R-DPSO* algorithm proposed in [9]. The *DPSO* [7] is an evolutionary algorithm that extends the well-known *PSO* [6] using natural selection, or survival-of-the-fittest, to enhance the ability to escape from local optima.

Since the *R-DPSO* approach is an adaptation of the *DPSO* to real mobile robots, four general features are proposed: *i*) a novel "punish"-"reward" mechanism to emulate the deletion and creation of robots; *ii*) an obstacle avoidance algorithm to avoid collisions; *iii*) an enforcing multi-hop network connectivity algorithm to ensure that the *MANET* remains connected throughout the mission; *iv*) a novel methodology to establish the initial planar deployment of robots preserving the connectivity of the *MANET* while spreading out the robots as most as possible.

These features are further explored in the following sections.

A. "Punish"-"Reward" Mechanism

In the common *DPSO*, "punish" means the deleting of particles and swarms, while "reward" means the spawning of new particles and swarms. In order to adapt *DPSO* to mobile robotics, the deleting and spawning of a robot are modelled by the mechanisms of *social exclusion* and *social inclusion*, respectively.

The *R-DPSO* is then represented by multiple swarms, *i.e.*, multiple groups of robots that altogether form a population. Each swarm individually performs just like an ordinary *PSO* in search for the solution and some rules governs the whole population of robots.

If there was no improvement in a swarm's objective over a period of time, the swarm is punished by excluding the worst performing robot, which is added to a socially excluded group. The worst performing robot is evaluated by the value of its objective function compared to other members in the same swarm. In other words, if the objective is to maximize the fitness function, the robot to be excluded will be the one with the higher fitness value.

Those socially excluded robots, instead of searching for the objective function's global optimum like the other robots in the active swarms do, they basically randomly wander in the scenario. This approach improves the algorithm making it less susceptible of becoming trapped in a local optimum. Note, however, that they are always aware of their individual solution and the global solution of the socially excluded group.

B. Obstacle Avoidance

A new cost or fitness function is defined in such a way that it would guide the robot to perform the main mission while avoiding obstacles. For this purpose it is assumed that each robot is equipped with sensors capable of sensing the environment for obstacle detection within a finite sensing radius . A monotonic and positive *sensing function* that depends on the sensing information (*i.e.*, distance from the robot to obstacle) is defined. In most situations the *sensing function* can be represented as the relation between the analog output voltage of distance sensors and the distance to the detected object.

The susceptibility of each robot (Figure 1) to the main objective and to obstacle avoidance need to be established and depends on several conditions related with the main objective (*i.e.*, minimize a cost function or maximize a fitness function) and the sensing information (*i.e.*, monotonicity of the sensing function)

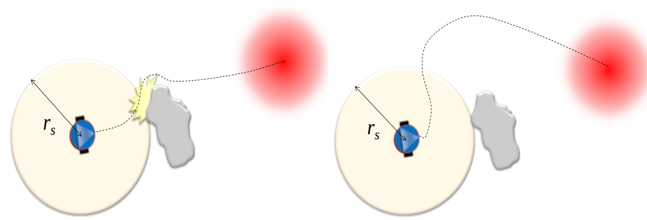

Figure 1. Illustrative example of obstacle avoidance behaviour of a robot with different obstacle susceptibility weights.

C. Enforcing MANET Connectivity

Robots' position need to be controlled in order to maintain the communication based on constraints such as maximum distance or minimum signal quality. The way network will be forced to preserve connectivity depends on the characteristics of the communication (*e.g.*, multi-hop, biconnectivity).

Assuming that the network supports multi-hop connectivity, the communication between two end nodes (*i.e.*, robots) is carried out through a number of intermediate nodes whose function is to relay information from one point to another (note that any robot may be used as a relay node independently of their swarm). Considering that nodes are mobile, it is necessary to guarantee the communication between all nodes. The nodes' position (*i.e.*, robots' position) are updated by means of the enforcing *MANET* connectivity algorithm are further described in [10].

D. *Initial Deployment*

This approach tries to get the benefits of a random planar deployment of robots eliminating the disadvantages inherent to it and taking into account the communication constraints using a deployment strategy based on the *Spiral of Theodorus* (*aka*, square root spiral) which is composed of contiguous right triangles (formerly called rectangled triangles) with each cathetus (*aka*, leg) having a unit length of 1 [11]. Each of the triangle's hypotenuses gives the square root to a consecutive natural number.

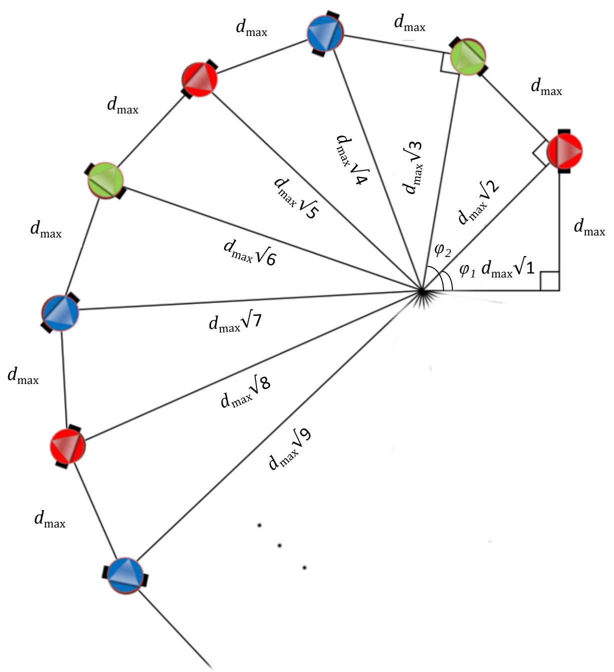

Figure 2. Initial deployment based on the spiral of Theodorus of a population of robots using the R-DPSO algorithm.

Since this approach uses the spiral of Theodorus to carry out the initial deployment of robots, two general adjustments need to be considered: *i*) the initial position of each robot is set at the further vertex of the centre of the spiral for each right triangle with a random orientation and also a random swarm in the specific situation of the *R-DPSO* algorithm; and *ii*) the size of the cathetus is set as the maximum communication range (instead of having the unit length 1) consequently changing the triangles' hypotenuses to the product between the maximum communication range and the square root of the consecutive natural number. These assumptions make it possible to have an initial deployment of the robots in an area that depends on both the number of robots and the communication constraints (Figure 2).

III. Multivariate Analysis of Variance

The significance of the maximum communication distance and the number of robots (independent variables) on the global solution and the runtime (dependent variables) was analyzed using a two-way *MANOVA* after checking the assumptions of multivariate normality and homogeneity of variance/covariance. The assumption of normality of each of the univariate dependent variables was examined using univariate tests of *Kolmogorov-Smirnov* (*p-value* < 0.05). Although the univariate normality of each dependent variable has not been verified, since and using the Central Limit Theorem (*CLT*) [12] [13] this statement was assumed [14] [15]. Consequently, the assumption of multivariate normality was validated [14].

The assumption about the homogeneity of variance/covariance matrix in each group was examined with the *Box's M Test* (M = 6465.13, F(69; 5368369.62) = 92.98; p-value = 0.001).

Although the homogeneity of variance/covariance matrices has not been verified, the *MANOVA* technique is robust to this violation because all the samples have the same size [14].

When the *MANOVA* detected significant statistical differences, we proceeded to the commonly-used *ANOVA* for each dependent variable followed by the *Tukey's HSD Post Hoc*. The classification of the size effect (*i.e.*, measure of the proportion of the total variation in the dependent variable explained by the independent variable) was done according to Maroco [14].

This analysis was performed using *IBM SPSS Statistics* for a significance level of 5%.

IV. Experimental Results

In this section, it is explored the effectiveness of the *R-DPSO*, while performing distributed unsupervised learning with local and global information, under communication constraints while increasing the population of robots.

The number of robots is set equal to the number of particles and the number of nodes in the network. Robots are deployed in the search space in a spiral manner (as previously described) where the radius depends on the maximum communication distance and the number of robots in the population. Since the *R-DPSO* is a stochastic algorithm, every time it is executed it may lead to different trajectory convergence. Therefore, multiple test groups of 100 trials, of 300 iterations each were considered. It will be used a minimum, initial and maximum number of 1, 3 and 6 swarms (represented by different colors in Figure 2), respectively, independently of the population of robots. The search space is represented by a Gaussian distribution consisting on a function of two variables of the search space, *x* and *y*-axis, which represents the position of the robot in meters. The particles will then move in an outdoor scenario of 300 x 300 meters where the *z*-axis represents the value of the objective function. In order to improve the interpretation of the algorithm performance, results were normalized in a way that the objective of robots team is to find the optimal value of 1 while avoiding obstacles and enforcing the *MANET* connectivity (*cf.*, Section II). Trying to maintain the network connectivity by only taking into account the communication range does not match reality since the propagation model is more complex – the signal depends not only on the distance but also on the multiple paths from walls and other obstacles.

However, in simulation, the communication distance is a good approach and it is easier to implement.

The maximum communication distance will then vary depending on the chosen wireless protocol. Four conditions were described: *1*) Existence of a communication infrastructure (*i.e.*, without communication constraints); *2*) *WiFi*; *3*) *ZigBee*; *4*) *Bluetooth*. Table 1 depicts the maximum communication distance adapted from a comparison between the key characteristics of each wireless protocol in [16]. The mean between the minimum and maximum range shown in [16] was considered as the maximum communication distance .

TABLE I. TYPICAL MAXIMUM COMMUNICATION DISTANCES OF THE WIFI, ZIGBEE AND BLUETOOTH PROTOCOLS

	No Limit	*WiFi*	*ZibBee*	*Bluetooth*
[m]		100	55	10

The number of robots will vary from 3 robots to 33 robots with incremental steps of 6 robots, *i.e.*, in order to understand the performance of the algorithm while changing the population size and the maximum communication distance.

Since these simulation experiments represent a search task, it is necessary to evaluate not only the completeness of the mission but also the speed. Therefore, the performance of the algorithm will be evaluated through the analysis of the final global solution of the population and the runtime of the simulation. If the group cannot find the optimal solution, the runtime is considered to be the simulation time (*i.e.*, 300 iterations).

A two-way *MANOVA* analysis was carried out to assess whether the factors on this study have a statistically significant effect on the team's performance. The *MANOVA* revealed that the maximum communication distance had a small effect and significant on the multivariate composite (*Pillai's Trace* = 0.75; F(6; 4752) = 30.974; *p-value* = 0.001; *Partial Eta Squared* = 0.038; *Power* = 1.0). The number of robots also had a small effect and significant on the multivariate composite (*Pillai's Trace* = 0.080; F(10; 4752) = 19.706; *p-value* = 0.001; = 0.04; *Power* = 1.0). Finally, the interaction between the two independent variables had a small statistically significant effect on the multivariate composite (*Pillai's Trace* = 0.032; F(30; 4752) = 2.55; *p-value* = 0.001; = 0.016; *Power* = 1.0).

After observing the multivariate significance in the maximum communication distance and the number of robots, an univariate *ANOVA* for each dependent variable followed by the *Tukey's HSD Test* was carried out.

For the maximum communication distance, the dependent variable final global solution presents statistically significant differences (F(3, 2376) = 45.185; *p-value* = 0.001; = 0.054; *Power* = 1.0) and the dependent variable runtime presents statistically significant differences (F(3, 2376) = 53.683; *p-value* = 0.001; = 0.063; *Power* = 1.0).

For the number of robots, the dependent variable final global solution also presents statistically significant differences (F(5, 2376) = 23.347; *p-value* = 0.001; = 0.047; *Power* = 1.0) and also the dependent variable runtime shows statistically significant differences (F(5, 2376) = 39.816; *p-value* = 0.001; = 0.077, *Power* = 1.0).

Using the Tukey's *HSD* Post Hoc, it is possible to verify where the differences between maximum distances of communication lie.

A. *Maximum Communication Distance*

Analyzing the team's final solution and the runtime variables, it appears that there are statistically significant differences between experiments without communication constraints and experiments using the *WiFi* protocol, the *ZigBee* protocol and the *Bluetooth* protocol.

TABLE II. TUKEY'S *HSD* POST HOC TEST TO THE MAXIMUM COMMUNICATION DISTANCE

	Final Solution	Runtime
No Limit vs WiFi	0.002*	0.854
No Limit vs ZigBee	0.001*	0.001*
No Limit vs Bluetooth	0.001*	0.001*
WiFi vs ZigBee	0.207	0.019*
WiFi vs Bluetooth	0.001*	0.001*
ZigBee vs Bluetooth	0.001*	0.001*

* The corresponding p-value for mean difference when it is significant at the 0.05 level

It is noteworthy that without communication constraints the algorithm produces better solutions. Also, using *WiFi* protocol produces better solutions than using the *ZigBee* protocol and, on the other hand, this last one produces better solutions than the *Bluetooth* protocol as expected.

In fact, using the *Bluetooth* protocol proves to be the "worse" communication protocol to employ.

B. *Number of Robots*

Analyzing both the final global solution of the team and the runtime variables, it appears that there are statistically significant differences between a population inferior to 15 robots and a population superior to 21 robots, not showing statistically significant differences for a population between 3 to 15 robots and 21 to 33 robots.

Note that the "worst" result is obtained using 3 robots but may not be considered significantly "worse" than using 9 or even 15 robots. This may be relevant since the increase in the number of robots result in an increase in the cost of the solution.

TABLE III. TUKEY'S *HSD* POST HOC TEST TO THE NUMBER OF ROBOTS

N	Final Solution	Runtime
3vs9	1.000	0.861
3vs15	0.151	0.182
3vs21	0.001*	0.001*
3vs27	0.001*	0.001*
3vs33	0.001*	0.001*
9vs15	0.249	0.844
9vs21	0.001*	0.001*
9vs27	0.001*	0.001*
9vs33	0.001*	0.001*
15vs21	0.004*	0.001*
15vs27	0.001*	0.001*
15vs33	0.001*	0.001*
21vs27	0.842	0.654
21vs33	0.785	0.076
27vs33	1.000	0.845

* The corresponding p-value for mean difference when it is significant at the 0.05 level

C. Discussion

As previously depicted in tables II and III, as figures 3 and 4 shows and using *Tukey's HSD Post Hoc* test, the conditions of the independent variables *N* and can be divided into different homogeneous subsets.

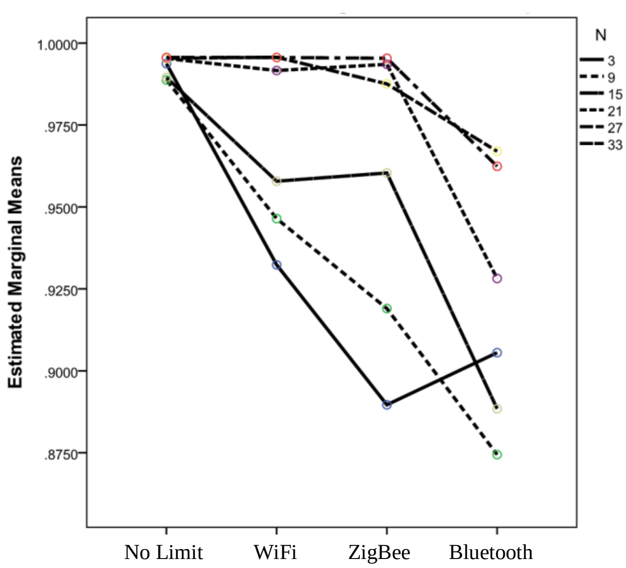

Figure 3. Estimated marginal means of the final global solution.

For instance, since there are no statistically significant differences between teams of 3, 9 and 15 robots in the analysis of both the final global solution and the runtime, this can be considered as a subset of *N*, *i.e.*, and . In other words, in an application were the cost of the solution needs to be considered, and since there are no significant advantages of having 15 robots instead of having just 3 or having 33 robots instead of 21, the choice would be using the minimum number of robots of each subset of *N*.

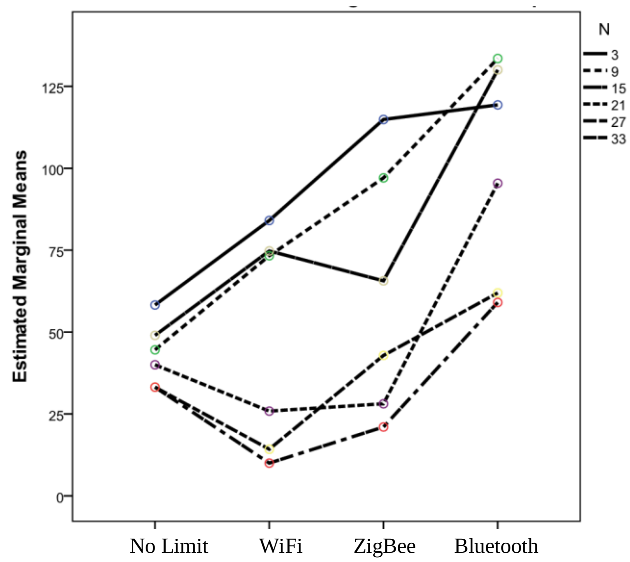

Figure 4. Estimated marginal means of the runtime.

The same analysis can be conducted for the maximum communication distance. However, in this specific situation, 3 subsets can be considered analyzing the statistically significant differences between the different values of , *i.e.*, , and . Put differently, the choice between wireless technologies, if unable to have a preexistent infrastructure (*i.e.*,), based on the maximum communication distance and ignoring other technical features, may be centered on the *WiFi* and *ZigBee* technologies.

V. CONCLUSION

This paper presented a methodology to evaluate the previously proposed algorithm, denoted as *R-DPSO*, which takes into account real-world multi-robot systems (*MRS*) characteristics.

Experimental results shows that the performance of the algorithm can be improved, thus decreasing the time needed to find the global optimum (*i.e.*, runtime), as the number of robots or the communication range increases. However, the choice on the number of robots and the wireless technology needs to take into account the global cost of the solution depending on the statistic significant differences between the independent variables.

As future work, a probabilistic model will be studied to optimize the swarm parameters such as the number of robots and the communication protocol in order to improve the overall runtime and find the optimal solution for a given scenario.

ACKNOWLEDGMENT

This work was supported by a *PhD* scholarship (SFRH/BD /73382/2010) granted to the first author by the Portuguese Foundation for Science and Technology (*FCT*) and the Institute of Systems and Robotics (*ISR*) also under regular funding by *FCT*.

REFERENCES

[1] D. Floreano and C. Mattiussi, *Bio-Inspired Artificial Intelligence: Theories, Methods, and Technologies*. Cambridge, MA: MIT Press, 2008.

[2] L. Marques and A. T. d. Almeida, "Finding Odours across Large Search Spaces: A Particle Swarm-Based Approach," in *Proc. 6th Intl. Conf. on Climbing & Walking Robots (CLAWAR)*, Madrid, Spain, 2004.

[3] A. Marjovi, L. Marques, and J. Penders, "Guardians Robot Swarm Exploration and Firefighter Assistance," in *Workshop on NRS in IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, St Louis, MO, USA, 2009.

[4] E. Bonabeau, M. Dorigo, and G. Theraulaz, *Swarm Intelligence: From Natural to Artificial Systems*. New York: Oxford University Press, 1999.

[5] G. Beni, "From swarm intelligence to swarm robotics," in *Proceedings of the Swarm Robotics Workshop*, Heidelberg, Germany, 2004, pp. 1-9.

[6] J. Kennedy and R. Eberhart, "A new optimizer using particle swarm theory," *Proceedings of the IEEE Sixth International Symposium on Micro Machine and Human Science*, pp. 39-43, 1995.

[7] J. Tillett, T. M. Rao, F. Sahin, R. Rao, and S. Brockport, "Darwinian Particle Swarm Optimization," *Proceedings of the 2nd Indian International Conference on Artificial Intelligence*, pp. 1474-1487, 2005.

[8] R. Rocha, J. Dias, and A. Carvalho, "Cooperative Multi-Robot Systems: a study of Visionbased 3-D Mapping using Information Theory," *Robotics and Autonomous Systems*, vol. 53(3–4), pp. 282-311, 2005.

[9] M. S. Couceiro, R. P. Rocha, and N. M. F. Ferreira, "A Novel Multi-Robot Exploration Approach based on Particle Swarm Optimization Algorithms," in *IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS'2011*, San Francisco, California, 2011 (Under Review).

[10] M. S. Couceiro, R. P. Rocha, and N. M. F. Ferreira, "Fulfilling Multi-Robot Ad Hoc Communication based on Robotic Darwinian Particle Swarm Optimization," in *IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS'2011*, San Francisco, California, 2011 (Under Review).

[11] H. K. Hahn and K. Schoenberger, "The Ordered Distribution of Natural Numbers on the Square Root Spiral," *The Journal of Business*, 2007.

[12] J. Maroco and R. Bispo, *Estatística aplicada às ciências sociais e humanas*, C. Editores, Ed. Lisboa, Portugal, 2003.

[13] A. C. Pedrosa and S. M. A. Gama, *Introdução Computacional à Probabilidade e Estatística*. Portugal: Porto Editora, 2004.

[14] J. Maroco, *Análise Estatística com ultização do SPSS*. Lisboa: Edições Silabo, 2010.

[15] M. H. Pestana and J. N. Gageiro, *Análise de dados para Ciências Sociais - A complementaridade do SPSS*, 5th ed. Lisboa, Portugal: Edições Sílabo, 2008.

[16] J. S. Lee, Y. W. Su, and C. C. Shen, "A comparative study of wireless protocols: bluetooth, UWB, ZigBee, and Wi-Fi," *Proceedings of the 33rd Annual Conference of the IEEE Industrial Electronics Society (IECON '07)*, pp. 46-51, 2007.