Abstract—A large amount of research focuses on experimentally optimizing performance of wireless solutions. Finding the optimal performance settings typically requires investigating all possible combinations of design parameters, while the number of required experiments increases exponentially for each considered design parameter. The aim of this paper is to analyze the applicability of global optimization techniques to reduce the optimization time of wireless experimentation. In particular, the paper applies the Efficient Global Optimization (EGO) algorithm implemented in the SUrrogate MOdeling (SUMO) toolbox inside a wireless testbed. The proposed techniques are implemented and evaluated in a wireless testbed using a realistic wireless conference network problem. The performance accuracy and experimentation time of an exhaustively searched experiment is compared against a SUMO optimized experiment. In our proof of concept, the proposed SUMO optimizer reaches 99.51% of the global optimum performance while requiring 10 times less experiments compared to the exhaustive search experiment.

Keywords: wireless experimentation, optimization, testbeds, SUMO.

I. INTRODUCTION

Wireless network solutions are utilized in many application domains such as sensor networks, Wi-Fi networks, body area networks, home automation and others. Solutions in such application domains often have several tune-able parameters. For example, Wi-Fi networks have parameters that can be tweaked at the physical layer (transmit power, channel, modulation), MAC layer (inter frame spacing, contention window), network layer (routing protocol, mobility, topology) and application layer (throughput, server configurations). Optimizing all or a subset of these parameters in order to find the optimum operating point, is time consuming since the design space grows exponentially for every investigated design parameter.

Often, these network problems are optimized using wireless network simulators with the added advantage of reduced set-up time, cost, higher degree of control and repeatability. However, wireless simulators also have a number of disadvantages. Results can be very different when executing identical experiments on multiple simulators. They are also inaccurate when modelling the underlying channel characteristics (i.e. antenna diversity) and incapable to address the difference between devices of the same type [1], which do have a considerable impact on the overall performance.

As a result, experimentally driven research is necessary to complement simulations [1]. Measurements and performance evaluations on a real-life testbed are gaining more attention as they account for channel characteristics and hardware imperfections. However, wireless testbeds also have limitations since they require more set-up time compared to their simulator counterparts. For example, when orchestration an experiment using the Orbit Management Framework (OMF), an experiment having N wireless devices adds an average delay of 5.17*N ms for each message sent [2]. In order to mitigate the time overhead, efficient optimization algorithms can be used that are best fitted to wireless testbeds. In this paper, we investigate one such algorithm called Efficient Global Optimization (EGO) [3] implemented in the SUrrogate MOdeling (SUMO) toolbox. EGO uses Kriging approximations to find optimal operation point(s) of a complex problem while minimizing the number of experiments needed. This way, the overall experimentation time is kept to a minimum [4]. In sum, this paper examines the strengths of the SUMO optimizer by applying it to a wireless network problem having multiple design parameters.

This paper presents the following novel contributions.
1. Integration of the SUMO toolbox in a wireless testbed.
2. Definition of wireless conferencing scenario which involves multiple design parameters and performance objectives.
3. Design of a generic stopping criteria that can be used in a variety of optimization problems.

The remainder of the paper is organized as follows. Section II explores related work on multi-objective optimization in wireless testbeds. The principles of SUMO optimization are explained in section III. In section IV, the SUMO optimizer is experimentally validated by optimizing a wireless conference network problem. The results of the experiment are discussed in section V. Finally Section VI concludes the paper.

II. RELATED WORK

Solutions of wireless network problems often involve multi-objective optimizers in order to optimize multiple design parameters. In literature, a wide range of multi-objective optimization algorithms exist. Exhaustive search approaches evaluate all operating points of a problem to select optimum settings from the design space. A generic Numerical calculation approach using MATLAB is presented in [5]. This algorithm exhaustively searches the design space and determine the optimum point that gives the highest performance objective.

Genetic Algorithms (GA) [6] are heuristic algorithms that mimic the process of natural selection. Starting from an initial population (called chromosomes), new generations are
produced which hopefully contain better chromosomes than the previous generation. The optimization process selects new off-springs according to a fitness function and the evolutionary iterations continue until a predefined stopping criterion is met.

Particle Swarm Optimization (PSO) [7] algorithm optimizes a problem by exchanging information with neighboring particles such that a single particle with given position and velocity parameters searches an optimum setting. PSO works according to a mathematical formula optimizing a population of particles and the optimization process stops when the improvement is below a given limit.

Differential Evolution (DE) [8], similar to GA, starts from a given population with a fixed number of randomly initialized vectors. In every iteration step, a new generation is produced by combining the vectors randomly in order to create a mutation. The newer generation mixed with the target vector is evaluated against an objective function and the selector decides whether or not it should compose the next generation.

Simulated Annealing (SA) [9] algorithm is based on the principle of freezing liquid when forming a crystalline structure such that with sufficient time the structure acquires a minimum energy state. In each iteration step, the newly generated point is checked against the current point based on a probability distribution scale proportional to the problem’s analogous temperature. The newly generated point is accepted when the total objective function decreases.

Table I shows the different multi-objective optimization algorithms applied in wireless network problems. All the optimization algorithms applied simulation as a validation method which has its own disadvantages as described in the introduction section. On the other hand, this paper investigates the SUMO toolbox to evaluate its suitability for wireless network optimization. The SUMO optimization toolbox is often used in electromagnetic [10] and aerodynamic [11] optimization problems. Even though we are validating the SUMO toolbox in a wireless testbed for the first time, previous comparisons on multi-objective optimizers [10] [11] favours the SUMO variants which our preference is based upon. These optimizer comparisons were not made on wireless network problems which might favour a different type but for now we leave this area as a future work.

III. SUMO OPTIMIZATION TOOLBOX

The SUMO optimization toolbox is an efficient implementation of the Expected Improvement (EI) criterion, popularized by Jones et al in [4]. The optimization algorithm starts from a well-chosen initial experimental design, and a global (but only locally accurate) Kriging surrogate model of the objective function is computed. Such Kriging models are part of a broader class of approximation methods, called the Gaussian Processes (GP), and have some interesting properties that can be exploited by the optimizer. Whereas the standard approximation methods predict only a single function value, GP methods can predict the uncertainty of a function value as the realization of a normally distributed random variable $Y(x) \sim N(\mu(x), \sigma^2(x))$, where $\mu(x)$ represents the predicted value for $f(x)$ and $\sigma^2(x)$ the prediction variance at an arbitrary point $x$ in the parameter space. Based on this random variable $Y(x)$, different statistical criteria (such as the Probability of Improvement (PoI) or EI) can be computed to quantify how interesting a new point in the design space is. In this work, we adopt the EI criterion which simultaneously balances exploration and exploitation [12] of the parameter space. This corresponds to the improvement that is expected to occur when compared to the optimum value obtained so far (i.e., $f_{\text{min}}$ or $f_{\text{max}}$). By picking additional points with the highest EI value in the parameter space, the optimization process is directed towards a configuration with optimal performance. For example, in the case of a minimization problem, it can be written in the form of an integral as in [10] where $\varphi(.)$ represents the probability density function of a random variable and $I(x)$ is the improvement function.

$$E[I(x)] = \int_{-\infty}^{f_{\text{max}}} I(x)\varphi(Y(x))dY$$

$E[I(x)]$ corresponds to the improvement that is expected to occur when compared to the optimal value of the objective. A detailed explanation can be found in Section II-B of [10].

IV. EXPERIMENTAL VALIDATION

This section verifies the use of the SUMO optimization toolbox by optimizing a wireless conferencing scenario inside a wireless testbed. First we give a description of the wireless testbed where experimental validation is carried out. Next the experiment scenario and the optimization processes are presented. Finally, we look at each individual performance objectives and discuss how they are combined into a single objective.

A. Wireless testbed

The wireless iMinds w-iLab.t testbed at Zwijnaarde (Ghent, Belgium), shown in Figure 1, is equipped with heterogeneous wireless devices used to conduct a variety of experiments. It has 60 nodes each consisting of an embedded Zotac PC having two Wi-Fi interfaces, a sensor node, a Bluetooth dongle and a wired control interface connected to the testbed management framework. Furthermore, the testbed is equipped with advanced spectrum sensing devices. These include Universal Software Radio Platform (USRP), IMEC Sensing Engines, and Wireless open Access Research Platform (WARP) boards.

B. Experiment scenario

A wireless conferencing scenario composed of a wireless speaker and multiple microphones is shown in Figure 2. The
Fig. 2: Left: wireless conferencing scenario consisting of 8 listeners, 1 speaker, and 1 interferer. Right: mapping of the conferencing scenario to the testbed nodes. The transmission range of the speaker and interferer is indicated.

wireless speaker broadcasts a speaker’s voice over the air and the multiple wireless microphones receive the audio at the listener end. This type of wireless network is used in a multi-lingual conferencing room where the speaker’s voice is translated into different languages, multiplexed into a single stream and broadcasted to each listener. Often, the speaker’s audio quality is reduced by external interference and the surrounding environment is impacted by external interference. Thus, the main objective of the wireless conferencing scenario is to improve the received audio quality while keeping the transmission exposure at a minimum. To this end, the conferencing operator has the possibility to adapt the channel and power transmission parameters.

The experiment is composed of 1 interferer creating background interference and a System Under Test (SUT) having 1 speaker and 8 listeners. The speaker transmits a 12 second Wi-Fi audio stream and each listener calculates the average audio quality within the time frame. During this time frame, the interferer transmits a 10 Mbps continuous UDP stream on dual channels (i.e. 1 and 13) generated using the iperf application. The transmitters, receivers and interference generators are shown in Figure 2.

On the left hand side of Figure 2, the realistic wireless conferencing scenario is shown, where as on the right hand side, the experimentation scenario is mapped on the iMinds w-iLab.t testbed. All listener nodes (i.e. 38, 39, 40, 48, 50, 56, 57, and 58) are associated to the speaker access point (i.e. node 47). Background interference is created by the access point (i.e. node 49) using two separate Wi-Fi cards. The Wi-Fi card and driver used for this experiment are “Atheros Sparklan WPEA-110N/E/11n mini PCI 2T2R” and “Atheros ath9k” respectively. The SUMO toolbox runs on a dedicated PC that can communicate with all nodes of the experiment.

C. Optimization process

The optimization process is illustrated step by step in Figure 3. At (1) the controller is given a list of settings of the first experiments that needs to be configured on the wireless testbed. (2) Experiments are deployed on the wireless testbed using the requested settings, thus resulting in an initial sample set. (3) At the end of each experiment, the controller retrieves the evaluation criteria of the experiment. For the conferencing scenario, the evaluation criteria are the audio quality and exposure performances from all listeners. (4) An objective function is created by processing the evaluation criteria (see Section IV-D). (5) When the SUMO optimizer receives a sufficiently large dataset, it generates a surrogate model. (6) The next sample point with highest expected improvement is predicted. (7) The controller starts the next optimization experiment using the new design parameters. (8) Again, the evaluation criteria are retrieved and (9) the objective function

TABLE I: Design parameters, performance objective and validation method of different multi-objective optimization algorithms applied to a variety of complex wireless network problems

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Problem definition</th>
<th>Design parameters</th>
<th>Performance objectives</th>
<th>validation method</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Numerical calculation</td>
<td>Tuning of physical layer parameters in Wireless Sensor Network</td>
<td>Node hop distance, Transmit energy, Modulation schemes</td>
<td>Energy per Successful received Bit↓</td>
<td>simulation</td>
<td>[5]</td>
</tr>
<tr>
<td>GA</td>
<td>Maximizing sensing converge of wireless sensor network</td>
<td>Sensor positions</td>
<td>Relocation energy↓</td>
<td>simulation</td>
<td>[6]</td>
</tr>
<tr>
<td>DE</td>
<td>Radio Frequency Identifier network planning</td>
<td>Position, Angle, Transmit power</td>
<td>Coverage↑, Interference↓ and Cost↓</td>
<td>simulation</td>
<td>[8]</td>
</tr>
<tr>
<td>SA</td>
<td>Cognitive Radio system optimization</td>
<td>Transmit power, Modulation type</td>
<td>Power usage↓, Bit Error Rate↓ and Throughput↑</td>
<td>simulation</td>
<td>[9]</td>
</tr>
</tbody>
</table>

Fig. 3: The process of SUMO optimization in the wireless conference network problem. The different sequential steps are numbered from 1 to 11.
is calculated for the new design parameters. (10) Based on the current dataset, extended by one record, the surrogate model is updated and (11) a new sample is predicted. The optimization process continues until stopping conditions are met.

D. Performance objectives

Dual objectives are applied in the wireless conference network problem. The first objective is maximizing the received audio quality which is measured using the Mean Opinion Score (MOS). MOS is a subjective audio quality measure represented on a 1 to 5 scale (i.e. 1 being the worst quality and 5 being the best quality). To calculate the MOS score, the experiment described in Section V uses the ITU-T Perceptual Evaluation of Speech Quality (PESQ) P.862 standard. It calculates the PESQ score from packet loss, jitter and latency network parameters and maps it onto a MOS scale [13]. The second objective is minimizing transmission exposure. In [14] an in depth calculation of transmission exposure is presented. The exposure at a certain location is a combined measure of received power and transmit frequency. Transmission exposure is an important evaluation metric related to potential health issues, leading the regulatory bodies to set maximum limits.

As maximizing the combined performance objective is the goal, the weight of performance metrics needs to be defined depending on the problem type. For example, a person who wants to install a wireless conference in urban areas applies tighter exposure requirement than in rural areas. We would also apply high audio quality requirement in parliament auditoriums compared to office meeting rooms. However in our case, the aim is to validate the SUMO optimization toolbox and equal weights are applied on the normalized objectives. All performance objectives are normalized using maximum and minimum attainable values which are retrieved by doing an exhaustive searching experiment (see Section V-A).

V. RESULT AND DISCUSSION

This section analyzes the viability and efficiency of the SUMO optimization toolbox as used in validating the wireless conference scenario. First, an exhaustive search model is given in section V-A which is used as a reference for experiment comparison. Next, sensitivity of experiments to the choice of the initial sample size is discussed in section V-B. After that, a potential stopping criterion is analyzed in section V-C. Finally, the SUMO optimized experiment is compared against the exhaustive search model in Section V-D.

A. Exhaustive search model

In this section, we describe a reference experiment that is performed to generate an exhaustive search model of the wireless conference network problem. Neither SUMO nor any optimization algorithm is used to generate the model. The exhaustive search model evaluates all possible combinations of settings and will be used as a reference model for comparing SUMO optimization experiments. In total, 260 experiments (i.e. 13 Channels × 20 Transmit Power) were required. Interference is created continuously on channels 1 and 13.

Figure 4 show the outcomes of the exhaustive search model for both performance criteria. The exposure model of Figure 4(a) only considers the exposure from the speaker but not from background interference, since the goal of the SUT is to reduce its own exposure. The exposure objective degrades with increased transmission power independent of the used channel. In contrast, the audio quality objective increases with increased transmission power and the influence of the interference can be noted on multiple channels. There is an area on the non-interfered channels (i.e. 6 to 8) where adequate performance is observed also for lower transmit Power (i.e. 1dBm to 6dBm). This area is of interest because it represents a region where exposure is low. The worst performance from the audio quality model is shown between channels 2 to 4, 10 to 12 and transmit power 1dBm to 7dBm. Interestingly, this region is not located on channels where background interference is applied on but on the neighboring channels. This is due to the fact that the speaker and interferer nodes apply CSMA-CA on identical channels but not on neighboring channels which results in degraded performance [15].

The combined objective model from Figure 4(c) is a combination of the exposure model and the audio quality model shown in Figure 4(a) and 4(b) respectively. Figure 4(d) plots the combined objective model per transmission channel.

B. Initial sample size sensitivity

As explained in Section IV-C, a surrogate model predicts the next experiment with highest expected objective value. However, the creation of the initial model requires a set of initial sample points on the design space and their outputs. This section investigates how many initial samples are required before a usable surrogate model can be created.
The initial sample points for any problem have to be selected carefully such that the optimization process quickly converges to the optimum. If the number of initial sample points is large, the optimizer spends too much time during exploration work. On the other hand, considering few initial sample points leads to the risk of missing global optima and thus exploiting local optima instead. One way to address the trade-off between exploration and exploitation during optimization is by selecting an appropriate initial sample size. Usually this depends on the complexity of a problem’s global model. The more complex a problem’s global model is, the larger the initial sample size needed to have good surrogate model approximation and vice-versa. It was indicated in [16] that extreme points of a surface can be used to measure the complexity of a problem. These are the minimums, maximums and saddle points of a problem’s global model. Moreover, it is also indicated that by setting the initial sample size to the number of extreme points, an optimizer has a higher chance to arrive at the global optimum in short amount of time. This assumption only works if the problem’s extreme points are known beforehand. Most of the time this is not the case as we generally optimize unknown problems. Moreover, initial sample size selection depends on the problem type [16]. For our specific problem, setting the initial sample size to 8 points is found a good choice. The 8 initial sample points together with the corner points which the SUMO optimizer adds, sums up to 12 initial points in total.

In the analysis of the coming sections, we will each time analyse four different sampling methods to pick the 12 initial sample points from the design space. These are Latin Hypercube Sampling (LHS), Orthogonal sampling, Random sampling and Hammersley Sequence Sampling (HSS).

C. Stopping criteria

This section investigates the effect of an Objective Function Improvement (OFI) stopping criterion. The OFI stopping criterion looks at the relative difference in performance and stops the iteration when the standard deviation of the sorted last iterations falls below a certain threshold. The idea behind this concept is that the sorted objective function of a list of experiments ideally should approach a flat curve as the number of experiment iterations increases. The OFI stopping criterion has two parameters to set. These are the Standard Deviation WIDTH (STD-WIDTH) which sets the number of objective performance values in the standard deviation calculation and the Standard Deviation Threshold (STD-THLD) which is used as a lower limit for the stopping criterion. Figure 5 shows the standard deviation curve as a function of iteration count for STD-WIDTH 6 and 10. Calculation starts after the iteration count reaches STD-WIDTH.

As stated previously, the output of the plots for each standard deviation width approaches a flat curve when the optimization reaches the optimum. On the other hand, the randomness of the curves gradually decreases as the standard deviation width increases. This also increases the settling time until the lowest standard deviation value is reached. In addition, the benefit of the SUMO optimization toolbox is visually noticeable after the 12 initial experiments which is seen as a sharp declining curve. As the optimization continues, the standard deviation curve converges to a stable value.

There are two things we want the standard deviation curve to achieve. First, we want the curve to reach a stable value as fast as possible. This depends on the size of the optimum region in the problem’s global model. The wider this area, the sooner the optimization locates the optimum and the standard deviation curve converges to a stable value and vice-versa. Moreover, STD-WIDTH can assume the number of elements contained in the optimum region but the size of a problem’s optimum region is not known beforehand. A good value for STD-WIDTH from experience is to use half the elements of the initial sample size. Second, we want the curve to reach a very small stable value. In fact, this value never approaches to zero as the wireless medium has a certain level of non-determinism. A threshold can be estimated by doing repeatability test around the optimum region which again is also not known beforehand. The work around is to perform repeatability tests on the problem itself but without applying background interference.

D. Performance comparison

Now we compare the SUMO approach to the traditional experimentation that exhaustively searches all parameters. For the comparison, we have defined the parameters of the OFI
TABLE II: Duration Gain and Performance Gain of SUMO optimized experiments using 4 sampling methods

<table>
<thead>
<tr>
<th>Sampling Method</th>
<th>Iterations</th>
<th>Duration Gain</th>
<th>Performance Gain</th>
</tr>
</thead>
<tbody>
<tr>
<td>LHS</td>
<td>26</td>
<td>0.9242/0.9287 = 99.51%</td>
<td></td>
</tr>
<tr>
<td>RAND</td>
<td>14</td>
<td>0.6920/0.9287 = 74.5%</td>
<td></td>
</tr>
<tr>
<td>ORTH</td>
<td>20</td>
<td>0.7720/0.9287 = 83.13%</td>
<td></td>
</tr>
<tr>
<td>HSS</td>
<td>53</td>
<td>0.9188/0.9287 = 98.93%</td>
<td></td>
</tr>
</tbody>
</table>

We see from Table II that LHS is the best sampling method in terms of performance gain: it stops the experiment at the 26th iteration with a duration gain of 10 and a performance gain of 99.51%. On the other hand, RAND sampling method converges the quickest but at the expense of a lower performance gain (74.5%). This is because of poor initial sampling and it leads to a local optimum instead of the global optimum. This is also seen in Figure 5 such that bumps appear on the curve (iteration 25-32) had we continue the optimization.

VI. CONCLUSION

This paper investigated the feasibility of the SUrrogate MOdeling (SUMO) optimizer when used in experimental optimization of wireless network problems. SUMO is a powerful optimizer but a number of configurable parameters affect its efficiency. The sensitivity to initial sample size and parameters of Objective Function Improvement (OFI) stopping criterion are investigated in this paper. The initial sample size sensitivity exploits the exploration and exploitation balance of an optimization problem such that with few initial samples, an optimizer locates the optimum in a short period of time. On the other hand, OFI exposes the STD-WIDTH and STD-THLD parameters in order to fine tune the optimization performance. Four sampling methods (Latin Hypercube Sampling, Random sampling, Orthogonal sampling and Hammersley Sequence Sampling) were combined with the SUMO optimization toolbox to optimize the experiment until the OFI stopping criterion is met. In our proof of concept, the LHS sampling method outperforms the others: with only 26 iterations it runs 10 times faster compared to the exhaustive search experiment while achieving 99.51% of the global optimum performance.

ACKNOWLEDGMENT

The research leading to these results has received funding from the European Union’s Seventh Framework Programme FP7/2007-2013 under grant agreement no. 258301 (CREW project www.crew-project.eu) and no. 318273 (LEXNET project www.lexnet-project.eu).

REFERENCES


