An Adaptive Database Design Approach for Biosignal Data Storage and Retrieval

Jinghong Cox Chen Hewijin Christine Jiau

Department of Electrical Engineering,
National Cheng Kung University, Taiwan, R.O.C.
cox@nature.ee.ncku.edu.tw,
jiauhjc@mail.ncku.edu.tw

Abstract

In the traditional scenario of database usage, database usually plays a passive role as a backend for supporting user applications. This introduces a tight coupling from database to user applications, which is the source of high maintenance fee for the database, especially for databases developed using conventional normalization-based design approach. In the case of designing a database for biosignal data storage and retrieval, uncertainties both from the data provider and the data consumer make an even stricter situation for database design. This paper addresses the issues concerned with an adaptive database design approach capable of producing a survivable database design in facing the changing environment of biosignal domain, and presents two showcases with databases designed following our adaptive database design approach.

Keywords: Adaptive Database Design, Biosignal Database, Biosignals, COMIS.

1 Introduction

Biosignal is the signal of human body reflecting specific aspect of health status. Within a typical biosignal analysis scenario, a monitor is connected with several biosignal devices measuring specific biosignals from human body. Collected biosignal data will be processed by corresponding schemes of biosignal analysis and their results will be merged to have an information representation on monitor's display. After studying represented biosignal information, trained specialist will have the knowledge of specific aspect of health status and will be able to make further decisions for scientific research, drug trials, or healthcare.

Initially, the usage of biosignal was restricted inside the scope of hospital. This was partly because of the unapplicability of vast manufacturing for complex biosignal devices. With the advances in device manufacturing, volume and cost of current biosignal devices could now be as very small as electronic devices for daily use. The convenience of equipping biosignal devices helps in spreading the applications of biosignal. Actually, there are already many handy biosignal devices designed as our iPods that are wearable and will not bother our daily life one bit. You may not notice that a Walkman-looking device would help in recording your heartbeat rate and a ring on middle finger is responsible to logging your skin temperature.

Although we can have such handy biosignal devices doing their jobs just the same as those complex devices in the hospital, there are still missing parts in the typical scenario of biosignal analysis. Biosignal data collected using these wearable devices have to be sent back to hospital for further studying. Here come the problems: "*how do we know which biosignal data should be transferred to whom being responsible?*", "*what should be done if these biosignal data could not be processed immediately?*", and etc.. There is always a general answer to this kind of data sharing problems: a database. A biosignal database will play a role as a broker to fulfill missing parts in the typical scenario of biosignal analysis where wearable biosignal devices are considered as major data sources. That is, biosignals measured using biosignal devices will be sent back to hospital as biosignal data. Various biosignal data will be uniformly stored in the biosignal database waiting for further studying. Through suitable query mechanisms, biosignal analysts can obtain desired biosignal data from the biosignal database without hard-wired with various biosignal devices. This may slightly alter the user model (Figure 1) of biosignal analysis, but it would be a worth-taking trade-off to have the application terrain of biosignal broadened.

Until this research work is done, according to our surveys, there is not a viable solution for biosignal data storage and retrieval. Maybe there would be some databases customized for use in current biosignal analysis activities, they hardly take the characteristics of biosignal data into consideration. Trivial database designs usually cause a sacrifice for biosignal database extensibility. In this research work, we focus on proposing a database design approach that would hopefully lead to an adaptive solution followed with most future extensibility for biosignal data storage and retrieval. Before rushing into our kernel, in Sec. 2, we will review the characteristics of biosignal data and thereof the challenges of biosignal database design. Our approach bearing in mind the nature of biosignal data will produce an adaptive database design for biosignal data storage and retrieval, which is fully explained in Sec. 3. In Sec. 4, two showcases are presented to show advantages of database designs using our approach. Finally, we conclude our research work on developing an adaptive database approach in Sec. 5

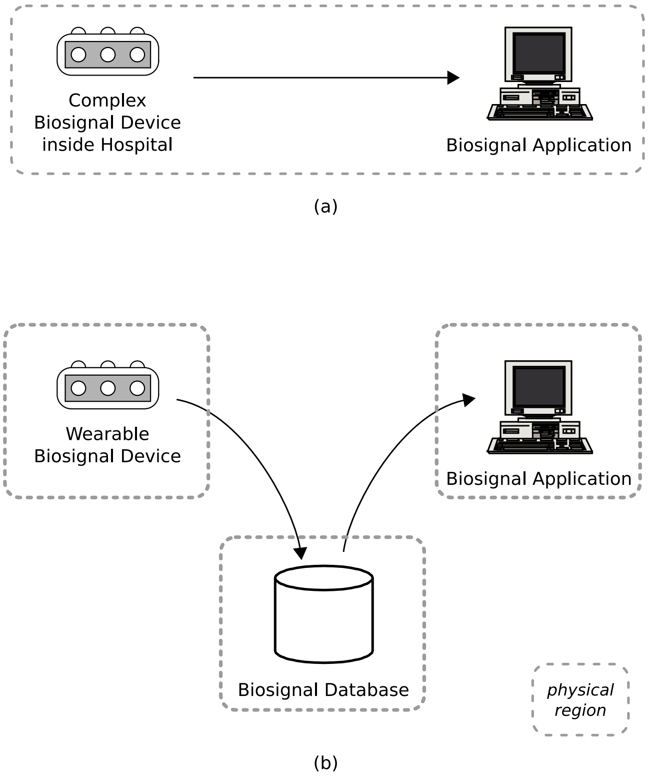

Figure 1. User models for (a) complex biosignal device inside hospital and (b) wearable biosignal device.

2 Characteristics of Biosignal Data and Challenges of Biosignal Database Design

Biosignals, also known as *Biomedical Signal*, are measurable signals of human bodies that are usually integrated to model specific health conditions. In this work, we especially focus on those biosignal data collected by wearable biosignal devices. Target biosignals to be measured by wearable biosignal devices are always the same as those measured using complex devices inside hospital. Major distinction relies on the user models, as illustrated in Figure 1, utilizing these two kinds of biosignal devices. More specifically, biosignal data collected by wearable biosignal devices have to be shaped as regular packets before being sent back to the biosignal database for storage, which is not usually the case for those collected using complex devices inside hospital. In the following paragraphs, we will use *biosignal data* in short for *biosignal data measured by wearable biosignal devices*.

2.1 Characteristics of Biosignal Data

Although particular biosignals will have their own distinct attributes, there are several common characteristics of biosignal data that are to be stored in the biosignal database. These common characteristics are concluded in the following list:

1. **Historicalness**: *Biosignal data can not be modified after they are generated.*
2. **Diversity**: *There is no single unified format for biosignal data.*
3. **Versatileness**: *The potential applications of biosignal are still evolving.*
4. **Imperativeness**: *Since biosignal reflects specific aspect of health status, collected biosignal data should be studied as soon as possible.*

In a working scenario of user model (b) illustrated in Figure 1, there will be tens of hundreds connections of wearable biosignal devices waiting for biosignal data storage service. Thus, the oftenness of demanding for maintenance routines (**Historicalness** and **Imperativeness**) in the biosignal database is understandably larger than usual scenarios of database usage. Considering the oftenness of maintenance routines and characteristics listed above, choosing a centralized architecture for biosignal database design is a better solution rather than distributed or federated architectures. So far, we can transcribe our design problems of biosignal database as: "*To have biosignal data stored in a centralized single database, but we don't have a unified format* *for biosignal data (**Diversity**) neither can we limit their possible applications (**Versatileness**)*"

2.2 Challenges of Biosignal Database Design

Challenges of biosignal database design come from the **Diversity** and **Versatileness** characteristics of biosignal data. It is impossible to fulfill both needs from data providers and data consumers without solving challenges caused by these two characteristics.

Heterogeneous data formats from biosignal data providers(Figure 2-(a)): Following conventional database design approach, the first job is to perform normalization on biosignal data. Once normalization job is done based on concrete biosignal data formats at hand, the database design is put to the dead-end that will never be adaptive to various biosignal data. This is because that conventional database design reflects normalized data structure in database logical view. Performing normalization on current biosignal data means to restrict the conformance of database logical view to a limited set of biosignal data formats. For the diversity characteristic of biosignal data, heterogeneous formats of biosignal data are hardly to be consolidated into one single database logical view. Somebody may argue that there are efforts on developing standards [1] for biomedical signal data formats. Yes, but they are still under development. And even one day we will have some standards for biosignals to be followed, issues on database design remain unsolved. This is because that *standards* may promise exchangeability or some level of interoperability, they never stand for being unchangeable themselves. Any slight change to these standards will cause certain risks on altering database design, this is crucial to database itself and other related entities.

Diverse application requirements from biosignal data consumers(Figure 2:(b)): Application requirements will also affect database logical view in conventional database design. What the database logical view represents usually have to be conformed to database users' needs. However, data normalization of conventional database design approach will fix database design, and there will have no space left for future extensibility.

Among the literatures, there are also some efforts on providing adaptive solutions for biomedical data storage [2, 3]. The work [2], entitled with *Database Design for Incomplete Relations*, is a supplementary solution for data normalization. Another work [3], entitled with *A Dynamic Clinical Dental Relational Database*, proposed a database design that stores dental epidemiological data according to their *Data Definition Structure* (DDS). But their assumptions never include the situation faced by biosignal database that challenges of uncertainty simultaneously come from both end of database context: the data provider and the data consumer.

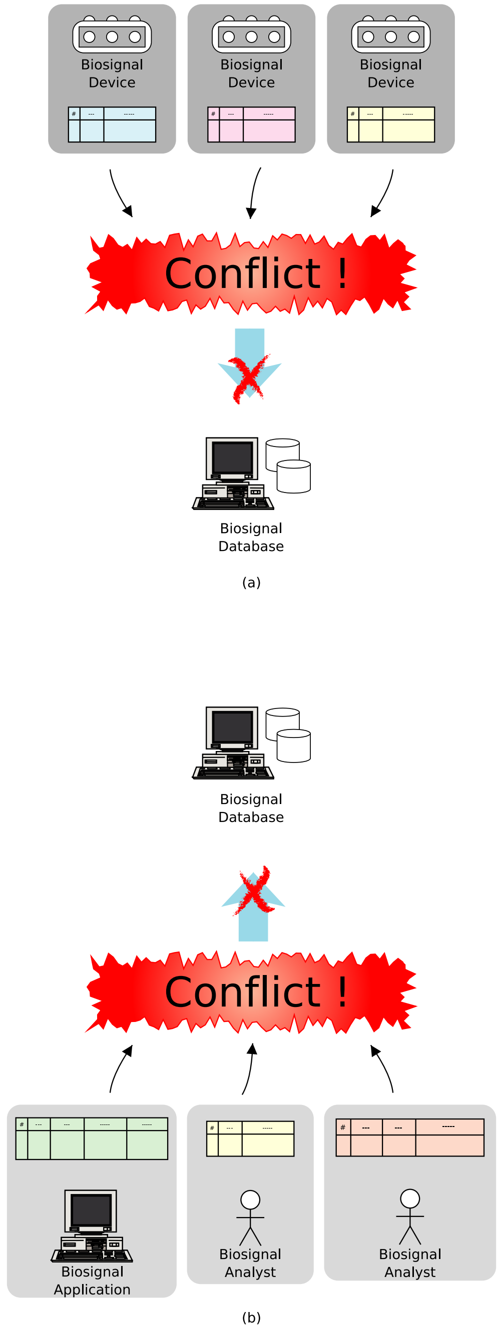

Figure 2. Challenges from (a) Data Providers and (b) Database Consumers.

3 Our Approach: Decoupling of Information Seeking Behavior from Database Logical View

Our approach for adaptive database design is consisting of three atomic steps of practices. Firstly, our approach begins in the decoupling of information seeking behavior. The following step is to decide a database logical view for the biosignal database. As for the final part of our approach, a new concept, *Biosignal Protocol*, for the logical organization of biosignal data is introduced for the integration of biosignal data into biosignal information.

3.1 Decoupling of Information Seeking Behavior

The information seeking behavior helps in constructing appropriate biosignal information based on biosignal data stored. Applying conventional approach for database design, the developer has to understand the nature and structure of biosignal data before he can decide database logical view. But, why does the biosignal database has to know the logical organization of biosignal data in order to store them? Reviewing user models presented in Figure 1, the role who must have sufficient knowledge of the structure of biosignal data to drive the typical scenario of biosignal analysis is *Biosignal Application* not biosignal database. Just like the role post offices play in the scenario of mail delivery, they do not have to know any detail of a package in order to deliver it.

In opposite to conventional database design approach that reflects data structure in database logical view, our approach leave database logical view neutral to the logical organization of data. The kernel of our approach is the decoupling of information seeking behavior from database logical view. We try to utilize minimum knowledge of biosignal data in the development of database logical view, but leave most clues only acknowledgeable to users of biosignal data. Thus, the accomplishment of information seeking behavior is not merely based on database logical view but gets majorly supported by our extra information-level mechanisms. We will have these information-level mechanisms further explained in Sec. 3.3.

3.2 Deciding Database Logical View

While deciding database logical view, conventional approach usually leads to a union of data diversity. Our approach, on the other hand, finds the minimum commonality among heterogeneous biosignal data formats according to our surveys [4–9] on biosignal data analysis and processing. The column fields of our database logical view will be neutral to specific needs of biosignal analysis. Every entry

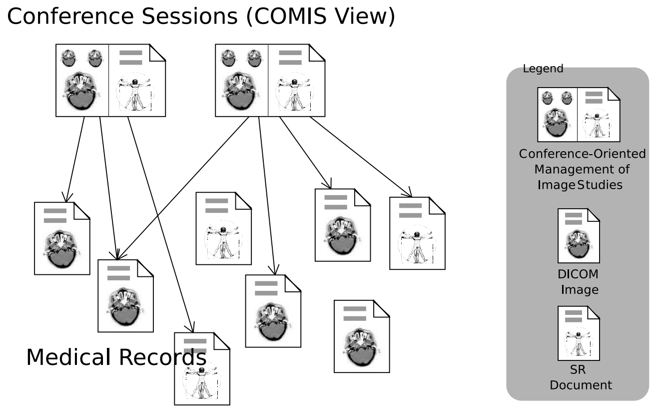

Figure 3. Conference-oriented management of image studies.

of biosignal data stored in the database should expose the database logical view comprising:

Device ID Every biosignal device connected to the biosignal database will be given a unique ID for identification.

Timestamp The exact time that specific biosignal data is recorded by the device.

Biosignal Type Specific type of biosignals that the biosignal data belongs to.

Biosignal Content Rest parts of an entry of biosignal data.

3.3 Biosignal Protocol for Modeling Biosignal Information

The last activity of our approach is to help the construction of appropriate biosignal information. In our previous work about the development of telemedicine systems [10], we have developed *Conference-Oriented Management of Image Studies* (COMIS) for the logical organization of medical data used during a telemedicine conference (Figure 3). While constructing biosignal information, we have found that concepts of COMIS are also applicable in biosignal domain. Thus, we tried to have a mapping from telemedicine domain to biosignal domain, and developed an extension of COMIS. The result extension for biosignal domain is given a domain-aware name as *Biosignal Protocol*.

Before a telemedicine conference can be held, leader of the conference have to provide a COMIS abstraction. A COMIS abstraction is consisting of all necessary medical records, including images, texts, or other media, which are intended to be discussed during the telemedicine conference. Biosignal protocol, similar to COMIS, facilitates the organization of biosignal data for the needs of a specific biosignal analysis activity. With information seeking behavior decoupled from database logical view of the biosignal database, biosignal protocol plays an important role in compensating semantical relations of biosignal data. Particular biosignal protocol description organizes semantically-related biosignal data into biosignal information needed for specific biosignal analysis activity.

The kernel spirit of the biosignal protocol concept is a redistribution of responsibility in the typical scenario of biosignal analysis. Figure 4 shows the result database system architecture using our approach compared to distributed and federated approach. We have mentioned that our approach will produce a better database design than other normalization-based approaches, such as distributed or federated approach, and Figure 4 would be an illustrative evidence of our points.

Basically, clients will not notice pros/cons among these three different database architectures during their usage experiences. The appropriate stage for database architecture developed using our approach is maintenance, especially the necessary costs of time and effort of various database maintenance activities. Maintenance costs are mostly related to the number of physical storages. Every physical storage depicted in Figure 4 may stand for different database logical view, which subsequently means different requirements for maintenance strategies or routines. Obviously, maintenance costs of distributed or federated architectures would be much higher than costs of the architecture introduced by our approach. Furthermore, if we consider the possibility of adding new formats of biosignal data, the situations of distributed and federated architectures will only getting worse, and ours remains. As stated in Sec. 3.1, we strongly believe that it is not necessary to have the nature and structure of biosignal data reflected in database logical view. If we implementing above believing by biosignal-protocol-based redistribution of responsibilities, maintenance works of resulted database would be surprisingly easy thereafter.

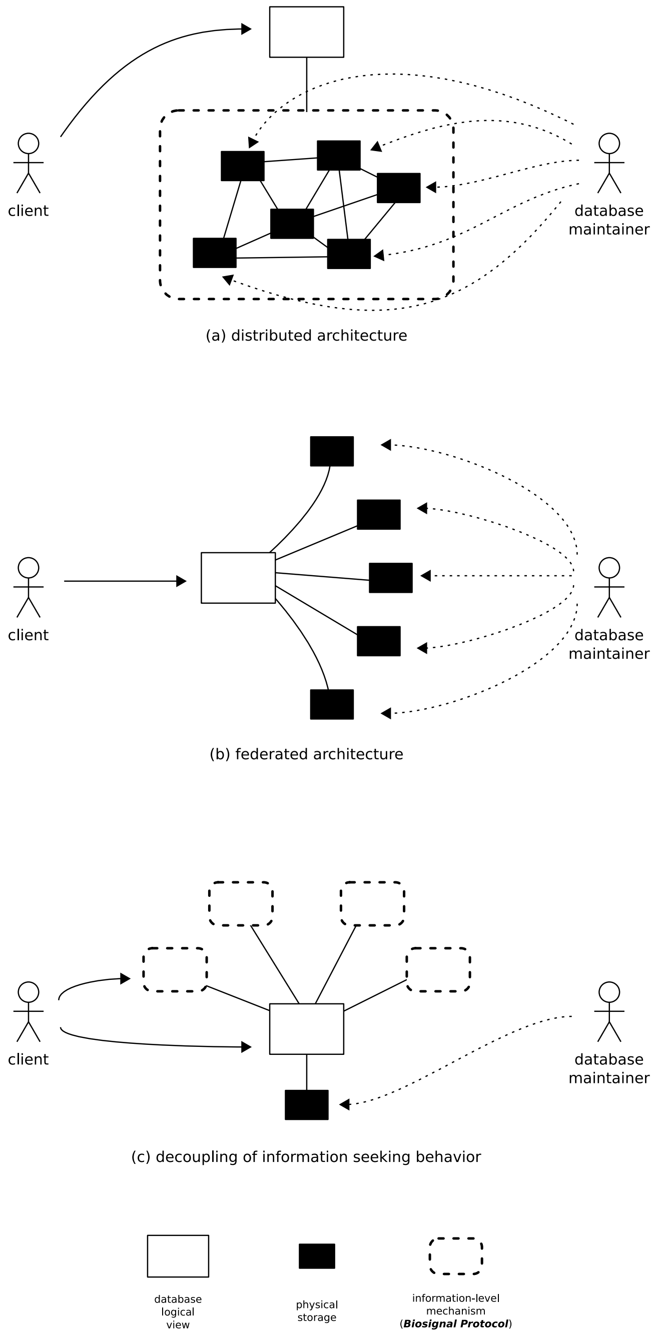

Figure 4. Database system architectures developed under (a) distributed, (b) federated, and (c) our decoupling of information seeking behavior design approach.

4 Showcases with Databases Designed Following Our Adaptive Database Design Approach

In this section, we are going to present two showcases with databases designed following our adaptive database design approach described in previous section. The first case is named as *Emotion Ring Application*, which is a real biosignal application with a biosignal device bundled as a commercial product. There was originally a database in *Emotion Ring Application*. However, considering the high maintenance costs brought by the evolution of bundled biosignal device, a refactoring to database design is inevitable for better adaptivity. We will have further explanation in Sec. 4.1. As for another showcase, *Smart Biosignal Database*, is a pure database system developed with intentions to facilitating multi-purposed biosignal data storage

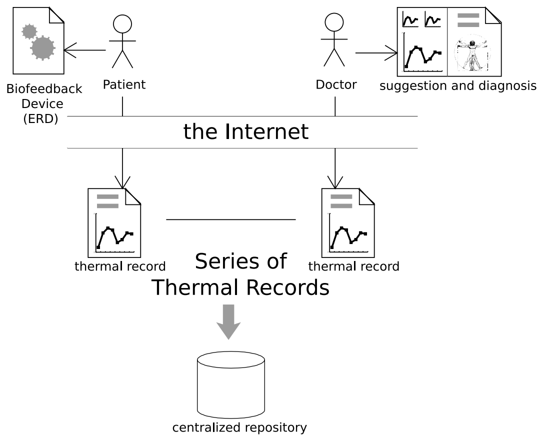

Figure 5. Conceptual architecture of the emotion ring application.

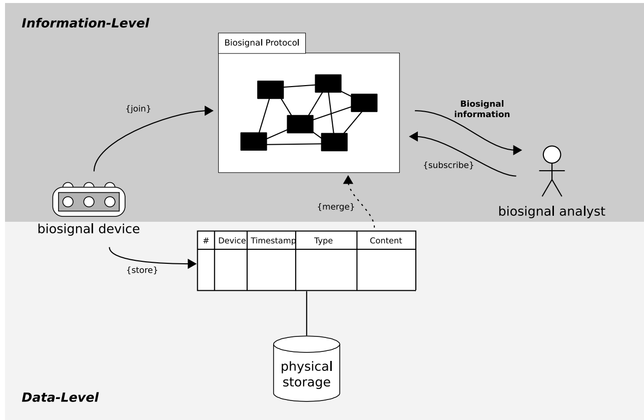

Figure 6. The conceptual architecture of the smart biosignal database.

and retrieval. *Smart Biosignal Database* is designed completely following our adaptive database design approach, which we will have detailed descriptions in Sec. 4.2.

4.1 Emotion Ring Application

Emotion Ring Application [11] (ERA) is a concrete biosignal analysis application. Based on the measurement of skin temperature, ERA assists in the biofeedback-based treatment of Melancholia. As shown in Figure 5, an accompanying device, *Emotion Ring Device* (ERD), is deployed to the patient for measuring instant changes of his/her skin temperature. Collected biosignal data of skin temperature will be sent back to a centralized repository of database for further study. The doctor in chief have to review a series of thermal records stored as biosignal data in order to make suggestions and diagnoses to current status of the patient.

The database design of ERA seems to be easy because there is only on biosignal type to be handled. However, the actual format of a thermal record changes with the evolution of the ERD, and user requirements of doctors are not usually the same for specific clinical scenarios. Thus, extra costs are usually spent in maintaining database structure for current needs rather than in the upkeep of rest parts of ERA. While introducing our database design approach to the refactoring for database of ERA, maintenance costs of database decrease significantly as expected. Database design have not to be altered with the evolution of ERD or with the change of user requirements of doctors.

4.2 Smart Biosignal Database: A Collaborative Project with I.I.I.

Smart Biosignal Database is a collaborative project with the Institute for Information Industry (I.I.I.). Foreseeing the popularity of biosignal applications in healthcare and related domains, the *Smart Biosignal Database* is intended to be a pioneer of biosignal data storage and retrieval facilities. To this end, the primary design goal of *Smart Biosignal Database* project is to establish a solid foundation for biosignal data storage and retrieval. So that future biosignal data with different formats from current ones could be stored adaptively without altering database design and unpredictable needs from biosignal analysis would be fulfilled based on the functioning of *Biosignal Protocol* mechanism mentioned in Sec. 3.3.

Figure 6 is the conceptual architecture of the functioning *Smart Biosignal Database* system. Biosignal data sent back by biosignal device would be simultaneously dispatched to be stored in physical storage and to be processed by the mechanism of biosignal protocol. On the other hand, a biosignal analyst can subscribe to existing biosignal protocol descriptions, which means to receive specific logical organization of biosignal data as desired biosignal information from *Smart Biosignal Database*. He can also describe special needs of biosignal analysis as a new biosignal description if he cannot find a satisfying one among existing protocols. During the construction of desired biosignal information, related biosignal data will be pulled out from physical storage of *Smart Biosignal Database* according to the details of specific biosignal protocol description.

Although *Smart Biosignal Database* is under its development, we have built a full-functioning prototype ready for testing of storing biosignal data of diverse formats and for integral testing with real-world biosignal devices. The final product of the *Smart Biosignal Database* project will be a multi-purposed database system developed completely following our adaptive database design approach explained in Sec. 3.

5 Summary and Conclusions

In common scenarios of database usage, databases always play a passive role that store what users want to store and return what users request for. This results in the strong coupling between database design and user applications. Once environment changes, database design has to be altered in order to satisfy new user needs. The maintenance fee of database is usually higher than rest parts of a system, and this is usually the case for database designed using conventional approach.

In this work, we have presented our adaptive database design approach that leads database design to an adaptive stage being neutral to the changes of environment. Maintenance fee of database is promised to be reduced based on the decoupling of information seeking behavior from database logical view. We also introduced two showcases advocating the remarkable advantage of reducing maintenance costs brought by our approach while comparing to conventional database design approach. In the near future, there maybe some standards developed to be followed by biosignal database design, just like the cases of DICOM [12, 13] for medical images and HL7 [14] for text-based medical records. Hopefully, our approach will remain surviving since compatibility to any biosignal standards would be a simple job based on the mechanism of biosignal protocol. Furthermore, With our strong believing, our approach is not only applicable in database design of biosignal domain but also can be developed as a general-purposed adaptive database design approach.

6 Acknowledgement

This research was supported by the Applied Information Services Development & Integration project of Institute for Information Industry and sponsored by MOEA, R.O.C..

References

[1] A. Värri, B. Kemp, T. Penzel, and A. Schlögl, "Standards for Biomedical Signal Databases," *IEEE Engineering in Medicine and Biology*, vol. 20, no. 3, pp. 33–37, May/June 2001.

[2] M. Levene and G. Loizou, "Database Design for Incomplete Relations," *ACM Transactions on Database Systems*, vol. 24, no. 1, pp. 80–126, March 1999.

[3] D. Taylor, R. N. G. Naguib, and S. Boulton, "A Dynamic Clinical Dental Relational Database," *IEEE Transactions on Information Technology in Biomedicine*, vol. 8, no. 3, pp. 298–305, Sept. 2004.

[4] D. A. Marcus, L. Scharff, S. Mercer, and D. C. Turk, "Nonpharmacological Treatment for Migraine: Incremental Utility of Physical Therapy with Relaxation and Thermal Biofeedback," *Cephalalgia*, vol. 18, no. 5, pp. 266–272, 1998.

[5] P. M. Marzuk, "Biofeedback for gastrointestinal Disorders: A Review of Literature," *Annals of International Medicine*, vol. 103, no. 2, pp. 240–244, 1985.

[6] C. Buby, L. F. Elfner, and J. G. M. Jr., "Relaxation Pretraining, Pulse Wave Velocity and Thermal Biofeedback in the Treatment of Essential Hypertension," *International Journal of Psychophysiology*, vol. 9, no. 3, pp. 225–230, 1990.

[7] M. Nakao, S. Nomura, and T. Shimosawa, "Blood Pressure Biofeedback Treatment of White-coat Hypertension," *Journal of Psychosomatic Research*, vol. 48, no. 2, pp. 161–169, 2000.

[8] P. M. Lehrer, D. Sargunaraj, and S. Hochron, "Psychological Approaches to the Treatment of Asthma," *Journal of Consulting & Clinical Psychology*, vol. 60, no. 4, pp. 639–643, 1992.

[9] G. D. Rose and J. G. Carlson, "â^The Behavioral Treatment of Raynaudâ^s Disease: A Review," *Biofeedback & Self Regulation*, vol. 12, no. 4, pp. 257–272, 1987.

[10] H. C. Jiau, J. C. Chen, K.-F. Ssu, and J.-M. Lin, "TeMeFr: Towards a Reuse-Based Development for Conference-Oriented Telemedicine Systems," *Journal of Information Science and Engineering*. Accepted for publication.

[11] M. C. Lai and H. C. Jiau, "Design and Develop an Interactive Biofeedback System," *Proceedings of the Twelfth Workshop on Object-Oriented Technology and Applications*, pp. 249–257, 2001.

[12] DICOM, "Digital Imaging and Communications in Medicine." URL http://medical.nema.org/.

[13] David Clunie, "DICOM Structured Reporting." URL http://www.dclunie.com/.

[14] Health Level Seven. URL http://www.hl7.org/.