

지식기반 회의관리 시스템 아키텍처에 관한 연구

김태현* · 송종철* · 김창수** · 송정영* · 김진수* · 정희경*

*배재대학교 컴퓨터공학과 · **청운대학교 인터넷학과

A Study on Knowledge based Conference Management System Architecture

Tae-hyun Kim* · Jong-Chul Song* · Chang-Su Kim** ·

Jin-Soo Kim* · Jeong-Young Song* · Hoe-Kyung Jung*

*Dept. of Computer Engineering, Paichai University · **Dept. of Internet, Chungwoon University

E-mail : *{kth0316, jysong, jskim, hkjung}@mail.pcu.ac.kr, *jcsong@iita.re.kr, **ddoja@chungwoon.ac.kr

요 약

현재 지식 기반 시스템의 한 분야로 회의시스템이 있으며 멀티미디어적 접근방식의 회의에 대한 연구가 활발히 이루어지고 있다. 그러나 멀티미디어적인 관점에서 연구되는 회의시스템의 경우 기존의 오프라인 회의의 장단점을 그대로 온라인으로 가져왔을 뿐 정작 중요한 회의 내용 및 방식에 대한 지식화 방안에 대해 고려하지 않고 있다

이에 본 논문에서는 기존의 멀티미디어적인 측면에서 회의 시스템을 구성하는 방식을 탈피해 회의 내용과 방식에 초점을 맞추어 기존 시스템들을 포함할 수 있는 지식기반 회의관리 시스템 아키텍처를 제안하였다.

ABSTRACT

Currently, There is a conference system in one field of knowledge base system, the study against a conference of multimedia approach method is become accomplished actively. However, Case of the conference system which is studied from multimedia viewpoint, only it will bring the merit and demerit of the off-line conference of existing with on-line, it isn't considering to make knowledge plan of important conference content and method.

This thesis proposes stand of Knowledge-base system architecture for it break from the method which composes a conference system from the multimedia side of existing and focus it will be able to contain existing systems to conference contents and method

키워드

온톨로지, 회의시스템, 지식기반, XML

1. 서 론

정보화 사회가 되면서 인터넷의 발달에 힘입어 정보가 기하급수적으로 늘어났고 이러한 정보가 가치를 가지고 있는 시대가 되었다. 정보화시대의 정보취득은 쉽지만 수집된 정보관리방법이 취약함에 따라 정보를 지식화하고 관리하는 지식관리 시스템이 등장하였다.

현재까지의 지식관리 시스템은 효과적으로 지식 프로세스를 지원하고, 다양한 비정형 지식들을 통합하고 연동하는 방향으로 연구 및 발전되어 왔다. 그러나 모든 지식들을 지식화 하기에는 모

호성에 따른 복잡도 증가로 인해 지식 표준화에 한계를 갖고 있다.

따라서 모든 정보를 지식화 하기보다는 전문가 시스템과 같이 특화된 분야로 접근하여 좀 더 구체적으로 정보를 전문적인 지식으로 만드는 지식기반 시스템이 연구되고 있다.

이러한 지식기반 시스템의 한 분야로 회의 시스템이 있는데, 본 논문에서는 기존 멀티미디어적인 측면에서 회의 시스템을 구성하는 일반적인 접근 방식을 탈피하고 회의 내용과 진행 방식에 초점을 맞추어 기존 회의 시스템들을 그대로 흡수 할 수 있는 지식기반 회의 관리 시스템 아키텍처를 제안한다.

II. 관련연구

2.1 회의 시스템

현재 회의 시스템 연구는 크게 두 가지 방향으로 연구되고 있다. 첫 번째 방법으로 멀티미디어에 기반을 두어 현재의 일반적인 오프라인 형태의 회의를 온라인상에서 사용하는 것이고, 다른 하나는 분석 통계에 따른 정보를 기반으로 회의의를 규정하고 그 정보를 활용하는 것이다.

2.2 온톨로지(Ontology)

온톨로지는 도메인 내에서 공유되는 데이터들을 개념화한 형식적이고 명백한 규정이며 이는 특정분야에서 사용되는 표준 어휘들의 모음이라고 할 수 있다[1]. 온톨로지는 어휘 사전 역할 외에 지식을 효과적으로 표현하기 위해 정보의 의미를 부여하고, 정보간의 관계를 설정한다. 따라서 온톨로지의 목적은 광범위한 도메인에 적용이 가능하도록 표준을 제시함으로써 웹 문서의 지식을 표현, 공유, 재사용하는 것이다.

2.3 기존 회의관리 시스템의 문제점

기존 멀티미디어 기반의 온라인 회의 시스템의 경우 시간간의 제약 없이 회의가 가능하다는 장점이 있으나 실제적인 회의에 대한 지식화가 불가능하다는 단점을 가지고 있다.

그리고 일반적인 지식 관리 시스템들의 경우는 언어처리 및 지식 표준화의 어려움으로 지식처리에 지능화 부분을 적용하는 연구는 미비한 실정이며 주로 지식베이스(knowledge base)를 구축하려 하고 있으나 회의 지식을 지식화 하기 위해서는 전문가에 의해 지식 분류가 수작업으로 이루어지므로 관점에 따른 오분류와 함께 상당한 관리 비용이 필요하다[2].

또한, 일반적인 회의 시스템의 경우 공유를 위한 표현방법이 제한적이며 회의 내용이 순차적으로만 관리되어 회의내용에 대한 조회나 검색 시 해당 내용을 모두 검색해야 하는 문제점이 있다[3].

III. 지식기반 회의 시스템 아키텍처

그림 1은 기존의 회의 시스템들의 문제점들을 보완한 지식 기반 회의관리 시스템 아키텍처이다.

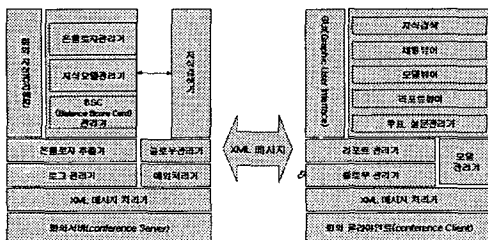


그림 4. 지식기반 회의 시스템 전체 아키텍처

전체적으로 회의 서버와 회의 클라이언트로 구성되며 서버와 클라이언트는 XML 메시지를 TCP/IP 소켓통신을 통해 교환한다.

3.1 회의 서버 아키텍처

회의 서버에서는 XML 메시지를 입력 받아 XML 메시지 처리기에서 해당 메시지의 내용을 파싱하여 메시지와 회의 정보 및 내용으로 분리한 후 정보들을 클라이언트 요청에 의해 전송한다.

로그 관리기는 서버에서 입력력되는 데이터에 대하여 접속, 에러, 정보 로그로 분리 기록한다.

온톨로지 추출기에는 특정한 카테고리 분류되는 문서들에 공통적으로 자주 나타나는 문장 형태와 키워드들을 온톨로지 패턴으로 만든다. 분류의 정확도는 온톨로지 패턴 매칭에 의해 문서를 분류하는 방법과 통계적인 문서 분류에 온톨로지 패턴들을 이용한 분류 방법을 접목시킴으로써 높이는 방법으로 제안하였다. 그림 2에 이를 나타낸다. 또한, 온톨로지 패턴 매칭은 빠른 처리속도를 보장함으로써 실시간 회의 시스템에 적합하다.

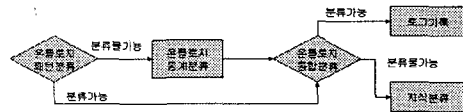


그림 5. 온톨로지 하이브리드 분류 방식

논문에서는 분류하려는 구문과 분류 대상 범주들을 온톨로지 벡터로 구성하고, 두 온톨로지 벡터를 비교하여 유사도가 가장 높은 온톨로지 구문을 분류하는 벡터 유사도 이용 방법[4,5]을 온톨로지 통계적인 분류 방법에 사용하였으며 구문과 카테고리의 벡터 표현에서 각 색인어의 가중치로는 주로 사용되는 역문헌빈도[4,6] 대신 색인어에 높은 가중치를 주는 역카테고리빈도[7]를 사용하였다.

예외 처리기는 회의 프로세스에서 예외 상황이 발생할 경우 이를 공지하고 공지를 받아 보고하고 관리하는 역할로 구분하여 예외처리를 진행한다.

지식검색기는 회의 지식 관리 엔진을 통해 온톨로지화된 지식들에 대한 검색으로 한정하며, 현재 회의 지식 관리 엔진을 통해 지식화 과정을 거친 데이터는 빠른 검색속도가 보장된다.

플로우 관리기는 회의 프로세스를 정의, 해석하고 회의의 태스크(Task) 작업 목록을 관리하여 회의진행을 제어한다.

3.2 회의 클라이언트 아키텍처

클라이언트의 XML 메시지 처리기는 서버의 XML 메시지 처리기에 해당 회의 내용과 회의 정보를 XML 구조로 전송한다.

모델 관리기는 서버측에서 받은 모델 정보에 뷰잉(viewing) 정보를 포함하여 실제 GUI를 표현하기 위한 정보들을 관리한다.

리포트 관리기는 지식화된 자료들을 텍스트 및 그래픽 등 다양한 형태로 리포팅할 수 있도록 관리한다.

플로우 관리기는 서버 플로우 메시지를 받아 전체 회의에 대한 진행 순서, 알림, 경고 메시지 등에 대한 처리를 담당한다.

GUI는 클라이언트 사용자 인터페이스로 일반적인 채팅 뷰어, 지식검색, 모델 뷰어, 리포트 뷰어, 투표/설문 관리기 등으로 구성되어 있다.

3.3 협업 아키텍처

그림 3은 일반적인 회의 형태를 개선하고자 제안하는 시스템의 협업 아키텍처를 표현한 것이다. 회의 참석자 역할은 동일하나 ICoMS(Intelligent Conference Management System)에서 서기의 역할과 함께 의장의 일부 역할도 함께 포함하고 있다. ICoMS에서 자동으로 처리되지 않은 것들에 대해서만 의장 및 서기가 직접 처리한다.

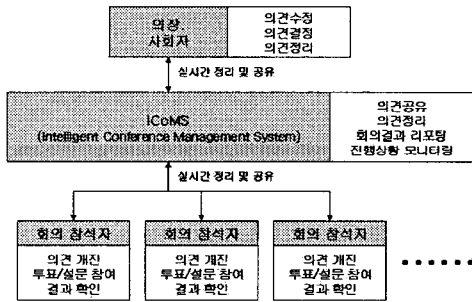


그림 6. 제안 시스템의 협업 아키텍처

IV. 시뮬레이션

4.1 시스템 전체 구성

지식기반 회의관리 시스템에서 서버환경은 펜티엄 III 1GHz, 윈도우 2000 서버 환경에서 자바를 이용하여 개발되었으며 클라이언트 환경은 펜티엄 III 1GHz, 윈도우 2000 워크스테이션에서 Visual C++를 이용한 ActiveX Control로 시뮬레이션 되었다.

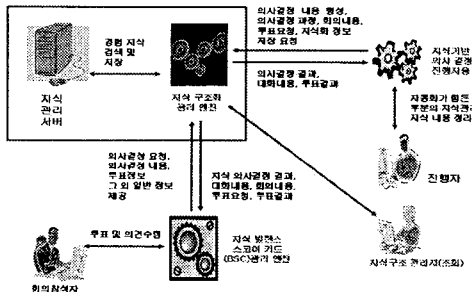


그림 7. 전체 시스템 흐름도

그림 4는 지식기반 회의관리 시스템의 전체 시스템 흐름도이다.

4.2 회의 서버 시뮬레이션

회의 서버에서는 모든 정보가 XML 메시지 처리기를 통해 회의 클라이언트와 통신하며, 들어온 메시지는 기본적으로 로그 관리기에서 로그 기록을 남기게 된다. 그런 다음, 지식 추출기가 자동으로 요약정리가 가능한 지식을 추출한 후, 회의 지식 엔진을 통해 종합적으로 체계화되고 관리된다. 이때 자동으로 정리가 안 된 부분은 예외처리기를 통해 별도 처리를 하며 지식 검색기를 통해 지식을 검색할 수 있다. 또한, 별도의 플로우 관리기를 두어 회의 진행 시에도 관리 될 수 있도록 하였다.

4.3 회의 클라이언트 시뮬레이션

회의 클라이언트는 실시간 온라인 회의에서 회의 내용을 시각적으로 표현하는 시스템으로, 패턴 매칭을 통해 지능적인 요약을 하고, 나머지 내용에 대해 회의 진행자가 직접 중요한 요점에 해당하는 부분을 드래그앤드롭(drag & drop) 방식으로 쉽고 간편하게 요약 정리한다. 그림 5는 진행된 회의내용과 이미 선택된 주제에 대해 한 화면으로 요약해서 보여준 화면이다.

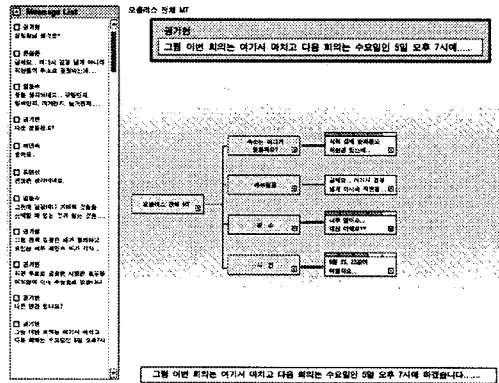


그림 8. 회의 내용 요약

4.4 평가 분석 및 결과

본 논문에서는 시뮬레이션을 위한 문서 집단으로써 한겨레신문(<http://www.hani.co.kr>) 웹사이트로부터 수집한 100개의 뉴스 기사를 이용했다.

그리고, 시뮬레이션에서 구문패턴 방식과 키워드 집합방식을 적용하여 분류율과 정확도를 구했다.

분류율과 정확도의 정의는 패턴에 의한 분류의 정확도를 Pp, 분류율을 Cp, 통계적 방법에 의한 분류의 정확도를 Ps, 분류율을 Cs 라고 할 때 다음과 같다.

$$\text{정확도}(P_p, P_s) = \frac{\text{정확하게분류된문자수}}{\text{분류된총문서수}}$$

$$\text{분류율}(C_p, C_s) = \frac{\text{분류된총문서수}}{\text{총문서수}}$$

분류율과 정확도의 문제점을 보완하기 위해 F-measure 척도를 사용하였다[4].

$$F\beta = \frac{(\beta + 1) \cdot \text{정확도} \cdot \text{분류율}}{\beta \cdot \text{정확도} + \text{분류율}}$$

F-measure에서 β 는 분류율과 정확도의 중요도를 반영하기 위한 가중치이며 본 논문에서는 0.1로 하여 정확도에 더 높은 가중치를 두었다.

표 1. 패턴에 의한 분류 결과

분류방법	표본 집단		
	분류율	대 분류	
		정확도	F-measure
구문패턴	36.7%	89.4%	88.15%
키워드집합 3개	35.8%	93.3%	92.35%
키워드집합 2개	44.1%	92.2%	91.21%
키워드집합 1개	64.1%	83.2%	82.96%

표 1은 표본 추출된 뉴스를 대상으로 한 분류 시뮬레이션 결과이다. 키워드 집합을 이용한 패턴 인식이 구문 패턴에 의한 방식보다는 6~4%정도 높은 정확도와 F-measure값을 나타내고 있으나 구문패턴의 경우 초기 도메인 온톨로지의 구축 정도에 따라 그 정확도는 많이 향상될 수 있다.

4.5 고찰

표 2는 일반 시스템과 제안 시스템을 비교한 것이다.

표 2. 일반회의시스템과의 비교

항 목	오프라인 회의	온라인회의	본 시스템
전체회의 시간	요약/정리 시간 별도 필요	요약/정리 시간 별도 필요	회의 진행 시간만 필요
정형 지식	별도 작업 필요	전체 백업만 가능	실시간 축적 가능
실시간 정리	수작업으로 가능	정형화된 단순한 형태로 지원	템플릿화된 다양한 형태로 지원
의견 조사	비 실시간	실시간	실시간
회의 내용 요약	회의 후 수작업	없음	지식화형 형태로 실시간 표현
회의 논제 저장	종이에 의한 저장	전체 저장	논제만 저장
회의 템플릿	없음	없음	다양한 템플릿 지원
시간간의 제약성	제약 있음	제약 없음	제약 없음
보고서 출력	별도 작업으로 가능	전체 보고서	실시간 요약 정리된 보고서
온톨로지 지식화 표현	수작업으로 가능	없음	시각적으로 표현
회의 히스토리	없음	없음	지원

기존의 회의 시스템과 본 논문에서 제안한 회의 관리 시스템을 비교해 볼 때 회의 지식에 대하여 온톨로지화하여 시스템에서 처리되는 데이터들의 체계성이나 논리성이 훨씬 강해졌다.

따라서 회의 진행 중 논의하게 되는 논제에 따른 의견에 대한 선호도 및 연계성 등이 일목요연하게 정리된다.

V. 결 론

본 논문에서는 기존의 온라인 회의 시스템의 단점을 보완한 지식 기반 회의 관리 시스템 아키텍처를 제안하며 통계적인 문서 분류에 패턴을 이용한 분류방법을 접목시켜 분류의 정확도를 높임으로서 진행자의 수작업 분류를 최대한 자동화하는 방안을 제안하며 온톨로지 지식 모델을 이용하여 좀 더 효율적인 지식기반 회의 관리 시스템을 구현하였다.

이는 기존의 회의시스템을 개선하여 회의 지식을 온톨로지화하여 언제든지 검색할 수 있으며 참석자들의 합리적인 의사 판단이 이루어질 것이다.

향후 과제로서 텍스트 기반이 아닌 멀티미디어 기반의 온톨로지화에 대한 연구와 둘 이상의 패턴이 매칭되어 모호성이 발생할 경우 어떻게 분류할 것인가를 결정하는 방법과 수작업으로 해왔던 패턴 찾기를 자동화하는 방법에 대한 연구가 진행되어야 할 것이다.

참고문헌

- [1] J. Davies et al, "On-To-Knowledge: Semantic Web Enabled KnowledgeManagement", In Web Intelligence. Ning Zhong (ed), Springer, 2003.
- [2] 문현정, "XML 구조적 특징을 이용한 온톨로지 기반의 지식 탐사 모델", 창원대 대학원, 2003
- [3] Daniel K.Power, " Decision Support Systems: Frequently Asked Questions", DSSResources.com, 2004
- [4] N. Kushmerick, E. Johnston, and S. McGuinness, "Information Extraction by Text Classification", IJCAI-01 Workshop on Adaptive Text Extraction and Mining", 2001
- [5] T. Lau and Y. Sure, "Introducing Ontology-based Skills Management at a Large Insurance Company", Proc. Modellierung, Germany, March 2002
- [6] Jerry R.Hobbs, Douglas Appelt, John Bear, David Israel, and Mabry Tyson "FASTUS: A System for Extracting Information from Natural-Language Text".
- [7] 강승식, "한국어 형태소 분석기와 한국어 분석 모듈(HAM: HangukAnalysis Module)", 국민대학교, 2003