ISSN 1870-4069

Implementation of Big Data in Intelligent Analysis of Data from a Cluster of ROVs Associated with System of Prevention and Reparation of Hydrocarbon Leaks to Optimize their Distribution in Gulf of Mexico

Raúl Mendez[1], Alberto Ochoa[2], Benjamín Zayas[3], Miguel Perez[4], Olivia Quintero[5]

[1] Instituto Tecnológico de Parral,
Mexico

[2] Universidad Autónoma de Ciudad Juárez,
Doctorado en Tecnología,
Mexico

[3] Instituto Nacional de Electricidad y Energías Limpias,
Gerencia de Control de Procesos,
Mexico

[4] Instituto Nacional de Electricidad y Energías Limpias,
Gerencia de Tecnologías de la Información,
Mexico

[5] Instituto Tecnológico de Parral,
Departamento de Sistemas y Computación,
Mexico

alberto.ochoa@uacj.mx

Abstract. The objective of this work is to present the multiplayer version of SimROV: a simulator of a ROV (Remote Operation Vehicle). Besides allowing sharing the same scenario between an instructor and different students through a network, the system include artificial intelligence algorithms. At the moment, these algorithms perceive the actions of students within the virtual environment and provide help for students to be guided to successfully achieve the goal assigned by the instructor. Among other things the trajectory of the movements of each ROV is verified, in this way when working on a specific maneuver such as closing a valve to stop a fuel leak, the trajectory of students is compared with the trajectory that is expected. It's worth adding that even without thinking about taking actions to respond to a contingency maneuver, SimROV can also be used for a trainee to learn how to use and control an ROV. This includes not only guiding or directing the ROV to a specific point, but knowing how to control it even in the presence of currents, which make it difficult to maneuver. The control of the arms also requires some expertise that can be obtained with the help of

SimROV, the difficulty that is presented here, is that you have to learn to use the controls to perform an action, which implicitly demands the realization of activities in parallel: move the different controls of each arm, in addition to guiding the ROV within the virtual environment.

Keywords: Virtual reality, artificial intelligence, rov multiplayer, simrov, harmony search algorithm.

1 Introduction

There are several applications for the operation of a cluster of ROVs including maritime leak monitoring in the distribution of oil in underwater scenarios, which is why each ROV for underwater work is equipped with two arms and 2 video cameras. It is worth adding that even without taking actions to respond to a contingency maneuver, SimROV (ROV Simulator) can also be used for a student to learn how to use and control a ROV. This includes not only guiding or directing the ROV to a specific point but also knowing how to control it even in the presence of marine currents, which increase difficult to maneuver it. In addition, a group of ROVs will allow generating a large set of data through the use of sensors. Control of the arms also requires some expertise that can be obtained with the help of SimROV, The difficulty presented here is learning how to use the controls to perform each of the actions to be performed, which implicitly demands carrying out activities in parallel: controlling each arm movements, in addition to ROV navigation within the virtual environment. In this research a Metaheuristic is used to find the most adequate way of dealing with an underwater leakage by a cluster of seven ROVs working collaboratively.

All this process is depicted in Fig. 1. Detecting each ROV position and the movements of their arms will be helpful to identify possible areas for improvement in the student's learning process. This also can feed the student model so that SimROV always knows the progress of the students and their current status, a key aspect in the

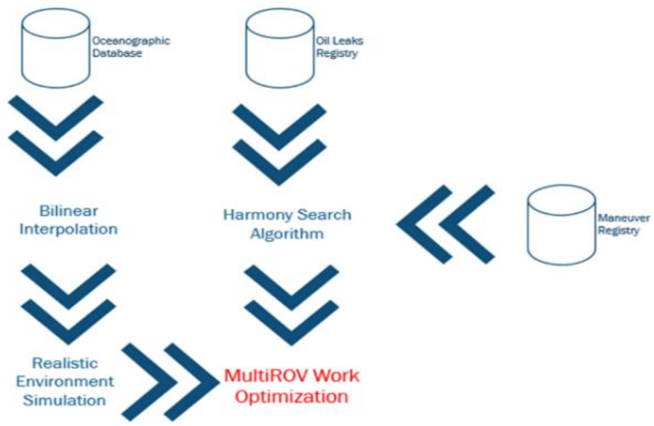

Fig. 1. Optimization of MultiROV.

Fig. 2. Instructor entering current.

Fig. 3. Instructor inserting leakage.

training models by the national petroleum agency in Mexico.The rest of the paper is organized as follows.

Section 2 includes some concepts of simulators using a ROV to model a set of solutions. Section 3 presents the SimROV architecture including the proposal for a ROV Cluster. Section 4 discusses the main achievements of the development of this multiplayer ROV version for SimROV. Section 5 details a multicriteria analysis of their behavior, Section 6 includes some conclusions and the future work, followed by a list of references.

2 Simulation of Underwater Scenarios Associated with an ROV

A simulator for training is an intelligent system, which allows to acquiring some specific type of skills [1].

A literature review on the subject of simulators distinguishes and specifies between two types of simulators:

a) The simulators oriented to the design of facilities (SODF)

b) Training simulators, whose main purpose is training (TS).

In turn, within the TS there are two types, depending on the type of action they generate:

i) Total-range simulators (TRS), which completely reproduce the operating environment. These simulators are associated with a replica of the control room (CR), both in appearance and functions, and seek through the use of multimedia processes and sensors to specify a behavior associated with their learning.

ii) The simulators that reproduce the behavior of the process (SRBP), but do not reproduce the CR, but that allow to act on the simulation in time of execution, and manage to evaluate the experience of its use, with specifications of its implementation.

This research work is intended analyze data generated by both SimROV and a database (DB) from nullschool.net. SimROV could be considered as a SRBP, since is a simulator of Remotely Operated Vehicles (ROVs) [2].

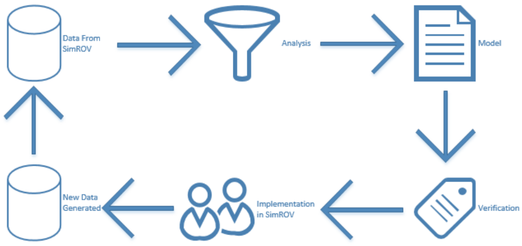

Fig. 4. Data Processing**.**

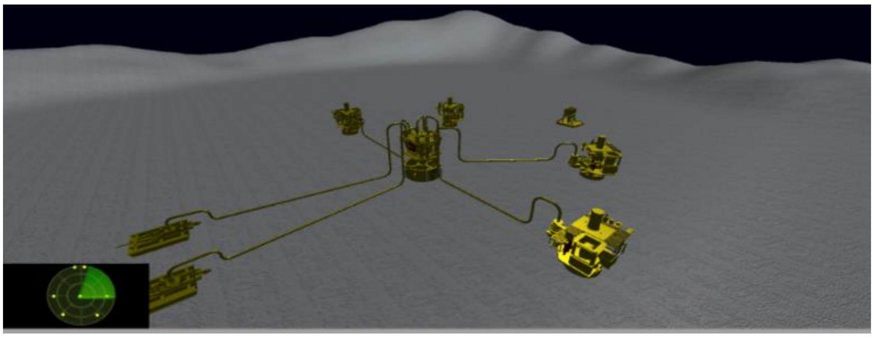

Fig. 5. SimROV environment.

Thus SimROV, is used to train the future ROV's operators. Including different parameters SimROV provides a simulation as realistic as possible.

In SimROV, (as mentioned) the trainees are learning how to operate a ROV, whose purpose is repair oil leakages at submarine pipelines. For this, the trainees and instructors log into the simulator as operator and instructor mode, respectively. Instructors have privileges to introduce leakages and marine currents whereas operators can close valves. [3] These functionalities are shown both in Fig 2. as Fig 3.

The only missing thing for SimROV to be considered as a full range simulator (TRS) is the control cabin.

However, in this multiplayer version of SimROV is possible to perform multiple actions by means of a cluster of ROVs using a framework associated with an innovative Metaheuristic under a Harmony Search Algorithm model. The controls of an actual ROV are based on joysticks and other similar motion controls, while a previous version of SimROV used only keyboard and mouse. The way in the data is processed with this metaheuristic is explained in Fig 4.

Regarding the metaheuristic, is an important part of this research, a software-assistant was developed in R. This language was selected after a comparative evaluation with Matlab [4], where R ended up being the most appropriated for the purpose established because is more portable and GPL-licensed [5], this translates into a considerable development cost reduction. This will be discussed in detail in section V. Considering that the development of the control capacity of the device associated with training operations of the ROVs cluster is being improved, SimROV contains the student model in order to personalize the instruction, besides being equipped with artificial intelligence algorithms to monitor and guide the student in his actions during his response to contingency situations, which are introduced by the instructor to monitor his performance as time progresses and the operator's learning curve increases.

3 Simrov and Why Use A ROV’s Cluster instead of Single One

Working in underwater environments require better skills than maneuver in land or even in the air. In submarines oil ducts, besides dealing with leaks, the operator has restricted vision (because the darkness of the deeps where the oil ducts are located), and might need to fight against forces coming from many sides of the ROV caused by the submarine currents, and so on. This is why is considered a priority that the operators working together during their training. Thus while some are repairing a leak, others would be working with the valves and the rest would be checking for new leaks in the surrounding oil ducts, that might be originated by oil pressure because of the leak repairing work activities.

3.1 Work Together? That is the Question

Working together is a big improvement for the simulation and therefore in the real world. Many people working together could be a mess for the instructor, if each one has a different opinion, a different strategy or simply different expertise. SimROV stores so much data about the work did in the different sessions that not take advantage of it would be a waste a time and resources. Once put this data on the table, is time to show the proposal for SimROV. It is proposed here to develop a software (SW) which calculates the optimal movements to be realized by the trainees. This SW, makes an analysis of the actual conditions of the system (valve's leaking, ROV's positions, arms movements, etc.) contrasting them with the data stored in previous sessions where students were trained. Based on this analysis, it proposes the best maneuver, which theoretically would take less time to the team. With this implementation, is expected that the team not only acquire the maneuver skills, but also learn to implement the best version of them.

3.2 Preparing SimROV for a More Accurate Environment

In order of have a more realistic simulator, it was used a Database (DB) with recordings of real Gulf of Mexico’s currents (speed and angle in certain coordinates). This data was obtained from nullschool.net, project created by Cameron Beccario, which in turn refers to OSCAR (Ocean Surface Current Analyses Real-time, a NASA funded research project and global surface current database). From this DB the DB the specific point the instructor wants can be calculated. For this, the instructor can provide the coordinates where the simulation should be recreated. All this because the procedure should change according the environment state.

For the creation of this DB, it was used a sample of 150 points in the Mexican oil exploitation zone, in 99 dates separated by 5 days one from each other (whose distribution are represented in Fig. 6, where are shown only the means of all the records

Fig. 6. Sampling of the marine currents Source.

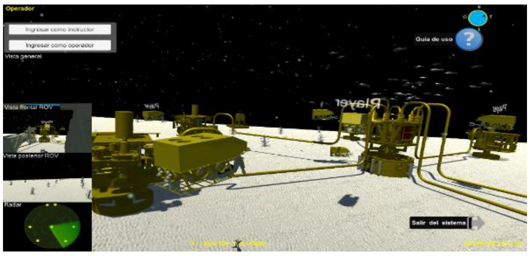

Fig. 7. Example of multiple ROV working together.

in the point). With this information the team expect make SimROV more realistic, adapting the conditions the instructor provides to the actual conditions presents in the real Gulf of Mexico. The first step, is make a bilinear interpolation for the speed with the surrounding points in the DB. This mathematical method was selected due its simplicity compared to other methods like bicubic interpolation or Bézier Surface and its accuracy in comparison with the most popular K Nearest Neighbors (KNN)). The second step, is calculate from the speed the exerted force by water currents in Newtons, while the pressure of a fluid in a pipe is measured in megapascals. To get this value is used the drag force formula, represented in equation 1:

$$F = \frac{1}{2}\rho v^2 C_D A \,, \tag{1}$$

where:

1. P is the density of the fluid. In this case is water, ergo 1000 kg/m^3.
2. v is the flow speed.
3. C_D is the drag coefficient for the object. Since the ROV is basically a cube. In order to simplify the calculations, this is going to be used as 0.8.
4. A is the surface area normal to the fluid flow.

The final step is to calculate the direction, for which is going to be used a bilinear interpolation too.

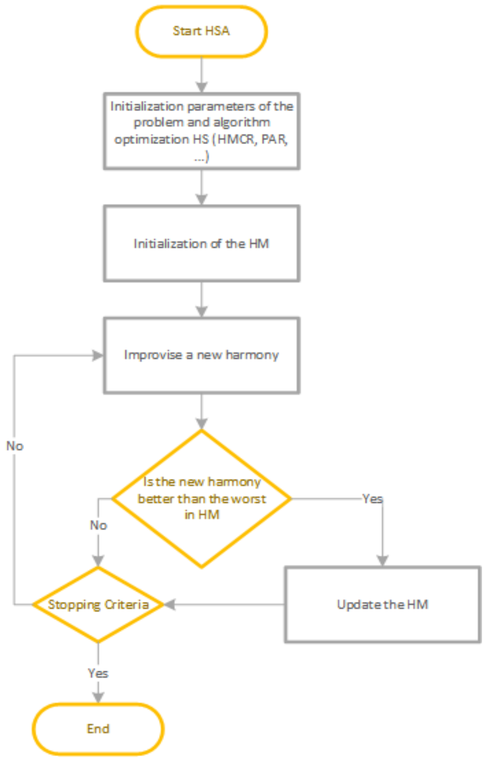

Fig. 8. Harmony Search Algorithm flowchart.

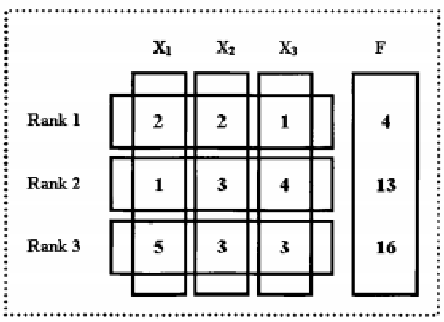

Fig. 9. Initial Harmony Memory Structure Source: [3].

4 Achievements of MultiROV

As mentioned, is really hard to work in submarines environments, besides that the ROV has limited resources (such only two arms) with this improvement in SimROV, is expected to see a significant reduction both in time and cost.

5 Multi-Criteria Analysis

For the creation de decision maker would be considered the work realized for the teams that used previously SimROV, analyzing their maneuvers (which include the movements of the ROV and the actions did with the arms) and the environment conditions.

5.1 Harmony Search Algorithm

The harmony search algorithm (HSA), inspired by musical performance process, consists of three operators: random search (RS), harmony memory considering (HMC) rule and pitch adjusting (PA). Standard HSA applies the HMC rule for both exploration and exploitation phases, RS for the exploration phase and PA rule for the exploitation phase. Indeed, the HMC rule acts as an exploration operator during the initial iteration because of large variations in the harmony memory (HM). However, as the iteration continues, this operator behaves as an exploitation agent for the HAS. [6] Since it first appearance in 2001, it has been applied to solve many optimization problems including function optimization, engineering optimization, water distribution networks, groundwater modelling, energy-saving dispatch, truss design, vehicle routing, and other [7].

In order to use the memory more effectively, it is typically assigned as a parameter $r_{accept} \in [0,1]$, called harmony memory accepting or considering rate (HMCR). If this rate is too low, only few best harmonies are selected and it may converge too slowly. If this rate is extremely high (near 1), almost all the harmonies are used in the harmony memory, then other harmonies are not explored well, leading to potentially wrong solutions. Therefore, typically, we use $r_{accept=0.7\ 0.95}$ $r_{accept} = 0.7\ 0.95$. [7]. The algorithm originally proposed in [8] (and graphically represented in Fig. 8) was this:

Step 1. Initialize a Harmony Memory (HM) with randomly generated solutions.

Step 2. Improvise a new harmony from HM.

Step 3. If the new harmony is better than minimum harmony in HM, include the new harmony in HM, and exclude the minimum harmony from HM.

Step 4. If stopping criteria are not satisfied, go to Step 2.

An example of the structure of Harmony Memory (HM) is shown in Fig. 9.

For the implementation of the current solution of the SimROV's problem was used the package metaheuristicOpt, which is free under the GPL-2 and GPL-3 license. This library includes many optimization algorithms, among which is Harmony Search (HS), this version of the algorithm originally proposed by Geem recommend certain values for the original values of HSA, for each of them are an explanation in [9]. Even so, it was decided make some empirical test to get the more optimal parameters given the characteristics of the problem. Once analyzed the algorithm, is moment to dive into the problem and the solution. Looking in the literature for similar problems and the way they were solved, is justly in the original article of HSA where are three examples of the applications for the algorithm.

In this case which cares is the first one, where like the SimROV's problem, is presented a problem with a discrete solution. In that case, HSA is applied to Traveling Salesman Problem (TSP) connecting points in a plane that represents the route for the salesman. Extrapolating the solution, the maneuvers of the ROVs are interpreted as linear movements, where it has a start and end point in the 3-dimensional space. Also, these maneuvers have registered the time it took and the valve that was repaired. As mentioned above, the library used is based in [9], and since the results were acceptable.

It was chosen use the default parameters included in the package. Even so, could be a theme for future research that born from this the optimization of the algorithm's parameters.

5.2 Cost Function or Time Function

The cost function, which is evaluated multiple times by HSA, consists in a function, which receives 3n variables, where n is the number of valves to repair. The variables follow the following structure:

a) $x_i$is the number of valve to repair,
b) x_{i+1} is the rov assigned to repair the valve x_i,
c) x_{i+2} is the priority of the valve to being repaired. Is important to remark that this number is not specifically sequential, instead of it, it only represents which one goes before which.

All this considering:

a) $i \in \{3N\}$,
b) $1 \leq x_i \leq 6$,
c) $1 \leq x_{i+1} \leq 7$,
d) $1 \leq x_{i+2} \leq 6$.

Once defined the variables, the function starts the evaluation, looking in the database for the optimal maneuver. The algorithm for determinate the optimal maneuver is the following:

1) Select all the previous maneuvers that repairs the specified valve.
2) Determinate the speed of the maneuver, for which it calculates the Euclidean distance traveled divided by the duration of it. This operation gets the unitary speed:
$$\mathrm{v} = \frac{d_E(m_0, m_f)}{t} \, .$$
3) Calculate the "adjustment time", that is to say, the estimated time that would takes to the ROV goes from its start position to the maneuver's start position $\Delta\mathrm{t} = \frac{d_E(P_r, P_m)}{v}$.
4) Sum the adjustment time plus the maneuver's duration and get the maneuver that has the minor estimated total time $t_e = \Delta t + t$.
5) Repeat the process for each valve. If a ROV repairs multiple leakings, the times are sum.

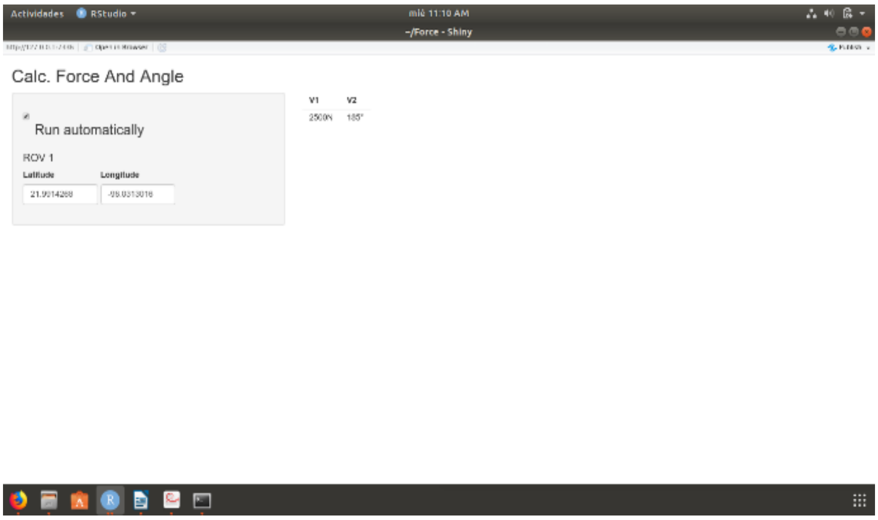

Fig. 10. Calc. of currents.

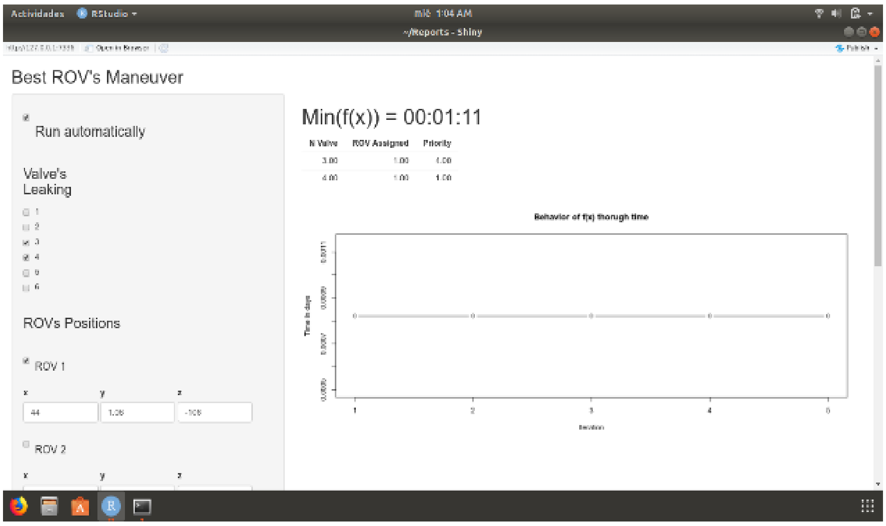

Fig. 11. SW for best maneuver.

6) Once all the valves' repairing time are calculating gets the higher which is interpreted as the global and is returned to the HAS.

Is important to remark that some parts of the algorithm are made in R, while others are implicit in the SQL queries or in the views made to simplify them. The tools developed for solve the optimization problem consists in two GUIs. The first one is in charge of calculate the velocity, therefore, the drag force of the submarine current in the given point (Fig 10). This receive the coordinates and the date of the year from which apply the methods and formulas mentioned previously.

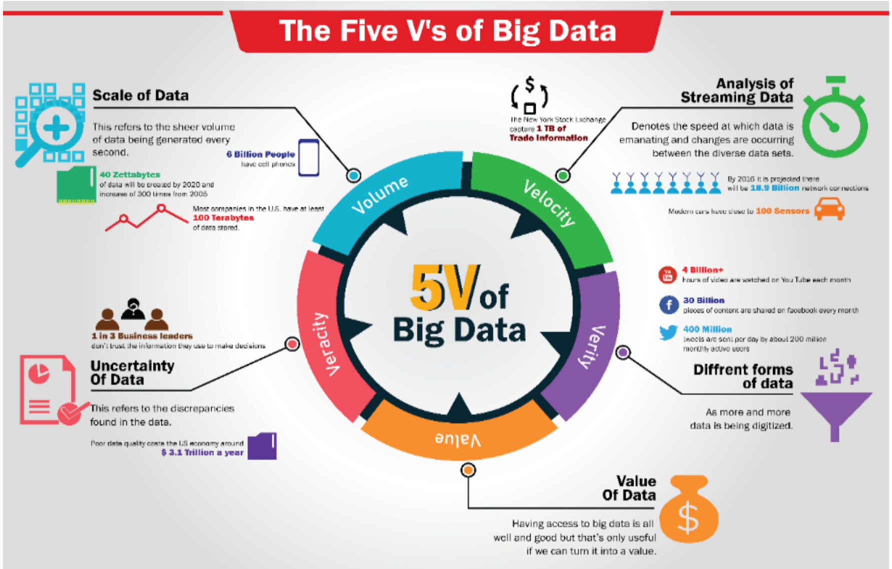

Fig. 12. Considerations for Big Data.

The other GUI is made for the user input the number of the leaking valves and the coordinates of the actives ROV. After the processing, is shown the minimum time, the maneuvers that the operators should make and the behavior of the f(x) value during the iterations of the algorithm (Fig 11).

5.3 Big Data

Big Data can be considered as a trend in the advancement of technology that has opened the door to a new approach to understanding and decision making, which is used to describe the huge amounts of data (structured, unstructured and semi-structured) that it would be too long and expensive to load a relational database for analysis. Thus, the concept of Big Data applies to all information that cannot be processed or analyzed using traditional tools or processes. In general terms, Big Data and the processes that this technique represents has a wide spectrum of potential applications. The biggest challenge for investment in Big Data occurs in relation to projects related to decision-making on a large amount of data in decision-making, definition of strategies and obtaining better experiences on consumer consumption actions. The challenge of Big Data is to capture, store, search, share and add value to the little used or inaccessible data to date. The volume of data or its nature is not relevant. What matters is its potential value, which only new technologies specialized in Big Data can exploit. Ultimately, the goal of this technology is to provide and discover hidden knowledge from large volumes of data.

In our investigation in particular, it will be used for the correct decision making in the discovery of major affectations derived from a fuel leak in an underwater scenario [10].

6 Results

An example of the work made by both SW is shown below. Where we enter a coordinate pair and the SW realize the interpolation for calculate the corresponding drag force. The instructor will enter this into SimROV and voila, the simulation is a bit more near to reality. In the side of the optimization of the maneuvers, we have a similar GUI, where enter the data (in this case, the leaking valves, the active ROVs and their corresponding coordinates in the space). After run the optimization, we get the minimum time, the order on which the valves should be repaired to get that time approximately and ends with the behavior of f(x), this more focused on watch how the algorithm work, ergo for technical purposes.

7 Conclusions and Future Research

Due to this work was possible implement an innovative algorithm, which met expectations in its work getting the optimal values. In some cases, even better than the expected due to human mistakes. Also, is important emphasize that we hope the SW would be important part in the improvement of the ROV operator's training procedure. Besides all the proposals for future research we would like to make a really big one: all this data could be used not only to train humans; this could be part of the supervised learning for Autonomous Underwater Vehicles (AUVs). Touching a related point, the research has a big trouble about time. Due that, was not possible use a big amount of data in the analysis of maneuvers. Since generate it would take a huge amount of time.

Also, as was mentioned in some points of the article, is from this site where can born many research branches some that we proposed are:

1. Optimization of the HSA's parameters for this specific problem.
2. Development of a version of HSA specially designed to discrete problems, which was not found during the time this research was made.
3. Implantation inside the SimROV.
4. Improvement of the SimROV to handle better the work with the valves (even that, the actual functioning of the solution proposed here are prepared to work with this).

8 Future Research

Could be a great future research to join the SW made as product of this research with SimROV in one unique, which putting together the best part of both, would result in a powerful learning tool. Moreover, even more can be found from this one. An important aspect to highlight will be to improve the design of experiments using the ad hoc validity of an orthogonal design.

References

1. Gopinath, C., Sawyer, J.E.: Exploring the learning from an enterprise simulation. Journal of Management Development, 18(5), pp. 477−489 (1999)
2. Beltran-Labra, R., Pérez-Ramírez, M., Ochoa-Zezzatti, C.A., Ontiveros-Hernández, N.J.: Simulador ROV multiplayer para escenarios petroleros submarinos. Research in Computing Science, 140, (1), pp. 66−77 (2017)
3. Olguín-Tolentino, N.D., Ingram-Ramírez, J.M.P.R., Nava-Ayala, M.Y.E.R.: Prototype of a ROV simulator using dynamic animation and Virtual. CIINDET (2016)
4. Ozgur, C., Colliau, T., Rogers, G., Hughes, Z., Myer-Tyson, E.: MatLab vs. Python vs. R. Journal of Data Science, 15, pp. 355−372 (2017)
5. R-Project: https://www.r-project.org/Licenses (2019)
6. Kim, J.H.: Harmony search algorithm: A unique music-inspired algorithm. In: 12th International Conference on Hydroinformatics (2016)
7. Yang, X.S.: Harmony search as a metaheuristic algorithm. In: Music-Inspired Harmony Search Algorithm: Theory and Applications, Berlin, Springer, pp. 1−14 (2009)
8. Geem, Z.W., Kim, J.H.: A new heuristic optimization algorithm: Harmony search. Simulation, 76(2), pp. 60−68 (2001)
9. Mahdavi, M., Fesanghary, M., Damangir, E.: An improved harmony search algorithm for solving optimization problems. In: Applied Mathematics and Computation, Iral, Elsevier, pp. 1567−1579 (2007)
10. Fragoso, R.B.: IBM. https://ibm.com/developerworks/ssa/local/im/que-es-big-data/index.html (2019)