International Journal of Engineering Sciences & Emerging Technologies, October 2013.
ISSN: 22316604 Volume 6, Issue 2, pp: 133-141 ©IJESET

CELLULAR AUTOMATA FOR ADAPTIVE WEB PORTAL STRUCTURE IN EGYPTIAN UNIVERSITIES

Yasser F. Hassan
Department of Mathematics & Computer Science,
Faculty of Science, Alexandria University, Egypt

ABSTRACT

In the last few years, intelligent software applications are increased in complexity and in stakeholders' expectations, principally due to new internet-centric application areas with better performance. The modeled development process in this paper involves simulated processes of evolution, learning and self-organization. University portal provides a single point of access to university-based resources which are authorized to use. The research explores and adaptive portal website link-structure considering the website as interconnecting cells in cellular automata model browsed from the root domain. The potential of cellular automata model guides the factor adaptation and evaluated and emergent computation adequately extracts the main website profiles in terms of their internal structure.

KEYWORDS- *University Portal, Web Software, Cellular Automata, Adaptive, Egypt*

I. INTRODUCTION

A university portal is a web-based gateway that provides seamless access to a variety of backend resources [1] [6]. University portal provides end-users with a customized view of software and hardware resources that specific to their particular domain. The main challenge for building university portal is to make up-to-date building information more accessible and reflect the increasing services [2]. Universities want to see a control portal that works with customers (students, staff, or visitors) to help them achieve a building portal that meets both their expectations and building standards designed for the public good.

The word "Portal" adapted from a term referring to a seaport which implies a gateway for the delivery of goods from place to place. In web context, the term "portal" describes an entry point for direct access to websites, connection, content, commerce and community, which are the core functions of a portal. Although non-English speaking online populations are growing rapidly, support for non-English web content is much weaker than for English content [5]. A report published in 2010 shows that the majority (65.4%) of the world's online population consists of non-English speaker [7]. Moreover, that population was estimated to grow significantly in the near to 820 million while the size of English-speaking online population was predicated to remain at 300 million. For instance, there are more than 3.5 million internet users in the Arab world where the growth of Arabic web content is estimated to double every year.

The functions of university websites are redefined due to advancements in information technology. It became more dynamic and responsive on the demands of the academic community. Teachers and students can communicate through the university web site. A university web portal integrates information, contents and enterprise applications. University portal usability requires a deep understanding of the way in which customers use and interact with. It includes page layout, content design, graphic design and accessibility standards. A study of the service quality of university portal suggested that portals should attract a large number of visitors and consistently deliver quality service. The actual measure of success is dependent on the use of the portal. Alexandria University (AU) in Egypt is an institution of higher learning that utilizes the capabilities of web portal. AU has alumni, academic, student, and faculty portals. The AU portal supports e-learning and provides library online resources from university's publications and theses.

A Cellular Automaton (CA), as the term is used in this paper, is a discrete state system consisting of a countable network of identical cells that interact with their neighbors. This network can take on any number of dimensions, starting from a one-dimensional string of cells. The cellular automata model is perhaps the simplest, general model available. It is simple in that the basic units are small, local, finite state machines (cells). It is general since: cellular automata model is support universal computation, and the rule represents a general form of local interaction. The two-dimensional cellular automata model is a grid of squares, each square is having surrounded adjacent neighbors. A cell occupying a square is born, lives, or dies based on the number of living neighbors it has.

Cellular automata model is considered as a dynamical system in which space and time are discrete, each entity or cell is in one of k states at any given time, and all cells change their states at the same instant according to local interaction rules. The rule of local interaction change the cell's state based on its current state in addition to the state of the cells that adjacent to it [Hassan, 2011]. The main difficulty with the cellular automata approach seems to lie with the extreme low-level representation of the interactions, and the important issue is the coding of the problem onto the cellular automata structure, which is non-trivial and may be complex.

The paper is organized as follows: it proposes to begin in section 2 with the formalism necessary to describe the basic design of new model. Section 3 details some illustrative applications of new approach. Finally, the paper is concluded in section 4.

II. THE FORMAL DEFINITIONS OF PROPOSED MODEL

Two main steps are often involved in building a web portal (1) creating a domain-specific collection of web content and (2) supporting analysis of usage of these contents to better support users' tasks. The proposed model is based on a server-side approach – computational components reside on the system's server rather than user's own personal computer (see Figure 1). Adaptive web-portal systems are based on the user's characteristics and collecting data about users browsing behaviors including visited nodes to determine their characteristics (See Figure 2).

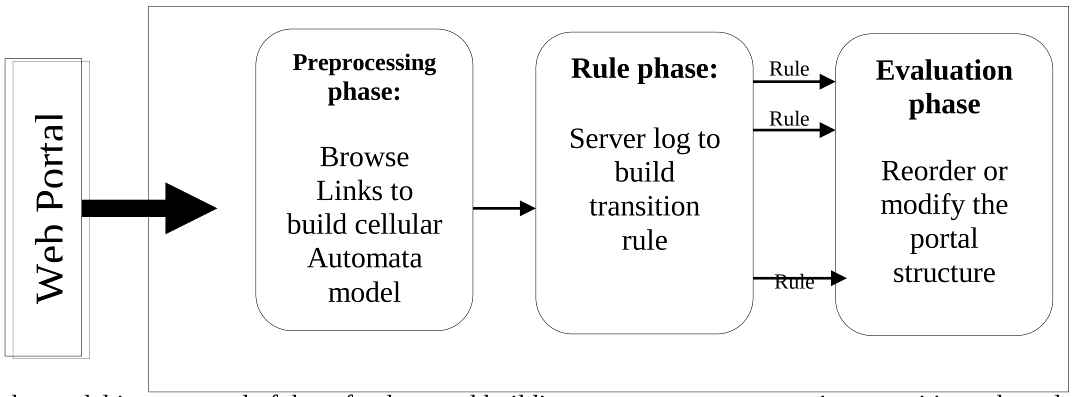


Figure 1: the model is composed of three fundamental building processes: pre-processing, transition rule and evaluation phase.

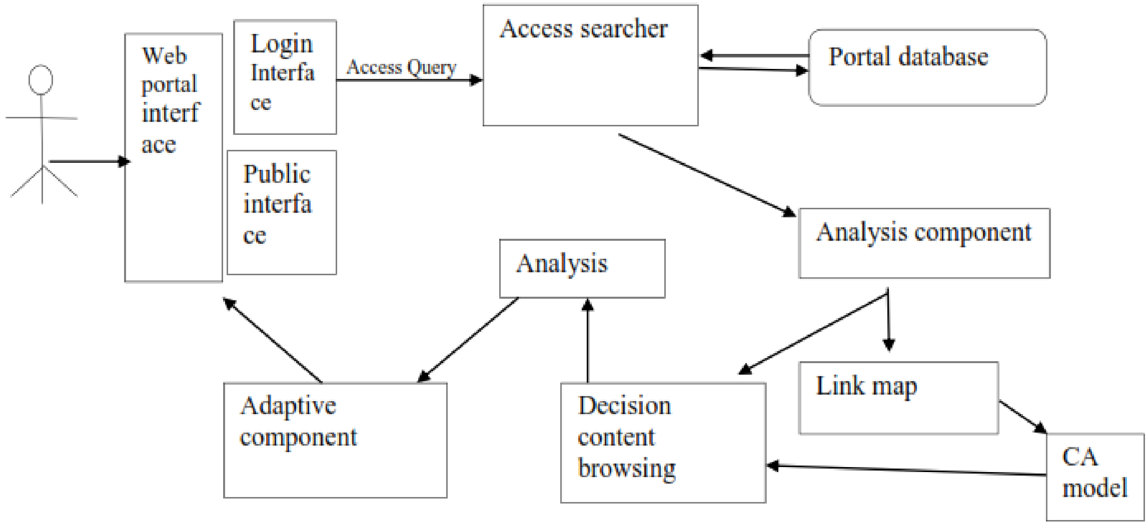


Figure 2: the proposed portal model.

2.1. Cellular Automata Web Link Map

The definition of software adaptation is: the process of operating software under specified conditioned, observing the results and making evaluating of software. The behavior of a system is represented by a set of interacting objects or components where each set of objects share the same attributes, operations and semantics. The model will be considered as two-dimensional cellular automata model which consists of array of cells x (i, j). The use of a cell as intelligent component link provides an even greater amount of flexibility to the ability and configuration of the system itself. Each software component contains the same or different test rule according to which web component links are updated in a synchronous and local manner. The neighborhoods of the web component x (i, j) where i and j are the row and column indices are the von-Neumann neighborhood of radius r, i.e.

$$N_r(i,j) = \{x(m,n) : |m-i| + |n-j| \le r\} \qquad (1)$$

where N_r is the neighborhood relation function of distance r. it assign r = 1, i.e. the neighbors of a web component x are the eight adjacent web component plus the web component x itself. The web component at cell $x(i,j)$ is called grid web component if and only if all neighborhood web component $x(m,n) \in N_r(i,j)$ exist. Otherwise $x(i,j)$ is called boundary cell. At any given instant, the web component is assumed to be in one state of the set of states: S. Then abstractly, each web component state function is

$$f : S \times A \rightarrow S \qquad (2)$$

where A = {a1, a2,…} is a set of actions. The intuition is that web component decides what action to perform based on its state and its environment states. I.e. the web component takes its current state and an action and maps them to a set of states S; those that are called result from performing action a ∈ A in state s∈S.

Once the automaton was embedded in the grid, the web component (cell) began to follow the rule that stored to it. Single web component cannot do much without interacting with other web component, and it has no concept of the whole. Yet, in combination it can play its part in producing complex results as emergence from local interactions. Briefly, each cell (software component) from the grid able to:

- Exercise a degree of freedom in its operations. It takes initiative and exercises a non-trivial degree of control over its own actions.
- Collaborate and exchange information with other web component in the grid (software

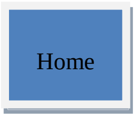
system) to assist other web component in improving their quality of decision making as well as its own.

- Change its state and the states of its immediate neighbors.
- Neither access directly to other web component in the grid except its immediate neighbors.

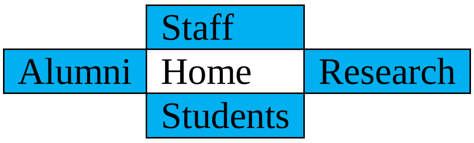
2.2. Multi-layers Cellular Automata

To build a links map, a model of multi-layers cellular automata will be used which is called Incremental Cellular Automata (ICA). Each link (web component) is represented as a cell in ICA grid in high level and the document in links will be represented as cells in low (second) level of the model. The ICA model begins with one (or more) cell as in Figure 3(b). cell and its neighbors will be added to the model at a time as in Figure 3 (c).
When ICA model is applied to web application (university portal) as Alexandria University web site "www.alexu.edu.eg", and it starts with menu item "Staff" see Figure 4.

(a)

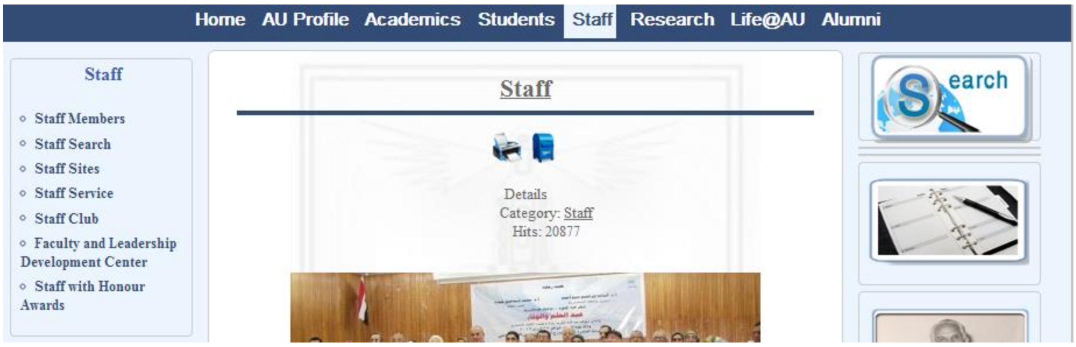


(b)



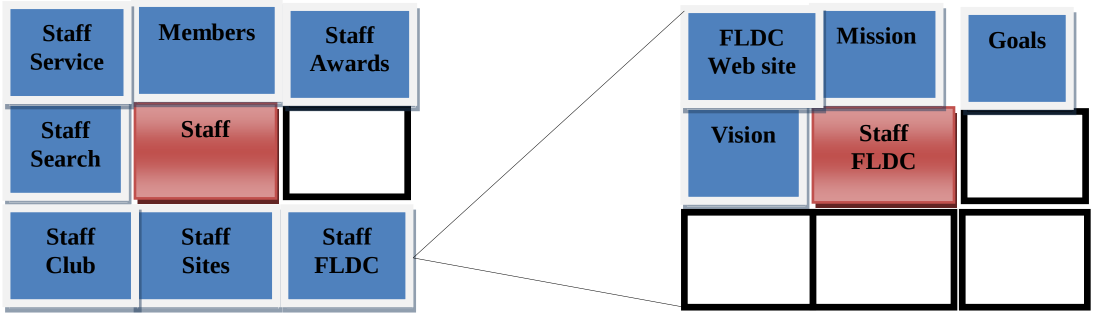
(c)

Figure 3: (a) The AU portal Home, (b) the ICA at time = 0 consists of one cell, (c) the ICA at time = 1.



(a)

(b)



(c)

Figure 4: a) AU portal one link, (b) start system with one cell, (c) system at time = 1, white cell is empty cell

Empty cell has state = 0, added cell has state = 4 (most important choice or first choice), and cell has state = 3 (70% important), state = 2 which means 40% important choice and cell state = 1 which means 10% important rate. Each links has attached cellular automata layer contains all insides links (options) as shown in Figure 5 & 6. The grid will start with set of states for cells (links) from the university itself to give percentage for importance of each link (see Figure 7).

AU profile	Academic	Alumni
Research		Social
Staff	Students	login

(a)

1	2	4
2		4
3	2	1

(b)

Figure 5: the start grid (t=0); a) the links, b) the state for each link

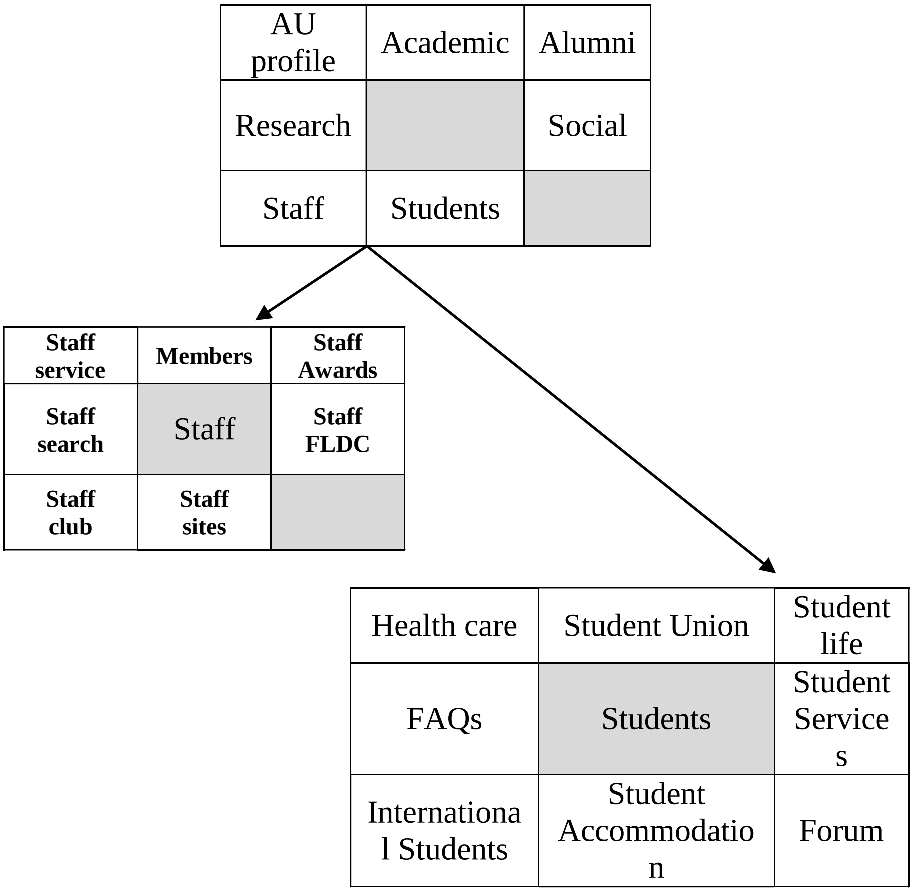


Figure 6: ICA model for links in AU portal in Figure 5.

Let the transition rule for determining the importance of links in the system is mentioned as: the state s of the link in the cell c is changed according to the rule:

$$s_{t+1}(c) = s_t(c) + \delta, \quad \text{-------- (3)}$$

where st+1 is the importance of link in cell c at time step (t+1), st is the importance at time step t, and δ is a parameter takes value $-k \le \delta \le k$. If st = 0 then the cell is empty. According to the previous rule, the state of each cell is increasing or decreasing according to the state of the portal, i.e. according to the value of δ, which can be obtained from the formula:

$$\delta = \begin{cases} \mu, & link\_is\_accessed \\ m, & otherwise \end{cases} \quad \text{----------- (4)}$$

where link is accessed is the flag if the cell is accessed this time, the quantity μ and m can be calculated as:

$$m = \begin{cases} s_t - 1 & if \;\; Access\_neighbors \\ s_t & if \;\; otherwise \end{cases} \quad \text{--- (5)}$$

$$\mu = \begin{cases} s_t + 1 & if \;\; state_t < 4 \\ s_t & if \;\; otherwise \end{cases} \quad \text{--------------- (6)}$$

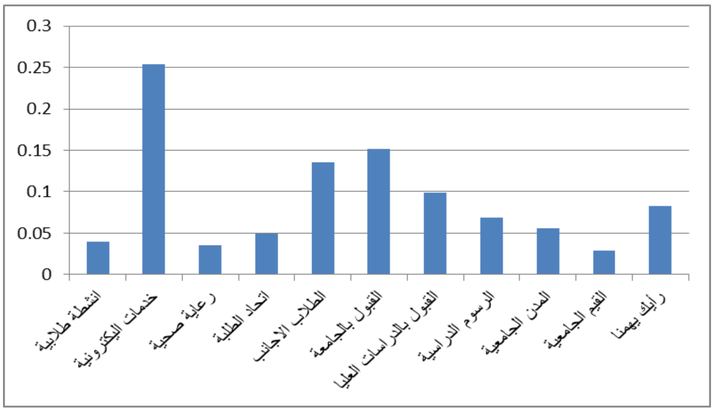


Figure 7: Distribution of selection ratio of links in one page (Student)

III. SIMULATION RESULTS

The case study to demonstrate the efficiency and potentiality of proposed model is based on the portals for Egyptian universities project funded by the Egyptian ministry of high education. The proposed method will be applied to 58 Egyptian University Portal web sites. All of them are included in the Webometrics Ranking of World Universities (www.webometrics.info), where more than 6000 universities all over the world are sorted according to size and visibility. Table 1 lists the root domains of the considered web sites. They cover almost the whole range of Webometrics ranking, and exhibit a variety of size in term of domains and web pages. Table 2 summarizes some descriptive statistics for Alexandria University Portal.

In this study, a log analyzer has been utilized to pre-process the web server access-log files. A typical web server access log file includes the information such as the IP address/host name of the site visitor/user, the URL of requested page, the date and time of the request, the size of the data requested and the HTTP method of request. Additionally, the log contains the user agent string describing the hardware and/or software the visitor was using to access the site, and the referrer field which specifies the web page by which the client reached the current page.

TABLE 1: list of considered web sites portals

Cairo University	Alexandria University	Zagazig University
American University in Cairo	Kafrelsheikh University	Benha University
Mansoura University	Ain Shams University	Assiut University
German University in Cairo	Helwan University	Al-Azhar University
Al Azhar Al-Sharif Islamic Research Academy	Arab Academy for Science & Technology and Maritime Transport	MISR University for Science & Technology
Minia University	Fayoum University	El Asher University
University of Tanta	South Valley University	Beni-Suef University
Université Française d'Egypte	Pharos University in Alexandria	Higher Technological Institute
Université Senghor d'Alexandrie	MISR International University	Egypt-Japan University of Science & Technology
Minufiya University	October 6 University	Port Said University
Modern University for Technology and Information	Delta University for Science & Technology	Modern Sciences & Arts University

Sohag University	Suez Canal University	Nile University
Military Technical College	British University in Egypt	Nahda University
Sinai University	Future University	Ahram Canadian University
Damanhour University	El Shorouk Academy	Akhbar El Yom Academy
Damietta University	Aswan University	
Canadian International College	Heliopolis University (Sekem University)	International Academy of Media Science
Integrated Thebes Academy for Science	Modern Academy for Engineering and Technology	Sadat Academy for Management Sciences
Egyptian Russian University	Science Valley Academy	Workers University
Alamein University	Higher Institute of Engineering	

TABLE 2: SELECTED SUBSET OF INDICATORS FOR ALEXANDRIA UNIVERSITY

Factor	Value
Domain	www.alexu.edu.eg
Subdomains	29
Number of pages in main domain	8545
Average number of pages in all subdomains	257

Figs. 8 and 9 show the particular case of the domain and page network, respectively, corresponding to the particular case of the University of Alexandria. We start "Academic" option with state = 2 (40% important) and after 4000 time it will be most important option.

Profile **1**	**Academic** **2**	**Alumni** **4**
Research **2**	News 4	Life 4
Staff **3**	Students 2	Login 1

(a)

News **4**	**Academic** **4**	**Staff** **4**
Life **4**	Profile 3	Alumni 2
Research **1**	Students 1	Login 1

(b)

Figure 8: the results of the ICA for Alexandria university portal structure. (a) Start state of ICA for AU home, (b) the ICA states after 4000 iterations

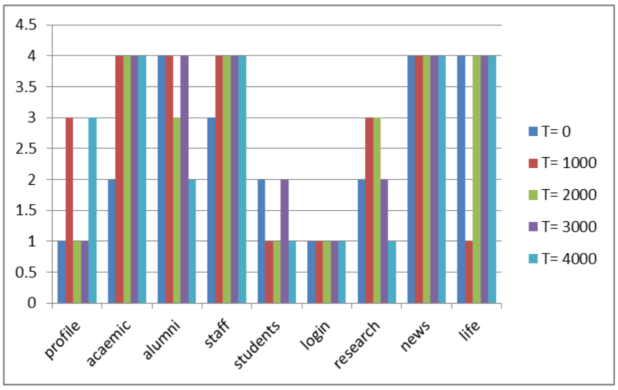


Figure 9: the results of the ICA for Alexandria university portal structure.

Although, Web structure has frequently been studied, comparatively little is known at the website level concerning its structure as an information organization and access mechanism. Here, an exploratory study for the identification of website link structure using cellular automata is proposed. For this purpose, the hypertext structure of institutional website has been extracted both at a domain and at a page level. A huge number of indicators related to different features of the derived networks can be computed. However, due to the exploratory nature of this study, it is difficult to select a subset of indicators to perform cellular automata structure analysis, and the alternative of considering all possible subset of indicators is computationally prohibitive. As a solution, a decision table of an optimum subset of indicators using a multi-objective fit function is proposed. The obtained result will provide new insights about web site patterns and highlights the utility of rough sets as a tool for new knowledge discovery.

IV. CONCLUSION

This paper has described a new approach to building cellular automata systems that can be applied to process test level of online university portal while still enabling good rule sets to be learnt in a practical manner. The proposed model has provided a direct approach to studying how dynamical systems perform emergent computation; that is, how the interaction of simple components with local information storage and communication gives rise to coordinated global information processing. Whether in real-life situation, the topology of the interconnections that gives meaning to the term immediate neighbor can change frequently. Although every link participating in the system must be able to communicate with its immediate neighbors, the system itself should not depend on knowledge of the overall system topology. The state transition within each cell could be identical throughout the system or unique to each web component. In practice, the state transitions within the cells can most conveniently be viewed as shared by all cells, but with local adaptations as a function of either static or dynamic local conditions. Because the model uses processes, which are fully incremental, it is able to show how a modular interaction can modulate and guide an underlying evolutionary process. Unfortunately, in its present manifestation the model has many shortcomings. As a computational implementation, it is not as robust as one would like. However, it is hoped that all these deficiencies will be remedied in the ongoing development of this work.

REFERENCES

[1] Luca, V., Epicoco, I., Lezzi, D., and Aloisio, G., A web API framework for developing grid portals, procedia computer science 4, pp. 392-401, 2011

[2] N. Thomas-Alvarez and L. Mahdjoubi, Testing the effectiveness of a web-based portal system for the building control sector, Automation in construction 29, pp. 196-204, 2013

[3] Yasser Hassan, Rough Set Adaptive in the Model Based of Cellular Automata and Multi-Agents, Journal of Emerging Trends in omputing and Information Sciences, Vol.2 No.8, 2011

[4] Yasser Hassan, rough sets for adapting wavelet neural networks as a new classifier system, Springer applied intelligence journal (35), pp 260-268, 2011.

[5] Chung, W., Bonillas, A., Lai, G., Xi, W., and Chen, H., Supporting non-English web searching: an experiment on the Spanish business and Arabic medical intelligence portals, Decision support systems, 42, pp. 1697-1714, 2006.

[6] Bringula, R., Influence of faculty- and web portal design- related factors on web portal usability: A hierarchical regression analysis, computer & education 68, pp. 187-198, 2013

[7] Internet world stats, http://www.internetworldstats.com, 2013