The Recommender System Component Report for Project Actibetes

Hao Ding
Department of Electrical and Electronic Engineering
Imperial College London
Email: hd1812@ic.ac.uk
Supervisor: Dr. Yiannis Demiris

***Abstract*—Chronic disease is the major cause of mortality and it has been leading to larger burden globally. Exercise appears to have various impacts on glucose level control. The 'Actibetes' aims to assist the self-management of patients with diabetes mellitus through proactively providing suggestions on their daily exercise. The recommender system component is back-end, which collaborates systematically with the front-end mobile application and the database. The recommender system uses the multiple models in behaviour modelling and uses the competitive parallel structure to suggest the suitable amount of exercise to patients. My work mainly involves architecture design, model implementation and interface design. This component report includes the literature research, current progress and future plan.**

I. Introduction

For patients with diabetes mellitus, lifestyle choice is considered as an important factor in disease management. Exercise is generally believed to benefit glycemic control by increasing energy expenditure and improving tissue sensitivity to insulin[1], [2]. The recommender system aims to tackle the challenges in diabetes mellitus management by delivering the exercise suggestions to patients. As the back-end of the 'Actibetes', the recommender engine operates together with the user interface and the database. It takes the user information from the front-end, trains and runs models, and returns the suggested exercise level to users. A parallel adaptive multi-model architecture will be used for action selection and recommendation[3]. The recommender engine is likely to show satisfying performance in diabetes management tasks owing to its adaptivity and proactivity that are important in chronic disease healthcare [4], [5].

II. State-of-the-art

A. M-Health solution and Diabetes Management Applications

Mobile healthcare (m-Health) is never a new idea and currently there is a wide range of m-health products available in the market. The majority of them focus on logging and tracking user data, while other applications not only track user data, but also deliver solutions or suggestions to users. The 'HidrateSpark' smart water bottle reminds users to drink water and assist them to meet their customised daily water requirement. The 'Pip' stress reliever solution includes gamification to encourage users to manage their stress levels. Our 'Actibetes' project follows the same idea and suggests the exercise level which is believed to best benefit the patient glycemic control.

Regarding the existing diabetes management applications, there are a variety of them available, including 'mySugr', 'Diabetes Pilot Pro', 'Glooko', 'Diabetes in Check', etc. Their functionalities usually include data recording, food database, reminder, etc. Though some of them provide functions related to exercise, for example, activity tracking and connecting to fitness device, none of them provides suggestions after receiving user data. A project in Jamaica has concluded that guidance is important in chronic disease healthcare [5]. Therefore, the proactive nature of the recommender system is believed to be one of the key points that are likely to differentiate 'Actibetes' to other products in the market.

The adaptive nature of 'Diabetes' is another advantage. It corresponds to the adaptive user model in the recommender system and it will be discussed in depth later. Research has shown that the adaptive mechanism is preferred in human-robot interaction [4]. However the adaptive user modelling can be hardly seen in current m-Health market. Therefore, the 'Actibetes' is likely to shown better performance and encourages more user engagement in the long run since the recommender system adapts to the changes in patient over time and provides the personalised suggestions.

B. The HAMMER Architecture

1) Overview: The recommender engine component will be based on the Hammer architecture, which is a hierarchical multi-model structure developed in Imperial College London. 'Hammer' is the abbreviation of the 'Hierarchical Attentive Multi Models for Execution and Recognition' [3]. It utilises the motor control system and arranges models hierarchically. Models imitate the behaviours of demonstrators, in this case, patients and doctors, and compete with each other in performance. In the 'Actibetes' project, models will only be organised in parallel which allows different models that perform exercise suggestions to run simultaneously, while only the suggestion from the best model is delivered to the user-end.

2) Inverse model and forward model: In the motor control theory, inverse models output the motor commands which lead to or approach the predefined target state[6]. Forward models predict the upcoming state following the inverse model commands. It allows fast error detection and adjustment[6]. In the scenario of healthcare, inverse models take the patient states, for example, age, mental condition, etc, and the desired target state, for example, the ideal glucose level. Inverse models then generate action suggestions that are likely to bring user close to the target state. Forward models read output commands from inverse models, for example, taking pills, and predicts the patient states by simulating the response of patients.

The basic unit in the Hammer architecture is an inverse-forward model pair. Commands generated from an inverse model is fed into forward models directly, allowing forward models to predict next state of this particular command. This design performs the imitation learning and can be used in two phases. In the learning phase, inverse models learn the behaviours of doctors and forward models learn the patient responses. The predicted states from forward models are compared with the actual state and the errors between them are used to adjust the forward models and modify the confident levels of inverse models. In the execution phase where multiple inverse models are present, the prediction mechanism of the forward model allows the outcome of various commands to be simulated and compared with each other and thus, the command which is most likely to lead to the desired outcome will be selected and delivered to the user[7]. Also, forward models continue to imitate the patient behaviour, leading to the adaptive nature of the system. Such adaptive mechanism is likely to lead to better performance in long-term chronic disease healthcare where patient physical conditions are likely to change overtime [4].

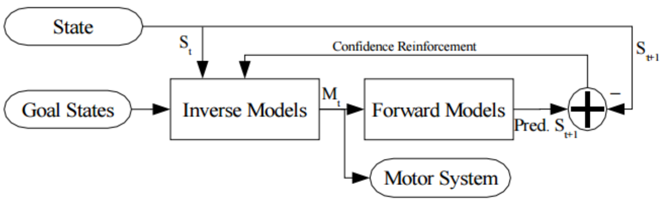

Fig. 1. the basic unit of the HAMMER architecture[8]

3) Multi-model nature and Hierarchy: The inverse-forward pairs can be implemented into the hierarchical structure and multiple model pairs are operating parallel. The target state and the user state are sent to inverse models simultaneously, which allows inverse models to produce a number of actions. All the generated actions are fed into forward models to predict the upcoming user state correspondingly. Since the 'Actibetes' project only focuses on the parallel structure of Hammer, the vertical hierarchical feature will not be explained further in this report.

C. User Modelling and Learning Algorithms

Machine learning is becoming increasingly popular in user modelling field[9] and user preference learning is one of the most common techniques to learn about new users. A range of algorithms are available and have different advantages. Bayesian network are commonly used in distinguishing user abilities and learning styles[10], [11], [12], [13]. Feedback based on learning style has been shown to contribute to better learning outcomes[13]. Support vector machine has been used to recommend learning resource to user[14]. Regression models has been used to learn the user preference on performance feedback[15].

Regarding the knowledge based user modelling, many research has been carried out on modelling patients with diabetes mellitus. A recent paper has suggested to use Gaussian process to model the glucose response to physical activity data for type 1 diabetes, considering insulin and food intake [16]. Another study based on it has included exercise effects and combines the knowledge-based modelling with the support vector regression to perform predictive metabolic modelling[17].

The machine learning in user modelling has less requirement on healthcare knowledge but it requires large, clean and labelled data set for model training[9]. Some machine learning model may involve risk in under or over-fitting, leading to a larger error rate which cannot meet the requirement in healthcare scenario. Instead, knowledge-based model is likely to achieve accurate and meaningful models, though the models can be less adaptive and potentially require complicated variable types and professional data measurement which is not always feasible.

III. Implementation Plan

A. Design Overview

The figure 2 has shown the systematic design of the 'Actibetes'. The recommender system is the back-end of project, receiving data from mobile application, operate the models, and return the exercise suggestions back to users.

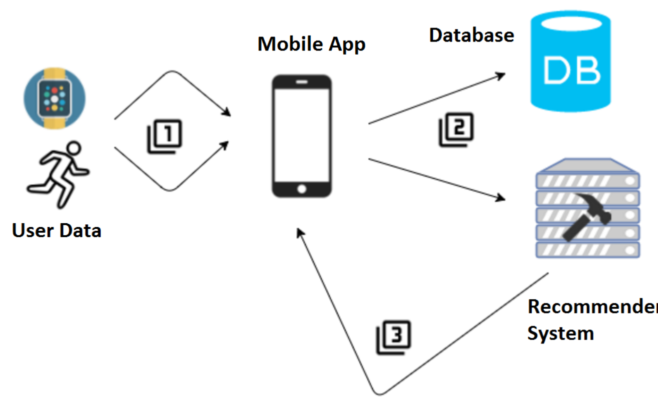

Fig. 2. Systematic design of all components in project 'Actibetes'[18].

The work involved in recommender system component mainly consists of the following:

- Interface Design
- Architecture Design
- Model Implementation

B. Interface Design

The recommender system will be also been wrapped into a online server, for example, the Google App Engine, using REpresentational State Transfer (REST) API, allowing the browser to communicate with the recommender through sending JSON object. REST API is considered as a convenient method owing to its advantages in resource management and simplicity in implementation. Figure 3 has shown the system operation mechanism of the major parties in a RESTful service. The REST API allocate server resource to different entry points and performs actions in response to user requests.

As a starting point, the recommender system has been wrapped into a local Tomcat server. The sample REST interface has been implemented and simple demonstrations have been done to show the communication between browser and recommender system through the REST API.

C. Architecture Design

The main architecture of the recommender engine is based on the HAMMER architecture. Currently, the skeleton of the HAMMER

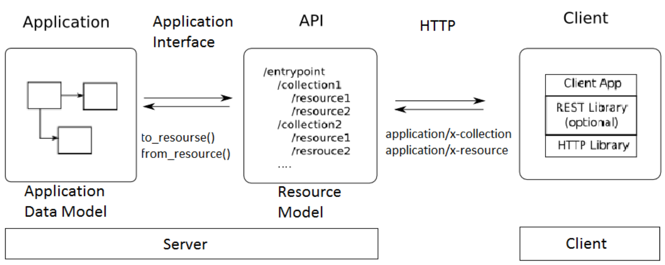

Fig. 3. The interaction between RESTful API components [19].

architecture is nearly accomplished, written in JAVA. Though some advanced features such as hierarchy requires more work to finalise, it is sufficient for the 'Actibetes' project since it only requires the parallel multi-model structure. Demonstrations on the parallel and sequential operation have been done. Fig 4 has shown the simplified version of the UML diagram of HAMMER implementation structure. The detailed UML diagram is also available on Github and can be provided on request.

The 'Hammer' class is the main interface for programmers and it provides many functions such as specifying the 'Core', updating state information and specifying model implementation. The 'Core' is the control hub of the operation, providing access to various threads, models and functions. The 'Core' contains a number of threads, each consists of an inverse-forward model pair. The relation between threads is implemented as a property called 'dependencies', which specifies the execution sequence of threads. The 'World' class simulates the user state change for early-stage testing purpose. Note that to compromise various data types, general types are used in the current implementation, allowing more freedom in real world application. A simple demonstration which includes several dummy models has been done to shown the parallel operation of the Hammer architecture.

D. Model Implementation

Regarding forward models, which are essentially glucose simulators, linear models will be implemented due to the strong linear relationship between exercise, insulin dose, meal ingestion and glucose concentration. So far, a simple linear model has been trained in Matlab. This model is a 1st-order ARMA model which quantifies the effect of past and current activities and combines with the previous glucose levels to predict the current glucose concentration. The training data comes from the UCI Machine Learning Repository[21]. This data set includes the glucose measurement and activity recordings for 70 patients over several months. The simple linear model assumes that the effect of activity happens immediately and decays linearly. Also, the unknown glucose level is assumed to be zero for simplicity. These assumptions are not realistic and are only for an intuitive overview of the data set. The model has the R-square of 0.693, which means around 70 percent of data variation can be explained by the model. The p-value obtained after applying the model on the test set is close to zero, indicates the model is significant. Some coefficient are listed here for a better intuitive understanding of the model.

The coefficients listed are not accurate enough due to the weak assumptions made, but but the values match the expectation to some extent. For example, the intercept point is the glucose level when all variables are zero and it is expected to be the normal glucose concentration while the value in the model is close to that. The

Variable Name	**Coefficient**	**P-value**
(Intercept)	233.04	1.8×10^{-41}
Regular Insulin Dose	-16.622	1.6×10^{-19}
Lunch	6.8745	0.24398
Less-than-usual exercise	-7.5973	0.16728

TABLE I
EXAMPLE VARIABLES AND VALUES

negative sign of the coefficient for regular insulin dose indicates that insulin dose reduces the glucose level. This dummy model proves the feasibility of the linear model to some extent and is expected to show better performance after proper data cleaning and feature engineering.

The inverse model implementation is a classification problem in machine learning. Based on the current glucose level and the effect of past activities, suggestion on exercise will be generated, in this case, the recommender outputs three different levels of exercise without any quantitative measure. As a starting point, possible machine learning algorithms include decision forest and support vector machine. Inverse models are awaiting implementation and it will be first investigated in Matlab using the 'Classification Learner' before translating into JAVA code. There are a number of Java libraries available for general purpose machine learning[22].

E. Future Work

Currently most unfinished work is in interface design and model implementation. First of all, in the next few weeks, I will upload the architecture to the Google App Engine and test the communication with iOS mobile application. Secondly, the current forward model needs improvement. The assumptions on the activity effect highly depend on the activity type. For example, regular insulin dose has an onset action of half an hour to one hour, peak effect in one to two hours and lasts four to six hours. While the ultralente insulin dose has an onset of 4 to 6 hours, a peak of 14 to 24 hours and a duration of 28 to 26 hours. Similarly, other activity effects over time can be estimated better by giving more realistic assumptions. Thirdly, inverse models are awaiting implementation. A variety of classification algorithms are available to be tested on the data set. Thirdly, the parallel model pairs need to be implemented and tested with the Hammer architecture. Fourthly, the systematic operation with other components requires to be tested.

F. Challenges

Many potential challenges have been foreseen so far.

First of all, a large and labelled data set is necessary for training and testing. Currently the dataset comes from the UCI Machine Learning Repository[21]. It may be sufficient for forward model training but training accuracy for inverse models is questionable since the assumption that all activities are suggested by healthcare professionals is be necessary.

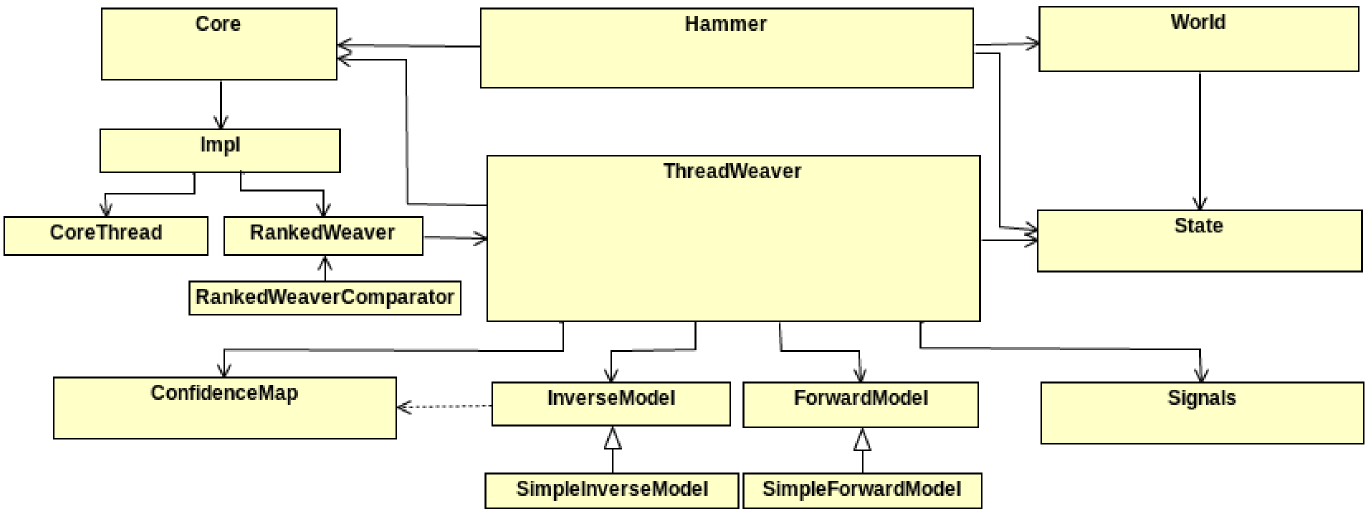

Fig. 4. the UML diagram showing the relation between building blocks in the HAMMER architecture[20].

Secondly, the current dataset only have limited variable types and only differentiate exercise into three levels qualitatively. It might affect the accuracy of the model and how to quantify the exercise level will have impact on model performance as well. Therefore, a better dataset is beneficial overall if available. I have considered using well-constructed model to simulate training data but so far most professionals I have contacted do not have reliable models in simulating glucose level with activity data.

Thirdly, the performance of the system certainly depends on other components as well. Therefore, the communication between the recommender system and other components needs testing. How fast the recommender system reacts and how much data it receives has impact on the overall system.

IV. Conclusion

This report has introduces the recommender system as a component of the 'Actibetes' project. It explains the current progress, future work and challenges. It is exciting to see the development of the 'Actibetes' project, in particular, the collaboration between the recommender system, mobile application and the database. Currently the project progress is positive and many approaches as well as ideas are proved feasible and awaiting to be implemented.

References

[1] N. G. Boulé, E. Haddad, G. P. Kenny, G. A. Wells, and R. J. Sigal, "Effects of exercise on glycemic control and body mass in type 2 diabetes mellitus: a meta-analysis of controlled clinical trials," *Jama*, vol. 286, no. 10, pp. 1218–1227, 2001.

[2] N. B. Ruderman and S. H. Schneider, "Diabetes, exercise, and atherosclerosis," *Diabetes Care*, vol. 15, no. 11, pp. 1787–1793, 1992.

[3] Y. Demiris and B. Khadhouri, "Hierarchical attentive multiple models for execution and recognition of actions," *Robotics and autonomous systems*, vol. 54, no. 5, pp. 361–369, 2006.

[4] A. Sekmen and P. Challa, "Assessment of adaptive human–robot interactions," *Knowledge-Based Systems*, vol. 42, pp. 49–59, 2013.

[5] Y. Tateda, T. Kawamura, T. Yoshida, and T. Yamanaka, "Health education as part of health promotion and prevention of chronic lifestyle diseases in an international cooperation project," *International Congress Series*, vol. 1267, pp. 51–58, 2004. ID: TN-sciversesciencedirect-elsevierS0531-5131(04)00072-X.

[6] R. P. Cooper, "Forward and inverse models in motor control and cognitive control," 2010.

[7] D. M. Wolpert and M. Kawato, "Multiple paired forward and inverse models for motor control," *Neural networks*, vol. 11, no. 7, pp. 1317–1329, 1998.

[8] M. Johnson and Y. Demiris, "Abstraction in recognition to solve the correspondence problem for robot imitation," *Proceedings of TAROS*, pp. 63–70, 2004.

[9] G. I. Webb, M. J. Pazzani, and D. Billsus, "Machine learning for user modeling," *User modeling and user-adapted interaction*, vol. 11, no. 1-2, pp. 19–29, 2001.

[10] N. Li, W. Cohen, K. R. Koedinger, and N. Matsuda, "A machine learning approach for automatic student model discovery," in *Educational Data Mining 2011*, 2010.

[11] A. Bunt and C. Conati, "Probabilistic student modelling to improve exploratory behaviour," *User Modeling and User-Adapted Interaction*, vol. 13, no. 3, pp. 269–309, 2003.

[12] S. Schiaffino, P. Garcia, and A. Amandi, "eteacher: Providing personalized assistance to e-learning students," *Computers & Education*, vol. 51, no. 4, pp. 1744–1754, 2008.

[13] S. M. Parvez and G. D. Blank, "Individualizing tutoring with learning style based feedback," in *Intelligent tutoring systems*, pp. 291–301, Springer, 2008.

[14] A. Al-Hmouz, J. Shen, J. Yan, and R. Al-Hmouz, "Enhanced learner model for adaptive mobile learning," in *Proceedings of the 12th international conference on information integration and web-based applications & services*, pp. 783–786, ACM, 2010.

[15] M. SUAREZ, "Predicting student's appraisal of feedback in an its using previous affective states and continuous affect labels from eeg data,"

[16] J. J. Valletta, A. J. Chipperfield, and C. D. Byrne, "Gaussian process modelling of blood glucose response to free-living physical activity data in people with type 1 diabetes," in *Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE*, pp. 4913–4916, IEEE, 2009.

[17] E. I. Georga, V. C. Protopappas, and D. I. Fotiadis, "Predictive metabolic modeling for type 1 diabetes using free-living data on mobile devices," in *Wireless Mobile Communication and Healthcare*, pp. 187–193, Springer, 2010.

[18] T. R.-P. T. A. Alexander Benoit, Hao Ding, "Actibetes: An activity recommender system for patients with diabetes," 2016.

[19] G. Jansen, "The job of the api designer," Feb. 2016.

[20] H. Ding, "An adaptive recommender engine for m-health solutions to chronic disease," 2016.

[21] M. Lichman, "UCI machine learning repository," 2013.

[22] josephmisiti, "Awesome machine learning github repository," 2016.