# Predicting Performance of SLAM Algorithms

Matteo Luperto[1], Valerio Castelli[2], Fabio Bonsignorio[3] and Francesco Amigoni[2]

[1]*Università degli Studi di Milano, Italy*
[2]*Politecnico di Milano, Italy*
[3]*Heron Robots, Italy*

**Abstract**
In this position paper, we claim that the availability of quantitative predictions of performance of robots in unseen environments can significantly contribute to the progress of autonomous robotics. Moreover, we overview a methodology that we adopted to provide such predictions in the case of Simultaneous Localization And Planning (SLAM)[0].

**Keywords**
Simultaneous Localization and Mapping (SLAM), prediction, autonomous robots

## 1. Predicting Robot Performance

Imagine a scenario in which a search and rescue human team has to deploy an autonomous robot system during a rescue operation in a large site, like an industrial plant, and has to choose, between different configurations of the robot system, the one that maximizes the efficiency in finding possible victims. Consider another scenario in which a company that develops autonomous robot systems for assistance to elderly people has to configure a new installation in a newly built residence for retired people in order to make the robots moving as safely and as robustly as possible. Finally, consider a third scenario in which a consumer has just bought a new autonomous cleaning robot for the office that he/she manages and has to manually set the best configuration of parameter values. Currently, to address all the above scenarios (and several others) and make the decisions they involve, autonomous robots developers and users mostly rely on trial-and-error approaches and on subjective previous experience. Trial-and-error approaches can be significantly time consuming and expensive. Subjective previous experience is something that could be expected from the experts (like the search and rescue human team) but not from end consumers.

One fundamental problem that characterizes the above settings is the difficulty in *predicting* the performance of a robot system in a new scenario. Indeed, beyond informal experience of researchers and developers, there is not any assessed methodology to *quantitatively* estimate the expected performance of a given robot system in a given scenario before actually running the system in the scenario. Methods for performance evaluation work almost exclusively *ex*

[0]A preliminary version of this paper appeared as [1].

*Artificial Intelligence and Robotics (AIRO) Workshop at AIxIA 2021*
matteo.luperto@unimi.it (M. Luperto); valerio.castelli@mail.polimi.it (V. Castelli); fabio.bonsignorio@heronrobots.com (F. Bonsignorio); francesco.amigoni@polimi.it (F. Amigoni)
© 2021 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).
CEUR Workshop Proceedings (CEUR-WS.org)

CEUR-WS.org/Vol-3162/short9.pdf

*post*, namely after the actual operations of the robots.

In this position paper, we argue that the development of innovative methodologies for predicting the performance of autonomous robots in new scenarios could represent a major contribution to the advancement of the field. In a sense, these methodologies will enable a sort of *predictive benchmarking* of autonomous robot systems, meaning that one could be able to anticipate, up to some degree of confidence, the performance results obtained by a system on a benchmark. There have been some attempts, especially in the military field, to predict robot performance in previously unseen scenarios. An example is [2], where a model to assess the traversal cost of a natural outdoor environment for an autonomous vehicle using A* planning is proposed. Considering autonomous wheeled robots, in [3] the authors introduce a method to estimate the traversal time of a path by using its length, smoothness, and clearance as features of a non-linear regressor. However, to the best of our knowledge, there is currently no assessed general methodology able to predict performance of autonomous robots in new environments.

## 2. Importance of Prediction

Methodologies to quantitatively predict robot performance are expected to be highly significant for at least three aspects.

First, the availability of reliable predictions on performance of autonomous robot systems can help developers to make informed decisions at design time, reducing expensive trial-and-error iterations. Consider the following scenario. A urban search and rescue team has to deploy a multirobot system in a building in which a gas leakage occurred in order to assess the presence and the location of possible human victims. The multirobot system is composed of several communicating mobile robots that should explore the initially unknown building and report to the search and rescue team a map of the site with the locations of the victims, so that the human rescuers can enter the building and quickly reach the victims, reducing risks for their safety. Assume that the multirobot system can have two different configurations A and B. In configuration A, the robots are equipped with powerful laser range finders (e.g., with field of view of 360 degrees and range of 70 m), have an extended communication range (e.g., 100 m without obstacles), but are slow (e.g., they cannot move faster than 0.5 m/s). In configuration B, the robots are equipped with less powerful laser range finders (e.g., with field of view of 200 degrees and range of 15 m), have a limited communication range (e.g., up 15 m without obstacles), but can move faster (e.g., they can move at 2 m/s). Currently, there is no principled way for the search and rescue team to choose *a priori* (before actually deploying the robots) the best configuration between A and B, beyond informally relying on past experience. Methodologies to predict robot performance will allow the team to make an informed decision about which configuration is better to use. This kind of predictive benchmarking perfectly fits with the Ethically Aligned Design approach of the IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems[1].

Second, knowledge of environments in which a given autonomous robot system is expected to perform well and in which it is expected to perform bad can be used to establish the boundaries of its applicability. This can enable the *generalization* of results obtained by robot systems [4],

[1] https://ethicsinaction.ieee.org

which currently represents a rather limited practice in many areas of robotics, hindering the possibility of extending the results obtained in a setting to other settings. For instance, the performance measured in the lab environments in which most robot systems are developed and preliminarily tested are usually little informative about the performance of the same robot systems in other environments.

Finally, the social impact of the methodologies to predict robot performance can be in the direction of an improved social trust in autonomous robots. Quantitative assessments of the performance of robot systems that can be reliably generalized can help to improve the degree of confidence and acceptance of such robot systems. This concern is addressed, among others, by the Foundation for Responsible Robotics[2]. For example, consumers will be more inclined to trust an autonomous robot system for domestic assistance if its performance measured in other houses quantitatively suggests a good level of performance also in their houses.

## 3. A Proposed Methodology

We started to develop a methodology for quantitatively predicting the performance of a Simultaneous Localization And Mapping (SLAM) [5] system in *indoor* environments. A SLAM algorithm integrates data collected by the robot's sensors in a map representing the environment (e.g., as a grid map or as a graph of poses). Here, we outline the main elements of our methodology (see [6] for more details).

We consider a SLAM system, denoted with $s$, composed of a mobile robot platform equipped with sensors for detecting the surrounding environment (a laser range scanner with a field of view of 270°, an angular resolution of 0.5°, and a maximum range of 30 m in our case) that is running a SLAM algorithm (GMapping [7] in our case). Given a SLAM system $s$, the methodology consists of a sequence of steps.

At first, we perform *data collection*. Specifically, given a set of environments $E$ and $s$, we collect the performance $P(s, e)$ of $s$ in each environment $e \in E$. Performance $P()$ is calculated as the localization error according to [8], which requires to know the actual trajectory followed by the robot (ground truth). Note that, since performance varies over different runs of the same robot system $s$ in the same environment $e$ (e.g., due to noise in movements and perceptions), $P(s, e)$ involves multiple runs, which we perform in simulation and with real robots.

Then, for each environment $e \in E$, we extract a set of features $F(e)$ which characterize $e$. In this step, called *feature extraction*, we obtain the features from knowledge available *a priori* (e.g., obtainable from floor plans of buildings) and not from the maps built by the robots. We consider geometrical features (like area and perimeter of $e$) and structural features (like those derived from the Voronoi skeleton of $e$ [9]).

The third step is *model learning*, in which a model that relates $F(e)$ to $P(s, e)$ (for any $e \in E$) is built. In general, a model correlates a subset of features of $F(e)$ to $P(s, e)$. In our case, we found a strong correlation (using linear regression) between the length of an exploration path over the Voronoi skeleton of $e$ and the performance $P()$. Next section discusses numerical results.

Such model is used to do *performance prediction*, namely to calculate the expected performance $\hat{P}(s, \epsilon)$ of $s$ in an unseen environment $\epsilon$ (namely, $\epsilon \notin E$). Our performance prediction extracts

[2]https://responsiblerobotics.org

features $F(\epsilon)$ from $\epsilon$ and finds $\hat{P}(s, \epsilon)$ exploiting the model built in the previous step.

In principle, the above sequence of steps can be repeated when the set $E$ of visited environments grows, resulting in increasingly accurate performance predictions.

The proposed methodology is based on two main assumptions:

(a) models relating $F(e)$ to $P(s, e)$ for all environments $e \in E$ can be found and
(b) a set of features $F(\epsilon)$ can be obtained without actually deploying the robot in the new environment $\epsilon$.

We practically validated assumption (a) in the case of GMapping and in the case of another SLAM system based on KartoSLAM [10] (results are not discussed here). Hence, we are confident that similar models, perhaps not linear and possibly based on other techniques (like neural networks), can be found also for other SLAM systems and, more generally, for other robot systems. Note that, if a large amount of data is available, feature extraction and model learning steps can both be addressed using learning techniques like deep neural networks, reducing the human effort required in adapting the proposed method to new applications. Assumption (b) basically implies that some knowledge about $\epsilon$ is available before robots are actually deployed in $\epsilon$. For example, such knowledge could consist in the floor plan of $\epsilon$ (e.g., obtained from urban administrative offices or from evacuation maps) from which $F(\epsilon)$ can be extracted. Note that knowing the floor plan of $\epsilon$ does not amount to knowing the map of $\epsilon$, because unforeseen obstacles that might occupy the free space, like, for instance, furniture, are not included in the floor plan.

## 4. Preliminary Results

In this section we present a sample of results we obtained with the proposed approach.

We collect data on performance of GMapping on a set $E$ of 100 indoor environments covering a wide range of building types (schools, offices, university campuses, and others), sizes, and shapes (to avoid overfitting). Environments are taken from those used in [11], a subset of 11 of the 20 environments of [11], a set of 25 floor plans that represent buildings of the MIT university campus from [12], and our own dataset of 64 floor plans representing real world buildings [13]. Simulations are performed with Stage, using the ROS GMapping[3] and Navigation packages[4] for SLAM and robot movement, respectively. Given an environment $e$ and a starting pose for the robot (close to the center of the environment, the same for all runs), a run $R_e$ evolves by exploring $e$ using the frontier-based exploration paradigm defined in [14]. Full details about data collection are reported in [15].

We extract a set of features $F(e)$ that characterize an environment $e$ using the Voronoi graph built on $e$ [16]. Intuitively, the edges of the Voronoi graph represent the clearest paths (i.e., the paths furthest from obstacles) that a robot could follow in the environment. We exploit the Voronoi graph to obtain an approximation of the path a robot follows to build the complete map of an initially unknown environment, considering the limitations of the robot sensors. In

[3] http://wiki.ros.org/gmapping
[4] http://wiki.ros.org/navigation

**Table 1**
Performance of the best single-feature linear models for the components of the localization error based on features $F_e$. NRMSE is expressed in percentage with respect to the range of the predicted variable, while RMSE is expressed in m for the translations and in rad for the rotations. *tra* and *rot* are translational and rotational error, as in [8].

| | feature | $R^2$ | RMSE | NRMSE |
|---|---|---|---|---|
| *tra* | VTD | 0.830 | 0.145 | 7.92% |
| *rot* | VTD | 0.723 | 0.004 | 11.15% |

practice, given the floor plan of $e$, represented as a grid map (image), we extract the Voronoi graph from the dual Delaunay triangulation, as specified in [9].

Given the Voronoi graph of an environment $e$, we calculate two quantities, the *Voronoi traversal distance* ($VTD_e$) and *Voronoi traversal rotation* ($VTR_e$), that are the overall distance and the overall amount of rotation, respectively, that a robot incurs to cover the Voronoi graph of $e$.

We build a model of the relationships between performance $P(s, e)$ (localization error, calculated as in [8]) and features $F(e)$ for all the environments $e \in E$. To assess the quality of a possible model, we compute the overall root mean square error (RMSE) and the average coefficient of determination ($R^2$) on the training set $E$ using k-fold cross validation (with a number of folds equal to 5). We also consider the normalized (scale-independent) version of the RMSE: $\text{NRMSE} = \frac{\text{RMSE}}{y_{\max} - y_{\min}}$, where $y_{\max}$ and $y_{\min}$ are the maximum and the minimum observed values of the target variable $y$, respectively.

Table 1 shows the best model, in terms of highest $R^2$ value for the localization error we are interested in predicting, and the corresponding results, which show a good performance of our approach. For more details, please refer to [6].

In conclusion, we think that the methodology depicted above can be generalized to different robot systems $s$ and to different performance measures $P$, enabling the possibility of getting quantitative predictions of the expected performance of robot system $s$ in new, unseen, environments.

## References


[1] F. Amigoni, V. Castelli, F. Bonsignorio, M. Luperto, Predicting robot performance: Why and how, in: Presented at the IJCAI-ECAI/ICML/AAMAS Federated AI for Robotics Workshop (FAIR), 2018.

[2] S. Young, T. Mazzuchi, S. Sarkani, A framework for predicting future system performance in autonomous unmanned ground vehicles, IEEE T Syst Man Cy-S 47 (2017) 1192–1206.

[3] P. Regier, M. Missura, M. Bennewitz, Predicting travel time from path characteristics for wheeled robot navigation, in: Proc. ECMR, 2017, pp. 1–6.

[4] F. Amigoni, M. Luperto, V. Schiaffonati, Toward generalization of experimental results for autonomous robots, Robot Auton Syst 90 (2017) 4–14.

[5] S. Thrun, W. Burgard, D. Fox, Probabilistic Robotics, The MIT Press, 2005.

[6] M. Luperto, V. Castelli, F. Amigoni, Predicting performance of SLAM algorithms, 2021. arXiv:2109.02329.

[7] G. Grisetti, C. Stachniss, W. Burgard, Improved techniques for grid mapping with Rao-Blackwellized particle filters, IEEE T Robot 23 (2007) 34–46.
[8] R. Kümmerle, B. Steder, C. Dornhege, M. Ruhnke, G. Grisetti, C. Stachniss, A. Kleiner, On measuring the accuracy of SLAM algorithms, Auton Robot 27 (2009) 387–407.
[9] A. Bowyer, Computing Dirichlet tessellations, Comput J 24 (1981) 162–166.
[10] K. Konolige, G. Grisetti, R. Kümmerle, W. Burgard, B. Limketkai, R. Vincent, Efficient sparse pose adjustment for 2D mapping, in: Proc. IROS, 2010, pp. 22–29.
[11] R. Bormann, F. Jordan, W. Li, J. Hampp, M. Hägele, Room segmentation: Survey, implementation, and analysis, in: Proc. ICRA, 2016, pp. 1019–1026.
[12] E. Whiting, J. Battat, S. Teller, Generating a topological model of multi-building environments from floorplans, in: Proc. CAADFutures, 2007, pp. 115–128.
[13] M. Luperto, A. Quattrini Li, F. Amigoni, A system for building semantic maps of indoor environments exploiting the concept of building typology, in: Proc. RoboCup, 2013, pp. 504–515.
[14] B. Yamauchi, A frontier-based approach for autonomous exploration, in: Proc. CIRA, 1997, pp. 146–151.
[15] F. Amigoni, V. Castelli, M. Luperto, Improving repeatability of experiments by automatic evaluation of SLAM algorithms, in: Proc. IROS, 2018, pp. 7237–7243.
[16] S. Schwertfeger, A. Birk, Map evaluation using matched topology graphs, Auton Robot 40 (2016) 761–787.