Conditional Random Fields for Web User Task Recognition based on Human Computer Interaction

Anis Elbahi
Research Unit MARS,
Department of computer sciences
FSM - Monastir 5019 Tunisia.
Elbahi.anis@gmail.com

Mohamed Nazih Omri
Research Unit MARS,
Department of computer sciences
FSM - Monastir 5019 Tunisia.
MohamedNazih.Omri@fsm.rnu.tn

ABSTRACT

In this paper we apply the Conditional Random Fields approach for modeling human navigational behavior based on mouse movements to recognize web user tasks. In fact, inferring activity of web users is an important topic of Human Computer Interaction. To improve the interaction process, many studies have been performed for understanding how users interact with web interfaces in order to perform a given activity. The Experimental evaluation and analysis of the results of the model we present in this paper demonstrate the efficiency of our model in human tasks recognition.

Keywords

Pattern recognition, Conditional Random Fields, E-learning Activity Recognition, Mouse movement tracking, Human Computer Interaction.

1. INTRODUCTION

The study of the activity of web users is an important topic of HCI. For years, various techniques have been used in this field, such as eye movements tracking [1], mouse tracking [7] and click-through analysis [10]. Understanding navigational behavior of users can improve interfaces usability, provide assistance for users with disabilities and others applications such as e-learning. On the one hand, the activity of mouse cursor can be easily captured and recorded. On the another hand, analysis of cursor behavior can provide high quality clues of a spontaneous, precise, direct and unbiased trace of user behavior. Such trace can be considered as a good indicator of the user reasoning strategy during a web activity. In this paper, we used the CRF approach [11] in order to recognize the tasks of web users, based on their navigational behavior using mouse movement.

2. ANALYSIS OF USER NAVIGATIONAL BEHAVIOR USING MOUSE MOVEMENT TRACKING

For each task (information searching, mail sending, downloading), users perform basic operations such as keyboard events, moving a cursor, clicking and pressing a button.

Using a cursor pointing device during web activities, users “draw” their navigational behavior. Mouse movement tracking has been evaluated as an alternative to eye tracking for determining attention on the web page. Therefore, various studies have been achieved in this context such as the study of Chen et al. [3] who have found that mouse and eye movements are strongly related and that 75% of mouse saccades move to significant regions of the screen where eye gaze are moved and they have been confirmed that mouse data can be used to infer the intent of user. So, mouse movements are explored to infer the user tasks during e-learning activity [5] and to provide insights into the intention behind a web search query [7]. Authors of [12] presented a user re-authentication approach using behavioral biometrics provided by mouse dynamics and in reference [8] Heimgartner identify users only by analyzing their interaction behavior mainly based on mouse events.

Elbahi et al. [16] presented a new possibilistic approach based only on mouse behavior for user task identification. Many other researches [4,18] have proposed different models based on possibility theory, on bayesian and semantic networks to recognize the goal of the users.

Obviously mouse movement tracking is a very effective technique, easy to use, freely available and does not disturb user behavior.

In this paper, we propose a new CRF model to automatically recognize web user tasks based on mouse trajectory recorded data.

3. THE USER TASK AS A SEQUENCE OF FIXED AREAS OF INTEREST

Each web interfaces can be described as a set of significant regions called Areas Of Interest (AOI) which can be manually specified or automatically discovered [9]. During a task, users move the cursor across the web interfaces and fix various AOI. Figure 1 presents an example of a sequence describing fixed AOI during "logging into Gmail account" task.

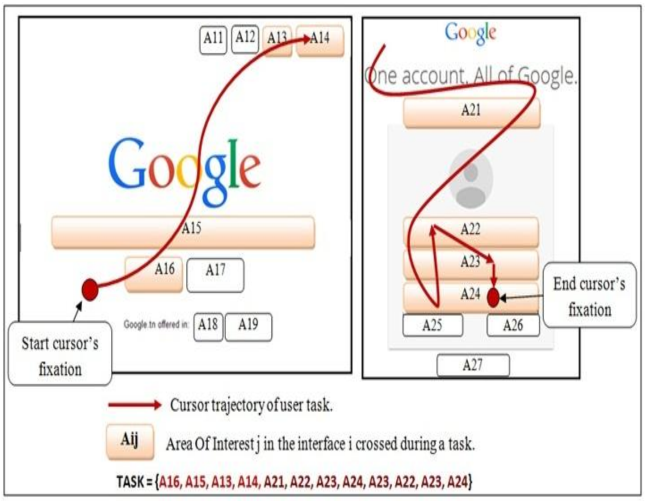

Figure 1. Example of user task defined as a sequence of fixed AOI.

Thus, each user task can be defined as a temporal sequence of fixed AOI during a period of time T. $TSK_i = \{AOI_1, AOI_2, \ldots, AOI_T\}$. Despite this clear definition of task, their automatic recognition is very challenging to solve.

The automatic task identification can improve the general interaction process by giving help in real time to unfamiliar users, helping users with disabilities and improving interfaces usability.

4. CRF: A BRIEF PRESENTATION

Hidden Markov Models [13] have been widely used for modeling and labeling stochastic sequences. In spite of their efficiency, CRF theory [11], have been proposed to alleviate HMM assumptions. Therefore, various studies have been successfully achieved for modeling and labeling sequences using CRF [11, 17].

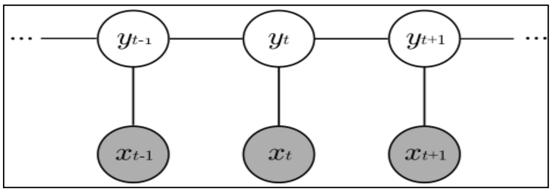

Figure 2. CRF linear-chain graphical representation

Figure 2 shows that CRF model involve hidden and observable variables at each time step and the edges between nodes are not oriented, making the CRF an undirected graphical model.

Due to the discriminative nature of CRF, it becomes possible to represent much more knowledge in the model using feature functions. With CRF we try to maximize the conditional probability distribution P(Y|X) represented as follows:

$$P(Y|X) = \frac{1}{Z(X)} \exp\left(\sum_{t=1}^{T} \sum_{k=1}^{N} \lambda_k f_k(y_{t-1}, y_t, X) + \sum_{t=1}^{T} \sum_{k=1}^{N} \mu_k g_k(y_t, X) \right) \quad \textbf{(1)}$$

where:

- Z(X) is a normalization factor used to ensure that outcome of P(Y|X) is a probability,
- $f_k(y_{t-1}, y_t, X)$ and $g_k(y_t, X)$ are features functions that return a real value.
- λ_k and μ_k are weights of each feature function,
- T is the length of the sequence X,
- N is the number of features functions,

CRF are designed to estimate the model parameters using an iterative gradient method such as BFGS algorithm and to perform the inference process using Viterbi algorithm. For CRF parameters estimation, we use a training set defined by: $D = \{x^{(i)}, y^{(i)}\}_{i=1}^{|D|}$, where each $x^{(i)}$ is a sequence of inputs and each $y^{(i)}$ is a sequence of desired predictions.

The estimation of weights of feature function (ϴ) is performed by maximizing the conditional log-likelihood of annotated sequences of D.

$$L(\Theta) = \sum_{i=1}^{|D|} \log P(y^i \,|x^i)$$

For more details about CRF, reader can see [11, 15].

5. CRF FOR USER TASK MODELING

5.1 The user task modeling

As shown in figure 3 the "Equation Grapher" simulator[1] interface is described as a set of areas of interest AOI={A,B,C,D,E,F,G,H} judged by an expert as frequently pointed regions during users tasks.

Like presented previously, each task can be defined as a finite, temporal, stochastic sequence of AOI set by a user during a period of time.

[1] Phet available on : http://phet.colorado.edu

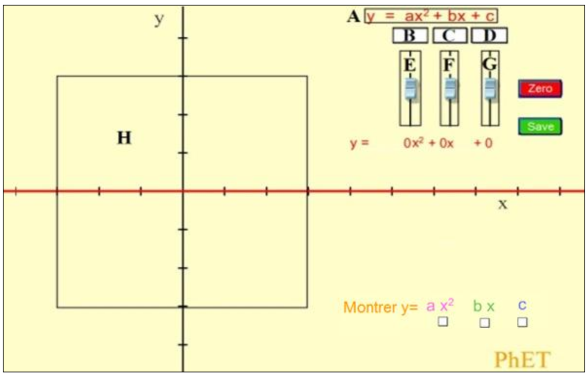

Figure 3. Areas Of Interest in "Equation Grapher" interface.

To define each user task (sequence of fixed AOI), coordinates of mouse cursor have been recorded at each time slice Δt.

5.2 The proposed model

The model structure is described by:

- TASKs= {tsk_1, tsk_2, …, tsk_M} : set of M labels concerning M tasks of users.
- AOI={aoi_1, aoi_2, …, aoi_N} : set of N areas of interest of the web interface that can be pointed by users during tasks.
- X={aoi_k_1,…, aoi_k_t, …, aoi_k_T} : the sequence of observations describing AOI fixed by mouse cursor during a task for a period of time T, with 1≤k≤N.
- Y={tsk_i_1,…, tsk_i_t, …, tsk_i_T} : the label sequence, with 1≤i≤M.

Each sequence of observations (sequence of fixed AOI) given to CRF model must be entirely labeled using a single tag corresponding to performed task. Graphically, our model can be presented as follows:

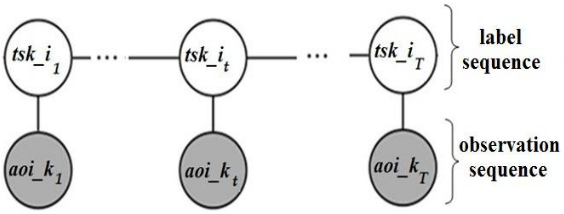

Figure 4. Graphical representation of proposed CRF model.

Let F={f_1,f_2,…,f_n} be a set of features functions. Each function $f_j(y_{t-1},y_t,X,_t)$ looks at a pair of adjacent labels (y_{t-1} and y_t) and all the observation sequence (X) at each time step ($_t$).

In order to validate the proposed model, we used CRF++ tool, by which we can define templates to automatically generate a set of features functions. Next, we present some examples of used features functions generated using CRF++ templates.

Template1 : U00 :%x[0,0] generate a set of functions like:

$$f_1(y_{t-1},y_t,X,_t) = \begin{cases} 1 \text{ if } y_t=tsk\_1 \text{ and } x_t=aoi\_2 ; \\ 0 \text{ otherwise} \end{cases}$$

The function f_1 return 1 if the current label (y_t) is tsk_1 and the current observation (x_t) is aoi_2 else f_1 return 0.

Template2 :U01 :%x[-1,0]/%x[0,0]/%x[1,0] generate a set of functions like:

$$f_2(y_{t-1},y_t,X,_t) = \begin{cases} 1 \text{ if } y_t=tsk\_2 \text{ and } x_t=aoi\_1 \\ \text{ and } x_{t+1}=aoi\_2 \text{ and } x_{t-1}=aoi\_4; \\ 0 \text{ otherwise} \end{cases}$$

The function f_2 return 1 if the current label (y_t) is tsk_2 and the current observation (x_t) is aoi_1 and next observation (x_{t+1}) is aoi_2 and previous observation (x_{t-1}) is aoi_4 else f_2 return 0.

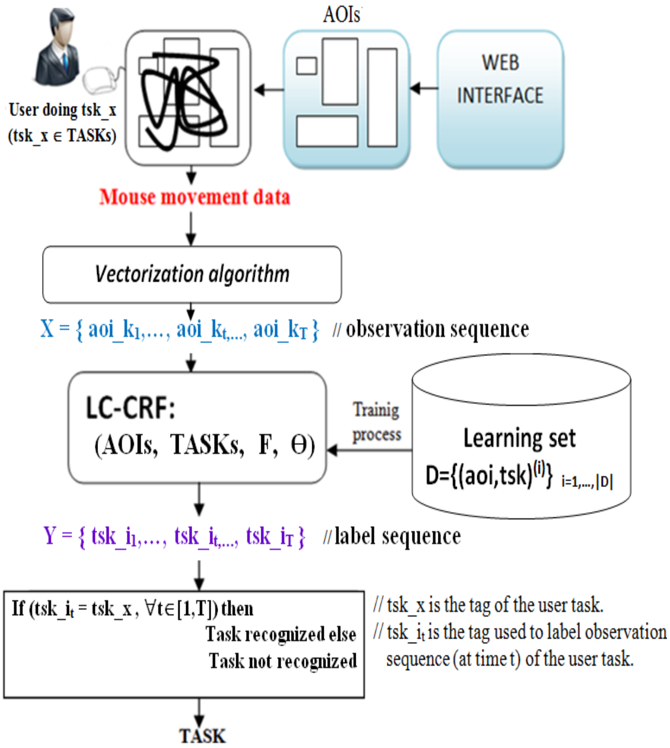

Figure 5. CRF model for task recognition using mouse movement data.

Once the model structure and features functions was defined, we train the model using labelled observation sequences. Each one corresponds to a single task and entirely labeled using a single tag. The figure 5 summarizes our task recognition approach.

6. EXPERIMENTS AND DISCUSSION

6.1 Experimental settings

We prepared a "training and test" set based on real manipulations, which consists of three tasks performed by students using the "equation grapher" interface. During each task, the student is asked to perform a graphical representation and to keep in memory the shape of the drawn curve. The three tasks are:

Task1(DEG2): representation of a quadratic equation of the form $ax^2+bx+c=0$ (with a, b and c $\neq 0$).

Task2(DEG1): representation of a quadratic equation of the form $ax^2+bx+c=0$ (with a=0 and b, c $\neq 0$).

Task3(INT): a student is asked to discover (and keep in memory) intersection coordinates of a quadratic equation of the form $ax^2+bx+c=0$ (with a,b,c $\neq 0$) and a quadratic equation $ax^2+bx+c=0$ (with a,c=0).

So, the three tasks are very similar and complex to distinguish based only on cursor trajectory.

For sequence of observations preparation, each user perform only one task and we use OGAMA tool [14] for recording mouse cursor coordinates. Based on the obtained data we produce an observation sequence corresponding to the performed task. Once, the 51 observation sequences are prepared and labeled, we estimate the parameters of the CRF model. The used sampling technique is LOOCV (Leave One Out Cross Validation).

6.2 Experimental results and discussion

Although, three tasks are quite similar, experimental results, presented in table 1, showed that CRF model make out 88,23% as recognition rate. These results show that the proposed model have a good ability in user task recognition.

Task Type	Samples	Error	Recognition rate
Task1	17	3	82,35%
Task2	17	2	88,23%
Task	17	1	94,11%
Total	51	6	**88,23%**

Table 1. Recognition rate of the proposed model.

In order to explain obtained results, remember that each task is defined as a finite, temporal, stochastic sequence of AOI fixed during a period of time T. Thus, each task is described by a sequence of observations $X=\{aoi\_k_1, \ldots, aoi\_k_t, \ldots, aoi\_k_T\}$. Therefore, in order to recognize a given task, it is necessary to take into consideration all focused AOI during a task.

Due to primary advantage of CRF approach which is the relaxation of the independence assumption, CRF model can take into account more complex dependencies between variables. So, all focused AOI during a task can be taken into consideration by CRF model.

For this reason, CRF presents high performance in user's tasks recognition.

TASK / AOI	DEG1	DEG2	INT	TASKX	TASKY
A	1,02	1,81	0,00	0,00	0,00
B	2,57	**15,18**	10,60	6,91	**18,86**
C	**14,74**	6,33	9,14	16,31	6,93
D	**14,91**	6,47	10,03	6,01	**21,24**
E	3,61	**28,84**	**13,64**	**9,85**	**29,51**
F	**18,57**	11,05	6,83	4,93	2,84
G	**21,84**	4,57	7,82	6,33	0,53
H	3,71	7,93	**24,51**	34,04	4,47

Table 2. Average Mouse Fixations (AMF) rate per AOI during tasks.

Table 2 shows the average cursor fixations in some AOI for three tasks. The same table also presents average cursor fixations in each AOI during two tasks TASKX and TASKY. TASKX is INT task and correctly recognized by CRF model while TASKY is INT task and judged by the model as DEG2 task. Likewise, Table 2 shows that area A is rarely fixed during three tasks, therefore the area A can be considered as unimportant item of interest [6] which do not attract the user cursor during interaction.

Task DEG2 is too dependent to areas E (28.84%) and B(15.18%) and task DEG1 is too dependent to areas G(21.84%) , F(18.57%), D(14.91%) and C(14.74%) while most used AOI for task INT are H(24.51%) and E(13.64%). These results show that each type of task attracts user attention into well defined regions in the interface. So, during each task mouse movements can be used to describe the strategy of user. A deeper analysis of mouse movements can give insights of the cognitive processes of the user during a task [2].

During TASKX which is INT task was correctly recognized by CRF model, in fact, mouse fixations rate of TASKX show that user focuses on areas E(9.85%) and H(34.04%) which are more relevant for INT task than DEG1 and DEG2 tasks.

TASKY which is of type INT, but recognized by the model as DEG2, the user usually focuses on relevant AOI for task DEG2 E(29.51%) and B(18.86%) and ignores area H(4.47%) considered as important for the task INT.

Knowing that all users have successfully performed the required tasks, we can see that during TASKY, the user adopts a different strategy to perform an INT task. This explains the failure of the model in recognizing TASKY because the CRF model adjusts its configuration based on the strategy of the group. To perform a given task, a human may adopt a strategy which is quite different of the one adopted by the majority of users; this task may be the cause of CRF failure. In fact a normal realization of a given task result in a normal use of important AOI which are relevant for this task and ignoring of important areas, or overusing of unimportant areas, should be considered as an indicator of different user strategy during task realization.

7. CONCLUSION

During interaction process, the analysis of mouse movement of the user can tell us about the user's task, AOI that have a user's attention high attraction and ignored AOI. Also, the analysis of cursor behavior can give insights about the strategy adopted by the majority of users and the particular user's strategy during a given task. In this work, we used CRF approach in order to recognize tasks performed by users. Experimental results show the good performance of the proposed model in user task recognition mainly based on mouse movements. Also, results show that each task type have a great impact on mouse behavior because the cursor is more attracted by some AOI than others according to each type of task.

8. REFERENCES

[1] A. Bulling, J. A. Ward, H. Gellersen and G. Troster, "Eye movement analysis for activity recognition using electrooculography," IEEE Transactions on Pattern Analysis and Machine Intelligence, 33(4), 2011, pp. 741-753.

[2] A. Nijholt and D. Tan "Brain-Computer Interfaces and human-computer interaction," In Brain-Computer Interfaces-Applying our Minds to Human–Computer Interaction, Springer - London, 2010 , pp. 3-19.

[3] M. C. Chen, J. R. Anderson and M. H. Sohn "What can a mouse cursor tell us more?: correlation of eye / mouse movements on web browsing," In CHI'01 extended abstracts on Human factors in computing systems, March 2001, pp. 281-282.

[4] M. N. Omri, C. Tijus, S. Poitrenaud, and B. Bouchon-Meunier, " Fuzzy sets and semantic nets for on line assistance," 11th Conference on Artificial Intelligence for Applications, 1995, pp. 374-379.

[5] A. Elbahi, M. A. Mahjoub and M. N. Omri, "Hidden markov model for inferring user task using mouse movement," In Information and Communication Technology and Accessibility (ICTA), IEEE, October 2013, pp. 1-7.

[6] A. Poole, and L. J. Ball, "Eye tracking in HCI and usability research," Encyclopedia of human computer interaction, 1, pp. 211-219. 2006.

[7] Q. Guo and E. Agichtein, "Exploring mouse movements for inferring query intent," In Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval, July 2008, pp. 707-708.

[8] R. Heimgärtner, "Identification of the User by Analyzing Human Computer Interaction," In Human-Computer Interaction. Ambient, Ubiquitous and Intelligent Interaction. Springer - Berlin Heidelberg, 2009, pp. 275-283.

[9] S. Mathe, and C. Sminchisescu, "Action from still image dataset and inverse optimal control to learn task specific visual scanpaths," In Advances in Neural Information Processing Systems, 2013, pp. 1923-1931.

[10] J. Huang, R. W. White and S. Dumais, "No clicks, no problem: using cursor movements to understand and improve search," In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems , May 2011, pp. 1225-1234.

[11] J. Lafferty, A. McCallum and F. C. Pereira "Conditional random fields: Probabilistic models for segmenting and labeling sequence data,", 2001.

[12] N. Zheng, A. Paloski and H. Wang, "An efficient user verification system via mouse movements," In Proceedings of the 18th ACM conference on Computer and communications security, 2011, pp. 139-150.

[13] L. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," In Proceedings of the IEEE, 77(2), 1989, pp. 257-286.

[14] A. VOßKÜHLER, V. Nordmeier, L. Kuchinke, and A. M. Jacobs, "OGAMA (Open Gaze and Mouse Analyzer): open-source software designed to analyze eye and mouse movements in slideshow study designs," In Behavior research methods, 40(4), 2008, pp. 1150-1162.

[15] C. Sutton and A. McCallum, "An introduction to conditional random fields for relational learning," Introduction to statistical relational learning, 2006, pp. 93-128.

[16] A. Elbahi, M. N. Omri and M. A. Mahjoub, "Possibilistic reasoning effects on hidden markov models effectiveness," In 2015 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), August, 2015. (accepted).

[17] F. Fkih and M. N. Omri. "Complex terminology extraction model from unstructured web text based linguistic and statistical knowledge," International Journal of Information Retrieval Research (IJIRR), 2(3), 2012, pp. 1-18.

[18] M. N. Omri, "Fuzzy knowledge representation, learning and optimization with bayesian analysis in fuzzy semantic networks," In 6th International Conference on Neural Information Processing, 1999. Proceedings. (ICONIP'99), 1, 1999, pp. 412-417.

 ISBN 978-80-86943-67-1