Budget-Aware e-Learning Systems on Cloud Computing Environments: A Genetic Approach

Gamal A. Ebrahim

Abstract—**Cloud computing can be considered as one of the most promising environments for e-learning systems. Unfortunately, services provided by the cloud are not always free. Hence, the primary goal of this paper is to select learning resources that best fit learner's budget dedicated for the e-learning class. In other words, learner gets the best e-learning curriculum that he can afford. The problem is modeled as a constrained optimization problem and a proposed solution based on genetic algorithms is introduced. Simulation study shows that the proposed solution provides results identical to optimal solutions in most cases.**

Index Terms—**E-Learning systems, cloud computing, genetic algorithms, learning objects.**

I. Introduction

Cloud computing [1]-[4] is introduced mainly to promote collaboration. It offers users three usage scenarios; infrastructure as a service (IaaS), platform as a service (PaaS), and software as a service (SaaS). These scenarios provide the required services to users with much lower cost in a pay-per-use basis compared with having the services locally. Cloud computing supports learning anywhere concept of the e-learning systems. In other words, it reduces the requirements at the learner's side to just a web-browser on a computing device. Unfortunately, it imposes a cost of using computing resources on clients (educational institutes, individual learners … etc.). This cost mainly covers the licensed software cost that supports the e-learning system at the cloud in addition to the cost of utilizing the hardware resources of the cloud.

Several research efforts have been accomplished on e-learning systems on cloud computing environments. The study in [5] concludes that e-learning on cloud computing environments can reduce the cost and enhance the learning process management. In contrast, the study in [6] verifies that cloud computing can be used to build the next generation platform-independent e-learning systems. Meanwhile, the study in [7] shows that the use of cloud computing improves periodical IT tasks. On the other hand, the study in [8] concludes that utilizing cloud computing as an e-learning ecosystem provides a fault tolerant system against hardware/software failures. Additionally, the study in [9] shows that e-learning systems on cloud computing environments could render the cost needed for building an e-learning system affordable.

A major problem of using e-learning systems on cloud computing environments is to select the best curriculum that fits the client's budget. This does not mean the cloud sacrifices the curriculum quality but means it selects the curriculum that utilizes certain licensed software with certain needs of processing and storage at the cloud to fit the client's budget. However, the integrity and quality of the curriculum are maintained. This problem is tackled in this paper based on genetic algorithms. Mainly, because genetic algorithms have powerful capabilities to search in huge search spaces [10]-[13]. They have been adopted in the study in [14] to find the best set of curriculums that represents a learning path for the learner based on his pre-test score. However, this study does not consider the impact of cloud computing on e-learning systems. More importantly, it deals with the curriculum as a one atomic entity. In general, curriculums should be divided into a set of learning objects [15]. Then, the most appropriate learning objects for the learner should be selected.

The rest of this paper is organized as follows: section II presents the mathematical model of the problem. Then, section III details the proposed algorithm followed by section IV that provides the results of an experimental study of the proposed algorithm. Finally, section V concludes this paper.

II. Mathematical Model of the Problem

Assume there is a cost that will be incurred by the learner for his e-learning class. This cost corresponds to the services provided by the e-learning system for the content provider, the software licensing, and the running time and storage used by the learner at the cloud. Moreover, the learner seeks a specific set of learning objects that provides him the required curriculum corresponding to his learning study level. The problem of seeking a curriculum corresponding to learner's study level has been studied in the past. However, the complexities of using e-learning systems on cloud computing environments impose several challenges. First, the cost of using the resources at the cloud must be paid. Second, there could be many contents that can be utilized to deliver certain curriculum with each one of them has its own cost, capabilities, benefits given to the learner … etc. Third, there should be a certain budget dedicated by the client to the e-learning class. Forth, there could be certain preferences in the delivered contents from the learner's point of view. For example, some learners need to have much graphics and animations in their delivered contents. Finally, there could be several ways to provide the curriculum by composing them from their elementary learning objects. Unfortunately, learning objects may not be compatible from the learner's perspective. For example, one learning object may

Manuscript received August 12, 2011; revised October 3, 2011.

Gamal A. Ebrahim is with the Computer and Systems Engineering Department, Faculty of Engineering, Ain Shams University, Cairo, Egypt (e-mail: gamal.ebrahim@eng.asu.edu.eg).

DOI: 10.7763/IJIET.2012.V2.208

extensively use graphics while others use illustrations by text. If these kinds of learning objects are utilized together, they will lead to abrupt transitions among the corresponding topics. Consequently, the ability of the learner to capture the required knowledge can be directly affected by utilizing incompatible learning objects. In addition, learning objects may be incompatible from the technical point of view. Consequently, the major objective in this paper is to find a set of compatible learning objects that represents certain curriculum and does not exceed the intended budget.

This problem can be represented as a constrained optimization problem. Assume the curriculum consists of k learning objects. Each learning object i has a cost c_i where this cost represents the total cost associated with using the learning object in the e-learning class. Additionally, each learning object i has a rank r_i that should be determined by independent experts in the field and reflects its pedagogical contents. Learning objects must satisfy a minimum acceptable rank to participate in curriculums. Furthermore, each learning object has a set of interoperability indices between the learning object and any other learning object that can participate with it in a curriculum. These indices can be determined based on an independent evaluation of learning objects and the authoring strategy used in creating them. In addition, assume the learner allocates a budget B for his e-learning class. Hence, the problem now is to find a set of interoperable learning objects that fulfills the learner's study level, maximizes the overall curriculum rank, and fits the allocated budget. In other words, this problem can be mathematically formulated as follows:

$$maximize \quad R = \sum_{\forall i} r_i . \tag{1}$$

$$subject to:$$

$$\sum_{\forall i} c_i \le B. \tag{2}$$

$$\forall \text{ learning object } i, \;\; l_i = L. \tag{3}$$

$$I = \left(\frac{1}{k \cdot (k-1)} \right) \cdot \sum_{\forall i, j, i \neq j} interoperability\_index(i, j) \ge T. \tag{4}$$

where R is the overall curriculum rank, l_i is the study level of learning object i, L is the learner's study level, I is the average interoperability among the learning objects, and T is the average interoperability threshold.

III. Proposed Algorithm

The proposed solution for the problem defined by equations (1)-(4) is to implement a genetic algorithm that searches for the best combination (in terms of the overall curriculum rank) of learning objects that constitute the required curriculum. This solution must satisfy the constraints given by equations (2)-(4). Each potential solution for this problem is represented by a chromosome of k genes, where k is the number of learning objects for the required curriculum. Each gene represents the index of the selected learning object. While, the whole chromosome represents the set of learning objects corresponding to a potential curriculum. Fig. 1 shows the genetic representation of this problem. The pool of valid values for each gene excludes any learning object that does not satisfy equation (3). Mainly, because the genetic algorithm investigates only curriculums that have learning objects corresponding to learner's study level.

Index of Learning Object 1	Index of Learning Object 2	…	Index of Learning Object k

Fig. 1. Genetic representation of the problem.

A. Genetic Operators and Fitness Function

Both crossover and mutation genetic operators are adopted in the proposed genetic algorithm. Single-point crossover is utilized with the crossover-point selected randomly. On the other hand, mutation is accomplished by selecting any gene in a chromosome randomly and changing its value from the pool of valid values for this specific gene. In contrast, the fitness function utilized in the genetic solution is the overall rank R of the chromosome given by equation (1). Meanwhile, (μ + λ) selection criterion is adopted in the proposed algorithm.

B. Penalty Functions

Some of the chromosomes of the proposed genetic algorithm may have costs larger than the learner's budget. In addition, the average interoperability of a chromosome may not satisfy equation (4). Hence, two penalty functions are adopted. The first is the cost penalty function while the second is the interoperability penalty function. Consequently, the fitness of a given chromosome is penalized by dividing it by a constant C_c if the total cost of the curriculum exceeds B. On the other hand, if the average interoperability of a chromosome is smaller than T then it is penalized by dividing its fitness by a constant C_T. Adopting these penalty functions reduces the probability of selecting such chromosomes while forming the next generation during the selection process.

IV. Experimental Results

In this section, the results of a simulation study are presented. The simulation setup utilized in the experiments consists of a set of learning objects with each group of them can participate in a specific topic for the required curriculum. The number of learning objects in each topic is uniformly distributed in the interval [1, 10]. The interoperability index between learning objects i and j is uniformly distributed in the interval [0, 1.0]. T is set to 0.5 and learning objects are assumed to have the same level as the learner's study level. In contrast, each learning object is assigned a rank in the interval [0.4, 1.0] where 0.4 is the minimum rank allowed for any learning object to participate in an e-learning curriculum. In addition, the cost associated with each learning object is uniformly distributed in the interval [1, 5]. Mutation rate is set to 5% and crossover rate is set to 80%, while C_c and C_T are set to 10. The proposed algorithm is compared with the optimal solution, which is computed by enumerating all possible combinations of learning objects and selecting the best of them according to equations (1)-(4).

In the first experiment, the budget is changed from 10 to 25 units while k is set to 5. Table I shows the results of this experiment. As shown in Table I, the proposed algorithm is able to compute solutions identical to the corresponding optimal solutions. It also shows that at a budget of 25 units, it returns the same results as when the budget is 20 units. This is mainly because any improvement in the rank will lead to a budget more than 25 units or there may not exist any curriculum with better overall rank than this one. The search space is increased in the second experiment by setting k to 8 and the maximum number of learning objects in each topic to 20. Table II shows the results of this experiment. As shown in Table II, there is no solution can be found at a budget of 10 units. This is mainly because each learning object has a cost uniformly distributed in the interval [1, 5] and each curriculum has 8 learning objects, hence, on the average the curriculum cost is 24 units. Consequently, both algorithms are not able to find a suitable solution for a budget of 10 units. In contrast, both algorithms provide either identical results or very close results in the rest of the cases.

In the third experiment, k is changed while the budget is kept constant at 30 units and the maximum number of learning objects in each category is set to 10. Table III shows the results of this experiment. As shown in Table III, the proposed algorithm is able to find solutions either identical to optimal solutions or very close to them. In contrast, when k increases, the chance of not finding any solution increases as shown when k = 15 in this table. The reason behind this observation is that the total cost of any valid curriculum in this case exceeds the allocated budget. Consequently, no solution can be found in this case.

TABLE I: CURRICULUM CHARACTERISTICS WHEN K = 5 AND CHANGING B

Budget	Proposed Algorithm			Optimal Solution		
	R	Cost	*I*	*R*	Cost	*I*
10	3.948	9.857	0.710	3.948	9.857	0.710
15	4.052	13.136	0.654	4.052	13.136	0.654
20	4.114	15.804	0.636	4.114	15.804	0.636
25	4.114	15.804	0.636	4.114	15.804	0.636

TABLE II: CURRICULUM CHARACTERISTICS WHEN K = 8, MAXIMUM NUMBER OF LEARNING OBJECTS IN EACH TOPIC IS 20, AND CHANGING B

Budget	Proposed Algorithm			Optimal Solution		
	R	Cost	*I*	*R*	Cost	*I*
10	No curriculum that fits the budget can be found					
20	6.742	18.667	0.604	6.785	19.868	0.582
30	6.946	24.647	0.594	6.946	24.647	0.594
40	6.946	24.647	0.594	6.946	24.647	0.594

TABLEIII: CURRICULUM CHARACTERISTICS WHEN B = 30 UNITS, MAXIMUM NUMBER OF LEARNING OBJECTS IN EACH TOPIC IS 10, AND CHANGING K

k	Proposed Algorithm			Optimal Solution		
	R	Cost	*I*	*R*	Cost	*I*
3	2.491	8.835	0.847	2.491	8.835	0.847
6	4.593	19.668	0.575	4.593	19.668	0.575
9	7.010	28.524	0.506	7.010	28.524	0.506
12	9.003	29.829	0.513	9.481	29.936	0.547
15	No curriculum that fits the budget can be found					

V. CONCLUSION

In this paper, a new algorithm that allows the learner to find the best curriculum that fits his budget is introduced. This algorithm is proposed for cloud computing environments that have the concept of pay-per-use. Genetic algorithms are adopted in the proposed solution. The proposed algorithm is able to find the best set of learning objects that constitute the required curriculum. The returned curriculums have the best available experts' rank and fit learner's budget. Simulation results show that the proposed algorithm is able to find solutions very close to optimal solutions and in most cases identical to them.

REFERENCES

[1] L. M. Vaquero, L. Rodero-Merino, J. Caceres, and M. Lindner, "A break in the clouds: Towards a cloud definition," *ACM SIGCOMM Computer Communication Review*, vol. 39, no. 1, pp. 50-55, 2009.

[2] R. Buyya, C. S. Yeo, and S. Venugopal, "Market-oriented cloud computing: vision, hype, and reality for delivering IT services as computing utilities," in *Proc. of 10th IEEE International Conference*

on High Performance Computing and Communications (HPCC'08), 2008, pp. 5-13.

[3] K. Keahey, R. Figueiredo, J. Fortes, J. Freeman, and M. Tsugawa, “Science clouds: Early experiences in cloud computing for scientific applications,” in *Proc. of Cloud Computing and its Applications (CCA’08)*, 2008.

[4] C. Hoffa, G. Mehta, T. Freeman, E. Deelman, K. Keahey, B. Berriman, and J. Good, “On the use of cloud computing for scientific workflows,” in *Proc. of 4th IEEE International Conference on e-Science (ESCIENCE'08)*, pp. 640-645, 2008.

[5] Z. Guoli and L. Wanjun, “The applied research of cloud computing platform architecture in the e-learning area,” in *Proc. of 2nd International Conference on Computer and Automation Engineering (ICCAE)*, vol. 3, 2010, pp. 356-359.

[6] M. Al-Zoube, “E-learning on the cloud,” *International Arab Journal of e-Technology*, vol. 1, pp. 58-64, 2009.

[7] F. Doelitzscher, A. Sulistio, C. Reich, H. Kuijs, and D. Wolf, “Private cloud for collaboration and c-learning services: From IaaS to SaaS,” *Journal of Computing - Cloud Computing*, vol. 91, pp. 1-20, 2011.

[8] B. Dong, Q. Zheng, J. Yang, H. Li, and M. Qiao, “An e-learning ecosystem based on cloud computing infrastructure,” in *Proc. of 9th IEEE International Conference on Advanced Learning Technologies*, 2009, pp. 125-127.

[9] P. Pocatilu, F. Alecu, and M. Vetrici, “Using cloud computing for e-learning systems,” in *Proc. of 8th WSEAS International Conference on Data networks, Communications, Computers (DNCOCO'09)*, 2009, pp. 54-59.

[10] S. N. Sivanandam and S. N. Deepa, *Introduction to Genetic Algorithms*, Berlin Heidelberg, Springer-Verlag, 2010.

[11] T. D. Gwiazda, *Genetic Algorithms Reference*, Poland, Tomasz Gwiazda, 2006.

[12] R. L. Haupt and S. E. Haupt, *Practical Genetic Algorithms*, 2nd ed., New York, Wiley-Interscience, 2004.

[13] F. Rothlauf, *Representations for Genetic and Evolutionary Algorithms*, 2nd ed, Berlin Heidelberg, Springer-Verlag, 2010.

[14] M. J. Huang, H. S. Huang, and M. Y. Chen, “Constructing a personalized e-learning system based on genetic algorithm and case-based reasoning approach,” *Journal of Expert Systems with Applications*, vol. 33, pp. 551-564, 2007.

[15] M. Weller, “Learning objects and the e-learning cost dilemma,” *Journal of Open Learning*, vol. 19, pp. 293-302, 2004.

Gamal A. Ebrahim was born in Cairo, Egypt in 1971. He received his B.Sc. and M.Sc. degrees in 1994 and 2000 respectively from the Faculty of Engineering, Ain Sham University, Cairo, Egypt. Additionally, He received his Ph.D. degree in Electrical and Computer Engineering in 2004 from University of Miami, Coral Gables, FL, USA.He has been a faculty member at the Computer and Systems Engineering Department, Faculty of Engineering, Ain Shams University since Oct. 2005. He was a post-doctoral associate at University of Miami, Coral Gables, FL, USA from Mar. 2007 until Jan. 2009. His major research interests include intelligent systems, e-learning systems, software engineering, data mining, VoIP SPAM identification, computer communication networks, and active networks.