

# 분산 네트워크 환경에서의 웹뷰 구체화

서건형

경인여자대학 컴퓨터정보기술학부

(jhsuh@yoda.kic.ac.kr)

A study of WebView materialization on distributed network environment

Jin Hyung SUH

School of Computer & Computer Information, Kyungin Women's College

요약

웹에서의 실시간 데이터 제공이 요구되고 있는 상황에서 웹에서의 적절한 데이터 갱신은 매우 중요한 사항이다. 단일 시스템 환경에서의 동적인 웹 데이터 갱신에 대하여는 이미 상당한 연구가 있어왔으나, 실제 네트워크 환경에서 분산되어 있는 데이터 갱신은 아직 많은 연구가 이루어져 있지 않다. 본 논문에서는 기존 연구 중 웹뷰 구체화에 대한 이론을 정리하고, 네트워크에 산재하여있는 데이터를 수집, 가공하여 제공하는 분산 네트워크 환경에서 데이터 갱신을 효율적으로 할 수 있는 방법에 대하여 살펴보고자 한다.

## 1. 서론

데이터베이스에서의 뷰 (view) 는 분산되어 있으면서 서로 연관성이 있는 (loosely coupled) 데이터베이스에 저장되어 있는 관계와 유추된 데이터베이스 테이블 등을 이용하여 원하는 데이터를 요구하는 사용자에게 보여주는 것으로 보여줄 때 마다 반복하여 계산을 하게 된다. [gupta95] [nick01] 이에 반하여 웹뷰 (WebView)는 사용자가 요구할 것으로 예상되는 결과를 데이터베이스에 저장되어 있는 자료 또는 기본 데이터 테이블을 이용하여 미리 계산하여 저장하고 있다가 사용자에게 제공되는 것으로 특히 HTML 형태의 페이지로 보여주는 것을 의미한다. 즉 기본적으로 가지고 있는 자료 (base data)로부터 사용자가 작성한 질의어를 통하여 얻어낸 결과 (query result)를 요구한 사용자에게 볼 수 있는 웹의 형태로 보여주는 페이지 [alex00], [alex02] 또는 웹 페이지에 대한 재사용을 도와주는 그래픽화된 도움방법이라고 할 수 있다. [andy99] 그러나 이 웹뷰는 새로운 데이터가 계속하여 동적으로 생성되어 생성된 데이터를 저장하기 위하여 저장되는 장소가 계속하여 확장되어야 한다는 문제점을 내포하고 있어 이를 해결하기 위하여 웹뷰 구체화 (WebView Materialization)라는 방법을 사용하게 된다. 다음으로 동적으로 데이터를 계속하여 생성하기 때문에 데이터에 대한 신뢰성 또는 확실성들을 따지는 데이터의 질

(Quality of Data; QoD)에 있어 많은 문제점을 내포하고 있다. 이러한 상황에 대하여 해결할 수 있는 방안으로 제시한 웹뷰 구체화 방법으로 구체화를 하지 않는 경우 (no materialization) 데이터베이스 서버에서 구체화를 시키는 경우 (materialize inside DBMS) 그리고 마지막으로 웹/응용 서버에서의 구체화를 시키는 경우 (materialize inside web/application)가 있으며 이중 웹/응용 서버에서 구체화 방법이 서버에 부리를 주지 않으면서 사용자가 요구하는 데이터를 제공할 수 있다. [alex00]

## 2. 웹뷰 갱신과 온라인 웹뷰 선택

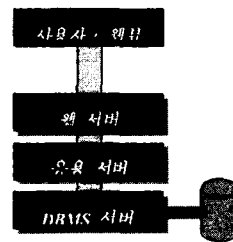
웹뷰 구체화에서 확인하여야 하는 것으로 계속하여 갱신되는 데이터의 질이 높은 수준을 요구함으로써 갱신되는 데이터가 과연 얼마만큼 변경되어야 하는가에 대하여 살펴볼 필요가 있다. 갱신되는 데이터가 일부분만이 계속하여 갱신될 경우와 상당부분의 데이터가 갱신될 경우 사용자에게 의하여 신뢰할 수 있는 데이터가 과연 얼마나 될 것인가와 갱신될 대상의 데이터를 어느 수준까지 갱신할 것인가에 대하여 확인할 필요가 있다. 이러한 데이터의 갱신을 유지할 수 있는 전략으로 데이터의 질 감지 갱신 스케줄링 알고리즘 (QoD-Aware; QoDA update scheduling algorithm) [alex01-1]를 언급할 수 있는데, 이 알고리즘은 저장되어 있는 기본 베이스 테이블의 갱

신 계획을 일치시키고, 일반적인 구조 하에서 구체화 뷰에 대한 갱신을 확인될 때까지 지연시키는 방법을 사용하게 되며, 일반적으로 사용할 수 있는 FIFO 스케줄링 방식에서의 데이터 전송의 질보다도 20배 정도 데이터 전송의 질을 향상시킬 수 있는 것으로 제시되었다. [alex01-2] 그러나 이 경우 속도에 중점을 두고 있어 사용자에게 제공되는 데이터의 질을 확인할 수 있는 방법에 미흡하여 높은 데이터 전송을 지원하는 데이터라고 하더라도 사용자가 요구하는 데이터가 아닐 경우에 이 데이터를 지연시키기 보다는 다시 복귀시키는 방법에 대하여 확인할 필요가 있다. 즉, 지속적으로 사용자에게 새로운 데이터를 제공하기 위하여 생성된 데이터 중 어떤 데이터를 사용자에게 보여줄 것인가를 선택하여야 하는 온라인 뷰 선택 (online view selection) 문제에 당면하게 된다. [alex00] 즉, 데이터의 질과 데이터 전송의 질을 최대화하지 않으면 사용자가 요구하는 데이터를 원하는 시간과 화면에 전송할 수 없다. 웹뷰 구체화에 있어 온라인 뷰 선택은 어느 한쪽의 질을 강화하게 되면 다른 한쪽의 질은 떨어지게 되어있다. 즉, 데이터의 질을 강조하게 될 경우는 전송 속도가 급격히 저하되어 사용자가 원하는 데이터를 시간 내에 보여줄 수 없으며, 전송의 질을 강조하게 되면 사용자에게 보여줄 데이터의 신뢰성이 급격하게 저하되어 사용자가 원하는 데이터를 보여줄 수 없게 된다. 이 경우 적용할 수 있는 알고리즘으로 점진적인 온라인 뷰 선택 알고리즘 (adaptive online view selection algorithm) 또는 이익 기반 온라인 선택 (benefit-bed online selection; BOS) [alex01-2]을 제시할 수 있는데, 이 알고리즘은 기본적으로 데이터 전송의 질이 보장된다는 가정하에서 최대의 데이터의 질을 제공하는 데이터를 구체화한 웹뷰를 선택하게 된다. 즉, 웹의 작업 환경에서 최저의 부하를 제공하게 되는 구체화된 웹뷰를 사용자에게 제공하게 된다

### 3. 네트워크 환경에서의 온라인 웹뷰 선택

기존 연구에서의 웹뷰는 데이터가 산재되어 있지 않은 상태에서 동적인 갱신되는 내용을 질의어를 보낸 사용자에게 제공하게 된다. 그러나, 데이터가 분산된 네트워크 환경에서 우리가 확인하여야 하는 부분은 바로 어디에서 웹뷰 구체화를 할 것이냐를 먼저 결정하여야 한다. 웹 구체화가 실행될 위치 즉, 어떤 클라이언트에서 요청할 콘텐츠를 어떻게 생성할 것인가에 대하여 확실한 결정을 하고 있어야 한다. 다음으로 어떤 항목에 대하여 구체화를 할 것인가 (뷰 선택) 즉, 다양한 리소

스에서 제공되는 분산된 정보 중 어떤 것이 가치 있는 것인가를 판단하여 사용자에게, 제공하여야 한다. 마지막으로 어떠한 순서로 제공되는 내용을 갱신할 것인가 (갱신 스케줄링)를 결정하여야 한다. 갱신 순서가 정하여져 있을 경우에는 문제가 없으나, 순서가 없을 경우는 단일 네트워크 환경에 있어서도 어떤 순서에 의하여 갱신할 것인가를 결정하는 것이 매우 어려운 문제이다. 역시 전송 조건에 있어 다른 위치에 있는 서버와 연계되어 사용자가 요구하는 데이터를 전송하게 되는 분산 네트워크 환경에서는 더욱 큰 문제로 해당 네트워크의 상태에 따라 효율적인 웹뷰 생성에 많은 차이가 있고, 이에 따라 실제 비용에 있어서도 상당한 차이가 날 수밖에 없다. 또한 사용자의 성향에 따라 해당되는 환경이 변화될 수 있음도 고려해야 한다. 위와 같은 문제를 해결하기 위하여 일부 웹 캐싱 (web caching)을 사용할 수 있으나, 실제 웹 캐싱은 일반적인 HTML 문서에서 적용되고 있으며, 부분적인 데이터의 갱신이 아닌 페이지 자체를 갱신하는 방법을 사용하고 있으며, [mark96] [alex02-2] 이를 개선하여 동적으로 웹의 내용을 동적으로 전송할 수 있으며, 정해진 기간까지의 재사용을 기본적인 목표로 하는 방법도 제안되어 있다. [alex02-2]



<단일 환경에서의 웹뷰>

네트워크 환경에서 제공된 데이터에 대한 온라인 웹뷰 선택은 다중 응용 서버와 단일 DBMS 서버 환경, 단일 응용 서버와 다중 DBMS 서버 환경, 그리고 다중 응용 서버와 다중 DBMS 서버 환경으로 나눌 수 있다.

#### (1) 다중 응용 서버와 단일 DBMS 서버 환경

사용자의 질의의 결과는 네트워크에 산재하여 있는 다중 응용 서버에 의하여 제공되며, 응용 서버에서 사용되는 데이터는 단일 DBMS에서 제공된다. 다중 응용 서버의 계산 결과 중 사용 가능한 결과를 선택하여 사용자의 웹뷰에 제공하거나, 사용자의 웹뷰에서 어떤 방식으로 주어진 결과를 보여줄 것인가를 결정하여야 한다. 또한 응용 서버의 결과를 네트워크를 통하여 전송되었을 경우 이에 따

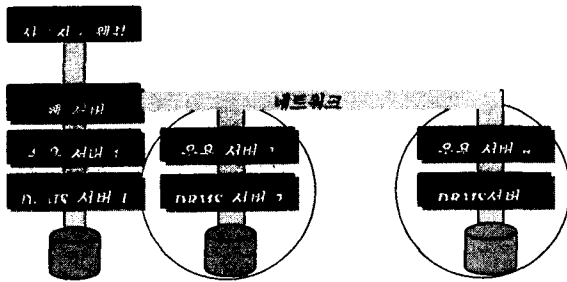
르는 네트워크 지연시간으로 데이터가 도착하는 순서에 의하여 사용자에게 어떻게 임시로 저장되고 이를 표시될 것인가에 대하여 확인하여야 한다.

(2) 단일 응용 서버와 다중 DBMS 서버 환경

일반적으로 웹 환경은 단일 리소스에 저장되어 있는 데이터베이스 데이터 또는 테이블을 사용하여 사용자의 질의에 대한 결과를 보여주는 것이나, 다양한 리소스에 저장되어 있는 데이터베이스 데이터 또는 테이블을 사용할 경우, 선택하여야 할 리소스의 위치 결정 문제와 여러 개의 데이터를 동시에 사용하여 할 경우 발생할 수 있는 어떤 데이터를 어떤 환경에서 얼마나 사용할 것인가를 결정하여야 하는 또는 동일한 데이터를 경쟁적으로 사용하기 위한 데이터의 경쟁관계 (data competition) 관계에 대하여 확인할 필요가 있다.

(3) 다중 응용 서버와 다중 DBMS 서버 환경

다중의 응용 서버와 DBMS 서버 환경은 위의 두 가지 환경을 혼합한 경우로 위의 두 가지 환경에서의 다중 환경을 고려한다. 그러나, 다중 환경에서 발생할 수 있는 문제점은 위의 두 가지 환경에서 발생할 수 있는 문제점이 증첩되어 나타나게 되어, 데이터와 전송의 질에 있어 더 많은 문제가 발생할 수 있다. 즉, 데이터의 갱신에 있어 데이터의 질을 최대로 하면서 선택적으로 웹뷰를 구축화시킬 수 있는 방법을 제시할 수 있어야 한다.



<네트워크 환경에서의 웹뷰 - 다중 응용, DBMS 서버>

4. 결론

본 논문에서는 기존의 연구에서 제시되었던 웹뷰에서의 문제점과 해결하여야 할 점 그리고 네트워크의 활성화와 더불어 논의되는 있는 분산 네트워크 환경에 웹뷰를 적용하였을 경우에 대하여 알아보았다.

본 논문에서 제시된 세가지 온라인 웹뷰 환경은 하나의 결과

에 대하여 정상적인 결과로 인식하는 환경에는 적합하지 않다. 또한 완만하게 연관되어 분산되어 있는 네트워크 환경에서 다양한 리소스로부터 데이터를 전송 받아 이를 사용자에게 제공하는 데이터베이스 뷰가 정적인 데이터를 보여주는 반면 동일한 환경에서 적용되는 웹뷰는 동적인 데이터를 사용자에게 보여지게 되어 제시된 환경에 대한 상세한 분석, 검증 그리고 이에 대한 시뮬레이션 결과와 네트워크의 단결로 인한 오류 발생 [alex02-1] 에 대한 연구도 필요로 한다.

참고문헌

[alex00] Alexandros Labrinidis and Nick Roussopoulos, "WebView Materialization" In Proc. of the ACM SIGMOD Intl Conf on Management of Data, Dallas, Texas, USA, May 2000.

[alex01-1] Alexandros Labrinidis, and Nick Roussopoulos, "Update Propagation Strategies for Improving the Quality of Data on the Web". In Proc. of the 27th Intl Conf. on VLDB'01, Rome, Italy, Sep. 2001.

[alex01-2] Alexandros Labrinidis, and Nick Roussopoulos, "Adaptive WebView Materialization". In the Proc. of the Fourth Intl Workshop on the WebDB'2001, Santa Barbara, California, USA, May 2001

[alex02-1] Alexandros Labrinidis, and Nick Roussopoulos, "Online View Selection for the Web", <http://www.cs.umd.edu/Library/TRs/CS-TR-4343/CS-TR-4343.ps.Z>

[alex02-2] Alexandros Labrinidis, "WebView Materialization for Scalable Web Servers", Colloquia, University of California Santa Barbara, May 2002

[andy99] Andy Cockburn Saul Greenberg Bruce McKenzie Michael Jasonsmith Shaun Kaaste, "WebView: A Graphical Aid for Revisiting Web Pages", <http://www.cpsc.ucalgary.ca/projects/group/lab/papers/1999/99-WebView.Ozchi/webview-ozchi.pdf>

[gupta95] Ashish Gupta, Inderpal Singh Mumick, "Maintenance of Materialized Views: Problems", *Techniques and Applications, Data Engineering Bulletin*, 18(2):3-18, June, 1995

[mark96] Evangelos P. Markatos and Antonis Danalis: Web Caching. *Enterprise Networking: Multilayer Switching and Applications*, editors V. Theoharakis and D.N. Serpanos, pp. 383-408, Idea Group Publishing

[nick01] Nick Roussopoulos, Yannis Kotidis, Alexandros Labrinidis and, Yannis Sismanis, "The Opsi Project: Materialized Views for Data Warehouses and the Web". In Proc. of the 8th Panhellenic Conf. on Informatics, Nicosia, Cyprus, Nov. 2001.