An improved position prediction algorithm based on Active LeZi in Smart Home

Hongqing Fang
College of Energy & Electrical Engineering
Hohai University
Nanjing 210098, China
fanghongqing@sohu.com

Jinjin Ruan
College of Energy & Electrical Engineering
Hohai University
Nanjing 210098, China
ruanjinjin@163.com

***Abstract*—In smart home environments, each inhabitant has their own habits, these habits have certain periodic and time-varying characteristics. According to these characteristics, an improved Time-varying LeZi algorithm (TALZ) is proposed, which is based on Active LeZi algorithm (ALZ). Compared to other models, the results show that TALZ has smaller time complexity, higher prediction accuracy of laboratory data acquisition, and it is suitable for real-time, accurate position prediction in smart home environments.**

Keywords- smart home; postion prediction; LeZi algorithm

I. INTRODUCTION

As the world's population ages, the number of elderly people living alone will enormously increase. With aging, health related problems also increase. Hence, the number of elderly people in need of assistance to have a normal life will increase rapidly. The number of care-givers will be outnumbered by the increasing elderly population. With the advent of smart home technologies, people with cognitive impairments can lead independent lives in their homes for longer duration of time. Smart homes can assist their residents by acting as a cognitive prosthesis, by handling various appliances/objects and also by facilitating emergency communication. Furthermore, cognitive health assessments performed in clinical settings do not always provide an adequate representation of a patient's behavior. Real life assessments of Activities of Daily Living (ADLs) can provide a better understanding of the subject than assessments performed in a clinical setting.

The recent emergency of the popularity of smart home environment [1-2] is the consequence of a convergence of technologies in machine learning and pervasive computing as well as the increased accessibility of robust sensors and actuators. Furthermore, smart home environment is developed to assist with valuable functions such as remote health monitoring and intervention. The need for the technology is underlined by aging population, the cost of health care and the importance that individuals place on remaining independent in their own homes. Individuals need to be able to complete the Activities of Daily Living to function independently at home.

The CASAS smart home project is a multi-disciplinary research project at Washington State University, focused on the creation of an intelligent home environment. The approach is to view the smart home as an intelligent agent that perceives its environment through the use of sensors, and can act upon the environment through the use of actuators. The research goals of the CASAS smart home project are to enhance and improve quality of life, prolong stay at home with technology-enabled assistance, minimize the cost of maintaining the home and maximize the comfort of its inhabitants. In order to meet these goals, the smart home must be able to reason about and adapt to provide information.

One of the research fields in some home is position prediction of inhabitants, and many various position prediction algorithm have been put forward, such as LZ78, LeZi-Update, Active LeZi (ALZ) [3-5], as well as another predict algorithm base on Markov model. In this paper, an improved Time-varying LeZi algorithm (TALZ) is proposed, which is based on Active LeZi algorithm. Compared to some predict algorithms mentioned above, TALZ not only inherits the perfect advantage of Active LeZi algorithm, but also has a corresponding improvement in the position prediction accuracy.

II. DATA COLLECTION

The smart home environment testbed for this research is located on the Washington State University campus and is maintained as part of the ongoing CASAS smart project [4]. As shown in Figure 1, there are three bedrooms, one bathroom, a kitchen, and a living/dining room. The smart apartment is equipped with motion and temperature sensors as well as analog sensors. The motion sensors are distributed approximately throughout the space on the ceiling. In addition, temperature sensors and analog sensors provide readings for hot water, cold water, and stove burner use. Voice over IP captures phone usage using Asterisk software and contact switch sensors are used to monitor usage of a cooking pot, the phone book, and the medicine container. Sensor data for activity recognition is captured using a customized sensor network and then stored in a SQL database. To maintain privacy, participant names and identifying information are

Identify applicable sponsor/s here. *(sponsors)*

978-0-7695-4719-0/12 $26.00 © 2012 IEEE
DOI 10.1109/CSSS.2012.433

removed and encrypt collected data before it is transmitted over the network.

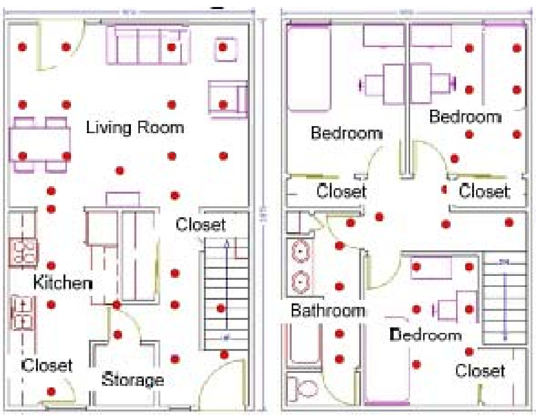

Fig. 1 The smart apartment testbed.

III. Position Prediction Algorithm

A. LZ78 model

LZ78 is the name for the lossless data compression algorithm published in papers by Abraham Lempel and Jacob Ziv in 1978 [3]. LZ78 algorithms have two parts: encode and decode. LZ78 model only use the encode part, which achieves compression by replacing repeated occurrences of data with references to a dictionary that is built based on the input data stream. Each dictionary entry is of the form dictionary[...] = {index, character}, where index is the index to a previous dictionary entry, and character is appended to the string represented by dictionary[index].

Good compression algorithm is also a good prediction algorithm. LZ78 model use the vocabulary entry to establish the decision tree.

The pseudo code representation of LZ78 algorithm is as below:

```
{
loop
wait for next symbol v
if ((w.v) in dictionary):
    w = w.v
else
    add (w.v) to dictionary
    w = null
    increment frequency for every possible prefix of phrase
forever
}
```

Consider of the standard test sequence of input symbols x = “aaababbbbbaabccddcbaaaa”, a LZ78 parsing of this string of input symbols would yield the following set of phrases: “a,aa,b,ab,bb,bba,abc,c,d,dc,ba,aaa”. Computing the frequency for every possible prefix of phrase, Fig.2 shows the decision tree of LZ78.

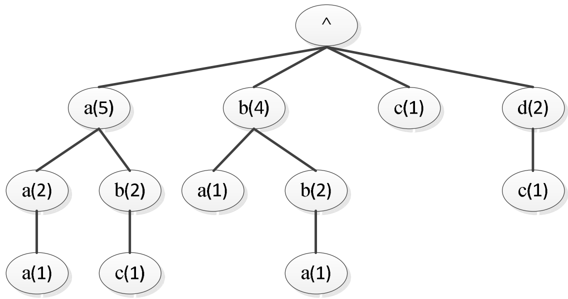

Fig. 2 Decision tree of LZ78

B. ALZ model

Although the using of the LZ78 model can also has the function of prediction, but the shortcomings is also quite obvious. There is a main disadvantage for LZ78 model, i.e., when an input string is parsed by LZ78, all the information across phrase boundaries is lost. The lost information is significant for the prediction performance, which results in lower prediction accuracy.

In order to solve this issue, researchers improved the LZ78 model; the new algorithm is called Active LeZi (ALZ) [5]. Compared with LZ78 model, ALZ adds a sliding window to get more information. Since the lost information in LZ78 can be preserved, the prediction accuracy has also been a correspondingly increased. The pseudo code representation of ALZ is as below:

```
{
initialize Max_LZ_length = 0
phrases = 0
loop
wait for next symbol v
if ((w.v) in dictionary):
        w := w.v
else
        add (w.v) to dictionary
if(length(w.v) > Max_LZ_length)
         Max_LZ_length = length(w.v)
        w := null
add v to window
if (length(window) > Max_LZ_length)
        delete window[0]
update frequencies of all possible contexts within window that includes v
forever
}
```

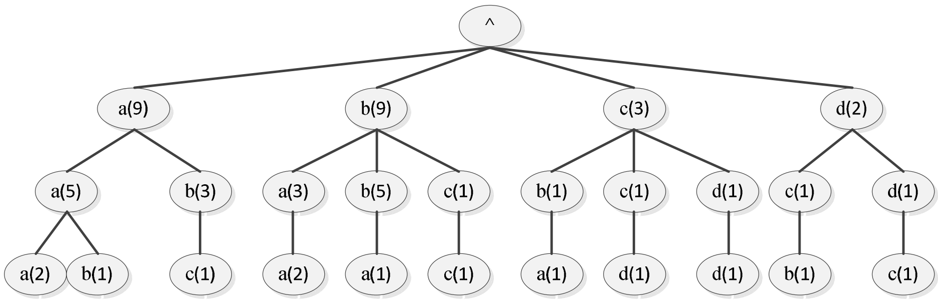

Fig. 3 Decision tree of ALZ

Using the same example, ALZ establishes a new decision tree. Compared with the tree of LZ78, the new tree has more nodes, which is shown in Fig. 3.

C. Time-varying LeZi model (TALZ)

As known, almost all of the predict algorithms are based on past information to predict, include the position prediction. Human daily habits have a certain periodicity, such as, almost every day some inhabitants get up at 7 o'clock in the morning, go to work at 8 o'clock, as well as have dinner at 18 o'clock in the afternoon. This periodicity gives researches the abilities to predict the behavior of inhabitants at some point, and then to predict the position of an inhabitant in smart home.

Since different behaviors happen in different time and may have a similar activity path, which will cause interference. This interference also results in lower prediction accuracy.

In order to eliminate this interference, an improvement of the ALZ model is proposed. The sample information with the limitation of time is split. ALZ establishes one tree with all sample information, but TALZ establishes 24 trees since one day has 24 hours, which considers the time factor. The sample information of each hour is used to build a decision tree. ALZ establishes a large tree with all sample information, however, TALZ divides the large tree into 24 small trees, which is shown in Fig. 4.

TALZ reduces the interference that comes from different activities, and then it reduce the size of decision tree, furthermore, it reduces the computing time of prediction.

The pseudo code representation of TALZ is as follows:

```
{
initialize Max_LZ_length = 0
phrases = 0
wait for time symbol t
if( t in dictionary)
go to t_tree
else
creat a new_tree
 go to new_tree
 loop
 wait for next symbol v
 if ((w.v) in dictionary):
     w := w.v
 else
     add (w.v) to dictionary
 if(length(w.v) > Max_LZ_length)
   Max_LZ_length = length(w.v)
 w := null
 add v to window
 if (length(window) > Max_LZ_length)
     delete window[0]
 update frequencies of all possible contexts within window that includes v
 forever
 }
```

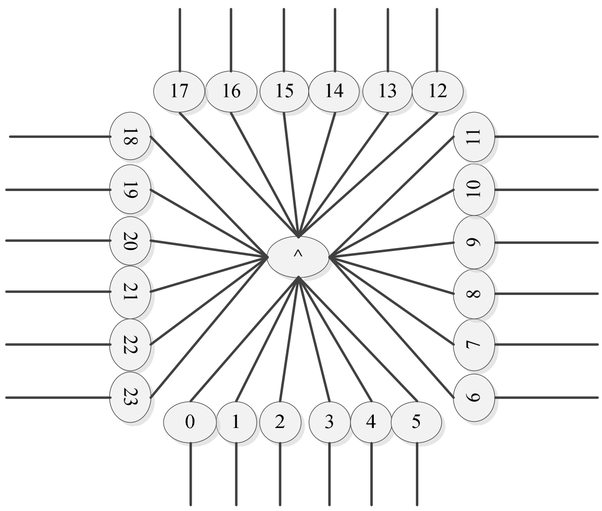

Fig. 4 Indicate tree of TALZ

IV. TEST RESULTS

The samples have been collected in the CASAS smart apartment testbed for 57 days, resulting in total of 647,485 collected motion sensor events which are used to predict. Fig. 5 shows the comparison results of three prediction models. It can be seen that the prediction performance of TALZ and ALZ are quite good than LZ78. The accuracy of LZ78 is the lowest, and the accuracy of ALZ has improved significantly, further, TALZ achieves the highest accuracy. The reason can be explained that TALZ inherits ALZ's all advantage and reduces the interferences that come from different activities.

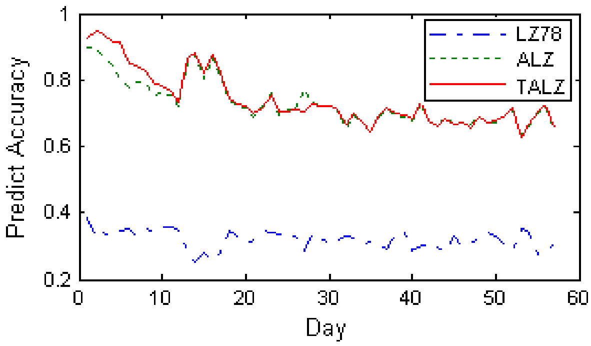

Fig. 5 Comparison results of three prediction models

V. CONCLUSION

In the present research, the models of position prediction of inhabitants in smart home have been discussed. The results show that Time-varying LeZi model has the best prediction accuracy.

ACKNOWLEDGMENT

This work was partially supported by Qing Lan Project, Jiangsu Province, China and the data were collected from the smart home testbed located on the Washington State University campus.

REFERENCES

[1] L. Liao, D. Fox and H. Kautz. Location-based activity recognition using relational Markov networks. In Proceedings of the International Joint Conference on Artificial Intelligence. 2005. 773–778.

[2] D. Cook and P. Rashidi. The resident in the loop: Adapting the smart home to the user. IEEE Transactions on Systems, Man, and Cybernetics, 2009, 39(5): 949-959.

[3] J. Ziv and A. Lempel, "Compression of Individual Sequences via Variable Rate Coding," IEEE Trans. Information Theory, 1978, 24(5): 530–536.

[4] A. Bhattacharya and S. K. Das. LeZi-Update: An Information-Theoretic Framework for Personal Mobility Tracking in PCS Networks. Wireless Networks J., 2002, 8(2–3): 121–135.

[5] K. Gopalratnam and D. Cook. Online Sequential Prediction via Incremental Parsing: the Active LeZi Algorithm. IEEE Intelligent Systems, 2007, 52-58.