

Design and Implementation of SQL Audit Tool for Database Performance

Chen Liu[†] · Taewoo Kim^{**} · Baowei Zheng^{***} · Jeongmo Yeo^{****}

ABSTRACT

Information system audit is the requirement to configure the information system successfully. However, there is a problem with reduction of efficiency in the audit work when the system capacity is growing up. In the subsidiary field of information system audit, there is the same problem. In this paper, we will focus on the database audit, and implement an SQL audit tool in order to improve the performance of a database. During the designing phase of the SQL audit tool which aim to improve the performance of a database, we have analyzed the requirements of SQL audit work. Based on the aforementioned them, the process of the SQL audit tool has been designed with SQL audit features in accordance with information audit process. During the implementation phase, we have implemented 3 main function modules according to the output of the design phase. The main modules we implemented are the audit job definition module, audit job executing module and result reporting module. With the implemented tool, it applied to an Electric Power Corporation project in China and compared with other tools that are able to use SQL audit. In this paper, the implemented SQL audit tool is able to perform the general SQL audit work and to reduce the cost of the audit work for database performance and to raise the accuracy of result and to apply the extended inspection rule which need to use SQL parsing.

Keywords : SQL Audit, Audit Tool, Database Performance, SQL Audit Tool, Database Audit

데이터베이스 성능을 위한 SQL 감리 도구의 설계 및 구현

Chen Liu[†] · 김 태 우^{**} · Baowei Zheng^{***} · 여 정 모^{****}

요 약

정보시스템 감리는 정보시스템을 성공적으로 구축하기 위한 요건이지만 정보시스템의 규모가 커질수록 감리작업 수행의 효율성이 저하되는 문제가 있다. 정보시스템 하위 감리 영역에도 동일한 문제가 존재하는데 본 논문에서는 데이터베이스 성능을 위한 SQL 감리 도구를 구현하여 데이터베이스 감리의 효율성 문제를 다루고자 한다. 본 논문에서 데이터베이스 성능을 위한 SQL 감리 도구를 설계하는 과정에서 SQL 감리작업에 필요한 요구사항들을 분석한다. 이를 바탕으로 SQL 감리 도구의 프로세스를 정보시스템 감리 절차에 따라 SQL 감리의 특성에 맞추어서 설계한다. 이후 데이터베이스 설계 방법론을 적용하여 SQL 감리 수행과정의 모든 정보를 저장하는 통합 레퍼지토리로도 설계하였다. 구현하는 과정에서는 설계된 내용에 따라 감리작업 정의모듈, 감리작업 수행 모듈, 감리작업 결과 보고 모듈로 나누어 구현하였다. 구현한 도구를 활용하여 중국 모 전력 공사 프로젝트에 적용하여 보았고, SQL 감리에 활용할 수 있는 도구들과 비교해 보았다. 본 논문에서 구현한 SQL 감리 도구는 전반적인 SQL 감리 작업을 수행 할 수 있으며 데이터베이스 성능을 위한 감리 작업에 소요되는 비용을 줄이고 결과의 정확성을 높일 수 있을 뿐만아니라 SQL 파싱을 통해 검사규칙을 확장하여 적용할 수 있다.

키워드 : SQL 감리, 감리 도구, 데이터베이스 성능, SQL 감리 도구, 데이터베이스 감리

1. 서 론

세계화된 정보사회에서는 IT의 중요성이 점차 증대되고 있다. 정보와 정보 기술을 효율적으로 관리하는 것은 조직

의 성공과 생존을 좌우할 수 있는 핵심요소가 되었다. 성공적인 기업이 되기 위해서는 정보시스템에 관련된 모든 것을 강력하게 통제를 할 수 있어야 한다[1]. 그렇기 때문에 2007년 이후 공공부문 정보화 사업뿐만 아니라 규모가 큰 기업에서도 정보시스템 감리에 투자하는 비중이 커지고 있다. 감리작업의 중요성이 대두되고 있지만 기업에서는 업무시스템의 규모가 커지면서 새로 오픈하는 많은 시스템들이 다양한 문제점을 안고 오픈하게 되는 경우가 많다. 이러한 시스템들은 운영 단계에서 문제를 해결하여 성능을 향상시킬 수 있지만 이러한 시스템은 이미 악영향을 미쳤을 뿐만아니라

※ 이 논문은 부경대학교 자율창의학술연구비(2015년)에 의하여 연구되었음.
† 준 회 원 : 부경대학교 컴퓨터공학과 박사과정
** 비 회 원 : 부경대학교 컴퓨터공학과 석사과정
*** 비 회 원 : 부경대학교 정보공학과 공학박사
**** 정 회 원 : 부경대학교 컴퓨터공학과 교수
Manuscript Received : January 4, 2016
First Revision : April 4, 2016
Accepted : April 4, 2016
* Corresponding Author : Jeongmo Yeo(yeo@pknu.ac.kr)

추가적으로 더 많은 비용을 투자해야 한다. 이러한 이유에도 불구하고 감리수행에 소요되는 시간이 많이 걸리기 때문에 시스템을 사용하기 전에 감리 작업을 생략한다.

감리 작업에는 많은 제약사항으로 인해 시간이 많이 소요된다. 본 논문에서는 감리의 분류 중 데이터베이스 성능을 위한 감리를 다루며 그 중 데이터베이스 성능에 가장 많은 영향을 미치는 SQL에 대해 다루고자 한다. 현재 데이터베이스 성능을 위한 SQL 검사 작업에는 많은 제약사항이 있다. 첫째, SQL에 대한 검사 작업은 감리원의 많은 경험과 전문적인 지식, 그리고 대상 업무의 특성을 고려한 분석/설계 등 상황에 대한 이해가 바탕이 되어야 한다[2, 3]. 둘째, 기존에 있는 도구들은 검사를 위한 자료 수집 단계에서는 활용되고 있지만 SQL 검사단계에서 적용하는데는 한계가 있다. 그렇기 때문에 검사하는 작업은 대부분 Eye-Check 방식으로 수행하고 있다[2]. 하지만 이 방식은 검사하는 SQL의 양이 증가할 수록 비용이 많이 소요되고 정확성도 보장하지 못한다. 이러한 문제점을 해결하기 위해 본 논문에서 데이터베이스 성능을 위한 SQL 검사도구를 설계 및 구현하고자 한다.

2. 관련 연구

2.1 데이터베이스 성능

데이터베이스 성능이란 사용자가 느낄 수 있는 응답 시간 및 시스템 처리 능력을 말한다[4].

데이터베이스 성능관리는 시스템의 효율 및 응답속도를 최적의 상태로 유지 및 제공하기 위하여, 낮은 성능 현상을 보이는 요소를 찾아 성능 개선을 수행하거나, 성능 분석을 통해 문제점을 발견하여 개선하는 것이다[5]. 그와 관련하여 데이터베이스 성능 영향 요인과 관련된 연구에서 시스템 성능 저하의 요인은 Application & DB Design(70%), DB Config. Design(20%), System Design(15%), H/W Resources(5%)으로 나타났다[4].

데이터베이스 액세스 효율 향상을 위한 핵심요건에는 SQL 최적화, 인덱스 최적화, 실행계획 최적화, 직렬/병렬처리 최적화, 연결 방식 최적화 그리고 부분범위 처리 방식이 있다. 이 중 본 논문에서는 SQL 최적화 방식에 대해 다루고자 한다.

2.2 감리 검사항목

데이터베이스 성능을 위한 감리 검사 항목은 데이터베이스 성능에 영향을 줄 수 있는 요인을 중심으로 정해질 수 있다. 이러한 성능 영향 요소를 분석함으로써 검사에 필요한 항목을 도출할 수 있다. 기존에 많은 자료 및 연구에서도 제시한 바가 있고 실무에서의 시스템 특성에 따른 조사를 통해서도 추가로 도출할 수도 있다.

1) 감리에 대한 표준 가이드

정보시스템 감리 수행 가이드[6]에서 정보화 사업 유형별

표준 점검항목 부분을 포함하고 있다. 이것은 정보기술아키텍처구축, 정보화전략계획 수립, 정보시스템 개발, 데이터베이스 구축, 정보시스템 운영 및 유지보수사업 등에 대한 감리 시 점검할 것을 권장하는 사항들을 망라한 것이다[6]. 가이드에서 제시한 검사항목들을 살펴보면 모든 사업 유형에서 어느 정도 데이터베이스에 관련된 항목들이 포함되어 있다. 특히 데이터베이스 구축 사업 중 다양한 측면에서 데이터베이스 감리 검사항목을 제시하였다. 그렇지만 정보시스템 감리 수행 가이드에서는 검사항목의 유형만 기재되어 있어 적용시에는 상세화하여 사용할 필요가 있다.

2) 감리에 대한 연구 논문

데이터베이스 감리에 관한 연구들이 많이 이루어져 있기 때문에 감리 수행 가이드에서 제시하는 검사항목을 기준으로 상세화하여 구체적으로 수행하는 방법을 제시하고 있다. 예를 들어, [7]는 데이터베이스 중에서 실행한 질의어에 대해 문제를 유발할 수 있거나 성능에 악영향을 줄 수 있는 질의어를 추출하는 측면에서 여러 검사 항목을 제시하였다. 또한 [8, 9]는 데이터베이스 안에 보관하고 있는 데이터의 품질을 위한 많은 감리 검사항목을 제시하였다.

3) 기술 전문 서적

전문 서적[4, 10-12] 중에서는 실무 경험을 기반으로 정리한 방법론이 포함되어 있다. 방법론에는 주로 데이터베이스를 정확히 활용할 수 있는 방법이 기재되어 있다. 그 중에는 데이터베이스 성능을 위한 감리방법도 포함되어 있다.

정보 시스템 감리 수행 가이드에서 많은 점검 항목들을 제시 하였지만 이런 항목 들은 감리의 목적, 그리고 해당 사업의 규모나 특성 등을 고려하여 선별적으로 선택해야 한다[6]. 따라서 기존에 제시한 점검항목뿐만 아니라 현재 진행 중인 프로젝트에서 기업의 요구사항들을 수집하여 추가적으로 도출할 수 있다.

2.3 감리 지원도구

최근 정보시스템은 업무, 정보기술, 관리 세가지 측면에서 대단위로 복잡하게 연계되어 차세대 시스템들과 같이 개발이 이루어지고 있다. 시스템의 규모가 커지면 감리 작업의 양도 증가함에 따라 감리 작업을 효율적으로 수행하여 결과의 정확성을 보장하기 위해 수동 작업을 대신할 수 있는 지원 도구에 대한 필요성이 대두되었다.

[2]에 의하여 감리 지원 자동화 도구의 종류로는 범용 감리소프트웨어, 고급언어(SQL), 유틸리티 소프트웨어 그리고 전문가 시스템이다. 데이터베이스 감리의 특성을 고려하여 더 세부적으로 분류하면 Table 1과 같은 지원 도구들이 있다.

그렇지만 이러한 도구는 데이터베이스 감리 작업의 특성에 맞추어 만들어진 것이 아니라 데이터베이스 설계, 구축 및 운영 단계에서 특정한 작업을 위한 지원도구로써 만들어진 것이다. 그러므로 자료 수집 단계에서는 활용할 수 있지만 검사 작업에는 부적합하다.

Table 1. The auxiliary tools for DB Audit

Software	Description
ER-Win DA#	Data Model design tool
SQL Expert Oracle AWR	SQL monitoring tool
Orange for Oracle Toad IDO# PL/SQL Developer	Development and performance tuning tool
IDO#	Index design and tuning tool
Meta#	Data governance tool

3. 데이터베이스 성능을 위한 SQL 감리 도구의 설계 및 구현

데이터베이스 성능을 위한 SQL 감리란 데이터베이스 성능에 영향을 미치는 요인 중 SQL에 대한 정보를 수집하여 정해진 규칙의 위배여부를 검사하는 과정으로 본 연구에서는 이러한 과정을 지원해주는 도구를 설계하고 구현하였다.

3.1 데이터베이스 성능을 위한 SQL 감리 도구의 설계

1) SQL 감리 도구를 위한 요구사항 분석

데이터베이스 성능을 향상시키기 위해 지정한 시스템의 SQL 문제를 발견하고 개선해야 한다. 이러한 작업은 시스템의 개발단계에만 수행하는 것이 아니라 정보시스템의 크기가 증대됨에 따라 새로운 모듈이 계속 추가되는 경우에도 주기적으로 시스템의 문제를 검사하고 그 결과를 기록해야 한다. 그렇기 때문에 대상 데이터베이스에서 SQL관련 정보를 수집하고, 수집한 정보를 검사하여 그 결과를 다차원 분석에 활용한다. 또한 스케줄링을 이용하여 주기적으로 수집과 검사를 수행하도록 하기 위해 하나의 통합 레파지토리를 구축하여 모든 정보를 관리할 것이다.

2) SQL 감리 도구를 위한 프로세스 설계

본 연구에서는 데이터베이스 성능을 위한 SQL 검사 도구를 구현하기 위하여 요구사항을 분석하였다. 이를 바탕으로 정보시스템 감리 프로세스에 기반하여 SQL 검사 특징을 맞추어 대상 선정, 수집, 검사, 결과 분석, 문제점 개선으로 나누었고, 본 연구에서는 대상 선정부터 검사까지 설계하였다.

SQL 감리 도구 프로세스 절차에 따라 심사 담당자가 작업 범위를 선정하고 해당 범주에 대상 데이터베이스를 등록한다. 그리고 수행할 작업을 정의하여 등록한다. 그 후, 수집과 검사 작업을 수행하며 그 결과를 레파지토리에 저장한다. 검사가 완료되면 심사 담당자가 검사 결과를 분석하여 적절한 조치를 취하고, 최종적으로 그 결과를 담당자에게 전달한다. 또한 수집 및 검사 작업은 수행 시간 및 주기를

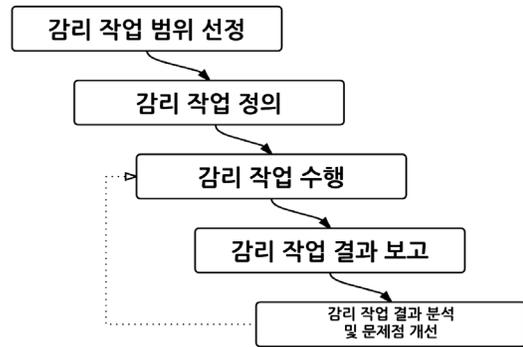


Fig. 1. The process of SQL audit

설정하여 지속적으로 심사 작업을 진행할 수 있다. 이러한 일련의 프로세스는 Fig. 1과 같다.

3) SQL 감리 도구를 위한 레파지토리 설계

앞에서 언급한 일련의 작업 과정의 모든 정보를 레파지토리에 저장한다. 저장되는 내용은 업무영역 및 작업 대상 데이터베이스 관련 정보, 대상 데이터베이스에서 수집한 데이터 사전 및 SQL관련 정보, 검사 및 통계 분석 결과 등이 있다. 이러한 정보를 포함한 모든 요구사항을 분석하여 데이터 모델링 과정[12]에 따라 논리 데이터 모델 설계, 데이터 표준화, 물리 데이터 모델 설계 과정을 수행하여 레파지토리를 구축하였다.

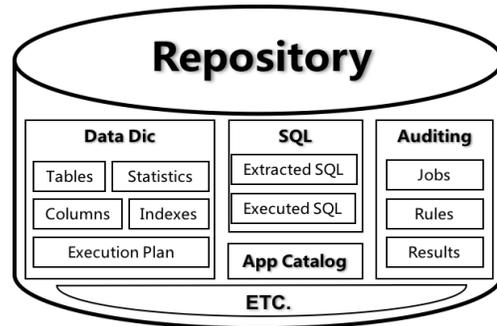


Fig. 2. The conceptual diagram of the repository

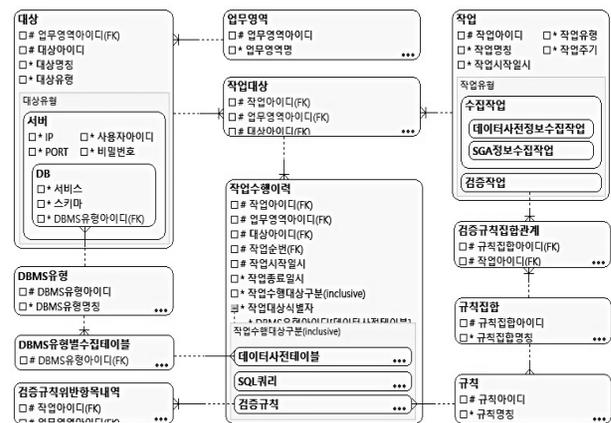


Fig. 3. Data Model 1: The part of main functions

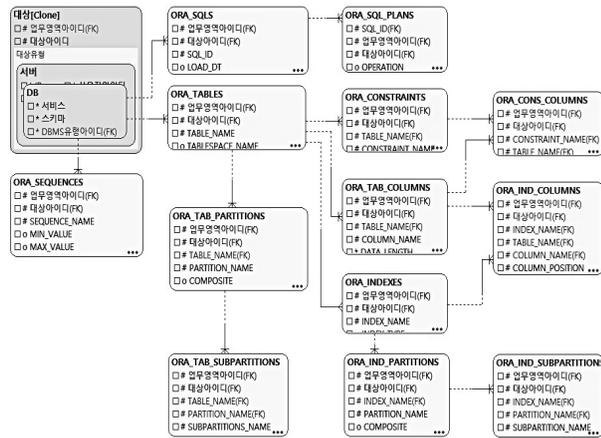


Fig. 4. Data Model 2: The part of collected information

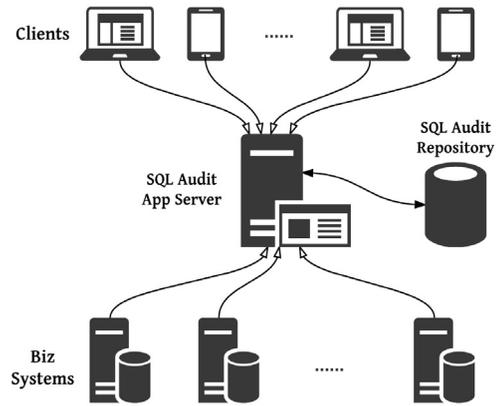


Fig. 6. The system architecture

3.2 데이터베이스 성능을 위한 SQL 감리 도구의 구현

1) SQL 감리 도구의 개발환경 및 프로그램 구성도

본 시스템은 Spring Framework[13] 기반으로 Fig. 5에서 나타난 라이브러리를 사용하여 개발하였다. 또한 SQL 쿼리를 파싱하기 위해서 ANTLR V4[14, 15]를 사용하였다.

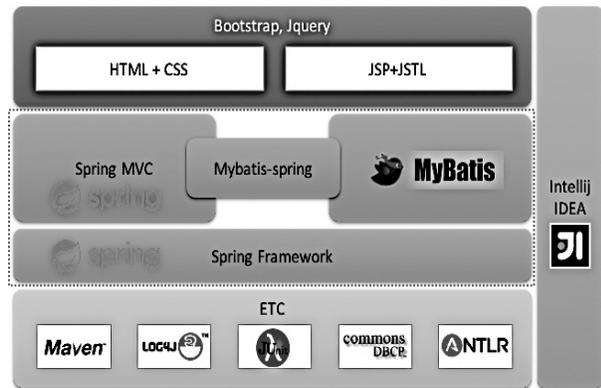


Fig. 5. The development environment

구현된 시스템의 서버 환경은 Table 2와 같다.

Table 2. The configuration of server

Hardware	CPU	Intel i5 4 Core
	RAM	8GB
	HDD	1TB
Software	Server	Apache Tomcat 8.0
	DB	Oracle 11g r2

본 연구에서 구현한 SQL 검사 도구는 지정된 서버에 설치하여 여러 개의 검사 대상 데이터베이스와 네트워크를 통해서 연결하고, 감리담당자는 클라이언트 브라우저를 통하여 서버에 접근한다. 서버는 데이터베이스와 연동하여 수집된 정보 및 검사 결과를 레파지토리에 저장한다.

2) SQL 감리 프로세스에 따른 기능 모듈 구현

데이터베이스 성능을 위한 SQL 감리 프로세스에 따라 Fig. 7과 같이 감리작업 정의모듈, 감리작업 수행모듈, 감리작업 결과 보고 모듈 세 부분으로 나눌 수 있다.

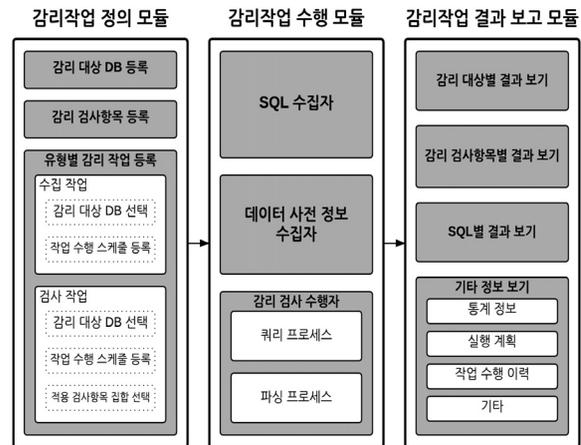


Fig. 7. The program architecture

첫째, 감리작업 정의모듈은 감리할 대상 SQL 및 근거 정보를 수집하여 검사하기 위해 대상 데이터베이스, SQL 감리 검사항목, SQL 감리 작업을 등록한다. SQL 감리 작업등록은 수집작업등록과 검사작업 등록으로 분류된다. 수집 작업 등록은 시스템에 등록된 데이터베이스중에 선택하고, 수집작업의 주기 및 시간을 정하여 등록한다. 검사작업 등록은 수집작업에 등록된 데이터베이스를 대상으로 검사에 적용할 검사항목을 선택하고 검사작업의 주기와 시간을 선정하여 등록한다.

둘째, 감리작업 수행 모듈은 크게 SQL관련 정보 수집과 검사로 나뉜다. SQL관련 정보는 데이터 사전과 SQL을 수집한다. 감리 검사 수행은 수집된 SQL을 쿼리 프로세스와 파싱 프로세스 방식으로 분류하여 수행한다.

셋째, 감리작업 결과 보고 모듈은 하나의 SQL 감리 작업의 결과를 다차원 분석이 가능하도록 감리 대상, 검사항목, SQL 그리고 기타정보 항목별로 분류하여 제공한다.

References

- [1] Hong-sup Jung, "Quality Improvement Plan through Analysis of Information System Audit Example," Master dissertation, Konkuk University, Seoul, KOREA, 2008.
- [2] Song-hae Kwak, "A Study on the Development of the Automation Tool for Improvement of Audit Productivity," Master dissertation, Konkuk University, Seoul, KOREA, 2009.
- [3] Seung-ryeol Jung and Jeon-guk Kim, "A Study on Collection and Evaluation of Audit Evidence," National Computerization Agency, 1999.
- [4] KoDB, "The Guide for Data Architecture Professional," KoDB, 2013.
- [5] Ministry of Information and Communication, "GuideLine for Performance Management of Information System."
- [6] National Computerization Agency, "The Guide for Information System Auditing," National Computerization Agency, 2013.
- [7] Jong-won Kim, "System Audit Improvement Through Identifying Database Query Audit Inspection Item," Master dissertation, Incheon National University, Incheon, KOREA, 2013.
- [8] Jae-Hwan Ko, Dong-soo Kim, and Ki-Joon Han, "Development of Automated Tools for Data Quality Diagnostics," *KITS*, Vol.11, No.4, pp.153-170, 2012.
- [9] Chang-gon Kim, "The Guide for Audit Inspection," National Computerization Agency, 2008.
- [10] Hwa-sik Lee, "New VLDB Solution," En-Core Consultant, 2005.
- [11] KoDB, "The Guide for SQL Professional," KoDB, 2013.
- [12] Hwa-sik Lee, Data Architecture solution, En-Core Consultant, 2003.
- [13] Rod Johnson, Juergen Hoeller, Spring Framework Reference Documentation [Internet], <http://docs.spring.io/spring/docs/current/spring-frameworkreference/htmlsingle/#overview-getting-started-with-spring>.
- [14] T. J. Parr and R. W. Quong, "ANTLR: A Predicated-LL(k) Parser Generator," *Software-practice and Experience*, Vol.25, No.7, pp.789-810, 1995.
- [15] T. J. PARR, "The Definitive ANTLR Reference: Building Domain-Specific languages," The Pragmatic Bookshelf, 2013.



Liu Chen

e-mail : liuchen@pukyong.ac.kr

2011년 부산외국어대학교 E-Business학과 (학사)

2013년 부경대학교 컴퓨터공학과(석사)

2013년~현 재 부경대학교 컴퓨터공학과 박사과정

관심분야: 데이터 아키텍처, 데이터 모델링, 데이터베이스 성능 튜닝, 데이터베이스 감리



김 태 우

e-mail : mtkim7895@pukyong.ac.kr

2015년 부경대학교 컴퓨터공학과(학사)

2015년~현 재 부경대학교 컴퓨터공학과 석사과정

관심분야: 데이터 모델링, 데이터베이스 성능 튜닝, 데이터베이스 감리



Baowei Zheng

e-mail : zhengbw@en-core.com.cn

2004년 중국 서안전자과학기술대학교 컴퓨터학과(학사)

2005년 위덕대학교 컴퓨터멀티미디어공학과 (공학석사)

2011년 부경대학교 정보공학과(공학박사)

2011년~현 재 (주)엔코아 차이나 법인장

관심분야: 데이터 모델링, 데이터 품질, 데이터 거버넌스, 데이터 분석, 데이터베이스 성능 튜닝



여 정 모

e-mail : yeo@pknu.ac.kr

1980년 동아대학교 전자공학과(학사)

1982년 부산대학교 전자공학과(공학석사)

1993년 울산대학교 전자 및 전산기공학과 (공학박사)

1986년~현 재 부경대학교 컴퓨터공학과 교수, (주)엔코아 사외이사

관심분야: 데이터 아키텍처, ITA/EA