Agent-based Simulation of Information and Communications Technology Practicum Courses for Engineering Students at Instituto Tecnológico Metropolitano

Julian A. Castillo, Cristian F. Gallego, Yony F. Ceballos, Carlos O. Zapata.
Engineering faculty
Instituto Tecnológico Metropolitano
Medellin, Colombia
{castillojulian, pharsat, fceball}@gmail.com; carloszapata@itm.edu.co

***Abstract*—In this paper, a simulation of enlisting and use of free practicum courses for engineering students from Instituto Tecnológico Metropolitano (ITM) is presented. Currently, this university is interested in getting better results to rank more competitively among other universities. Although the current courses are free, a high quality offer is evidently more expensive and we need to maximize the impact of the students in the labor market. In Colombia, many universities certify engineering students in computer sciences so, in order to attend the current needs of our institution and the labor market, we would like to contribute to decision-making processes through agent-based simulation regarding what Information Communications Technologies courses may fit best for the community and the industrial clusters ITM students are and will be working in.**

Keywords-Agent-based modelling; Computer engineering; Simulation; Netlogo.

I. INTRODUCTION

The ITM is a public university whose mission aims to provide charge-free education in different areas. Besides, ITM is in touch with the needs of the labor market. However, the current techniques to identify educational mechanisms to meet labor market needs are based on online surveys applied to students during practicum and to recently graduated bachelors. This may become a reason to use simulations that help determine how these different techniques of information gathering on the educational offer will allow this population to impact positively the labor market in the future, making possible to see devise educational mechanisms for students to learn about new technologies in the industry and be able to found cutting-edge enterprises.

ITM bachelors have particular features. They are educated through propaedeutic cycles and belong to low socioeconomic sectors of the city (see Figure 1). Therefore, the institution makes efforts to form competitive bachelors allocating resources for them to have access to free training during their practicum stages, so that they are able to improve their quality of life and learn the skills necessary to compete in equal conditions with students from other highly qualified universities. In order to contribute to decision-making processes aligned with the university' mission and its population, modeling and agent-based simulations are proposed in this research to identify where it might be necessary to provide student training and decide how the university should invest resources to meet the needs of students and local enterprises.

Figure 1. Valle de Aburrá Urban Area. Grey sectors represent bachelors geographical location in the Metropolitan area.

As the student and bachelor population of the institution is small and many simulation techniques usually need a huge amount of input data, we have used Netlogo. This software is used to provide low abstraction levels to characterize a small population in a simple and detailed manner, resulting in a reliable and easy way to validate a simulation [2]. Moreover, we have chosen this low-cost simulation tool that provides close-to-reality results regardless of small populations. Figure 2 displays the distribution of enterprises in the urban area of Medellin city, those who are the main information source for this model.

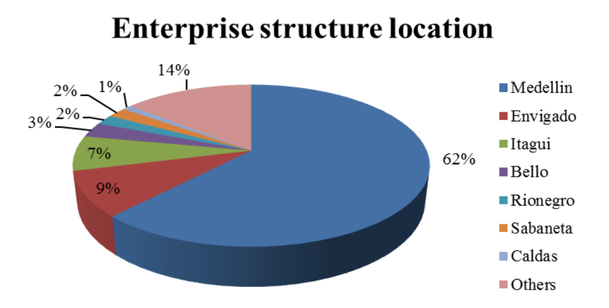

Figure 2. Distribution of enterprises in Valle de Aburrá urban area.

For this simulation, local industrial clusters were identified in the city [3], which generally hires bachelors from different higher education institutions. These clusters are leading organizations of software development groups. They are located in strategic coordinates of the city. Nearly

 ISBN: 978-1-61208-371-1

4-5 enterprises per cluster are development companies. From these, two of the most remarkable are the Tata Consultancy Services Ltd (TCS) cluster, located in Olaya Enterprise Center and Hewlett-Packard in Ruta N Cluster (see Figure 3).

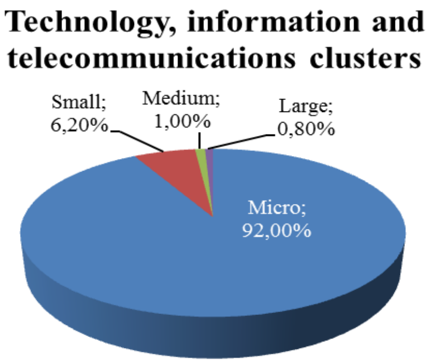

Figure 3. Typology of Enterprises in Medellin Urban Area [3].

For this simulation, we have also developed an agent-based model. The purpose was to know the relation between ITM bachelors' interaction in the mentioned market sectors (clusters) with the creation of training courses at ITM. These relations have given rise to a number course programs that respond to the wishes of students and companies to reach an intermediate point where the decisions in training and learning and training are the most appropriate for both agents.

Section III discusses the ODD (Overview, Design Concepts and Details) protocol [4] used to build the simulation. Then, we will present an overview of the description of the model. Next, in section IV, we will report the results of a set of preliminary simulation experiments. Finally, in the section V and VI, we will provide concluding remarks of our study and provide a summary of the future work that has been projected from this study.

II. JUSTIFICATION

The proposed research project is part of the modeling and simulation of systems by agents, which is an area that has been barely explored in Colombia due to its high computational requirements and random components [1].

This technique differs from other modeling techniques in the way that abstraction of the real system and consequently, the functional model, is constructed. In the models constructed using agent-based simulation, the basic components of the real system are explicitly and individually represented in the model. Thus, defining the basic components of the real system correspond to the boundaries that define the model agents, and the interactions that take place between the basic components of the real system correspond to the interactions that take place between agents in the model [5][6]. This direct correspondence contrasts with the traditional use of 'representative agents' and is able to increase the realism and scientific rigor of formal models built using this approach.

Agent-based systems are characterized by comprising several agents that are to a greater or lesser degree, autonomous, heterogeneous and independent, each showing their own goals and objectives, and are generally able to interact with each other and with their environment [7].

They are usually characterized by the existence of a large number of relatively simple agents that can evolve over time to adapt to new environmental conditions or to new systems objectives. The agent-based methods facilitate the study and modeling of complex systems from their component units, allowing to build experimental models of reality from a different point of view to the traditional: from the simplest to the most complex. Furthermore, an intrinsic agent's result is the idea of an emergency. Emergent phenomena are patterns that emerge from the decentralized interactions of individual simpler components [8].

What characterizes these emerging phenomena is that their presence or appearance is not apparent from a description of the system consisting of the specification of the behavior of the individual components and rules of interaction between them [9]. A typical example of an emerging phenomenon is the formation of distinct groups in the Schelling segregation model [10]; the emergence of clear patterns of segregation is not explicitly imposed in the model definition, but emerges from the local interactions of individuals with segregationist tendencies in weak times. Another example is the migration patterns of Sugarscape [11]. In the next section, we show the ODD protocol [12] used in agent-based models.

Given the small number of courses, the population size, geographic location and heterogeneous characteristics of the population, the simulation technique used is based on agents as it provides a complete mapping of the selection process.

III. SIMULATION PROCESS

We use an agent-based approach to build a simulation of the relationship between the ITM for bachelors with medium to low income levels in the Metropolitan Valley in Medellin – Colombia and practicum courses. This involves a relation made by the bachelor with the ITM and the company where the bachelor works for, the training courses that are made for them and how they make the decision based on a company request and their own preferences and if the course is available in the curricular offer. The process is described as the ODD protocol [4].

A. *Purpose*

The purpose of this model is to make decisions for the future position ITM a pioneer in the process of educating highly trained professionals that learn in relevant aspects needed in business and the Institution. This leads to making decisions based on the model to allow the institution to grow in terms of academic extension courses.

B. *State variables and scales*

Bachelors are located in the metropolitan area and are related as agents with a cluster. This determines their initial position as return values when the learning cycle ends. Each bachelor also has a preference for a specific course given by a vector of 10 positions on which each position determines their sympathy for the course of their position, e. g., if at the 3 position vector of sympathy there is a value close to 10, it means that their sympathy to take the course "3" is high; instead, if at position 4 there is a value close to zero, it means that their sympathy to take the course "4" is low. Both the student and the company for which they work (Cluster) have preferences and it is this relationship that determines the selection of a training course.

The cluster as agent has the ability to create its own sympathy for courses randomly, which is defined as a 10-position vector where each position represents a course. There is also a unique cluster number that identifies its own position.

The university as an agent is related to its two main sites that are identified by numbers. With one to two, each has a random vector of courses that are offered, which cannot be further repeated also defined in 16 weeks with a weekly intensity of 4 hours.

An array of preferences associated with each cluster and student is defined. After that, a new array is created using the addition from position to position of the elements of both arrays, of which a maximum value for the course is chosen. This value represents the most suitable for the cluster and student, simultaneously. To select this position, a run on the array of partial sums is made. The ITM has an estimated dropout rate with a maximum value of 30% of bachelors enrolled in training courses; this value is cumulative for each bachelor and accounts for a temporary variable that determines the possibility of dropping out from or a course.

The main attributes of these three classes of agents are shown in Figure 4.

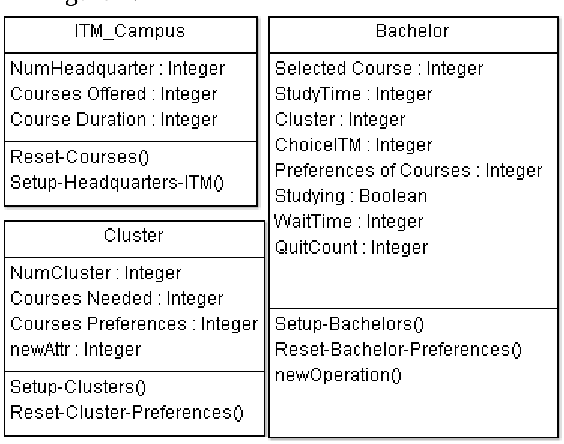

Figure 4. Class diagram agent.

All agents need to be initialized with values for every variable. The values are generated based on a set of hypothetical assumptions.

C. Process overview and scheduling

Description of the World: The world also known as 'work area' is defined by the location of the agents located in the Valle de Aburrá within the Department of Antioquia - Colombia. Valle de Aburrá has several locations that belong to the metropolitan area. This simulation considers the clusters closely located to the influence area of the ITM. The simulated clusters are located as seen in Figure 1.

According to the dictionary of technology companies registered in ClusterTic, these were referenced by 6 strategic points located spatially on xy coordinates on the map [1]. Each cluster can be identified as follows: 1 Cluster Ruta N, 2 Bello Cluster, 3 Poblado Cluster, 4 Olaya Centro Empresarial Cluster, 5 Candelaria or Downtown Medellín Cluster and 6 Envigado and Sabaneta Cluster. For the space, each patch is defined with a unit value and a 400 by 400 patches long map on which the image is overlaid with political delimitations of Valle de Aburrá. Each agent interacts in the map and each is spatially associated with each other. The loaded image below has been enhanced to visualize this datum using a filter process in 3 scales: gray, black and white (See Figure 5)

Agents Interaction: Agents that interact in the process are denominated: Bachelors, Clusters, and ITM's (two campuses). The central element on which the simulation is performed is the Bachelor. We aimed to simulate whether a bachelor takes a training course or not and whether the bachelor finishes it or drops it, Nonetheless, this information measurements are taken into a simulation process of 10 periods of 2 years interactions each.

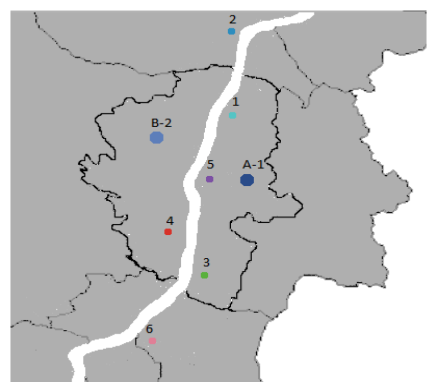

Figure 5. Location of cluster and university campus.

The main process is defined as the student's empathy to choose one training course mixed with the enterprise needs they belong to, and this process is conditioned to make a decision about the ITM course offers.

Bachelors: Each bachelor is born randomly in each cluster; the number assigned to them is according to a max number of bachelors determined by a random variable ranging from 100 to 500 agents. For each bachelor, it is defined an array of 10 random values for training course selection empathy of each training course [2]. These values have no relation between them.

Cluster: Each cluster is an agent. A cluster is created spatially approximating to the actual position in the metropolitan area for a total of 6 clusters. Each has a number that identifies it, and has a need for knowing of same training courses. For instance, when an enterprise recommends that a bachelor employee (bachelors) take a certain course, there is a preference for courses that would affect the decision of the graduates.

The ITM: It is composed of two branches, each of them represented as agents, and they are symbolized by an identification number (Figure 5 B2 is for the Robledo Section in the west of the metropolitan area and the A1 is for the Fraternidad placed in the east of the same area). Each ITM campus offers a random number of training courses from 1 to 10 as maximum; one ITM campus cannot have all courses [2]. The relationship between each entity is shown in Figure 6.

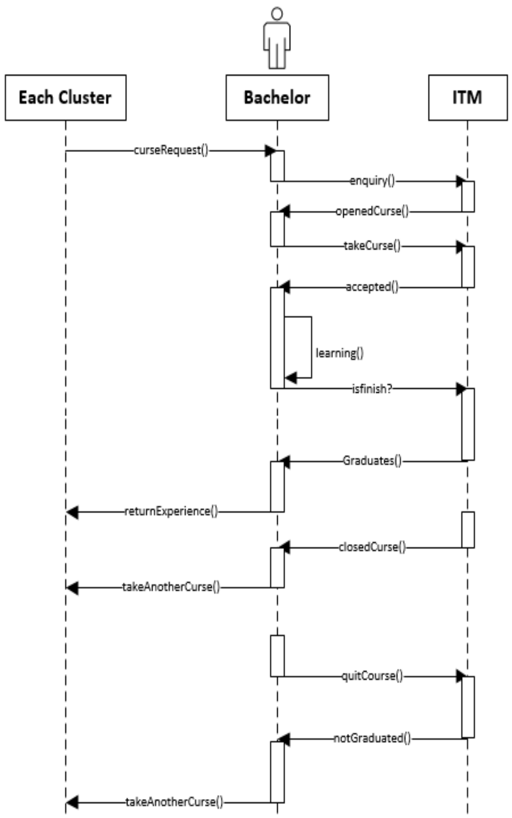

Figure 6. Sequence Diagram (Bachelor and training course flow).

Decision: As seen in the section of state variables and scales, the decision of a bachelor is taken according to the highest greater value of sympathy for a course added to the sympathy of the courses by the company. After this, the algorithm verifies that the student is not taking courses by measuring two bachelors' own variables, which are "chosen course" and "study time"; both must be zero in order to make the assignment of the course.

The decision is assigned to the variable "course chosen", having the decision taken involving both the company and the students. Then, the bachelor proceeds to verify in which ITM campus this training course is offered.

Training course in any ITM campus must exist in order to move a bachelor for training, which will take 16 weeks decreasing with each time lapse in the software (Tick) one to one.

In order to do the training when a bachelor is assigned to an ITM, is assigned the course time left for complete it (16 weeks, 16 ticks) this number decreases with each tick; when it reaches to zero, then the bachelor leaves the ITM returning to the original cluster. When a bachelor returns to their own original cluster, their training course preferences refresh.

There is a possibility of that a bachelor not to take a course, however is needed a required value of sympathy for its selection. This happens when the training course chosen by the bachelor and by the enterprise is not offered by any ITM campus. In this case, the bachelor awaits a maximum of 5 weeks; then, the bachelor refreshes their own preferences.

Every 20 weeks training courses are refreshed by the ITM, also it holds a 30% bachelor desertion rate. In order to simulate this variable, we defined it by a pseudo-random value between 0 and 1. If this value is lower than the exogenous variable "likelihood of attrition" (0.3), it is excluded from the course. After this, training courses preferences are refreshed.

For this process, we defined 104 weeks or two years, being the average period for an extension academic program.

D. Validation

The fitting analysis of the data obtained by the simulation is performed using the information contained in [2], compared to historical data from the year 2013. In addition, expert knowledge is used in order to validate the seasonality of the series obtained together with the characteristics that define each agent and the taking courses frequency.

IV. Results

The simulation setup was executed having the following parameters: the inclusion probability of each training course for bachelors was 50%, the probability of desertion was estimated on 15%, the amount of bachelors was 150, and the empathy probability with a training course was presented with 10%. For the simulation process, we set and run a 10 time round of simulations for periods of two years each one, meaning 104 weeks or ticks. For coding, the software Netlogo ® was used [13].

As seen in Figure 7 for the Robledo campus, there is an oscillation in the amount of bachelors in the training courses for all clusters. However, around week 54 there was a re-initialization of the bachelors who started doing another training course. This may have occurred because the courses are restarted at week 60 of the academic semester.

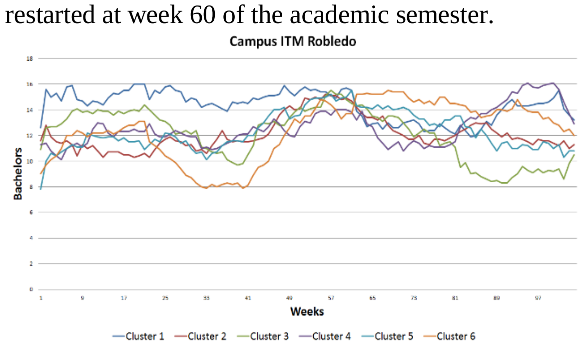

Figure 7. Bachelors taking a training course by cluster.

In the Figure 8, there is a change in the trend because in the description of the model courses are redefined based on information from the office of academic extension.

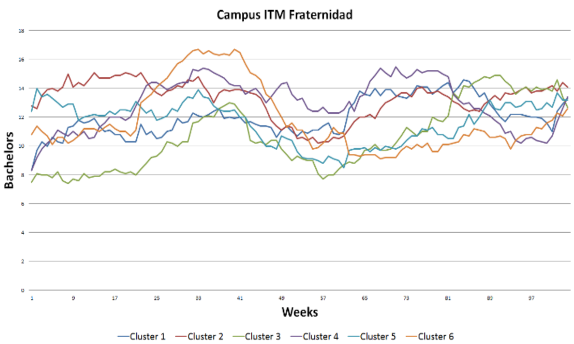

Figure 8. Bachelors taking a training course by cluster.

The graphs are related because the students in the courses belong to the same population and it is presented in Figure 7 as a valley and in Figure 8 as a peak and vice versa. This is caused by the courses offered per year at each campus are exchanged so that generates a migration of the number of students between campuses.

In Figure 9, we can see the student dropout, which is exacerbated in the valley of the simulation periods, as presented in the above figures is observed.

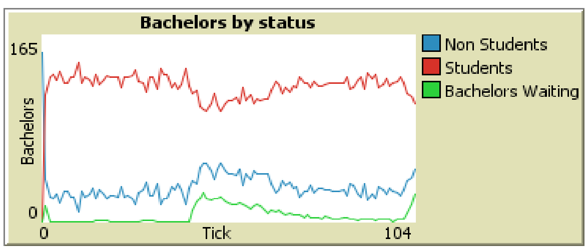

Figure 9. Bachelors taking a training course by cluster.

It is suggested and implementation of short courses are planned to capture students who drop out often, given the duration of 16 weeks of regular courses.

V. CONCLUSIONS

According to the results obtained for the simulation generates an inverse correlation in the number of students shown in both graphs, it means when there is increase of Fraternidad students will notice a decrease in the Robledo campus.

The variability of the courses influences the dispersion of students through them, and if the same training course is offered in both campuses the amount of bachelors decrease in the related University.

The flow on the increase and decrease of the number of students per campus is consistent with the distribution of students in courses on campus with a fluctuation of 20 weeks

To improve competitiveness in the local market of universities and the possibility for graduates to access a better labor compensation, will be promoted from the first half of 2015 a project based on this simulation, taking the information generated in the cluster and weeks in which the courses are taken. The offer will be screened with seasonality allowing a proper rotation of courses taught previously. Desertion of each student and students waiting will be considered for a course as well as seen in Figure 9. This is necessary to reduce costs because the university takes the fee of these courses.

VI. FURTHER WORK

For future work, the amount of students who change campus completing courses will be quantified. We will measured how many students quits courses, as is the average waiting time of a student with a chosen course without availability by the University, the preferred course selected by students and their relationship with their respective campus.

REFERENCES

[1] AMVA (Área Metropolitana del Valle de Aburra), "Area Digital - Area urbana de Medellin," Municipio de Medellin, 2014. Available in spanish http://www.clustertic.co/site/DescubraelemClusteremTIC/Qu%C3%A9esiClusteriTIC.aspx. [retrieved: august, 2014]

[2] I. Ramirez, "Characterization of college graduates ITM" "Caracterización de egresados de la institución universitaria ITM," Medellin, 2013. Available in spanish http://www.itm.edu.co/autoevaluacioninstitucional/acreditacion/imagenes/Anexos/73.InformesCaracterizacionEgresados20122.pdf [retrieved: august, 2014]

[3] AMVA (Área Metropolitana del Valle de Aburra), "Cluster Technology, Information and Communication" "Cluster Tecnología, Información y comunicación," Municipio de Medellin, 2014. Available in Spanish http://www.clustertic.co/site/Comunidad/Directoriodeempresas.aspx. [retrieved: august, 2014]

[4] V. Grimm, U. Berger, F. Bastiansen, S. Eliassen, V. Ginot, J. Giske, et al, "A standard protocol for describing individual-based and agent-based models," Ecol. Modell., vol. 198, no. 1–2, Sep. 2006 pp. 115–126.

[5] B. Edmonds, A. Building, and A. Street, "The Use of Models-making MABS actually work," Cent. Policy Model. pp. 15–32, 2000.

[6] Q. B. Le, S. J. Park, and P. L. G. Vlek, "Land Use Dynamic Simulator (LUDAS): A multi-agent system model for simulating spatio-temporal dynamics of coupled human–landscape system," Ecol. Inform., vol. 5, no. 3, May 2010, pp. 203–221.

[7] I. S. Torsun and I. Torson, Foundations of intelligent knowledge-based systems. Academic Press, 1995.

[8] J. H. Holland, Emergence: From chaos to order. Oxford University Press, 1998.

[9] N. Gilbert, "Agent-based social simulation: dealing with complexity," Complex Syst. Netw. Excell., 2004.

[10] T. C. Schelling, "Dynamics models of segregation," J. Math. Sociol., no. 1, pp. 143–186, 1971.

[11] J. M. Epstein and R. L. Axtell, Growing Artificial Societies—Social Science from the Bottom Up. MIT Press, 1996.

[12] J. G. Polhill, D. Parker, D. Brown, and V. Grimm, "Using the ODD Protocol for Describing Three Agent-Based Social Simulation Models of Land-Use Change," J. Artif. Soc. Soc. Simul., vol. 11, p. 3, 2008.

[13] U. Wilensky, "NetLogo: Center for Connected Learning Comp.-Based Modeling," Evanston, IL: Northwestern Univ., 1999.