

Implementation of Web Database for ECG

MA Paracha, SN Mohammad, PW Macfarlane, JM Jenkins

University of Michigan, Ann Arbor MI, USA and University of Glasgow, Scotland

Abstract

The collection over a period of years of a large number of electrocardiograms (ECGs) from normal subjects has allowed the development of age and gender based limits of normality. Ease of access to this digital library among collaborators would be beneficial for future research studies.

A friendly environment for retrieving the required information is a key component of the information system infrastructure of modern research projects. Due to the rapid growth of the global Internet, researchers are demanding an improvement of traditional techniques for storage of data in order to get required information in a convenient fashion. This feature is prototypical of web databases; thus they are widely used for their excellent capability of securely sharing data and providing fast access to on-line information.

We have developed a web database in order to facilitate access to a valuable, decades-long collection of data: over 3000 normal ECGs from neonates, infants, adolescents and adults along with information about their age and sex. Patient name, ID, and personal information are kept confidential. A mean P, QRS, T cycle together with the usual fiducial points, wave durations, and interval measurements are available.

This application is built around a three-tier architecture model using open source software. The base of the application (third tier) consists of the database management system. MySQL is used for this purpose. Built on the top of the third tier is the second tier that contains most of the application logic, which is achieved by using PHP: Hypertext Preprocessor (PHP). Apache web server is used to facilitate communication between tiers. The first tier (client) is typically web browser software that interacts with the application.

This system provides the facility to acquire ECGs by features of interest, i.e., age, sex, heart rate, etc. Security is afforded by restricting access to research collaborators only. This web database provides a user -

friendly environment for ease of data access, and has resulted in a significant improvement in the data retrieval rate. One early application has clearly demonstrated the divergence in selected ECG measurements in children and adolescents at about the time of puberty.

1. Introduction

A carefully crafted protocol to collect electrocardiograms (ECGs) from normal subjects was initiated at the University of Glasgow many decades ago, and the collection now contains over 3000 records of subjects from age 0 days to 82 years of age [1,2]. The recordings are 12-lead ECGs processed by the Glasgow program to identify fiducial marks (onset and offset of P and QRS, onset and offset of T waves) and provide interval and waveform measurements. The data were translated to vector representation for one particular study in children [3].

This collection has been the basis of numerous studies identifying ranges of *normal* measurements in a large heterogeneous population. This database of thousands of *normals* has been the subject of numerous studies and publications, and remains a valuable source of data for future studies.

A collaborative study between the Section of Cardiology, University of Glasgow and Medical Computing Laboratory, Electrical Engineering and Computer Science, University of Michigan and Ann Arbor Electrogram Libraries [4] provoked the creation of an improved method of data storage and a more efficient, friendly, and secure access to data. This led to the development of a web-database, which uses Database Management System for the storage of data and adds value to the data by turning data into information; and by turning this information into knowledge [5]. We have created a web-based relational database of these important data to provide powerful retrieval and analysis of data, and have designed an access mechanism that allows secure access to valid research users for future studies.

Other investigators have similarly designed database and expert systems for use by the cardiology community [6–10]. In comparison to traditional approaches for storage of data, using a web-based relational database has

the benefit of allowing simultaneous and platform-independent access to information. This affords retrieval and modification of data by all investigators regardless of their geographical location.

2. Materials and methods

2.1. Materials

The application was developed using open source software. The use of open source software for application development makes it cost effective and flexible.

MySQL is used as a database server. It is a multi-user, multi-threaded and robust SQL (Structured Query Language) database server. The MySQL database server embodies an ingenious software architecture that maximises speed and customisability [11]. The main features of MySQL are speed, robustness, and flexibility to store large objects such as ECGs. Furthermore the fact that MySQL is a freeware and its web interface can be built easily makes it the database server of choice.

Apache is used as a web server. It is a fast, stable and feature-full open source web server [12]. It is designed to run under multitasking operating systems and is capable of handling simultaneous requests.

PHP is used as a scripting language. PHP is a widely used open source scripting language. PHP is particularly suited for web database applications because it supports a wide range of databases and has tools that integrate the Web and database environments [13].

2.2. Methods

The implementation of our web database can be described as three-tiered architecture model. The base of the application is the database tier, comprised of the database management system, MySQL. Built on the top of the database tier is the middle tier that contains most of the application logic, which is accomplished by using PHP: Hypertext Preprocessor (PHP). Apache web server is used to facilitate communication between tiers. On top is the client tier, i.e., typical web browser software that interacts with the application.

2.2.1. First tier

The client in our three-tier architecture model is implemented as thin client. Web browsers are very thin clients and thus the first tier in our three-tier architecture model is just a web browser. Using a web browser as a thin client offers the advantage of easy deployment, the user who has a web browser can use our application without installing anything new. Also using web browser as a thin client makes our application platform independent. Thus a diverse range of geographically dispersed users can access useful information online using our application.

2.2.2. Second tier

The second tier handles most of the application logic. It mainly serves as an interface to connect the other two tiers. It consists of an Apache web server, PHP Zend engine (scripting engine) and PHP scripts.

The second tier is responsible for collecting the data from the user (a request is made to the web server, which is then passed to the Zend engine (scripting engine) through the Zend engine's web server interface). Based on the information provided by the user data query is prepared and sent to the database (An appropriate PHP script is retrieved from the disk which is then compiled and run by the Zend engine). Finally, the results are formatted and sent back to the user for display using the Internet (The web server interface returns the output to the web server which passes it back to the user.)

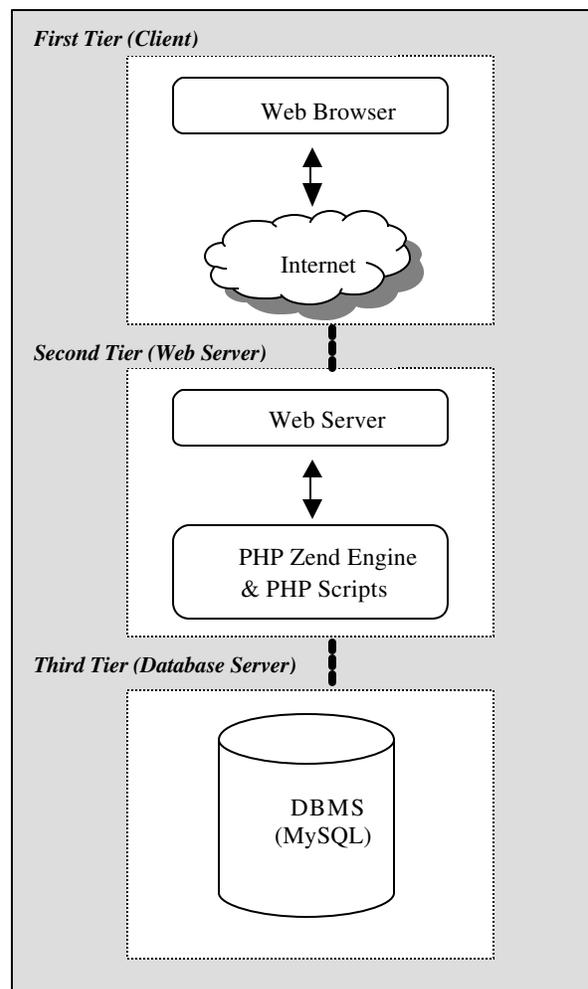


Figure 1. The three-tier architecture model of web database

2.2.3. Third tier

The third tier in our three-tier web database application is the Relational Database Management System (RDBMS). We have used MySQL RDBMS to provide the required functionality. The responsibilities of RDBMS include

- To effect efficient storage of data
- To provide data integrity constraints
- To provide concurrent multi-user access
- To guarantee authorized and secure access to the data
- To recover the database in case of a crash

3. Results

Plots were generated from the Glasgow Data using Matlab numerical analysis software package. For example, for each of the pediatric patient records (over 1800 entries total), derived XYZ-plots were generated in Matlab and the images were exported to the relational database linking the individual images to corresponding patients identified anonymously by number. In essence, a category “XYZ-plots” consisting of the images was created as a table in the database and mapped to the main patient record table using the primary key.

Figure 2 demonstrates how an individual patient’s record would appear using the scheme described above. In this case, we have four categories: individual lead plots, XYZ-plots, QRS complex, and T-wave with dV/dT plots. These categories appear as links from the patient record page that serves to retrieve the image itself.



Figure 2. Individual patient record view

In each of the categories, there need not be a one-to-one relationship to patient records. For example, a single patient may have multiple plots associated with the same category. If this is the case, the link can be programmed to retrieve all images in that category for the

corresponding patient. Retrieval of a single image from the “XYZ-plot” category for a unique patient is demonstrated in figure 3. If multiple images were associated for this particular patient in this category, they would simply appear one after the next on the same page.

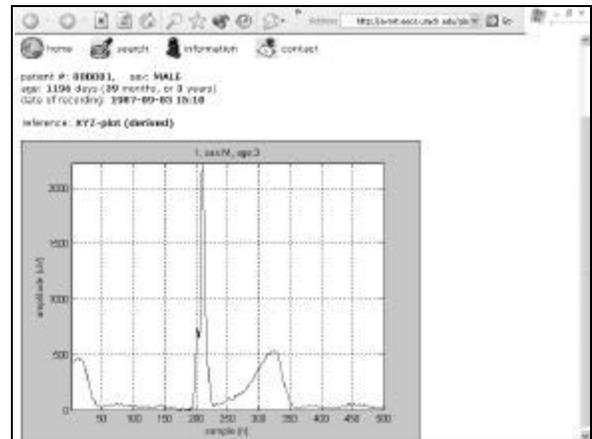


Figure 3. Patient-specific plot retrieved from database

Due to the sensitive nature of the data, sufficient security must be maintained on all access to the database. To this effect, user-level access is controlled through a login system; when an individual or an organization is granted access to the database, they are assigned a unique “username” and “password” which is required to access the system. Figure 4 shows the initial screen challenging the user for authentication.



Figure 4. System login, user-level security

Using the login scheme, different privilege levels can be granted to various users depending on their nature of work. For example, data maintenance staff may have privileges to add or modify patient records whereas systems administrators may have privileges to add,

modify or delete user accounts. General users may have read-only privilege on the data. Furthermore, general user accounts may be restricted to the pediatric or the adult dataset, or may have access to both. Data access licenses themselves are stored in a separate database, allowing the administrative staff to seamlessly grant and administer licenses.

As a value-added security feature, all accesses whether to user accounts by administrators, or to the database itself by the end-users, are logged. For example, in Figure 2, the users last login date and location are displayed; if the user has any doubts about whether unauthorized accesses have taken place using their username and password, they can contact the administrative staff who will have access to the entire login history for that particular user.

The complete system functions seamlessly from the end-users perspective, providing them with an easy-to-use secure interface to a wealth of data. Enabling the system to deliver content in an elegant format allows the end-user to spend less time retrieving the required data and more time on its applications in their respective research (14). Effectively, the system discussed above is more “knowledge-delivery” than “data-delivery.”

4. Conclusions

The prototype database system was used for an investigative study to determine the age at which sex differences in ECG measurements, not present in newborns or infants, but well known in adults, begin to appear. Investigators worked remotely from Glasgow, Scotland and Michigan USA [15]. Electronic access to data is now easily afforded via web-based libraries. Security of access is an important component if such research databases are to become realizable. Our web database design incorporates ease of use, and simultaneous access by multiple investigators in a collaborative research study. We believe that web databases are the means for facilitating interaction despite geographical distance.

Acknowledgements

This work was supported in part by National Science Foundation grant BES9825125, Guidant Corporation, and an equipment grant from Intel Corporation.

References

[1] Macfarlane PW, Lawrie TDV (1989): The normal electrocardiogram and vectorcardiogram. In *Comprehensive Electrocardiology: Theory and Practice in Health and Disease*, 1st ed. Vol. 1, ed. PW Macfarlane, TDV Lawrie, pp. 407-57, Pergamon Press, New York.

[2] Macfarlane PW, McLaughlin SC, Devine B, Yang TF. Effects of age, sex, and race on ECG interval measurements. *Journal of Electrocardiology*. 27 Suppl:14-9, 1994.

[3] Edenbrandt L, Houston A, Macfarlane PW. Vectorcardiograms synthesized from 12-lead ECGs: a new method applied in 1792 healthy children. *Pediatric Cardiology*. 15(1):21-6, 1994 Jan-Feb.

[4] Ann Arbor Electrogram Libraries, Ann Arbor MI.

[5] John van den Hoven. Adding value to data. *Information-Systems-Management* June 2002;19(3):89-92. Boston: Auerbach.

[6] Magdalena JR, Soria E, Serrano A, Calpe J, Guerrero J, Martinez M. NEMESIS: A new telemedicine approach to cardiologic software. *Computers in Cardiology* 1999;26:37-40.

[7] Fayn J, Rubel P, Girard P, Ravel C, Förlini MC, Boissel JP. CAVIAR, a serial ECG processing and management system for pharmacological drug trials. *Computers in Cardiology* 1991. Venice Italy: IEEE Computer Society Press. 1992:301-4.

[8] Combi C, Pincirolì F, Pozzi G. Temporal clinical data modelling and implementation for PTCA patients in an OODBMS environment. *Computers in Cardiology* 1994. Bethesda MD USA: IEEE Computer Society Press. 1994: 505-8.

[9] Nieberl J, Khor S, Khor M, Kekes E, Szaboki F. Twin expert systems for the complex analysis of conventional 12 lead electrocardiograms. *Computers in Cardiology* 1993. London UK: IEEE Computer Society Press. 1993:511-3.

[10] Yuanling-Zhu, Lu-KJ, Tjoa-AM, Wagner-RR, Al-Zobaidie-A. Performance modelling and metrics of database-backed Web sites. *Proceedings 11th International Workshop on Database and Expert Systems Applications*. IEEE Comput. Soc, Los Alamitos, CA, USA; 2000; 494-8

[11] MySQL web site: <http://www.mysql.com>

[12] Apache Software web site: <http://www.apache.org>

[13] PHP web site: <http://www.php.net>

[14] Soldar-G, Smith-D, Hamza-MH. Web based model for retrieval of scientific data semantics using DBMS as data repository. *Proceedings of the IASTED International Conference Applied Informatics International Symposium on Software Engineering, Databases, and Applications*. 2001 ACTA Press, Anaheim, CA, USA; 509-14

[15] Macfarlane P, Biersack M, Mohammad S, Jenkins J. Electrocardiogram-based separation of boys from girls. *Eur Heart*, 2003;24 (Abs Suppl):512.

Address for correspondence:

Professor Janice M Jenkins
Dept. of Electrical Engineering and Computer Science,
and Dept. of Biomedical Engineering
Medical Computing Laboratory
University of Michigan
Ann Arbor MI 48109-2122 USA
jenkins@umich.edu