



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2008-0000393
(43) 공개일자 2008년01월02일

(51) Int. Cl.

G06F 17/40 (2006.01)

(21) 출원번호 10-2006-0058178

(22) 출원일자 2006년06월27일

심사청구일자 2006년06월27일

(71) 출원인

세종대학교산학협력단

서울 광진구 군자동 98 세종대학교

(72) 발명자

김충곤

서울 관악구 신림12동 747~766

(74) 대리인

박윤희

전체 청구항 수 : 총 5 항

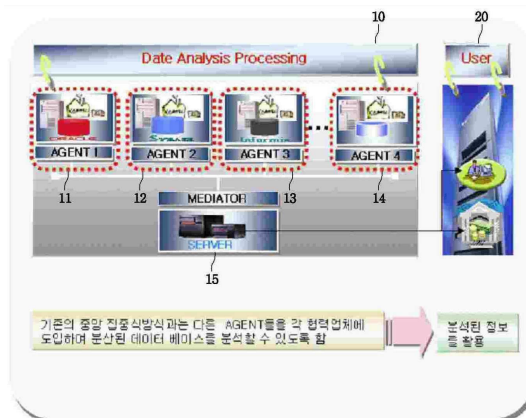
(54) 지식융합을 위한 분산형 데이터마이닝 에이전트시스템

(57) 요약

본 발명은 분산된 각 사이트 데이터들의 부분적인 분석을 하는 마이닝 에이전트들의 서로 간 통신을 통해 궁극적으로 선택된 모든 데이터를 종합 분석할 수 있는 지식융합을 위한 분산형 데이터마이닝 에이전트시스템에 관한 것이다.

이를 위한 본 발명은, 에이전트들을 각 협력업체에 도입하여 분산된 데이터베이스를 분석하는 데이터 분석프로세스(10)은 복수의 에이전트(11 - 14) 시스템과, 이 에이전트(11 - 14) 시스템이 상호 연결되는 서버의 메디에이터(15)로 구비하여, 상기 에이전트(11 - 14) 시스템이 네트워킹되는 메디에이터(15)의 서버는 분석된 정보를 활용하는 유저(20)의 컴퓨터와 상호 연결되고 있는 것을 그 특징으로 한다.

대표도 - 도2



이 발명을 지원한 국가연구개발사업

과제고유번호 10643

부처명 서울시

연구사업명 신기술 연구개발 지원사업

연구과제명 유비쿼터스 컴퓨터 환경을 위한 분산형데이터마이닝 기술

주관기관 세종대학교 산학협력단

연구기간 2005년12월1일 - 2007년11월30일

특허청구의 범위

청구항 1

에이전트들을 각 협력업체에 도입하여 분산된 데이터베이스를 분석하는 데이터 분석프로세싱은 복수의 에이전트 시스템과, 이 에이전트 시스템들이 상호 연결되는 서버의 메디에이터로 구비하여,

상기 에이전트 시스템들이 네트워킹되는 메디에이터의 서버는 분석된 정보를 활용하는 유저의 컴퓨터와 상호 연결된 것을 포함하는 것을 특징으로 하는 지식융합을 위한 분산형 데이터마이닝 에이전트시스템.

청구항 2

각 에이전트 시스템사이에 메디에이터의 서버를 네트워킹되도록 구성하여;

상기 메디에이터의 서버로 데이터 분석요청이 오면 해당 에이전트의 시스템에 수행명령하고, 한 에이전트 시스템에서는 주도적 에이전트선별을 위한 데이터편향성을 분석하고 분석결과를 메디에이터 서버로 통보하며, 다른 에이전트 시스템에서도 주도적 에이전트선별을 위한 데이터편향성을 분석하고 분석결과를 메디에이터 서버로 통보하며;

이 메디에이터 서버는 주도적 에이전트 시스템을 선별하고 모든 참여 에이전트 시스템에 통보하게 되면, 한 에이전트 시스템은 해당 단계에서의 의사결정나무를 부분적으로 구축하고 구축된 의사결정나무 정보를 메디에이터 서버에게 전달하며, 이 메디에이터 서버는 의사결정나무 정보를 그 이외 에이전트 시스템에 전달하고, 다른 에이전트 시스템은 전달된 의사결정나무 정보를 통한 의사결정나무를 부분적으로 구축한 것을 특징으로 하는 지식융합을 위한 분산형 데이터마이닝 에이전트시스템.

청구항 3

제 2 항에 있어서,

상기 각 에이전트 시스템에서는 각 사이트에 최종 의사결정나무를 구축하게 된 것을 특징으로 하는 지식융합을 위한 분산형 데이터마이닝 에이전트시스템.

청구항 4

메디에이터 서버는 분석작업을 수행할 모든 에이전트 시스템들을 인보우크시킴으로 에이전트 시스템들을 동작시키고, 인보우크된 에이전트 시스템들은 자신의 시스템상에서 마이닝을 위한 하나의 쓰레드를 수행시킴으로 자신이 분석해야 할 파일을 읽으며, 각 에이전트 시스템들은 자신이 분석할 파일에서 분류의 기준이 되는 값과 이 값으로 분류하고 있을 시의 엔트로피 값을 계산하여 엔트로피정보를 메디에이터 서버에게 전달하고;

상기 메디에이터 서버에서는 분석에 참여하는 모든 에이전트 시스템들의 엔트로피정보를 받아 이를 비교하여 승자 에이전트시스템과 패자 에이전트 시스템을 구별한 후 이를 각 에이전트 시스템에게 통보하고, 상기 승자 에이전트시스템으로 통보받은 에이전트 시스템은 자신의 분류 기준값에 따라 의사결정 나무(DT)의 일부를 구축한 후 이 구조 정보를 메디에이터 서버에게 전달하며;

구조정보를 받은 메디에이터 서버는 패자 에이전트 시스템들에게 이 정보를 전달하고, 이를 받은 패자 에이전트 시스템들은 이 정보를 토대로 자신의 데이터를 전달받은 구조 정보에 맞게 의사결정나무의 일부를 구축하며, 전술한 엔트로피의 계산 과정부터 의사결정나무가 완전히 구축되기까지 일련의 반복 작업을 하고, 의사결정나무가 최종으로 완성되면 마이닝 작업을 종료한 것을 특징으로 하는 지식융합을 위한 분산형 데이터마이닝 에이전트시스템.

청구항 5

제 4 항에 있어서,

상기 에이전트시스템 들간의 통신은 자바이고, 메디에이터 서버에서는 자바 애플릿인 것을 특징으로 하는 지식융합을 위한 분산형 데이터마이닝 에이전트시스템.

명세서

발명의 상세한 설명

발명의 목적

발명이 속하는 기술 및 그 분야의 종래기술

- <14> 본 발명은 지식융합을 위한 분산형 데이터마이닝 에이전트시스템에 관한 것으로서, 더욱 상세하게는 분산된 각 사이트 데이터들의 부분적인 분석을 하는 마이닝 에이전트들의 서로 간 통신을 통해 궁극적으로 선택된 모든 데이터를 종합 분석할 수 있는 지식 융합을 위한 분산형 데이터마이닝 에이전트 기술(Data Mining Agent Technology for Knowledge Convergence)에 관한 것이다.
- <15> 일반적으로, 비즈니스에서 고부가 가치를 창출할 수 있는 지식정보를 얻기 위해서는 다양하고 풍부한 데이터를 기반으로 한 분석이 필요하다.
- <16> 그러나, 여러 사이트에 분산되어 있는 다양한 데이터들을 통합하는데 많은 제약이 따른다.
- <17> 도 1 은 중앙집중식 데이터 마이닝 기술을 설명하기 위한 구성도이다.
- <18> 상기 데이터 마이닝(Data Mining)은 이전에는 발전되지 않았던 데이터들간의 상호관계를 분석하는 것이고, 예를 들어 만약 특정 상표의 테니스 라켓 판매기록이 충분히 분석되고 다른 시장데이터와 관련지어졌다면 같은 회사의 골프 장비구입과의 계절적인 상호관계를 밝힐 수 있을 것이다.
- <19> 그리고, 데이터 웨어하우스(Data Warehouse)의 구성 요소는 데이터베이스와 데이터 웨어하우스 관리도구로 구분되고 있다.
- <20> 상기 데이터베이스는 데이터 웨어하우스 구축의 핵심이며 데이터 웨어하우스 수반되는 대용량의 데이터를 의사결정과정에 연계시키는 가장 필수적인 요소이고, 상기 데이터 웨어하우스 관리도구는 오퍼레이셔널 데이터를 별도의 프로그래밍 없이도 자동적으로 변형시켜 데이터 웨어하우스를 도와주는 기술인 것이다.
- <21> 따라서, 상기 중앙집중식 데이터 마이닝 기술은 인터넷상의 실세계(Real World) 데이터(1), 데이터 웨어하우스(2), 분석 어플리케이션의 소프트웨어(3)및 유저(4)순으로 구성되는 바, 이는 한 업체에서 필요한 자료를 인터넷이나 수작업으로 데이터 웨어하우스로 수집하게 되고, 이렇게 수집된 데이터를 대용량 데이터 웨어하우스에 저장하게 되며, 이 대용량 데이터 웨어하우스를 분석하게 됨으로 분석된 정보를 이용하여 각 업체에서 사용하게 된다.
- <22> 따라서, 상기 중앙집중식 데이터 마이닝의 특징은 대규모 중앙식 데이터 웨어하우스를 구축하고, 데이터 이행(Migration)하며, 분석어플리케이션의 데이터 웨어하우스에 상주시키는 것이다.
- <23> 그러나, 대규모의 데이터 웨어하우스구축을 위한 고비용이고, 데이터 이행을 위한 통합지원, 비용소모 및 정보가 유출된다는 문제점이 있었다.
- <24> 네트워크 인프라가 잘 갖추어 졌음에도 정보 보안상의 문제로 인하여, 정보의 온라인 공유를 기피하고 있는 실정이다.
- <25> 사실상 정보는 각 사이트에 흩어진 정보를 통합적이고, 종합적으로 분석할 때 진정 가치 있는 정보가 될 수 있지만 현 실정은 정보 보안상의 문제로 인하여, 각 기업의 중요한 정보들은 통합적인 분석을 기피하고 있어, 이에 따른 국가 및 기업, 개인의 이윤이 축소되고 있는 실정이다.
- <26> 또한, 어디서나 종합적이고 지역적 차별이 없는 정보를 제공해야 하는 유비쿼터스 컴퓨팅 환경이 도래되면서, 데이터가 분산된 상태에서 데이터마이닝이 이루어져 지식정보를 얻을 수 있는 방법이 자연스럽게 요청되어, 차세대 데이터 분석 기술로 개발될 필요성이 높아지고 있는 것이다.

발명이 이루고자 하는 기술적 과제

- <27> 본 발명은 상기와 같이 중앙집중식 데이터 마이닝 기술이 가지고 있는 문제점을 해소하기 위해 발명한 것으로, 데이터 분석프로세싱에서 각 에이전트 시스템들을 각 협력업체 시스템에 도입하여 분산된 데이터베이스를 분석함으로써 유저 컴퓨터가 분석된 정보를 활용할 수 있어, 분산형 데이터마이닝으로써 철저한 정보보안과 데이터 마이닝의 부하분산을 기할 수 있는 지식융합을 위한 분산형 데이터마이닝 에이전트시스템을 제공하고자 함에 그 목적이 있다.

발명의 구성 및 작용

- <28> 상기 목적을 달성하기 위한 본 발명은, 에이전트들을 각 협력업체에 도입하여 분산된 데이터베이스를 분석하는 데이터 분석프로세싱은 복수의 에이전트 시스템과, 이 에이전트 시스템이 상호 연결되는 서버의 메디에이터로 구비하여, 상기 에이전트 시스템이 네트워킹되는 메디에이터의 서버는 분석된 정보를 활용하는 유저의 컴퓨터와 상호 연결되고 있는 것을 특징으로 한다.
- <29> 본 발명의 다른 구체적인 특징은, 각 에이전트 시스템사이에 메디에이터의 서버를 네트워킹되도록 구성하여; 상기 메디에이터의 서버로 데이터 분석요청이 오면 해당 에이전트의 시스템에 수행명령하고, 한 에이전트 시스템에서는 주도적 에이전트선별을 위한 데이터편향성을 분석하고 분석결과를 메디에이터 서버로 통보하며, 다른 에이전트 시스템에서도 주도적 에이전트선별을 위한 데이터편향성을 분석하고 분석결과를 메디에이터 서버로 통보하며; 이 메디에이터 서버는 주도적 에이전트 시스템을 선별하고 모든 참여 에이전트 시스템에 통보하게 되면, 한 에이전트 시스템은 해당 단계에서의 의사결정나무를 부분적 구축하고 구축된 의사결정나무 정보를 메디에이터 서버에게 전달하며, 이 메디에이터 서버는 의사결정나무 정보를 그 이외 에이전트 시스템에 전달하고, 다른 에이전트 시스템은 전달된 의사결정나무 정보를 통한 의사결정나무를 부분적으로 구축한 것을 특징으로 한다.
- <30> 본 발명의 또 다른 구체적인 특징은, 메디에이터 서버는 분석작업을 수행할 모든 에이전트 시스템들을 인보우크 시킴으로 에이전트 시스템들을 동작시키고, 인보우크된 에이전트 시스템들은 자신의 시스템상에서 마이닝을 위한 하나의 쓰레드를 수행시킴으로 자신이 분석해야 할 파일을 읽으며, 각 에이전트 시스템들은 자신이 분석할 파일에서 분류의 기준이 되는 값과 이 값으로 분류하고 있을 시의 엔트로피 값을 계산하여 엔트로피정보를 메디에이터 서버에게 전달하고; 상기 메디에이터 서버에서는 분석에 참여하는 모든 에이전트 시스템들의 엔트로피정보를 받아 이를 비교하여 승자 에이전트시스템과 패자 에이전트 시스템을 구별한 후 이를 각 에이전트 시스템에 통보하고, 상기 승자 에이전트시스템으로 통보받은 에이전트 시스템은 자신의 분류 기준값에 따라 의사결정나무(DT)의 일부를 구축한 후 이 구조 정보를 메디에이터 서버에게 전달하며; 구조정보를 받은 메디에이터 서버는 패자 에이전트 시스템들에게 이 정보를 전달하고, 이를 받은 패자 에이전트 시스템들은 이 정보를 토대로 자신의 데이터를 전달받은 구조 정보에 맞게 의사결정나무의 일부를 구축하며, 전술한 엔트로피의 계산 과정부터 의사결정나무가 완전히 구축되기까지 일련의 반복 작업을 하고, 의사결정나무가 최종으로 완성되면 마이닝 작업을 종료한 것을 특징으로 한다.
- <31> 이하, 본 발명의 바람직한 실시예를 예시도면에 의거하여 상세히 설명한다.
- <32> 도 2 는 본 발명의 실시예에 관한 분산형 데이터마이닝 기술을 설명하기 위한 구성도로서, 본 발명은 데이터 분석프로세싱(10)과 유저(20)로 구분되고, 이 데이터 분석프로세싱(10)은 복수의 에이전트(11 - 14) 시스템과, 이 에이전트(11 - 14) 시스템이 상호 연결되는 서버의 메디에이터(Mediator : 15)로 구성된다.
- <33> 따라서, 상기 에이전트(11 - 14) 시스템이 네트워킹되는 메디에이터(15)의 서버는 유저(20)의 컴퓨터와 상호 연결되고 있다.
- <34> 본 발명은 종래의 중앙집중방식과 다른 에이전트(11 - 14) 시스템들을 각 협력업체에 도입하여 분산된 데이터베이스를 분석할 수 있도록 함으로, 유저(20) 컴퓨터가 분석된 정보를 활용할 수 있다.
- <35> 따라서, 분산형 데이터마이닝으로 철저한 정보보안을 행하고 데이터 마이닝의 부하분산도 기할 수 있다.
- <36> 그러므로, 본 발명의 도입 효과는 최소 추가비용, 비용 절감, 정보유출 방지, 데이터 분석을 위한 참여용이, 중앙집중식이 불가능한 환경하의 데이터 분석가능, 종합적인 DM 분석가능, 데이터 활용도 증가를 도모할 수 있다.
- <37> 본 발명에 대한 특징은, 분산된 환경에서 독립적으로 수행되는 정보처리시스템들의 정보를 융합 및 분석하여 종합적이고 견고한 예측모델을 제공하고, 이 예측 모델과 분산되어 있는 데이터의 연관성을 고려한 예측 기능을 갖으며, 에이전트 기반 분산처리에 따른 빠른 정보를 추론하며, 유사 또는 이종정보 및 다양한 정보처리시스템의 정보를 융합 추론하고 있다.
- <38> 또한, 본 발명은 분산 처리시 정보 유출을 방지하고, 정보 융합시스템 간의 연동 확장이 가능하며, 의사결정나무 기술이고, 분산화 시스템을 활용하여 방대한 중앙데이터 관리시스템의 구축비용을 절감하며, 플랫폼과 상관없이 어떤 오퍼레이팅 시스템에서나 연동 가능한 시스템이며, 대용량의 데이터를 처리할 수 있다.
- <39> 이상과 같이 구성되는 본 발명은 한 곳에 데이터를 수집하여 종합적인 데이터 분석을 하는 중앙집중식 데이터마

이닝 기법이나, 각각의 사이트에서 데이터를 분석한 결과를 함께 취합하는 퓨전 방식과는 전혀 다른 새롭고 혁신적인 기술인 것이다.

- <40> 본 발명의 지식 융합과정은 여러 사이트에 있는 다양하고 종합적인 데이터 분석 과정 중에 단순히 각 사이트에 구축되어 있는 데이터베이스 내부구조에 대한 내용만을 교환함으로 가능하게 한다. 그러나, 일반 데이터 마이닝 기법은 데이터를 분석하기 위해 여러 사이트에 분산된 데이터를 한 곳에 수집하는데 어려움이 있다.본 발명은 방대한 데이터를 분산 처리함으로 분석 시간을 단축할 수 있다.
- <41> 도 3 은 본 발명의 지식융합을 위한 분산형 데이터마이닝 에이전트시스템을 설명하기 위한 분산형 데이터마이닝의 알고리즘이고, 각 에이전트(11, 12) 시스템사이에 메디에이터(15)의 서버를 네트워킹되도록 구성한 것이다.
- <42> 즉, 상기 메디에이터(15)의 서버로 데이터 분석요청이 오면 해당 에이전트의 시스템에 수행명령한다.
- <43> 에이전트(11) 시스템에서는 주도적 에이전트선별을 위한 데이터편향성을 분석하고, 분석결과를 중재자인 메디에이터(15) 서버로 통보한다.
- <44> 다른 에이전트(12) 시스템에서도 주도적 에이전트선별을 위한 데이터편향성을 분석하고, 분석결과를 중재자 메디에이터(15) 서버로 통보한다.
- <45> 상기 메디에이터(15) 서버는 주도적 에이전트 시스템을 선별하고 모든 참여 에이전트 시스템에 통보하게 된다.
- <46> 그러면, 상기 에이전트(11) 시스템은 해당 단계에서의 의사결정나무(DT)를 부분적으로 구축하고, 구축된 의사결정나무 정보를 중재자 서버에게 전달한다.
- <47> 이 메디에이터(15) 서버는 의사결정나무 정보를 그 이외 에이전트 시스템에 전달한다.
- <48> 다른 에이전트(12) 시스템은 전달된 의사결정나무 정보를 통한 의사결정나무를 부분적으로 구축한다. 따라서, 상기 각 에이전트(11, 12) 시스템에서는 각 사이트에 최종 의사결정나무를 구축할 수 있게 된다.
- <49> 데이터마이닝 알고리즘에 각 사이트의 데이터의 이동없이, 분산된 환경에서 각 사이트에 상주하는 에이전트(11, 12) 시스템들이 서로간의 통신을 할 수 있는 기능을 포함한다.
- <50> 알고리즘에서 예측모델을 구축하기 위해, 데이터 구조(실제 데이터와 무관함)에 대한 내용을 서로 교환할 수 있다. 데이터 구조에 대한 내용을 서로 공유함으로 실제 데이터 정보를 얻지는 못하지만 궁극적으로 지식 융합에 활용되는 공유 환경을 구성한다.
- <51> 가상적인 데이터 공유 환경에서 종합적인 데이터를 분석한 결과(지식정보)를 토대로 정보를 예측할 수 있다. 예측에 사용되는 데이터들도 가상적인 데이터 공유 환경에서 관련된 데이터와 함께 예측에 필요한 근거 자료가 된다. 데이터의 종합 분석의 결과뿐만 아니라 예측의 필요한 데이터가 관련 데이터와 함께 예측 대상이 됨으로 신뢰도가 높은 예측이 가능하다.
- <52> 도 4 는 분산형 방법을 통해 의사결정 나무가 구축되는 과정도이다. 상기 분산형 데이터의 의사결정나무 유도 알고리즘은 다음과 같다.
- <53> 메디에이터 서버는 분석작업을 수행할 모든 에이전트 시스템들을 인보우크(Invoke)시킴으로서 에이전트 시스템들을 동작시킨다. 인보우크된 에이전트 시스템들은 자신의 시스템상에서 마이닝을 위한 하나의 쓰레드를 수행시킴으로 자신이 분석해야 할 파일을 읽는다. 각 에이전트 시스템들은 자신이 분석할 파일에서 분류의 기준이 되는 값과 이 값으로 분류하고 있을 시의 엔트로피(Entropy)값을 계산하여 엔트로피정보(Entropy Information)을 메디에이터 서버에게 전달한다.
- <54> 상기 메디에이터 서버에서는 분석에 참여하는 모든 에이전트 시스템들의 엔트로피정보를 받아 이를 비교하여 승자 에이전트(Winner Agent)시스템과 패자 에이전트(looser Agent) 시스템을 구별한 후 이를 각 에이전트 시스템에게 통보한다. 상기 승자 에이전트시스템으로 통보받은 에이전트 시스템은 자신의 분류 기준값에 따라 의사결정 나무(DT)의 일부를 구축한 후 이 구조 정보를 메디에이터 서버에게 전달한다.
- <55> 구조정보를 받은 메디에이터 서버는 패자 에이전트 시스템들에게 이 정보를 전달하고, 이를 받은 패자 에이전트 시스템들은 이 정보를 토대로 자신의 데이터를 전달받은 구조 정보에 맞게 의사결정나무의 일부를 구축한다.
- <56> 전술한 바와 같이 엔트로피의 계산 과정부터 의사결정나무가 완전히 구축되기까지 일련의 반복 작업을 한다. 의사결정나무가 최종으로 완성되면 마이닝 작업을 종료한다.

- <57> 본 발명은 매우 첨단적이면서 비즈니스 분야의 경험을 많이 필요로 하는 내용이므로, 예컨대 대학의 연구실과 비즈니스 사업 경험이 풍부한 기업의 산학 협력을 통해 기술 개발을 추진하고 있다. 참여 업체는 이-비즈니스(E-Business)로 사업화에 성공한 기업임으로, 비즈니스 관련 경험들과 대용량 데이터베이스 환경에서의 시스템 테스트베드 환경을 제공해주고 있다.
- <58> 상기 참여업체로부터 고객관계관리 시스템 등과 같은 데이터마이닝 기술이 활용되는 기업형 애플리케이션의 노하우를 제공받고 있다.
- <59> 그러므로, 본 발명은 의사결정나무 기반 분산형 데이터 마이닝 알고리즘 개발 및 구현을 할 수 있다.
- <60> 분산형 데이터 마이닝 시스템 개발 및 구현은 마이닝 엔진과 예측 엔진은 C와 C++로 구현되고 있다. 또한, 웹 기반 인터페이스는 자바(Java)로 구현되고, 에이전트 시스템들 간의 통신 기술은 자바로 구현된다.
- <61> 도 5 은 본 발명의 지식융합을 위한 분산형 데이터마이닝 에이전트시스템의 개요도로써 분산형 데이터 마이닝 시스템의 개요인 것이다.
- <62> 메디에이터는 사이트A의 에이전트1, 사이트B의 에이전트2, 사이트C의 에이전트N들과 상호 연결되고, 상기 메디에이터는 웹서버를 통해 복수의 웹브라우저가 연결되며, 상기 에이전트1 - N는 각각 마이닝 엔진이 구성되고 있다. 여기서 점선라인은 콘트롤 메시지고, 실선은 예측 모델의 구조 정보인 것이다.
- <63> 도 6 및 도 7 은 본 발명의 지식융합을 위한 분산형 데이터마이닝 에이전트시스템의 실행화면들인 바, 도 6 은 분산형 데이터마이닝 시스템에서 데이터가 송수신되는 것을 보여주고, 도 7 은 실제 사용자가 데이터를 입력하고 데이터 마이닝을 실행하는 것이다.
- <64> 본 발명의 파급효과 및 활용 방안에서 파급효과는 기술적 측면의 파급효과에 있어 분산형 데이터마이닝 기술은 여러 데이터들을 분석한 결과를 필요로 하는 모든 분야에 적용 가능한 혁신적인 기술인 것이다.
- <65> 특히, 비즈니스 업계의 지식 융합기술 도입으로 인한 새로운 패러다임을 제시함으로, 침체된 데이터마이닝 기술의 활성화를 기대할 수 있다. 유비쿼터스 환경을 위한 종합적인 데이터 분석 기반 한 정보 제공 시스템 기반 구축이 가능하다.
- <66> 또한, 경제산업적 측면의 파급효과로는, 유비쿼터스 환경에서의 종합적인 데이터 분석 및 예측이 가능하고, 개인 정보 유출 없는 정보 공유로 인한 건전한 신용질서와 공정한 금융거래 관행을 확립한다.
- <67> 비즈니스 업계의 종합적인 데이터 분석을 통한 효율적인 업무 방향 제시 및 기업 이윤의 극대화, 분산화 시스템을 활용하여 방대한 중앙 데이터 관리 시스템의 구축 비용 절감, 침체된 데이터마이닝 기술의 활성화 기대, 정부 기관들의 종합적인 데이터 분석을 통한 효율적인 국가 경영(국가 기밀과 일반 정보들이 가상적으로 통합, 분석되어 정확한 정보 획득으로 인한 정부 기관의 업무 신뢰도 증가)이다.
- <68> 또한, 가상적 정보 공유를 통한 국가 간 경제 협력 및 정책 공유의 가능성을 제시하고, 기존의 다국적 기업들의 제품들보다 성능이나 기능면에서 우수하기 때문에, 다국적 기업들이 수익을 올림으로써 국내 자본이 외국으로 유출되는 것을 방지할 수 있다. 본 발명의 제품이 해외로 수출되어 기업형 애플리케이션으로 외화를 획득할 수 있는 기회가 될 수 있다.
- <69> 본 발명의 활용방안으로 본 기술은 종합적인 데이터 분석 결과를 토대로 중대한 의사 결정을 하는 기관 및 기업들 간에 사용되어 질 수 있다. 본 발명을 이용하여 활용될 수 있는 분야는 공공 기관 분야, 금융/보험/증권 분야, 일반 기업체, 의료기관 등이 있다.
- <70> 도 8 은 본 발명의 지식융합을 위한 분산형 데이터마이닝 에이전트시스템에서 공공 기관분야의 예시도이며, 도 9 는 본 발명의 지식융합을 위한 분산형 데이터마이닝 에이전트시스템에서 금융/보험/증권 분야의 예시도이고, 도 10 는 본 발명의 지식융합을 위한 분산형 데이터마이닝 에이전트시스템에서 일반 기업체의 예시도이다.
- <71> 공공 기관 분야는 국가 기밀정보들의 가상 정보 공유가 가능한 지식융합 기술을 통해 (종합적인 데이터 분석) 신뢰도 높은 지식정보들을 추론함으로 국가 기관의 업무 신용도 증가한다.
- <72> 한 예로 국가 기밀과 일반 정보들이 가상적으로 통합, 분석되어 정확한 정보 획득으로 인해 효율적인 국가 경영 가능. 관세청, 세무서, 국세청, 세관, 국정원 등과 같은 정보 보안을 중요시 여기는 국가 기관도 데이터 공유에 참여한다.
- <73> 일반 기업과 긴밀히 데이터의 교환이 필요한 정보통신부, 산업자원부 등과 같은 국가 기관들은 일반 기업에 데

이터를 요구하지 않고, 분산된 정보 데이터들의 종합 분석 가능하다. 도 8 은 공공 기관 분야에서의 사용 예를 보여주고 있다.

<74> 금융/보험/증권 분야에 있어, 금융결제 위원회, 보험감독원, 금융 감독원, 전경련, 농협 중앙회, 증권 감독원 등과 같은 회원사를 보유하고 있는 주관 기관들과 회원사들과의 정보 유출 없이 정보 공유로 인해 (한 예로, 금융감독위원회/금융감독원의 데이터와 각 은행들의 데이터들 간의 지식 정보융합을 통해) 건전한 신용질서와 공정한 금융거래 관행을 확립한다. 도 9 는 금융/보험/증권 분야에서의 사용 예를 보여주고 있다.

<75> 일반 기업체는 각 대기업들의 계열사 및 이하 하청업체들과의 정보 공유를 통한 안정적인 정보 관리, 각 업체들 간의 전사적 자원을 공유함으로써 효율적인 기업 업무방향을 제시한다. 비즈니스 업계의 종합적인 데이터 분석을 통한 기업 이윤 극대화한다. 도 10 은 일반 기업체에서의 사용 예를 보여주고 있다.

<76> 의료 기관은 환자들의 질병들에 대한 정보를 여러 의료기관들이 서로의 데이터를 가상 공유함으로써 국내에서 급성 유행병 등을 사전에 예방할 수 있으며, 특히 국제 의료 기관들 간의 협력을 통해 SARS와 조류독감 등과 같은 국제적인 질병을 예방할 수 있다.

<77> 기타 응용사례로는, 회사 (또는 그룹)내 단일 시스템, 회사 (또는 그룹)내 단일 시스템에서 서로 다른 정보 시스템 간, 회사 (또는 그룹)내 지역적으로 분산된 시스템, 회원 정보를 확보한 포털사이트 시스템, 동종 업체와 상위 조직 (예; 단위농협과 농협중앙회), 직접적인 협력 조직 (예; 은행과 금융감독원), 협력을 요하는 관계 (예; 은행과 국세청), 유사하나 관련이 없던 관계 (예; 은행과 생명보험사), 서로 관련이 적거나 없는 관계 (예; 관공서와 편의점). 그리고 위에 열거한 관계들의 조합을 통한 관계 등인 것이다.

<78> 본 발명의 지식융합을 위한 분산형 데이터마이닝 에이전트시스템에 대한 기술사상을 예시도면에 의거하여 설명했지만, 이는 본 발명의 가장 양호한 실시예를 예시적으로 설명한 것이지 본 발명의 특허청구범위를 한정하는 것은 아니다.

<79> 본 발명은 이 기술분야의 통상 지식을 가진 자라면 누구나 본 발명의 기술사상의 범주를 이탈하지 않는 범위 내에서 다양한 변형 및 모방이 가능함은 명백한 사실이다.

발명의 효과

<80> 이상 설명한 바와 같이 본 발명의 지식융합을 위한 분산형 데이터마이닝 에이전트시스템에 의하면, 데이터 분석 프로세싱에서 각 에이전트 시스템들을 각 협력업체 시스템에 도입하여 분산된 데이터베이스를 분석함으로 유저 컴퓨터가 분석된 정보를 활용할 수 있어, 분산형 데이터마이닝으로써 철저한 정보보안과 데이터 마이닝의 부하 분산을 기할 수 있는 지식융합을 위한 분산형 데이터마이닝 에이전트시스템을 제공하는 효과가 있다.

도면의 간단한 설명

- <1> 도 1 은 중앙집중식 데이터 마이닝 기술을 설명하기 위한 구성도,
- <2> 도 2 는 본 발명의 실시예에 관한 분산형 데이터마이닝 기술을 설명하기 위한 구성도,
- <3> 도 3 은 본 발명의 지식융합을 위한 분산형 데이터마이닝 에이전트시스템을 설명하기 위한 분산형 데이터마이닝의 알고리즘,
- <4> 도 4 는 분산형 방법을 통해 의사결정 나무가 구축되는 과정도,
- <5> 도 5 는 본 발명의 지식융합을 위한 분산형 데이터마이닝 에이전트시스템의 개요도,
- <6> 도 6 및 도 7 은 본 발명의 지식융합을 위한 분산형 데이터마이닝 에이전트시스템의 실행화면,
- <7> 도 8 은 본 발명의 지식융합을 위한 분산형 데이터마이닝 에이전트시스템에서 공공 기관분야의 예시도,
- <8> 도 9 는 본 발명의 지식융합을 위한 분산형 데이터마이닝 에이전트시스템에서 금융/보험/증권 분야의 예시도,
- <9> 도 10 은 본 발명의 지식융합을 위한 분산형 데이터마이닝 에이전트시스템에서 일반 기업체의 예시도이다.

<10> * 도면의 주요 부분에 대한 부호의 설명

<11> 10 : 데이터 분석프로세싱

<12> 11 - 14 : 에이전트

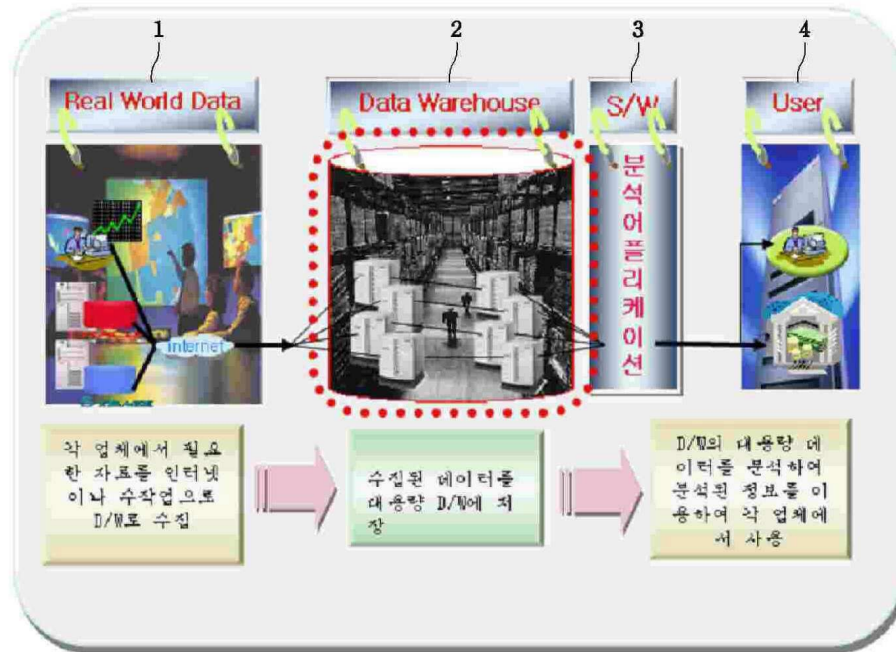
15 : 메디에이터

<13>

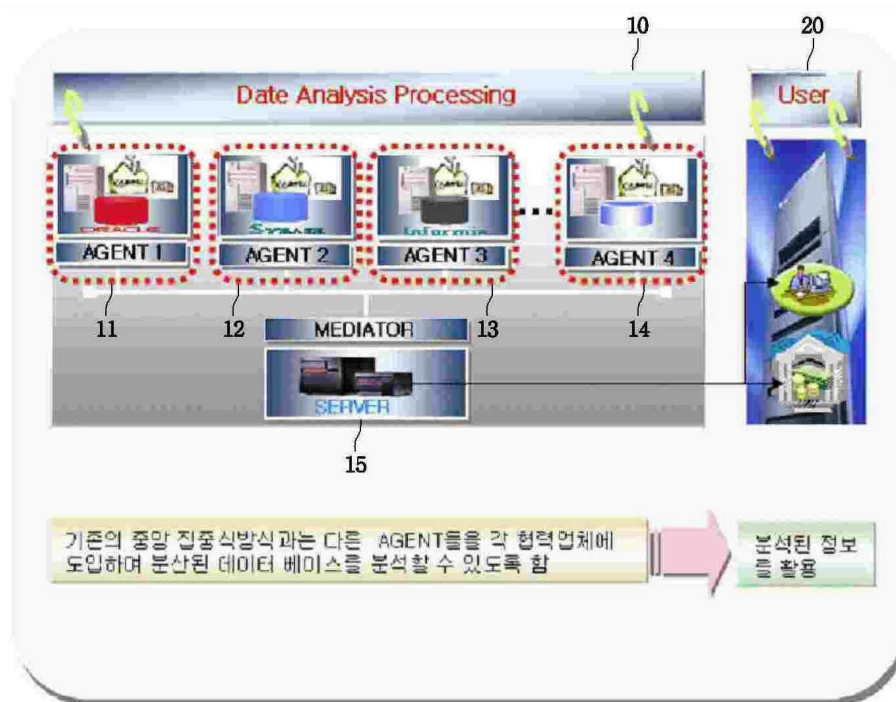
20 : 유저

도면

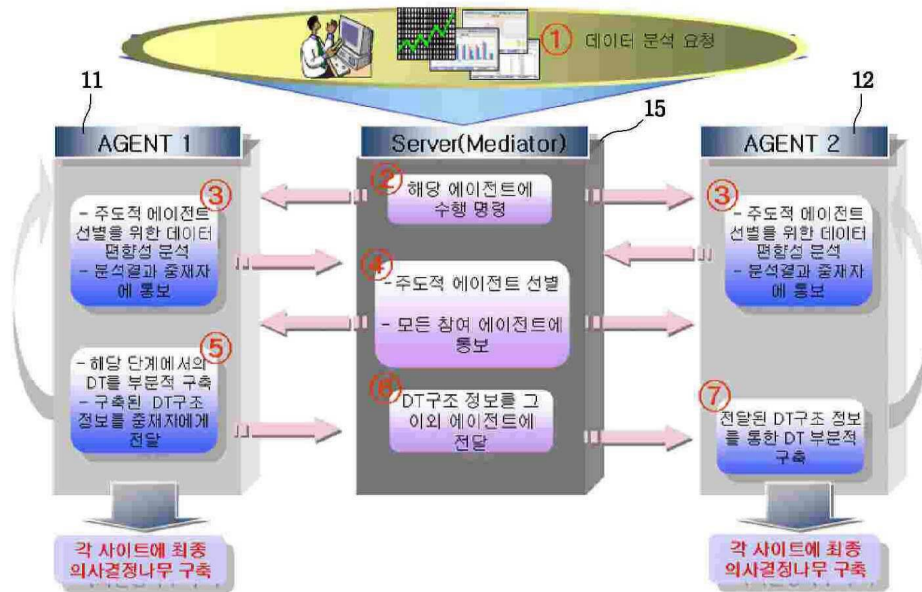
도면1



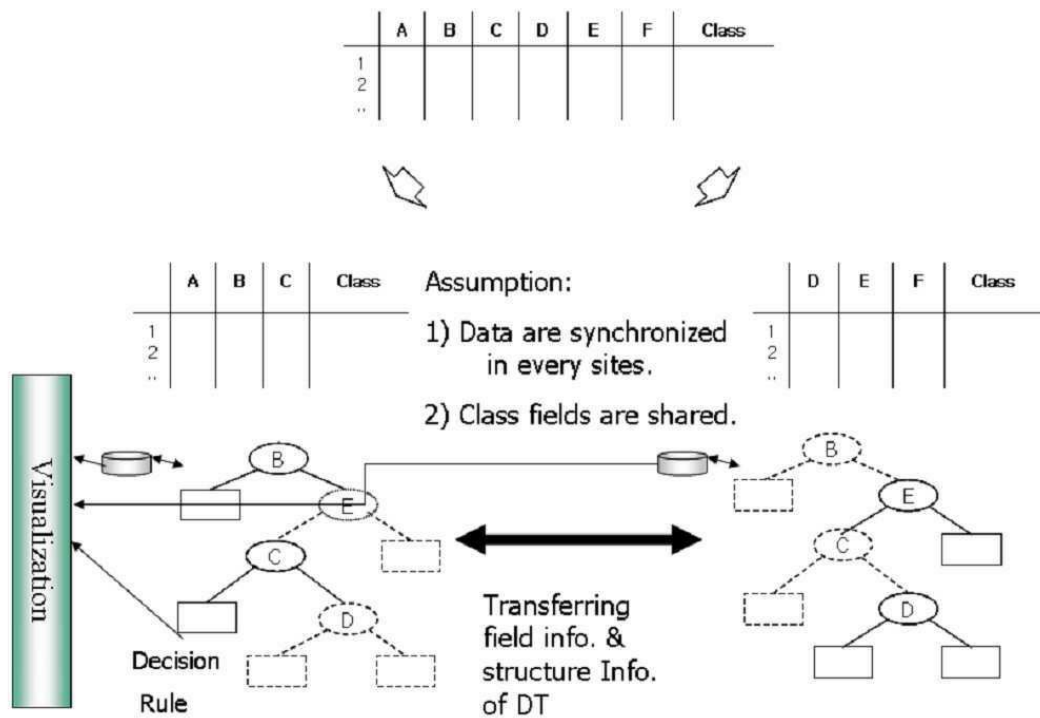
도면2



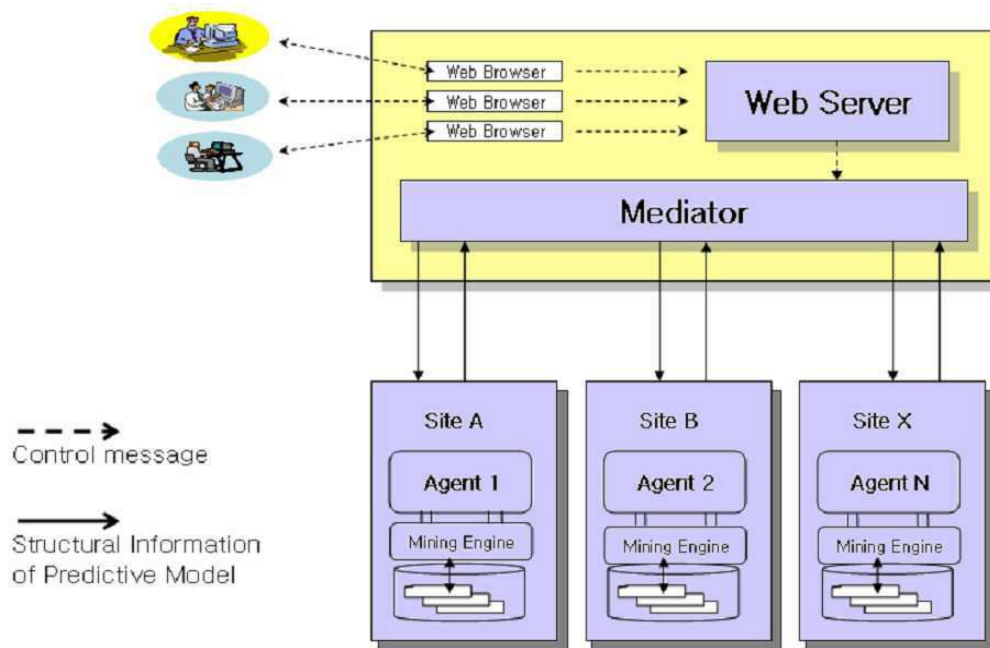
도면3



도면4



도면5



도면6

```

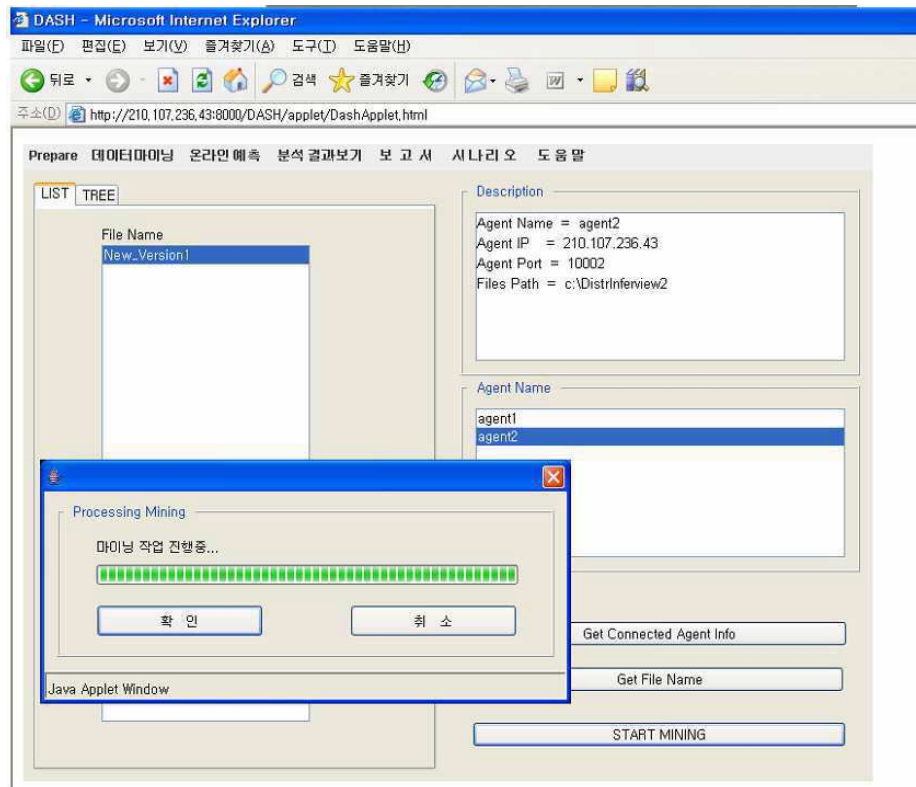
C:\Project\W\ASHWJ2sk1.4.2_02WinWJava.exe
Start Mediator
AgentInfo : //210.107.236.43:10001/agent1
AgentInfo : //210.107.236.43:10002/agent2
Binding mediator Successful!
Start mining thread!!
Start mining thread!!

C:\Project\W\ASHWJ2sk1.4.2_02WinWJava.exe
nt_flag before == 0
0 49
nt_flag after == 0
Decision Tree has been generated.
A set of rules is successfully generated.
The inference process has been successfully completed.
ProcessDone!!
!!!!!!!!!!!! In
RmiMediator =
AgentName = agent1
FileName = c:\
Just Before StartInferEngin d111
Please wait !!!
!!!!!!!!!!!! File
read Successfully!!
Data was successfully read.
nt_flag before == 0
0 49
nt_flag after == 0
Decision Tree has been generated.
A set of rules is successfully generated.
The inference process has been successfully completed.
ProcessDone!!
!!!!!!!!!!!! In
!!!!!!!!!!!! In setDone -> signal -> 1

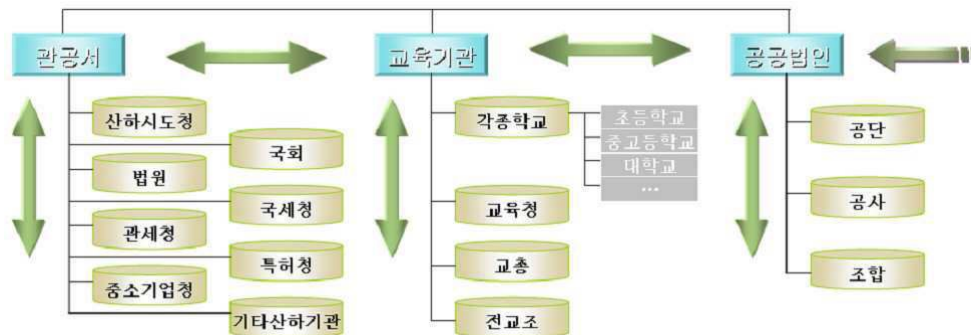
C:\Project\W\ASHWJ2sk1.4.2_02WinWJava.exe
nt_flag after == 0
Decision Tree has been generated.
A set of rules is successfully generated.
The inference process has been successfully completed.
ProcessDone!!
!!!!!!!!!!!! In setDone -> signal -> 1

```

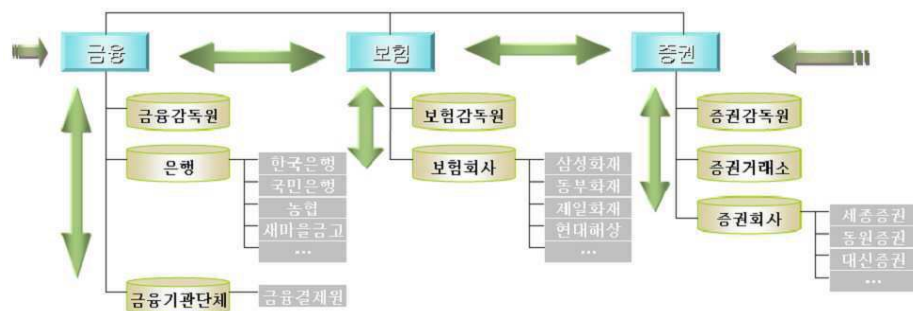
도면7



도면8



도면9



도면10

