

실시간 헬스케어 시스템을 위한 데이터 스트림 서버의 설계 및 구현

오택균*, 이연*, 배해영*

The Design and Implementation of the Real-time Data Stream Server for Continuity of Care Record

Zejun Wu*, Yan Li*, Hae-Young Bae*

요약

스마트폰 기반 개인 전자의료기록 관리 서비스는 병원과 의사와 환자가 병력을 공유하고 조회하여 환자의 병력을 효과적으로 관리하여 진단 효율을 높이는 서비스이다. 그러나 이런 전자의료기록 관리 서비스에서 발생하는 데이터는 도시와 같은 큰 범위의 모든 환자들의 병력 데이터를 전통 데이터베이스 관리시스템(DBMS)을 사용해 처리할 경우 빈번하게 발생하는 질의 처리 요청 등의 원인으로 병목현상이 발생 할 수 있어 실시간 데이터 스트림의 처리를 지원하는 데이터 스트림 관리시스템(Data Stream Management System)에 대한 연구가 진행되고 있다. 본 논문에서는 시스템 구조 및 연속 질의 처리기와 CCR 데이터 구조를 포함한 CCR 데이터 스트림 서버 시스템을 제안한다. 제안 시스템은 실시간적으로 발생하는 CCR 데이터 스트림을 처리 및 모니터링 하기 위한 시스템으로 사용자의 CCR 데이터 이력데이터에 대한 조회 및 실시간으로 발생하는 CCR 데이터에 대한 연속질의를 지원한다. 또한 개인 건강 정보는 지리적으로 떨어져 있는 헬스케어 프로바이더 사이에서 공유되어 환자의 병원이력정보를 관리한다. 또한 본 논문에서는 제안시스템에 기반한 아이폰 헬스케어 응용어플리케이션을 설계 및 개발하여 제안 시스템의 설계와 구축과정 및 성능을 보여준다.

▶ Keyword : CCR, 전자 의료 기록, 데이터 스트림 관리 시스템

Abstract

The EMR management services can monitoring the patients' record with any doctors in any hospital by using the internet and smartphones online. To handle the real time, multidimensional, continuous data, database management systems (DBMS) must cope with high insert rates for updates, however the traditional DBMS suffers from processing these kinds of data due to its serious design bottlenecks. So the researchers put forward to Data Stream Management System

• 제1저자 : 오택균 • 교신저자 : 이연 • 책임저자 : 배해영

• 투고일 : 2011. 10. 31, 심사일 : 2011. 11. 15, 게재확정일 : 2011. 11. 18.

* 인하대학교 컴퓨터정보공학과(Dept. of Computer & Information Engineering, Inha University)

※ 본 논문은 2011년 한국컴퓨터정보학회 하계학술대회 우수논문상을 받은 논문을 확장한 논문입니다.

※ This work was supported by INHA University.

(DSMS). In this paper we describe the real-time Data Stream Server for Continuity of Care Record (CCR) that including continuous query processor. This system is compiled with DSMS and DBMS in EMR system for processing and monitoring the coming CCR data stream, and also storing the processed result with high-efficiency. The system enables users not only to query stored CCR information from DBMS, but also to execute continue query on real-time CCR Data Stream, and health information can be transferred between different healthcare providers that would reduce medical error. At last, we develop a iPhone mobile application to test the proposed real-time data stream server.

▶ Keyword :CCR Data Stream Server, DSMS, EMR.

I. Introduction

The development of the smartphone and information management has enabled new applications, which include: Electronic medical record (EMR), intelligent transportation, environmental monitoring, telecommunications services, supply chain management, industrial production, etc. [1][2][3]. In these applications, the data which we called data stream is multidimensional, continuous, rapid, and time-varying, which should not only meet the needs of traditional ad-hoc query, but also continuous one for special.

As the EMR system for example, there is multiple of healthcare provider (HCPs) in a country or fixed area, and due to the health insurance coverage, cost and quality of care, the patients often be cared from different HCP [4][5], and at a period time like daytime there would be generated a large of Continuity of Care Record (CCR) data from a lot of devices and many kinds of operation to the CCR information, for example view, update, monitor, statistics or analysis CCR information data, even more these data contains geo-location information, according to these, it would be form a geospatial data stream to the server centre. However the traditional Database Management System (DBMS) suffers from processing the real time and complex application since it serious design bottlenecks [1]. As a result, researchers put forward to Data Stream Management System (DSMS) which focusing on

supporting continuous queries over massive data streams, also providing quality of service (QoS) guarantees. However, there are no components for process spatial data that include geo-location information and no result storage components in DSMS, while in practical analytical applications, such as CCR information monitor and transaction, it is necessary to analyze the area distribution and the process results for future. So it need add geo-location component to DSMS and combine with DBMS in EMR system that DSMS would process the geo-location data stream and DBMS would store the processed results. For these reasons, we design a system to processing the CCR data stream and transfer between different HCPs with high-efficiency.

The rest of this paper is organized as follows. Section II describes the related work to the EMR system. In Section III, we will discuss the system architecture, stream data query management and workflow. In Section IV, we implement the system followed by the related work and conclusion.

II. Related work

1. Continuity of Care Record (CCR)

It has been reported that between 44,000 and 98,000 patients in the world die from medical errors annually and many errors and adverse incidents occur as a result of poor quality data and information [6][7]. For the importance of the correct and continuity of care record [8][9], the Healthcare

Informatics committee of the American Standard for Testing and Materials (ASTM) standards development organization developed the Continuity of Care Record (CCR) standard (Figure 1) in order to improve continuity of patient care, to reduce medical errors, and to assure at least a minimum standard of health information transportability [10][11]. The CCR was designed to provide a snapshot of treatment and basic patient information and ease the transition of a patient from one HCP to the next cause. The CCR is a core data set of the demographic, medication lists, clinical and insurance information, medication lists and recent medical procedures that can be useful or possibly lifesaving, due to it expressed in the standard data interchange language known as XML, a CCR can potentially be created, read and interpreted by any EHR or EMR software application like Google Health (GH) and Microsoft Healthvault (HV)[12][13].

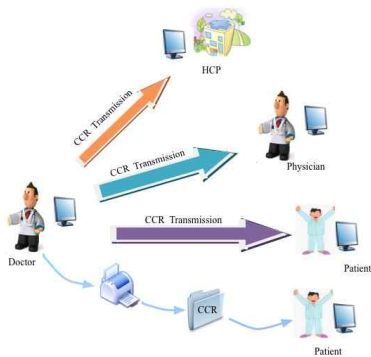


Figure 1: the transmission of the CCR Standard

And in 2009, the President Obama of USA launched an ambitious program to modernize the American healthcare system by making all health records standardized and electronic. The goal is to have all paper medical records—for the entire U.S. population digitize and available online. This way, each emergency room will be immediate access to a patient’s medical history, the effectiveness of medicines can be researched over large populations, and general practitioners and specialists can

coordinate their treatments. After two years the Center for Medicare and Medicaid Services (CMS) announced that the online Attestation System for the Medicare EHR Incentive Program had been launched on April 18, 2011. Eligible professionals and eligible hospitals will be able to use this online portal to self-attest to meeting the Meaningful Use criteria.

In our CCR Data Stream Server we combine the DSMS and DBMS for processing the CCR data information. According to the intergrade system, we can get the many benefits from the system like:

- The time to update the patient’s information and ask symptom will be minimized because all the information is stored in system and we can update these information on time;
- The next HCP will not need to search for or guess about the patient’s allergies and can be informed about the patient’s most recent healthcare and know who last treated the patient and what had done and cause of this the medical error will reduced;
- Some health care costs will not be needed due to the basic information and adverse reaction can be queried from system.

2. Data Stream Management System

The data stream model is the big difference between a traditional DBMS and a DSMS. Instead of processing a static query over a set of data that stored in advance on the disk, the DSMS process a continue query over the data stream which data elements arrive real-time and stay only for a limited time period in memory. So, the DSMS has to handle the data elements before the buffer is overwritten by new incoming data elements, and it cannot be retrieved again without storing it explicitly after the data elements had been processed. Since the coming data streams may not end, the continuous query results are often generated over a predefined window, and the window techniques are especially important for aggregation and join queries. Examples some major research prototypes system

include STREAM [14], Aurora[2], TelegraphCQ[15], Gigascope[16], NiagaraCQ[17], Coral8[18], Sensor Edge Server[21] and so.

STREAM built by Stanford is an all-purpose relation-based system with an emphasis on memory management. Aurora is a workflow-oriented system that allows users to build query plans by arranging boxes (operators) and arrows (data flow among operators) which built jointly at Brown University, Brandeis University, and MIT. TelegraphCQ built by UC Berkeley is a continuous query processing system that focuses on shared query adaptive and meeting the challenges that arise in handling large streams of continuous queries over high-volume data stream. Gigascope is a distributed network monitoring architecture that proposes pushing some query operators to the sources. NiagaraCQ, focuses in process a large volume of XML data streams transferred from Internet. Corral8 enables rapid application development and deployment of robust applications that derive insight from streaming event data, empowering instant responses to changing conditions. And the Sensor Edge Server developed by Oracle, which provides sensor streams process interface for several end-point applications and provides optional functions to maintain query results using traditional DBMS.

III. CCR Data Stream System Design

A CCR Data Stream Server is designed to manage the large CCR data stream from multiple clients with many kinds of operation. The system not only enable user to query the stored data from database, but also enable user to execute continue query from real-time CCR Data Stream and get the period query result with the help of DSMS.

1. Overview

The following figure shows the overview of the system. As shows in figure the clinical and CCR

data management is the main job of the system. The patients and HCPs can query data from system for the next diagnosis, and both of them can update, delete, export the CCR information to the file like PDF, XML, and HTML and then transfer between different HCPs since the data is according to the international standard. For the CCR data stream monitor, users can get the area distribution change information from the continue query result.

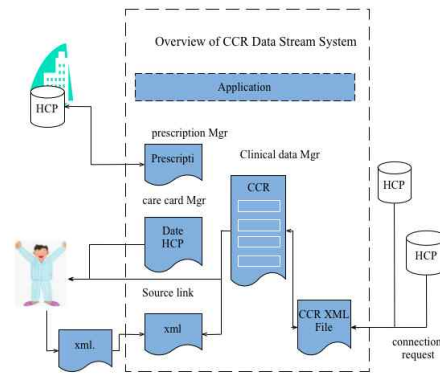


Figure 2: the main function of CCR Data Stream System

2. System architecture

The figure 3 shows the detail architecture of the CCR Data Stream System. As shown in the figure, the system is divided to three layers: the data access layer, business layer and data presentation layer.

2.1 Data Access Layer

This layer contains three modules for processing the coming CCR Data Stream: Data Adapter, DSMS and DBMS.

- Adapter is used to get CCR Data Stream and analyse the spatial information;
- The DSMS will process CCR Data Stream with high-effect and provide Continue Query;
- The DBMS will store the result from DSMS for the next time query or export.

2.2 Business Layer

This component contains three function module

(Filter Configure, Query Modify, and Dispatch Unit) and Query Execute Information (QEI).

- Filter Configure will receive the filter condition from Listener component and then send to DSMS in order to filter data stream;
- Query Modify is used to modify query information that in QEI;
- Dispatch Unit is a query schedule unit, it will decide the query information for DSMS and DBMS by schedule the query request;
- QEI is used to store the query information, and the data structure is called Query Data Structure;

2.3 Data Presentation Layer

In the presentation layer, the Listener part provide the GUI user interface for the patients, doctor, HCPs that include Receive Query and Return Result two parts module.

- Receive Query Module is designed to accept two types query, one type query is the from UI menu operation that include filter condition, another is one-time query that user self-defined. When the system received the query request, this module will change the request context, request time, and user information to the system executable query and then send to Query Manager.
- Return Module is designed to return data process result from DSMS and DBMS.

3. CCR Data Stream Query Processor

According to data stream characteristics, we use the Continuous Query Language (CQL) for data stream query in DSMS [3][14][16], CQL uses the similar SQL syntax to describe the events and events reflected processing rules. And continuous semantics was defined that “The result of a continuous query is the set of data that would be returned if the query were executed at every instant”.

For a lot of external events and internal objects are in the memory, the DSMS process the CCR Data Stream in memory, and after get results, we just store the interested output data to DBMS for future query or display. Our system’s CQL provides two inquire types query, one-time and continuous query. The one-time query just do once time over the unbounded data stream using limited resources, but the continuous query adapts the similar rules of the SQL engine, as long as the events or objects had been changed, it will execute inquires and the object is invoked to the corresponding call-back functions. The figure 4 illustrates the query processor model of the system.

In the query model a set of operator trees consist the query plan for continue query. And we provide some primitive operators that named filter, map, union, windowed join and windowed aggregate and spatial query operator. The windowed operators provide optional timeout and slack parameters that enable the system to deal with bounded disorder in the form of slow or out-of-order arrivals of data tuples, respectively. The spatial query operators provide the location query like find patients who are inside the HCP-A (Inside(h)), who are outside the HCP-A (Outside(h)), who stay in the HCP-A during t time (StayDuring(h,t)).

For example we can use the following CQL to get the patients’ information that stay the HCP-A during the 60 minutes.

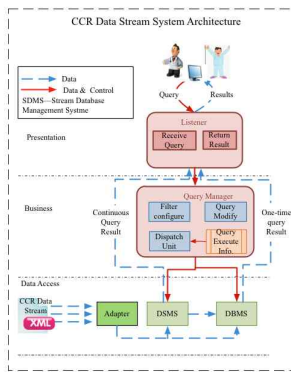


Figure 3: CCR Data Stream System Architecture

```

Select StayDuring('HCP-A',60)
From CCRDataStream1 as S
Where S.hop = 'HCP-A'
Begin (Now - 60)
End (Now)

```

The continuous query engine compiles a continuous query into a query plan and registers it with the scheduler for processing select-project-join queries and simple aggregate queries. The rewriting technique is used to deal with reordering a sequence of binary joins in order to minimize a particular cost metric [21]. At a high level, query optimizer work by enumerating a set of possible plans, assigning a cost to each plan based on estimated costs of each of the operators, and choosing the lowest-cost plan.

3.1 Query transformer

A parsed query represented by a set of query blocks that are nested or interrelated to each other input to the query transformer. The form of the query determines how the query blocks are interrelated to each other. The main objective of the query transformer is to determine if it is advantageous to change the form of the query so that it enables generation of a better query plan.

3.2 Estimator

The estimator includes three different types of measures: Selectivity, Cardinality and Cost. These measures are related to each other, and one is derived from another. The end goal of the estimator is to estimate the overall cost of a given plan. If statistics are available, then the estimator uses them from dictionary to compute the measures.

3.3 Plan generator

The main object of the plan generator is to try out different possible plans for a given query and pick the one that has the lowest cost. Many different plans are possible because of the various combinations of different access paths, join methods, and join orders that can be used to access and process data in different ways and produce the same result.

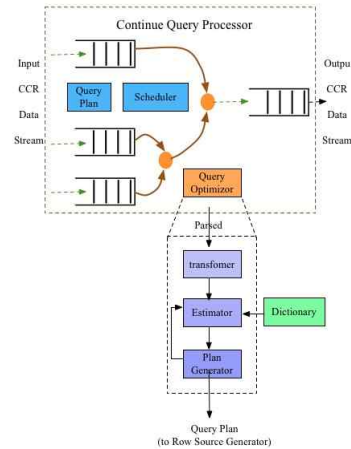


Figure 4: Continue Query Processor

4. Workflow

There are various steps are performed between CCR Data Stream System and User Client as shown in figure5. Step 1: the patient turns on the CCR Application when he/she gets disease or needs some test; Step 2: the doctor or HCP will select the patient from patient list that from CCR Data Stream System; Step 3: for example in the scenario the patient 1 is selected; Step 4: the system will send a permission request to the patient 1 for the next action; Step 5: if the patient accept the accessing permission, client application will create a connection session with the Server and the doctor's CCR application, else ignore the requirement; Step 6: when the session is created the doctor can view the patient's detail CCR information as a reference to next diagnosis in order to reduce medical error; Step 7.1: when the diagnosis is completed, the patient's CCR information is updated in the CCR Data Stream System; Step 7.2: it also will update the EMR (Electronic Medical Record); Step 7.3: finally this update will show in the patient's CCR application. So according the above steps, all of the participant can update the CCR information that contributes to avoid discontinuous records information.

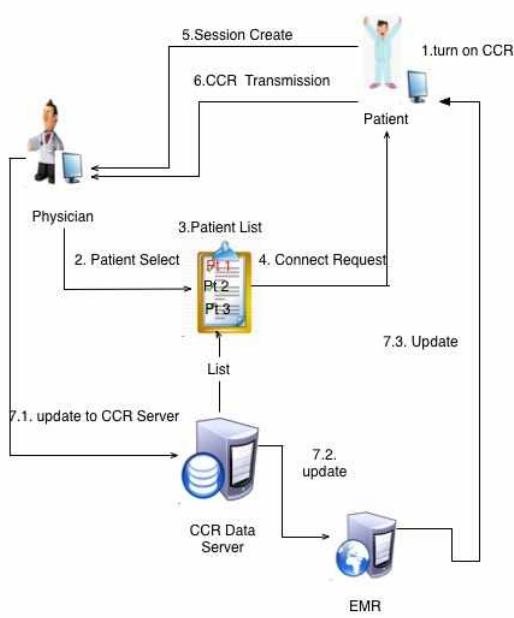


Figure 5: an example workflow of the CCR Data Stream System

5 Data Management Architecture

In the CCR Data Stream System side, we use the following format to store the processed coming data source in DBMS.

- 1) The Header section contains some “from” information that referring which patient, some “to” information that referring which HCP, CCR version information and document created date time;
- 2) The identifying section contains some information like patient’s name, ID card number, blood type, phone number, address, short description and so on that can be identify and distinguish from others patients;
- 3) The insurance section contains some information like insurance number, name, period of validity, contact method, description that patient subscribed;
- 4) The health section contains the following detail subsection that make up the main data set of the CCR. As the Advance directives for example.

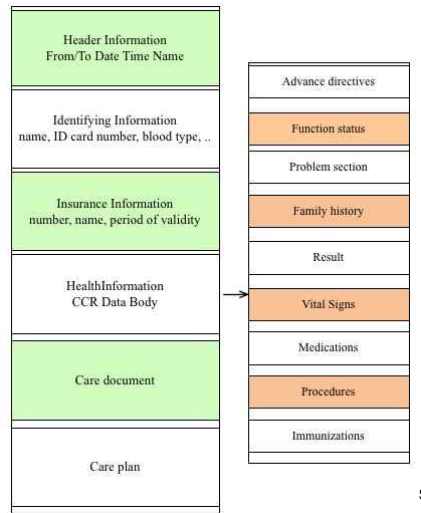


Figure 6: the format of stored CCR data in DBMS

- Advance directives and Alerts lists some health suggestion and also alerts that something can’t eat or drink or allergic. The field is described by element Id, but the value is described by text, since every alert have the different fields. And the table field also contains the initial date time, last update date time, tuple Id, and tuple status. The other subsection’s table field is also like this.
- 5) The care document section contains some special information like the disease management, institution, attachments and so on.
- 6) The care plan section contains some plan or schedule about immunization, procedures or test.

IV. Implementation

According to the Nielsen Company’s report, now the account of smartphones is more than 45.8% for the past 2010 in the worldwide and it also shows continuing growth. People like use the smartphone to record their lives, make plan, share the latest news, events, thinking and so on. So an application complaint with CCR standard requirement was designed and developed in the mobile device for

connecting to the CCR Data Stream System for viewing and updating CCR information.

Users are enabled to view, add or modify their personal health records, and export to some common human readable file type like PDF, HTML with entry level based on CCR template module at any time. Even more, users can do continue GEO query to get location information.

1. Application Interface

A base on http connection API framework is designed to create Space, Project, Stream, Schema, Query, and Query Plan etc. to the Server as shown in Figure 7.

For example we define a patient's CCR data stream, first create a schema, then covert the schema to a data stream use setSchemaToStream API.

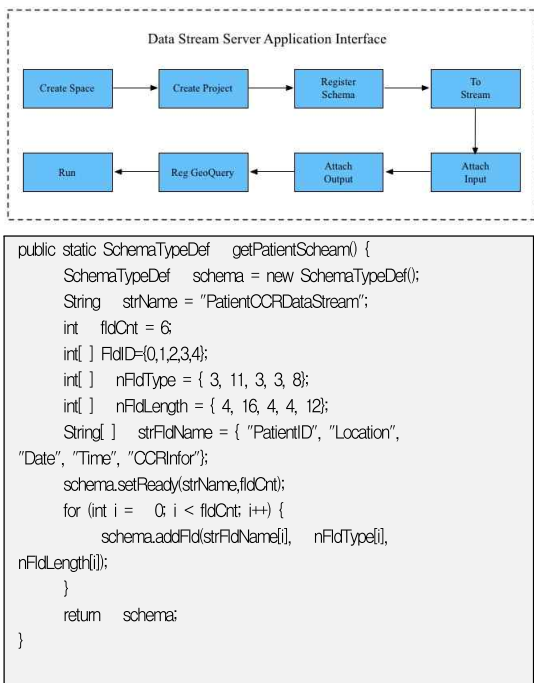


Figure 7 User Application Interface

2. Database Design

According the above data management architecture how to manage the CCR data set in high-efficiency is becoming the key point of the

system. After server times discussion finally we design all the CCR information tables have same fields, but connect to different data element and different value, so we can use the same table structure but present different information. As the advance directives CCR information for example, the relationship between other tables is shown in the figure 8. The advance directives table has 8 fields, the AdvanceDirective_ID is the primary key for the table; PHR_ID is a foreign key for identify from other patients; Data_Element_Id refer to the data element attribute tables that contains filed name described by English and the data type; the Value field contains the detail description text; the Initial_Update and Last_Update field means the create date time and update date time of the current tuple; the Status_Flag is a enumeration type that just include 1 and 0, 1 means the tuple's status is activate and 0 means the tuple's status is deactivate; finally since the rows of a advance advice is variable, so we need a Tuple_ID to distinguish from other advance advice.

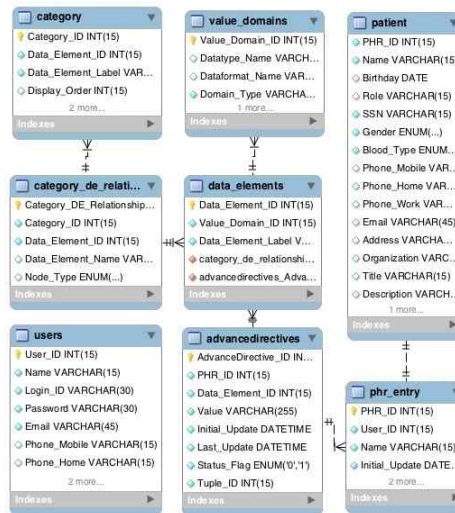


Figure 8: an example of the relationship between tables

According to the above database design, we can get CCR information of the Medications use the following SQL.


```

SELECT cdr.Data_Element_Name as label, cat.Value
FROM category_de_relationships as cdr
left JOIN Medications as cat
ON cdr.Data_Element_ID = cat.Data_Element_ID
AND cat.PHR_ID = 1
WHERE cdr.Category_ID = (select Category_ID
from category
where Table_Name='Medications' )
AND (cdr.Node_Type = 'Atomic'
or cdr.Node_Type = 'Choice')
AND cdr.Data_Element_Korean_Name != 'NULL'
AND cat.Tuple_ID = 1000;

```

After execute the query, database will return one advice CCR data for the patient 1 and the Tuple_ID is 1000 as shown in figure 9.

3. User Interface

The CCR Viewer application provides an easy but convince user interface for displaying, adding, modifying, deleting the CCR information. The main menu (figure 9.a) contains the CCR standard body information like advance directives, alerts, function status, problem, family and social history, results, vital signs, medications, procedures, immunizations, encounters, and all the body section is designed like the medications part (figure 9.b) which allow patients to update their CCR information by convince navigation and easy operation.



Figure 9: a). the main menu b). an example of the CCR Data Display

V. Conclusions

This paper discusses the CCR Data Stream System's design and implementation, the system is designed to manage the large CCR data from multiple clients at the same period, to integrate the different clinical records in one system in order to implement centralization of management, and to transfer CCR information between different HCPs, that will help minimize the time of updating information and asking symptom, also reducing the medical cost since avoiding some reduplicate test or operation would be reduced. The system combine the DSMS and DBMS to process the coming CCR Data Stream that include Geo Information, it not only enable user to query the stored data from database, but also enable user to continue query the real-time CCR Data Stream and get the period query result. With the help of CCR Viewer Application patients can view or update their personal health records, and export the CCR information to human readable file like PDF, Doc, and Html. This application also enables patients or HCP to compare records with standard records in order to monitor the changes of health and this point will be more deeply studied and researched.

Although it can bring a lot of benefits from using the system, it just allows the user who registered in the system to use the CCR application. The system design has no mechanism to ensure that patients and/or HCPs do not conceal information during data acquisition process. However even with the CCR international standard used in different HCPs, there are legitimate reasons why HCPs would not want patient's CCR information shared with different HCPs or patients. Nowadays, we just support mobile device accessing to system to update their information; future study and design will need to overcome the shortcoming.

References

- [1] Hetal Thakkar, Nikolay Laptev, "SMMF: a Data Stream Management System for Knowledge Discovery", ACM SIGMOD Record, 2011.
- [2] Majeed, Mahmood, Ubaid, "A burst resolution technique for data streams management in the real-time data warehouse", Emerging Technologies (ICET), 2011 7th International, P1-5, 2011.
- [3] Gert Brettlecker, "Reliable distributed data stream management in mobile environments", Web Information Systems Engineering, Volume 36, May 2011, Pages 618-643.
- [4] E. Reilent, I. Lööbas, "A. Kuusik and P. Ross: Improving the Data Compatibility of PHR and Telecare Solutions", IFMBE Proceedings 2011, Volume 37, Part 1, Part 9, 925-928.
- [5] Kyuchang Kang, Seonguk Heo, Changseok Bae and Dongwon Han: "Mobile Health Screening Form Based on Personal Lifelogs and Health Records", Lecture Notes in Electrical Engineering, 2011, Volume 107, Part 6, 557-564.
- [6] Stephen Kayb, Thomas Marley, Nicholas R. Hardiker, James J. Cimino: "Incorporating personalized gene sequence variants health knowledge into an EHR prototype based on the Continuity of Care Record standard", Journal of Biomedical Informatics, 2011.
- [7] WALTER V. SUJANSKY, MD, PHD, "The Development of a Highly Constrained Health Level 7 Implementation Guide to Facilitate Electronic Laboratory Reporting to Ambulatory Electronic Health Record Systems", Journal of the American Medical Informatics Association Vol. 16 No. 3, pp285-290, 2009.
- [8] David T. Liss, Jessica Chubak, "Patient-Reported Care Coordination: Associations With Primary Care Continuity and Specialty Care Use", Ann Fam Med July 1, 2011 vol. 9 no. 4 323-329.
- [9] Benjamin Schooley, "Improving IT Enabled Continuity of Care Across Pre-Hospital and Hospital Settings", Proceedings of the Sixteenth Americas Conference on Information Systems, Lima, Peru, August 12-15, 2010.
- [10] Health Level Seven Document Continuity of Care Document, <http://www.hl7.org/documentcenter/public/pressreleases/20070212.pdf>
- [11] ASTM E2369 - 05e1 Standard Specification for Continuity of Care Record (CCR), <http://www.astm.org/Standards/E2369.htm>
- [12] Google Health, <https://health.google.com>
- [13] Microsoft Healthvault, www.healthvault.com/
- [14] Arasu, A. and Babcock, B. and Babu, "STREAM: The Stanford Data Stream Management System", <http://infolab.stanford.edu/stream/2004>.
- [15] Sirish Chandrasekaran, "Telegraphcq: Continuous dataflow processing for an uncertain world", In CIDR, 2003.
- [16] Cranor, C., Johnson, T., & Spatscheck, "The Gigascope stream database", Data Engineering, 2003.
- [17] Jianjun Chen, "NiagaraCQ: A scalable continuous query system for internet databases", SIGMOD Rec. 29, May 2000.
- [18] Streambase, <http://www.streambase.com/>, 2005.
- [19] Coral8 Inc, <http://coral8.com/>.
- [20] Oracle Inc., http://download.oracle.com/docs/cd/B14099_19/wireless.1012/b13819/rfid.htm
- [21] Lukasz Golab and M. Tamer, "Issues in data stream management", ACM SIGMOD Record, 32(2), 5-14. ACM

저자 소개



Zejun Wu

2003 : Chongqing University of Posts and Telecommunications BEng degree

2010 : Inha University, information engineering department, Master course

Research Interests : Spatial database based on NoSql Cloud

Email : wuzejun3q@gmail.com



Yan Li

2002~2006 : Chongqing University
of Posts and Telecommunications, BEng
degree

2006~2008 : Inha University,
information engineering department,
Master degree

2008년~now : Inha University,
information engineering department,
Phd. course

Research Interests : Spatial database,
spatial datawarehouse, GIS, USN,
stream database system

Email : leeyeon@dblab.inha.ac.kr



Hae-Young Bae

1976: INHA University, Physical
Engineering, bachelor
degree

1979: Yonsei University, Computer
Science and Engineering,
master degree

1990: Soongsil University,
Computer Science and
Engineering, Ph.D

1982~now: worked as a professor
of Inha University and the
director of Intelligent GIS
Research Center.

2000~now: the Emeritus Professor
of Chongqing University of
Posts and Telecommunication,
China.

Research Interests : Database,
Spatial
database.

Email : hybae@inha.ac.kr