Volume 5, Issue 12, December 2015 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

Hidden Markov Model (HMM) Based Intention Prediction

Manisha Prakash Bharati
Compute Engineering Department
Maharashtra, India

Dr. Santosh Lomte
Computer Engineering & Dr. BAMU
Maharashtra, India

Abstract— In this paper, a survey of need to generate an intention prediction model of user interactions with systems is introduced which includes personal aspects, such as user characteristics, that can increase prediction accuracy. The model can be developed that is automatically trained according to the user's fixed attributes (e.g., demographic data such as age and gender) and the user's sequences of actions in the system. The generated model has a tree structure. Based on previous work done that also indicate the capability of the proposed method to discover the correct user intention model and increasing intention prediction accuracy compared with single HMM or CRF models.

Keywords— Hidden Markov model (HMM), user action Recognition, smart home systems, intention prediction, sequence learning.

I. INTRODUCTION

PREDICTING user intentions or actions when using a system can improve the services provided to users by adjusting the system services to their needs. The goal is to identify what the user wants to accomplish when performing a certain sequence of activities.

This goal is achieved by training a specific learning algorithm in order to generate models that will maximize the accuracy of the prediction. The challenge in predicting the intention of the user given a sequence of interactions can be categorized as a problem with a sequential problem space which is commonly solved using sequence learning algorithms.

The motivation of this paper is to achieve a higher level of compatibility between user intentions and the system he/she is using. This motivation derives from the fact that, in recent years, applications have continuously been gaining functionality and emerging features are complicating user interactions with applications.

Predicting user intentions can also be used to assist users to perform or complete tasks, to alert them regarding the availability of features, and to increase their general knowledge. Effective assistance will allow users to acquire the new skills and knowledge they need to easily operate an unfamiliar device or function.

Different users tend to select different sequences to accomplish the same goal. Specifically, our hypothesis is that the user's attributes and context (such as age or the operating system) indicate which sequence the user will eventually perform. This hypothesis is based on studies that examined the interaction of users with systems (such as web browsers) and found that user attributes affect the way of interaction. If a young male uses a system, it is possible to use the model to predict the goal that he intends to accomplish and provide him with a user interface (UI) more appropriate to his usage and intentions. Although methods exist to predict user intentions, they do not take into account user attributes that can increase the accuracy of prediction. In this paper, new method is developed, that examines how employing user attributes can improve the accuracy of intention prediction. The new method is termed an attribute-driven hidden Markov model tree (ADHMMT) and it builds a separate tree of hidden Markov models (HMMs) for each goal in the system. The tree is branched according to the user's attributes and each node in the tree consists of a single HMM. When the user performs a new task, we first employ his/her attributes to find the suitable node in each of the goal trees. Then, based on the actions the user has already performed in implementing this task, we anticipate the goal he/she is trying to accomplish and the action he/she intends to perform next.

The reason for using HMMs is that the sequence of actions that users perform is not always observable but can only be observed through another set of stochastic processes that produce the sequence of observations. For example, in a mobile device, the sequence of buttons that the user pressed is unknown (hidden) but the sequence of screens that was created from using these buttons is known (observable). Thus, we can learn the sequence of buttons from the sequence of screens using HMM.

The use of user attributes stems from personalization-related studies. It was shown that utilizing user demographic attributes for personalization and recommendation of relevant items to users increases the accuracy of recommendations since these attributes influence user's preferences, and interaction with systems. One study was able to predict user's attribute based on their browsing behavior. Thus, we utilize user attributes to predict user intentions assuming that since user attributes relate to user behavior, they naturally relate to the intentions preceding that behavior.

This paper is organized as follows: In Section II, we review related works on intention prediction and HMM. Section III formulates the problem. Section IV Applications of human action/intention recognition Section V Concludes the paper and presents suggestions for further research.

II. RELATED WORK

A. Intention Prediction

Intention prediction has become a very active research area recently. In this section, we summarize real-world applications that use intention prediction. Some applications treat the data as a sequence. One such example is the Lumiere project that leverages reasoning methods using Bayesian models to capture uncertainty about the goals of software users. The models can be employed to infer user needs by considering the user's background, actions, and queries. Horvitz *et al*. used Markov representation of the temporal Bayesian user-modeling problem by considering dependences among variables at adjacent time periods. The Lumiere project was first used in Microsoft's Office 97 product suite, containing the Office Assistant. Chen *et al*. [7] argue that although Office Assistance is probably the only one method, which has extensively studied user intention, its predictions can be further improved if semantic contexts are also used in addition to action sequences to mine user intentions.

There are few works about predicting user intentions as they surf the Web. Chen *et al*. [7] used a modified na¨ıve Bayes classifier algorithm to support incremental learning to model the user's intended action on a computer. They focused on predicting actions based on the features extracted from the user's interactions such as the user's typed sentences and viewed content.

Their goal was to predict the series of basic actions that the user will be performing in a system to accomplish his/her intention. Sun *et al*. presented a method to predict the user's browsing intention based on the web page sequences he/she had previously visited. Their method employs a multistep dynamic *n*-gram model and predicts the next action that lies on the optimal path that leads to the ultimate goal.

Task Predictor is a machine learning system that attempts to predict the user's current activity. It operates in the Microsoft Windows environment and collects a wide range of events describing the user's computer-visible behavior. As a first step, feature selection, a threshold to make classification decisions and naive Bayes is applied to decide whether to make a prediction.

Then, a discriminative, model linear support vector machine (SVM) is applied to make the prediction itself. In their work, they treated task prediction as a traditional supervised learning problem, ignoring the sequential aspects.

Few works attempt to identify the goal of user queries while searching the web. Rose and Levinson identified the goal of queries through manual, query-log investigation and concluded that the goals can be divided into categories in a hierarchical structure. Lee *et al.* Suggested automating the identification of user goals in a web query since in their study of human subjects, the majority of queries had predictable goals. Baeza-Yates *et al*. [1] try to infer the user's search intention by using supervised and unsupervised learning. Jansen *et al*. presented an algorithm that aims to automatically understand a user's intention by classifying queries as navigational, transactional, or informational.

Several research projects have focused on intelligent wheelchairs. Taha *et al*. Present a technique to predict the wheelchair user's destination to locations much further away than the immediate surroundings. The system relies on minimal user input obtained from a standard wheelchair joystick and a learned, partially observable Markov decision process. Jung *et al*.designed a user intention recognition module for an intelligent bed robot system. Since feature values are in sequence form, they used HMM. Feng *et al*. [8] developed a plan recognition-based method to predict the anomalous events and intentions of potential intruders using system call sequences as observation data. An algorithm based on a dynamic Bayesian network (DBN) with parameter compensation progressively accomplishes the prediction, while a recursive process identifies the intruder's hostile intentions. Yudhistira *et al*. presented an intelligent television set that can adaptively propose certain shows to the viewer based on previous shows that the user watched. An HMM models the user's preferences. All of the aforementioned studies focus on usage behavior and neglected user attributes in predicting intentions.

B. Hidden Markov Models

There are various methods of sequence learning including neural networks, DBNs, different Markov models, and others.

Each node in the hierarchical structure contains a model and an attribute. All nodes contain the same model but with different transition probabilities. (b) Structure of HHMM. The hierarchical structure is inside the states. Each hidden state can be HHMM as well. Problem is an HMM due to its ability to very efficiently calculate the probability of sequences in the given model and the existence of an efficient training algorithm (the Baum–Welch algorithm) [2]. As far as we know, no research until now has tried to increase prediction accuracy by utilizing user attributes. HMM is a type of DBN. It is a stochastic process with an underlying unobservable (hidden) stochastic process that can only be observed through another set of stochastic processes that produce the sequence of observed symbols. HMM can be viewed as a specific instance of a state-space model in which the latent variables are discrete. In HMM, the probability distribution of zn depends on the state of the previous latent variable $zn-1$ through a conditional distribution $p(zn \mid zn-1)$.

The following characterize an HMM: N, the number of states in the model. The individual states are denoted as $Z = \{z1, z2, ..., zN\}$ and the state at time t as qt. M, the number of distinct observation symbols per state. The observation symbols correspond to the physical output of the system being modeled. The symbols are represented as $X = \{x1, x2, ..., xM\}$.

The state transition probability distribution $A = \{ai,j\}$ where $ai,j = P[qt+1 = zj \mid qt = zi]$ $1 \leq i, j \leq N$.

The observation symbol probability distribution in state j, $B = \{bi(k)\}$ where $bi(k) = P[xk \text{ at } t \mid qt = zj]$,

$1 \leq j \leq N$, $1 \leq k \leq M$. The initial state distribution $\pi = \{\pi i\}$ where $\pi = P[q1 = zi]$, $1 \leq i \leq N$.

III. PROBLEM FORMULATION

In this section, several basic definitions are introduced followed by the problem formulation. In a typical sequence learning problem, a training set of sequences $S = \{s1, s2, ..., sm\}$ is given. Each sequence $sj \in S$ is an ordered set of nj elements (actions) $\{e1, e2, ..., enj\}$. G denotes the set of all possible goals (for example, in an email application G = {"Add Training recipient," "Send an email," "Read email"}). Each training sequence is associated with one goal and several characterizing attributes. The notation V denotes the set of input attributes
containing η attributes: $V = \{v1, ..., vn\}$. The domain (possible values for each attribute) of an attribute is denoted by dom(vi). User space (the set of all possible users) is defined as a Cartesian product of all the input attribute domains: U = dom($v1$) × dom(vi)× · · · ×dom(vn). The input in the problem consists of a set of m records and is denoted as Train = (_$s1$, $g1$, $u1$ _, . . . , _sm, gm, um_) where $sq \in S$, $gq \in G$, and $uq \in U$.
The notation L represents a probabilistic sequence learning algorithm such as HMM. Tree structure is used to improve the accuracy of training models. For each goal in the system, a different tree is generated. This structure is built using user attributes in an attempt to differentiate between various usage behaviors in the system. For example, employing the age, gender, and device can be used for differentiating between usage behaviors. In this structure, each node contains a model and an attribute vi that splits the node. Assuming this tree structure, the problem can be formally phrased as follows:
Given a sequence learning algorithm L and a training set Train with input sessions set S, users U, and goals G, the aim is to find an optimal set of trees (a tree for each goal). Optimality is defined in terms of minimizing prediction errors.
Where L is an algorithm to train HMMs. Particularly, the model λ in each of the tree's nodes is trained using the Baum–Welch algorithm [2] and the probability $p\lambda(G = gj \mid sq, uq)$ is estimated using the forward–backward algorithm.
paragraphs must be indented

IV. APPLICATIONS OF HUMAN ACTION/INTENTION RECOGNITION

As the past decade have witnessed a rapid development of computer vision, as one of the most promising applications of this area, video-based human action recognition has caught the attention of researchers from both industry and academia. Video based human action has a wide range of application mainly in three domains:

1. Security and surveillance: Intelligent security and surveillance are often found in occasion requiring high level of security, like banks, supermarkets and garages, human action recognition make it possible for surveillance system to implement real-time detection and recognition of suspicious action.
2. Interactive application: we always hope that computers in future are able to communicate with people in an easier way, for example, by understanding human action including gestures even expressions, users will enjoy a quite different and amazing using experience.
3. Content-based video retrieval: The explosion of multimedia data especially video, users make it so difficult to find specific video from mass data. Instead of raw video, users want to query the content-based video retrieval can find those videos with this specification action.

In order to improve the intelligent of home sensor network system, we concentrate on the research of video-base human action recognition for smart home, which make it possible for the system to analyze the surveillance videos automatically and alarm owners for abnormalities like burglary or falling down of aged people.
Since yamto [12] introduced HMM into human action recognition, HMM and improved models are widely used. The assumption of HMM is that current state of motion only depends on the previous state. HMM is comprised by two sequences of stochastic variables, one of which is a sequence of states cannot be observed, another is the sequence of symbols generated by states. State refers to the current attribute which can only be implied by observation, the representation of state. Human action recognition based on HMMs normally extract the sequence of feature vectors to represent action in the first step, followed by training parameters of HMMs through algorithms, finally classify the unknown motion sequence according the trained model.
Generally, human action is not corresponded to the assumption of HMM, the structure of traditional HMM also limits its application to single dynamic process, that's why traditional HMM provides poor performance in recognition of complicated actions. To address this problem, coupled Hidden Markov Model was proposed by Brand et al. as solution to interactive action recognition. Hierarchical HMM was proposed by Luhr and Nguyen to analyze action for a long time.

V. CONCLUSION AND FUTURE WORK

In conclusion HMM is very effective for modeling of temporal data and more robust than template matching. But the application of HMM is limited by its assumption, complicated model and numerous parameters.
Dynamic Bayesian networks construct more flexible and effective model for complex dynamic procedure, leading to successful application in human segmentation and recognition. Other probabilistic networks, for instance conditional random field has also achieved implementation for action models.Advances in human action or intentions recognition enables machine to understand and interact with human, thanks to the efforts of all researchers fascinated by this field. The development of home sensor network give rise to a growing demand for more intelligent and humanized system equipped with excellent action recognition technology. However, present action recognition still faces the challenges of both efficiency and robustness for practical application.
Many approaches assume that video has already been segmented into sequences that contain single action, thus action detection is ignored, which is unacceptable in real time recognition for home sensor network. Although some work

addressed this topic and it remains a challenge for action recognition in smart home. Normally action or intention recognition algorithms are tested on public datasets, for instance, UCF sport dataset, KTH human motion dataset, Hollywood rate of over 80 %, performance in realistic situation appears to be poor for complex and ever changing real environment.

Besides most existing work focuses an accuracy rate of recognition more than processing efficiency, while in home sensor network, response time should be regarded as a matter of primary. For instance, an old man passed out at home, smart video surveillance hos to respond as soon as possible. Thus, recognition efficiency remains a significant problem in home sensor network.

The problem of multi view which is also an important and inevitable problem to be considered in home sensor networks. The previous work done on view invariant problem in human action recognition set the stage for more researchers to encourage themselves in this area. Challenges lies in future in development of action recognition for home sensor networks in terms of real time performance, robustness to real world conditions, and complexity of sensor network architecture.

REFERENCES

[1] R. Baeza-Yates, L. Calderon-Benavides, and C. Gonzalez-Caro, “The intention behind web queries,” in *Proc. String Process. Inf. Retrieval*, Glasgow, Scotland, 2006, pp. 98–109.

[2] L. E. Baum and J. A. Egon, “An inequality with applications to statistical estimation for probabilistic functions of a Markov process and to a model for ecology,” *Bull. Amer. Meteorol. Soc.*, vol. 73, pp. 360–363, 1967.

[3] L. E. Baum and G. R. Sell, “Growth functions for transformations on manifolds,” *Pacific J. Math.*, vol. 27, no. 2, pp. 211–227, 1968.

[4] C. M. Bishop, *Pattern Recognition and Machine Learning (Information Science and Statistics)*. New York: Springer-Verlag, 2006.

[5] A. Bhargava and G. Kondrak, “Multiple word alignment with profile hidden Markov models,” in *Proc. Human Language Technol. Conf.*, Morristown, NJ, 2009, pp. 43–48.

[6] C. Chen, J. Liang, H. Zhao, H. Hu, and J. Tian, “Factorial HMM and parallel HMM for gait recognition,” *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 39, no. 1, pp. 114–123, Jan. 2009.

[7] Z. Chen, F. Lin, H. Liu, Y. Liu, W. Ma, and L. Wenyin, “User intention modeling in web applications using data mining,” *J. World Wide Web*, vol. 5, no. 3, pp. 181–191, 2002.

[8] L. Feng, X. Guan, S. Guo, Y. Gao, and P. Liu, “Predicting the intrusion intentions by observing system call sequences,” *Comput. Security*, vol. 23, no. 3, pp. 241–252, 2004.

[9] S. Fine,Y. Singer, andN. Tishby, “The hierarchical hiddenMarkovmodel: Analysis and applications,” *Mach. Learning*, vol. 32, pp. 41–62, 1998.

[10] N. Friedman, K. Murphy, and S. Russell, “Learning the structure of dynamic probabilistic networks,” in *Proc. 14th Conf. Uncertainty Artif. Intell.*, 1998, pp. 139–147.

[11] Proceedings of the 2012 International Conference on Information Technology ... By Wei Lu, Guoqiang Cai, Weibin Liu, Weiwei Xin

[12] Yamto J, Ohya J Ishii K(1992) Recognizing human action in time sequential images using hidden markov model. CVPR, pp 379-385.

