

PTL Volume 3(ISSN 2288-8195)

2015 (사)ICT플랫폼학회 추계학술발표논문집

- 일 시 : 2015년 11월 11(수) 13:00 ~ 18:00
- 장 소 : 서울 SETEC 중회의실 (학여울역 1번출구)
- 주 최 : (사)ICT플랫폼학회
- 후 원 : (주)오픈링크

	<p>Web상의 실시간 이미지 검색을 위한 알고리즘 연구 이지용(금오공과대학교), 추인오(구미전자정보기술원), 고재필, 이용환(금오공과대학교)</p> <p>영상 처리 기술을 활용한 차량 반대쪽 보행자 인식 시스템 구본근(한국교통대학교)</p> <p>안전 소프트웨어를 위한 소프트웨어 개발 프로세스 개선 이근상, 김영철()</p> <p>빅데이터와 동일한 효과를 위한 Temporal 데이터베이스의 구조적 데이터 처리 장우성, 박보경, 김영철()</p>	
<p>16:30 ~ 16:50</p>	<p>coffee break</p>	
<p>Session 2 (16:50 ~ 18:00)</p>	<p>가축질병 진단용 귓속삽입형 온도센싱 모듈 박형준, 류지형, 강현서, 김영선(한국전자통신연구원)</p> <p>농산물 콜드체인을 위한 센싱기반 실험적 분석 이민정, 이기현, 정경숙(농업기술실용화재단)</p> <p>스마트 콜드체인 IOT 하드웨어 설계연구 고대식(목원대학교), 박화세(대림대학교)</p> <p>Spark Streaming 기반의 콜드체인 위기 관리 시스템 설계 송석일,최도진(한국교통대학교),고대식(목원대학교),곽윤식(한국교통대학교)</p> <p>스마트 콜드체인 기술 개발 및 적용사례 윤정미, 이상학()</p>	<p>좌장 : 구본근</p>

빅데이터와 동일한 효과를 위한 Temporal 데이터베이스의 구조적 데이터 처리

Handling Structured data with Temporal Database for the same effect as Bigdata

장우성*, 박보경*, 김영철**

Woo-sung Jang*, Bo-kyung Park*, R. Youngchul Kim**

요 약

방대한 양의 데이터 및 처리 관점에서의 빅데이터의 효율성이 있다. 또한 예측/추론의 이슈에 장점이 있다. 기존의 구조화된 데이터를 재사용 관점에서 빅데이터와 대등한 효과를 얻도록 하자는 것이다. 기존의 관계데이터간의 Big table화를 통해 관련된 데이터의 locality를 기반으로 Temporal 데이터화 하고자 한다. 이를 위해, 시간 관련 정보와 행위를 지원되는 구조화 된 데이터화가 중요하다.

Key words : Big Data, Temporal Database, Structured Data

I. 소개

빅데이터는 IT를 활용한 다양한 산업분야의 전자화와 자동화가 보편화됨에 따라 비정형화 된 데이터 처리 및 분석에 필요한 대용량 데이터가 축적된다. 이러한 축적된 데이터의 유용성과 정제되지 않은 데이터의 활용이 더욱 중요한 가치가 존재할 수 있다. 보통 빅데이터는 기존 방식으로 저장, 관리, 분석하기 어려운 큰 규모의 데이터를 말한다. 즉, 빅데이터를 활용하면 데이터의 일정한 흐름을 포착해 미래의 예측이 가능하다[1]. 빅데이터는 기업의 소비자 수요 예측에서 데이터 분석을 이용한 치료 / 인간의 수명 예측 / 범죄예방 / 국가 안전 관리 등의 많은 분야의 영

역으로 넓혀지고 있다. 그러나 아직도 기존의 구조화 된 데이터의 관계 데이터베이스도 중요하다. 본 논문은 비구조화 된 데이터보다 구조화 된 데이터의 중요성에 초점을 둔다.

또한 1995년 Temporal 데이터의 저장 문제로 사장된 Temporal 데이터베이스를 예측과 추론 기법을 적용하도록 하고자 한다. 이를 통해 빅데이터와 같은 효과를 얻으려 한다. 1장은 소개, 2장은 Temporal 데이터베이스 및 구조화 데이터, 3장은 결론을 언급한다.

II. Temporal 데이터베이스

시간은 실시간에서 매우 중요한 정보지만. 기

* 홍익대학교 소프트웨어공학연구소(jang@selab.hongik.ac.kr)

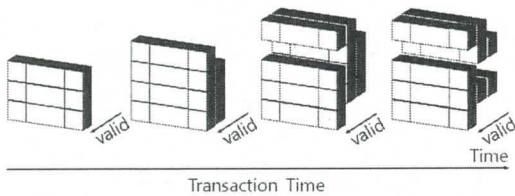
** 교신저자: 홍익대학교 소프트웨어공학연구소(bob@selab.hongik.ac.kr)

· 제1저자 (First Author) : 장우성

존의 데이터베이스에서는 실시간의 시간 데이터를 기록하는 용량이 부족과 Historical Query 지원 불가하다. 모든 정보는 현재의 오직 하나의 Snapshot 같은 데이터로 보유한다.

1. Temporal 데이터베이스

트랜잭션 시간(Transaction Time)과 유효 시간(Valid Time)을 지원하는 방법으로써, 각각의 튜플 내에 시간 개념을 추가하고, 변경된 속성값만 지속적 저장하고자 하고자 한다[2]. 이를 통해서 Retroactive/Postactive 한 변경 정보들을 유지한다.



[그림 1] Temporal 데이터베이스

그림 1은 History 데이터와 Temporal 데이터베이스의 Temporal Relation을 보여준다.

Name	Rank	Valid Time		Transc. Time	
		from	to	from	to
Young	Assoc	09/01/77	∞	08/25/77	12/15/82
Young	Assoc	09/01/77	12/01/82	12/15/82	∞
Young	Full	12/01/82	∞	12/15/82	∞
Tom	Full	12/05/82	∞	12/07/82	12/07/82
Tom	Assoc	12/05/82	∞	12/07/82	∞
Mike	Assist	01/01/83	∞	01/10/83	02/25/84
Mike	Assist	01/01/83	03/01/84	02/25/84	∞

[그림 2] Temporal Relation

그림 2는 기존의 관계데이터베이스의 유효 시간과 트랜잭션 시간을 추가한 Temporal 릴레이션을 보여준다. 예를 들면, 그림 2에서 Postactive 한 데이터 의미에서 Young이 1977년 9월 1일부터 부교수로 일을 시작한다. 데이터베이스에 입력한 시간은 1977년 8월 25일이다. Retroactive 데이터 의미에서는 1982년 12월 1일에 정교수로 승진한다. 그러나 데이터베이스는 1982년 12월

15일에 기록된다. Tom은 1982년 12월 5일에 정교수가 되고, 1982년 12월 7일에 데이터베이스에 입력된다.

```
range of f1 is faculty
range of f2 is faculty
retrieve (f1.rank)
  where f1.name = "Mike"
  and f2.name = "Tom"
  when f1 overlap begin of f2
  as of "12/10/82"
```

[그림 3] TQuel Query

그림 3은 1982년 12월 10일의 데이터베이스 상태에서 Tom이 도착했을 때 Mike의 직위를 질의하기 위한 Query이다.

Rank	Valid Time		Transc. Time	
	from	to	from	to
Assoc	09/01/77	∞	08/25/77	12/15/82

[그림 4] Query Result

그림 4는 그림 3에서 입력한 Query에 대한 결과는 부교수이다. 이 결과의 릴레이션은 Temporal 릴레이션이고, 1982년 12월 20일에 같은 질의를 한다면, 대답은 그 시간의 Retroactive 하게 기록된 사실 때문에 정교수로 대답할 것이다.

2. Temporal 데이터의 구조

시간 연속의 타입은 Step-wise, Discrete, Continuous 시간이며, 이들은 모든 History 데이터에서는 저장할 필요가 없다.

Relation	Symbol	Symbol for inverse	Picture Example
X before Y	<	>	XXX YYY
X equal Y	=	=	XXX
			XXXYYY
X meets Y	m	mi	XXX
X overlaps Y	o	oi	YYY
			XXX
X during Y	d	di	YYYYYY
			XXX
X begins Y	b	bi	YYYYY
			XXX
X finishes Y	F	fi	YYYYY

[그림 5] Temporal Relationship

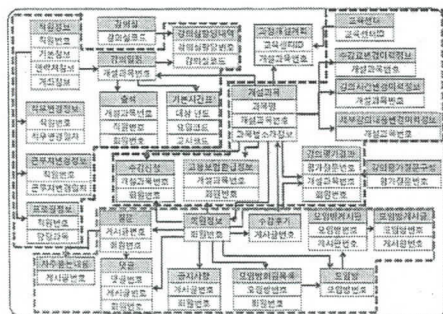
그림 5는 두 릴레이션 간의 시간 관계를 보여 준다.

- [a,b] BEFORE [c,d] IFF $b < c$
- [a,b] AFTER [c,d] IFF $a > d$
- [a,b] DURING [c,d] IFF $(a < c) \ \& \ (b > d)$
- [a,b] EQUIVALENT [c,d] IFF $(a=c) \ \& \ (b=d)$
- [a,b] ADJACENT [c,d] IFF $(c - b = 1) \ | \ (a - d = 1)$
- [a,b] OVERLAP [c,d] IFF $(a < d) \ \& \ (c < b)$
- [a,b] FOLLOWS [c,d] IFF $(a - d = 1)$
- [a,b] PERCEDES [c,d] IFF $(c - d = 1)$

[그림 6] when clause

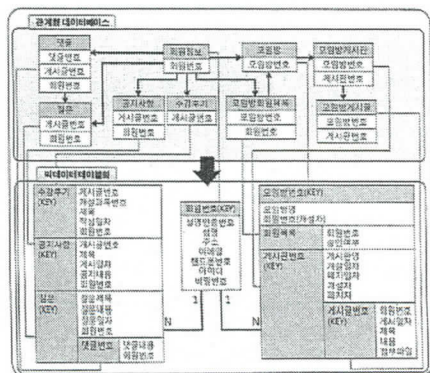
그림 6은 시간 Query에서 when 절의 시간 간격에 대한 표현식이다.

빅데이터화를 위한 관계 데이터 모델의 빅데이터 테이블화[3] 방법에 Temporal 데이터를 접목하면, 빅데이터와 같은 예측/추론을 가능하다고 생각한다. 기존 연구[3]에서 교육관리 시스템 구축을 위한 데이터 모델링 사례를 이용한다.



[그림 7] 교육관리 시스템 구축을 위한 데이터 모델링

그림 7은 기존의 관계데이터베이스 모델링이다.



[그림 8] 회원정보 테이블의 빅데이터 테이블화

그림 8은 그림 7을 “functional access path”기법을 이용한 빅데이터화이다.

이 두 기법을 접목하여, 빅데이터와 같은 예측과 추론이 가능하리라 본다.

III. 결론

기존의 구조화된 데이터를 재사용 관점에서 빅데이터와 대등한 효과를 얻도록 하고자 한다. 기존의 관계데이터간의 Big table화를 통해 관련된 데이터의 locality를 기반으로 Temporal 데이터화 하고, 기존 데이터를 재사용 가능하게 한다.

감사의 글

* 이 논문(저서)은 2015년 교육부와 한국연구재단의 지역혁신창의인력양성사업(NRF-2015H1C1A1035548)과 2015년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임(NRF-2013R1A1A2011601)

참고문헌

- [1] 윤홍근, “문화산업에서 빅데이터의 활용방안에 관한 연구”, 글로벌문화콘텐츠학회, no.10, pp.157-179, February 2013.
- [2] C.S. Jensen, J.Clifford, S.K.Gadia, A.Segev, R. T.Snodgrass, “A Flossary of Temporal Data base Concepts”, SIGMOD RECORD, Vol.21, No.3, pp.35-43, September 1992.
- [3] 양효석, 황준순, 강진희, 권하은, 박보경, 김동호, 박용범, 김영철, “빅데이터화를 위한 관계 데이터모델의 빅데이터 테이블화”, 한국정보과학회, pp.330-332, November 2013.



(사)ICT플랫폼학회 추계학술발표논문집