



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2019-0077693

(43) 공개일자 2019년07월04일

- (51) 국제특허분류(Int. Cl.)

G06F 16/00 (2019.01)

- (52) CPC특허분류

G06F 16/21 (2019.01)

G06F 16/182 (2019.01)

- (21) 출원번호 10-2017-0179093

- (22) 출원일자 2017년12월26일

심사청구일자 2017년12월26일

- (71) 출원인

경희대학교 산학협력단

경기도 용인시 기흥구 덕영대로 1732 (서천동, 경희대학교 국제캠퍼스내)

- (72) 발명자

취의남

경기도 용인시 기흥구 이현로29번길 86-21 201
동401호(보정동,e편한세상대림아파트)

김용현

경기도 용인시 기흥구 용구대로 1842, 107동 180
1호(보라동, 현대 모닝사이드2차)

고승완

대전광역시 서구 둔산로 155, 104동 209호

- (74) 대리인

김등용, 김홍석

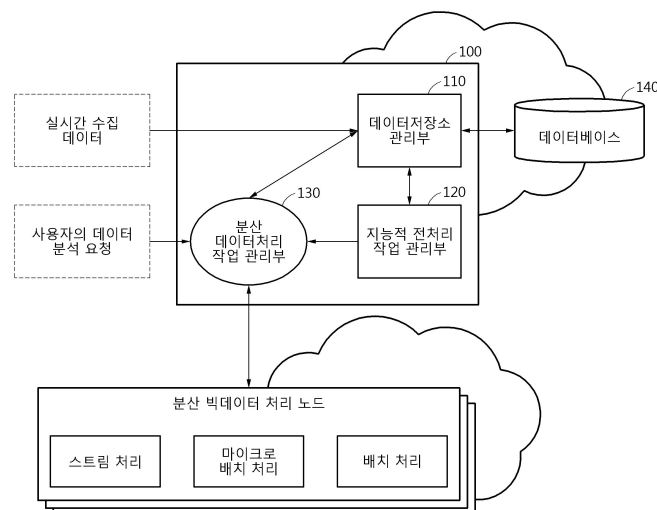
전체 청구항 수 : 총 17 항

- (54) 발명의 명칭 분산 클라우드 환경에서 실시간 데이터 처리의 지연 시간 최소화를 위한 지능적 데이터 전처리 시스템 및 방법

(57) 요약

본 발명은 분산 클라우드 환경에서 실시간 데이터 처리의 지연 시간 최소화를 위한 지능적 데이터 전처리 기술로서, 일실시예에 따른 지능적 데이터 전처리 시스템은 수집되는 데이터에 대한 메타데이터를 생성하고, 상기 데이터와 상기 메타데이터를 연관지어 데이터베이스에 저장하는 데이터 저장소 관리부, 상기 저장된 메타데이터에 대한 통계 분석을 처리하고, 상기 통계 분석 결과에 기초하여 상기 저장된 데이터에 대한 사전 데이터 전처리를 수행하는 지능적 전처리 작업 관리부, 및 상기 전처리가 처리된 데이터 중에서, 입력되는 분석 요구사항에 상응하는 데이터들의 분산 데이터처리를 수행하는 분산 데이터처리 작업 관리부를 포함할 수 있다.

대표도



이 발명을 지원한 국가연구개발사업

과제고유번호 1711055131
 부처명 과학기술정보통신부
 연구관리전문기관 정보통신기술진흥센터
 연구사업명 대학ICT연구센터육성지원사업
 연구과제명 실시간 모바일 클라우드 서비스 플랫폼 연구개발
 기 여 율 3/5
 주관기관 경희대학교 산학협력단
 연구기간 2017.01.01 ~ 2017.12.31

이 발명을 지원한 국가연구개발사업

과제고유번호 2017-0-00093
 부처명 과학기술정보통신부
 연구관리전문기관 정보통신기술진흥센터
 연구사업명 SW중심대학 지원사업
 연구과제명 SW중심대학(경희대)
 기 여 율 2/5
 주관기관 경희대학교 산학협력단
 연구기간 2017.05.01 ~ 2017.12.31

명세서

청구범위

청구항 1

수집되는 데이터에 대한 메타데이터를 생성하고, 상기 데이터와 상기 메타데이터를 연관지어 데이터베이스에 저장하는 데이터 저장소 관리부;

상기 저장된 메타데이터에 대한 통계 분석을 처리하고, 상기 통계 분석 결과에 기초하여 상기 저장된 데이터에 대한 사전 데이터 전처리를 수행하는 지능적 전처리 작업 관리부;

상기 전처리가 처리된 데이터 중에서, 입력되는 분석 요구사항에 상응하는 데이터들의 분산 데이터처리를 수행하는 분산 데이터처리 작업 관리부

를 포함하는 지능적 데이터 전처리 시스템.

청구항 2

제1항에 있어서,

상기 지능적 전처리 작업 관리부는,

상기 저장된 메타데이터에 대해 연관 분석과 클러스터링 분석을 통해 상기 컬럼들의 값 분포를 생성하여, 상기 통계 분석을 처리하고,

상기 통계 분석의 처리 결과에 따라 상기 저장된 데이터에 대한 전처리 우선 순위를 결정하며,

상기 결정된 전처리 우선 순위에 따라 상기 전처리를 수행하는 지능적 데이터 전처리 시스템.

청구항 3

제2항에 있어서,

상기 지능적 전처리 작업 관리부는,

상기 데이터가 양의 상관관계 또는 음의 상관관계를 일정 수치 이상 가질 경우 연관성이 있는 것으로 상기 연관 분석을 처리하는 지능적 데이터 전처리 시스템.

청구항 4

제2항에 있어서,

상기 지능적 전처리 작업 관리부는,

상기 연관 분석의 결과에 따라 연관성이 있는 컬럼에 대해 상기 클러스터링 분석을 처리하되, 클러스터 수를 증가시키며 평균 클러스터 간 거리($D_{Average}$)의 계산하고, 상기 계산된 평균 클러스터 간 거리($D_{Average}$)를 균등 분포 시 클러스터 간 거리(D_{Max})에 대비하여 상기 데이터가 분포된 정도를 판단하는 지능적 데이터 전처리 시스템.

청구항 5

제4항에 있어서,

상기 지능적 전처리 작업 관리부는,

인접한 두 클러스터의 중심점 간 차이 값들을 누적한 값과 클러스터의 개수간 비율에 기초해서 상기 평균 클러스터 간 거리($D_{Average}$)를 계산하는 지능적 데이터 전처리 시스템.

청구항 6

제4항에 있어서,

상기 지능적 전처리 작업 관리부는,

클러스터가 1개일 때의 최대 지름값과 전체의 클러스터 개수의 비율에 기초하여 상기 균등 분포시 클러스터 간 거리(D_{Max})를 계산하는 지능적 데이터 전처리 시스템.

청구항 7

제4항에 있어서,

상기 지능적 전처리 작업 관리부는,

상기 데이터가 분포된 정도에 기초하여 상기 데이터에 대한 전처리 우선 순위를 결정하고, 상기 결정된 전처리 우선 순위에 따라 전처리 후보를 결정하며, 상기 결정된 전처리 후보에 대해 순서대로 전처리를 수행하는 지능적 데이터 전처리 시스템.

청구항 8

제1항에 있어서,

상기 메타데이터는 상기 데이터의 컬럼 정보, 라인 수, 및 컬럼들의 값 분포 중에서 적어도 하나를 포함하는 지능적 데이터 전처리 시스템.

청구항 9

제1항에 있어서,

상기 분산 데이터처리 작업 관리부는,

처리 작업에 대한 워크플로우를 생성하고, 상기 생성된 워크플로우에 따라 작업들을 스케줄링 하되, 상기 스케줄링 하는 작업들을 스트림 처리, 마이크로 배치 처리, 배치 처리 중에서 어느 하나의 처리에 할당하는 지능적 데이터 전처리 시스템.

청구항 10

데이터와 상기 데이터에 상응하는 메타데이터를 기록하고 유지하는 데이터베이스;

상기 메타데이터에 기초하여, 상기 데이터의 연관성과 분포된 정도를 판단하여, 상기 데이터에 대한 전처리 우선 순위를 산출하는 전처리 우선 순위 산출부; 및

상기 산출된 전처리 우선 순위에 기초하여 상기 데이터에 대한 전처리를 수행하는 전처리부를 포함하는 지능적 데이터 전처리 시스템.

청구항 11

제10항에 있어서,

상기 전처리 우선 순위 산출부는,

상기 데이터가 양의 상관관계 또는 음의 상관관계를 일정 수치 이상 가질 경우 연관성이 있는 것으로 판단하고, 상기 연관성이 있는 데이터에 대해서만 분포된 정도를 판단하되, 상기 메타데이터에 포함된 클러스터 수를 증가시키며 평균 클러스터 간 거리(D_{Average})의 계산하고, 상기 계산된 평균 클러스터 간 거리(D_{Average})와 균등 분포시 클러스터 간 거리(D_{Max})의 차이에 대한 절대값을, 균일 분포 임계값(T_{cluster})에 대비하여 상기 데이터가 분포된 정도를 판단하는 지능적 데이터 전처리 시스템.

청구항 12

데이터 저장소 관리부에서, 수집되는 데이터에 대한 메타데이터를 생성하고, 상기 데이터와 상기 메타데이터를 연관지어 데이터베이스에 저장하는 단계;

지능적 전처리 작업 관리부에서, 상기 저장된 메타데이터에 대한 통계 분석을 처리하고, 상기 통계 분석 결과에 기초하여 상기 저장된 데이터에 대한 사전 데이터 전처리를 수행하는 단계;

작업 관리부에서, 상기 전처리가 처리된 데이터 중에서, 입력되는 분석 요구사항에 상응하는 데이터들의 분산 데이터처리를 수행하는 단계

를 포함하는 지능적 데이터 전처리 시스템의 동작 방법.

청구항 13

제12항에 있어서,

상기 사전 데이터 전처리를 수행하는 단계는,

상기 저장된 메타데이터에 대해 연관 분석과 클러스터링 분석을 통해 상기 컬럼들의 값 분포를 생성하여, 상기 통계 분석을 처리하는 단계;

상기 통계 분석의 처리 결과에 따라 상기 저장된 데이터에 대한 전처리 우선 순위를 결정하는 단계; 및

상기 결정된 전처리 우선 순위에 따라 상기 전처리를 수행하는 단계

를 포함하는 지능적 데이터 전처리 시스템의 동작 방법.

청구항 14

제13항에 있어서,

상기 통계 분석을 처리하는 단계는,

상기 데이터가 양의 상관관계 또는 음의 상관관계를 일정 수치 이상 가질 경우 연관성이 있는 것으로 상기 연관 분석을 처리하는 단계

를 포함하는 지능적 데이터 전처리 시스템의 동작 방법.

청구항 15

제13항에 있어서,

상기 통계 분석을 처리하는 단계는,

상기 연관 분석의 결과에 따라 연관성이 있는 컬럼에 대해 상기 클러스터링 분석을 처리하되, 클러스터 수를 증가시키며 평균 클러스터 간 거리($D_{Average}$)의 계산하는 단계; 및

상기 계산된 평균 클러스터 간 거리($D_{Average}$)를 균등 분포시 클러스터 간 거리(D_{Max})에 대비하여 상기 데이터가 분포된 정도를 판단하는 단계

를 포함하는 지능적 데이터 전처리 시스템의 동작 방법.

청구항 16

제15항에 있어서,

상기 통계 분석을 처리하는 단계는,

인접한 두 클러스터의 중심점 간 차이 값들을 누적한 값과 클러스터의 개수간 비율에 기초해서 상기 평균 클러스터 간 거리($D_{Average}$)를 계산하는 단계; 및

클러스터가 1개일 때의 최대 지름값과 전체의 클러스터 개수의 비율에 기초하여 상기 균등 분포시 클러스터 간 거리(D_{Max})를 계산하는 단계

를 더 포함하는 지능적 데이터 전처리 시스템의 동작 방법.

청구항 17

제12항에 있어서,

분산 데이터처리 작업 관리부에서, 상기 전처리가 처리된 데이터 중에서, 입력되는 분석 요구사항에 상응하는

데이터들의 분산 데이터처리를 수행하는 단계

를 더 포함하고,

상기 분산 데이터처리를 수행하는 단계는,

처리 작업에 대한 워크플로우를 생성하는 단계; 및

상기 생성된 워크플로우에 따라 작업들을 스케줄링 하되, 상기 스케줄링 하는 작업들을 스트림 처리, 마이크로 배치 처리, 배치 처리 중에서 어느 하나의 처리에 할당하는 단계

를 포함하는 지능적 데이터 전처리 시스템의 동작 방법.

발명의 설명

기술 분야

- [0001] 본 발명은 분산 클라우드 환경에서 실시간 데이터 처리의 지연 시간을 최소화하는 기술로서, 구체적으로는 분산 처리 환경에서 실시간 분석 요청의 요구사항에 따라 지능적으로 사전 전처리 작업을 스케줄링함으로써, 사용자 요청의 전체 처리 시간의 지연 시간을 최소화하는 기술적 사상에 관한 것이다.

배경 기술

- [0002] 데이터 전처리 과정은 데이터 분석을 위해서는 필수적으로 수행되어야 하는 단계로 데이터 전처리는 수집되는 데이터의 저장에서 실시간 빅데이터 분석에 이르기까지 다양한 처리 과정에서 데이터 정형화, 노이즈 제거, 샘플링 등으로 수행될 수 있다.
- [0003] 데이터 전처리를 위한 기존 기술들은 데이터가 수집될 때 데이터 전처리를 즉각 수행하기 위한 방법과 데이터의 분할 또는 병렬 처리를 통해 전처리에 소요되는 시간을 줄이기 위한 방법들이 주를 이룬다.
- [0004] 하지만 데이터 분석을 위해 요구되는 데이터의 크기가 급증하고 있기 때문에 데이터 전처리의 성능 향상으로는 실시간 처리를 요구하는 사용자들을 만족시키기에 한계점이 존재한다. 데이터를 처리함에 있어 사용자들의 요구사항은 크게 속도(처리 속도)와 품질(데이터의 크기)로 구분되고, 실시간으로 데이터가 수집되는 환경에서 사용자들이 서로 다른 요구사항으로 데이터 처리를 요청한다.
- [0005] 데이터 처리에 요구되는 데이터가 스트림 데이터(Hot Data)와 대용량 히스토리 데이터(Cold Data)로 구분될 때, 스트림 데이터의 경우 데이터의 크기가 작고 처리 속도가 빠르지만 데이터의 분석 결과에 있어 대용량 히스토리 데이터에 비해 품질이 낮게 판단될 수 있다. 반면, 대용량 히스토리 데이터는 크기가 크고 처리 속도가 느리지만 실시간으로 수집되는 데이터를 반영할 수 없는 문제점이 존재한다. 따라서 사용자의 분석 요구사항에 대해 두 가지 데이터가 모두 분석될 필요가 있을 경우, 기존 리액티브(Reactive) 방식의 데이터 처리 방법에서 매 요구사항에 대한 속도와 품질 요구사항의 최적점(Tradeoff)을 찾아 데이터 처리를 수행하기는 한계점이 존재한다.

선행기술문헌

특허문헌

- [0006] (특허문헌 0001) 한국특허출원 제2016-0008634호 "빅데이터 처리 시스템 및 처리 방법"
- (특허문헌 0002) 한국특허출원 제2015-0189903호 "멀티 센서 발생 데이터의 고성능 데이터 처리를 위한 하둡 기반의 데이터 분석 시스템"

발명의 내용

해결하려는 과제

- [0007] 본 발명은 빅데이터 분석을 요구하는 산업 전반에서 다양한 사용자 및 시스템 요소들을 고려한 최적의 워크플로우를 생성하고 사용자 요구사항을 충족시켜 서비스 품질을 높이는 것을 목적으로 한다.
- [0008] 본 발명은 통계적 분석 방법을 이용한 프로액티브(Proactive) 방식의 데이터 전처리를 수행하는 것을 목적으로

한다.

[0009] 본 발명은 사용자 요구사항에 따른 지능적 데이터 전처리를 수행하여 실시간 데이터 처리의 지연시간을 최소화하는 것을 목적으로 한다.

과제의 해결 수단

[0010] 일실시예에 따른 지능적 데이터 전처리 시스템은 수집되는 데이터에 대한 메타데이터를 생성하고, 상기 데이터와 상기 메타데이터를 연관지어 데이터베이스에 저장하는 데이터 저장소 관리부, 상기 저장된 메타데이터에 대한 통계 분석을 처리하고, 상기 통계 분석 결과에 기초하여 상기 저장된 데이터에 대한 사전 데이터 전처리를 수행하는 지능적 전처리 작업 관리부, 및 상기 전처리가 처리된 데이터 중에서, 입력되는 분석 요구사항에 상응하는 데이터들의 분산 데이터처리를 수행하는 분산 데이터처리 작업 관리부를 포함할 수 있다.

[0011] 일실시예에 따른 상기 지능적 전처리 작업 관리부는, 상기 저장된 메타데이터에 대해 연관 분석과 클러스터링 분석을 통해 상기 컬럼들의 값 분포를 생성하여, 상기 통계 분석을 처리하고, 상기 통계 분석의 처리 결과에 따라 상기 저장된 데이터에 대한 전처리 우선 순위를 결정하며, 상기 결정된 전처리 우선 순위에 따라 상기 전처리를 수행할 수 있다.

[0012] 일실시예에 따른 상기 지능적 전처리 작업 관리부는, 상기 데이터가 양의 상관관계 또는 음의 상관관계를 일정 수치 이상 가질 경우 연관성이 있는 것으로 상기 연관 분석을 처리할 수 있다.

[0013] 일실시예에 따른 상기 지능적 전처리 작업 관리부는, 상기 연관 분석의 결과에 따라 연관성이 있는 컬럼에 대해 상기 클러스터링 분석을 처리하되, 클러스터 수를 증가시키며 평균 클러스터 간 거리($D_{Average}$)의 계산하고, 상기 계산된 평균 클러스터 간 거리($D_{Average}$)를 균등 분포시 클러스터 간 거리(D_{Max})에 대비하여 상기 데이터가 분포된 정도를 판단할 수 있다.

[0014] 일실시예에 따른 상기 지능적 전처리 작업 관리부는, 인접한 두 클러스터의 중심점 간 차이 값들을 누적한 값과 클러스터의 개수간 비율에 기초해서 상기 평균 클러스터 간 거리($D_{Average}$)를 계산할 수 있다.

[0015] 일실시예에 따른 상기 데이터 저장소 관리부는, 클러스터가 1개일 때의 최대 지름값과 전체의 클러스터 개수의 비율에 기초하여 상기 균등 분포시 클러스터 간 거리(D_{Max})를 계산할 수 있다.

[0016] 일실시예에 따른 상기 지능적 전처리 작업 관리부는, 상기 데이터가 분포된 정도에 기초하여 상기 데이터에 대한 전처리 우선 순위를 결정하고, 상기 결정된 전처리 우선 순위에 따라 전처리 후보를 결정하며, 상기 결정된 전처리 후보에 대해 순서대로 전처리를 수행할 수 있다.

[0017] 일실시예에 따른 상기 메타데이터는 상기 데이터의 컬럼 정보, 라인 수, 및 컬럼들의 값 분포 중에서 적어도 하나를 포함할 수 있다.

[0018] 일실시예에 따른 상기 분산 데이터처리 작업 관리부는, 처리 작업에 대한 워크플로우를 생성하고, 상기 생성된 워크플로우에 따라 작업들을 스케줄링 하되, 상기 스케줄링 하는 작업들을 스트림 처리, 마이크로 배치 처리, 배치 처리 중에서 어느 하나의 처리에 할당할 수 있다.

[0019] 일실시예에 따른 지능적 데이터 전처리 시스템은 데이터와 상기 데이터에 상응하는 메타데이터를 기록하고 유지하는 데이터베이스, 상기 메타데이터에 기초하여, 상기 데이터의 연관성과 분포된 정도를 판단하여, 상기 데이터에 대한 전처리 우선 순위를 산출하는 전처리 우선 순위 산출부, 및 상기 산출된 전처리 우선 순위에 기초하여 상기 데이터에 대한 전처리를 수행하는 전처리부를 포함할 수 있다.

[0020] 일실시예에 따른 상기 전처리 우선 순위 산출부는, 상기 데이터가 양의 상관관계 또는 음의 상관관계를 일정 수치 이상 가질 경우 연관성이 있는 것으로 판단하고, 상기 연관성이 있는 데이터에 대해서만 분포된 정도를 판단하되, 상기 메타데이터에 포함된 클러스터 수를 증가시키며 평균 클러스터 간 거리($D_{Average}$)의 계산하고, 상기 계산된 평균 클러스터 간 거리($D_{Average}$)와 균등 분포시 클러스터 간 거리(D_{Max})의 차이에 대한 절대값을, 균일 분포 임계값($T_{cluster}$)에 대비하여 상기 데이터가 분포된 정도를 판단할 수 있다.

[0021] 일실시예에 따른 지능적 데이터 전처리 시스템의 동작 방법은 데이터 저장소 관리부에서, 수집되는 데이터에 대한 메타데이터를 생성하고, 상기 데이터와 상기 메타데이터를 연관지어 데이터베이스에 저장하는 단계, 지능적 전처리 작업 관리부에서, 상기 저장된 메타데이터에 대한 통계 분석을 처리하고, 상기 통계 분석 결과에 기초하여

여 상기 저장된 데이터에 대한 사전 데이터 전처리를 수행하는 단계, 및 작업 관리부에서, 상기 전처리가 처리된 데이터 중에서, 입력되는 분석 요구사항에 상응하는 데이터들의 분산 데이터처리를 수행하는 단계를 포함할 수 있다.

[0022] 일실시예에 따른 상기 사전 데이터 전처리를 수행하는 단계는, 상기 저장된 메타데이터에 대해 연관 분석과 클러스터링 분석을 통해 상기 컬럼들의 값 분포를 생성하여, 상기 통계 분석을 처리하는 단계, 상기 통계 분석의 처리 결과에 따라 상기 저장된 데이터에 대한 전처리 우선 순위를 결정하는 단계, 및 상기 결정된 전처리 우선 순위에 따라 상기 전처리를 수행하는 단계를 포함할 수 있다.

[0023] 일실시예에 따른 상기 통계 분석을 처리하는 단계는, 상기 데이터가 양의 상관관계 또는 음의 상관관계를 일정 수치 이상 가질 경우 연관성이 있는 것으로 상기 연관 분석을 처리하는 단계를 포함할 수 있다.

[0024] 일실시예에 따른 상기 통계 분석을 처리하는 단계는, 상기 연관 분석의 결과에 따라 연관성이 있는 컬럼에 대해 상기 클러스터링 분석을 처리하되, 클러스터 수를 증가시키며 평균 클러스터 간 거리($D_{Average}$)의 계산하는 단계, 및 상기 계산된 평균 클러스터 간 거리($D_{Average}$)를 균등 분포시 클러스터 간 거리(D_{Max})에 대비하여 상기 데이터가 분포된 정도를 판단하는 단계를 포함할 수 있다.

[0025] 일실시예에 따른 상기 통계 분석을 처리하는 단계는, 인접한 두 클러스터의 중심점 간 차이 값들을 누적한 값과 클러스터의 개수간 비율에 기초해서 상기 평균 클러스터 간 거리($D_{Average}$)를 계산하는 단계, 및 클러스터가 1개일 때의 최대 지름값과 전체의 클러스터 개수의 비율에 기초하여 상기 균등 분포시 클러스터 간 거리(D_{Max})를 계산하는 단계를 더 포함할 수 있다.

[0026] 일실시예에 따른 지능적 데이터 전처리 시스템의 동작 방법은 분산 데이터처리 작업 관리부에서, 상기 전처리가 처리된 데이터 중에서, 입력되는 분석 요구사항에 상응하는 데이터들의 분산 데이터처리를 수행하는 단계를 더 포함하고, 상기 분산 데이터처리를 수행하는 단계는, 처리 작업에 대한 워크플로우를 생성하는 단계, 및 상기 생성된 워크플로우에 따라 작업들을 스케줄링 하되, 상기 스케줄링 하는 작업들을 스트림 처리, 마이크로 배치 처리, 배치 처리 중에서 어느 하나의 처리에 할당하는 단계를 포함할 수 있다.

발명의 효과

[0027] 일실시예에 따르면, 빅데이터 분석을 요구하는 산업 전반에서 다양한 사용자 및 시스템 요소들을 고려한 최적의 워크플로우를 생성하고 사용자 요구사항을 충족시켜 서비스 품질을 높일 수 있다.

[0028] 일실시예에 따르면, 통계적 분석 방법을 이용한 프로액티브(Proactive) 방식의 데이터 전처리를 수행할 수 있다.

[0029] 일실시예에 따르면, 사용자 요구사항에 따른 지능적 데이터 전처리를 수행하여 실시간 데이터 처리의 지연시간을 최소화 할 수 있다.

도면의 간단한 설명

[0030] 도 1은 일실시예에 따른 지능적 데이터 전처리 시스템을 설명하는 도면이다.

도 2는 평균 클러스터 간 거리 측정 결과를 나타내는 도면이다.

도 3은 전처리 여부에 따른 성능 비교를 설명하는 도면이다.

도 4는 다른 일실시예에 따른 지능적 데이터 전처리 시스템을 설명하는 도면이다.

도 5는 일실시예에 따른 지능적 데이터 전처리 시스템의 동작 방법을 설명하는 도면이다.

발명을 실시하기 위한 구체적인 내용

[0031] 본 명세서에 개시되어 있는 본 발명의 개념에 따른 실시예들에 대해서 특정한 구조적 또는 기능적 설명들은 단지 본 발명의 개념에 따른 실시예들을 설명하기 위한 목적으로 예시된 것으로서, 본 발명의 개념에 따른 실시예들은 다양한 형태로 실시될 수 있으며 본 명세서에 설명된 실시예들에 한정되지 않는다.

[0032] 본 발명의 개념에 따른 실시예들은 다양한 변경들을 가할 수 있고 여러 가지 형태들을 가질 수 있으므로 실시예들을 도면에 예시하고 본 명세서에 상세하게 설명하고자 한다. 그러나, 이는 본 발명의 개념에 따른 실시예들

을 특정한 개시형태들에 대해 한정하려는 것이 아니며, 본 발명의 사상 및 기술 범위에 포함되는 변경, 균등물, 또는 대체물을 포함한다.

- [0033] 제1 또는 제2 등의 용어를 다양한 구성요소들을 설명하는데 사용될 수 있지만, 상기 구성요소들은 상기 용어들에 의해 한정되어서는 안 된다. 상기 용어들은 하나의 구성요소를 다른 구성요소로부터 구별하는 목적으로만, 예를 들어 본 발명의 개념에 따른 권리 범위로부터 이탈되지 않은 채, 제1 구성요소는 제2 구성요소로 명명될 수 있고, 유사하게 제2 구성요소는 제1 구성요소로도 명명될 수 있다.
- [0034] 어떤 구성요소가 다른 구성요소에 "연결되어" 있다거나 "접속되어" 있다고 언급된 때에는, 그 다른 구성요소에 직접적으로 연결되어 있거나 또는 접속되어 있을 수도 있지만, 중간에 다른 구성요소가 존재할 수도 있다고 이해되어야 할 것이다. 반면에, 어떤 구성요소가 다른 구성요소에 "직접 연결되어" 있다거나 "직접 접속되어" 있다고 언급된 때에는, 중간에 다른 구성요소가 존재하지 않는 것으로 이해되어야 할 것이다. 구성요소들 간의 관계를 설명하는 표현들, 예를 들어 "~사이에"와 "바로~사이에" 또는 "~에 직접 이웃하는" 등도 마찬가지로 해석되어야 한다.
- [0035] 본 명세서에서 사용한 용어는 단지 특정한 실시예들을 설명하기 위해 사용된 것으로, 본 발명을 한정하려는 의도가 아니다. 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 본 명세서에서, "포함하다" 또는 "가지다" 등의 용어는 실시된 특징, 숫자, 단계, 동작, 구성요소, 부분품 또는 이들을 조합한 것이 존재함으로 지정하려는 것이지, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성요소, 부분품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다.
- [0036] 다르게 정의되지 않는 한, 기술적이거나 과학적인 용어를 포함해서 여기서 사용되는 모든 용어들은 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에 의해 일반적으로 이해되는 것과 동일한 의미를 가진다. 일반적으로 사용되는 사전에 정의되어 있는 것과 같은 용어들은 관련 기술의 문맥상 가지는 의미와 일치하는 의미를 갖는 것으로 해석되어야 하며, 본 명세서에서 명백하게 정의하지 않는 한, 이상적이거나 과도하게 형식적인 의미로 해석되지 않는다.
- [0038] 이하, 실시예들을 첨부된 도면을 참조하여 상세하게 설명한다. 그러나, 특허출원의 범위가 이러한 실시예들에 의해 제한되거나 한정되는 것은 아니다. 각 도면에 제시된 동일한 참조 부호는 동일한 부재를 나타낸다.
- [0039] 도 1은 일실시예에 따른 지능적 데이터 전처리 시스템(100)을 설명하는 도면이다.
- [0040] 실시간으로 수집되는 IoT, 로그, 트랜잭션 등의 다양한 스트림 데이터들은 일반적으로 과거에 수집되어 이미 저장되어있는 데이터와 함께 처리/분석되어 비즈니스적 의미를 갖는다. 따라서 실시간 수집되는 스트림 데이터와 저장되어 있는 대용량의 데이터를 함께 처리할 경우 대용량 데이터 처리에 의한 전체 처리 시간의 지연 시간을 최소화할 필요가 있다.
- [0041] 이를 위해 지능적 데이터 전처리 시스템(100)은 분산 클라우드 환경에서 지능적 데이터 처리를 수행하며, 데이터 저장소 관리부(110), 지능적 전처리 작업 관리부(120), 및 분산 데이터처리 작업 관리부(130)를 포함할 수 있다.
- [0042] 먼저, 일실시예에 따른 데이터 저장소 관리부(110)는 수집되는 데이터에 대한 메타데이터를 생성하고, 데이터와 메타데이터를 연관지어 데이터베이스(140)에 저장할 수 있다.
- [0043] 데이터베이스(140)에는 실시간으로 수집되는 데이터와 각 데이터의 메타데이터가 저장된다. 이를 위해, 데이터 저장소 관리부(110)는 실시간으로 수집되는 데이터의 저장과 함께 메타데이터를 생성할 수 있다. 메타데이터는 데이터의 컬럼 정보, 라인 수, 및 컬럼들의 값 분포 중에서 적어도 하나를 포함할 수 있다.
- [0044] 일실시예에 따른 지능적 전처리 작업 관리부(120)는 저장된 메타데이터에 대한 통계 분석을 처리할 수 있다. 또한, 통계 분석 결과에 기초하여 상기 저장된 데이터에 대한 사전 데이터 전처리를 수행할 수 있다.
- [0045] 지능적 전처리 작업 관리부(120)는 저장된 메타데이터를 통계 분석하여 실시간 빅데이터 처리를 위한 사전 데이터 전처리 작업을 생성할 수 있다.
- [0046] 사용자가 요구하는 작업 완료 시간이 처리 예상 시간보다 짧은 경우 작업 실패 확률이 증가되어 원활한 서비스 제공이 어렵게 된다. 따라서 지능적 전처리 작업 관리부(120)에서는 데이터 저장소 관리부(110)로부터 전달되는 메타데이터를 통계 분석하여 데이터베이스(140)에 저장된 대용량 데이터를 사전에 전처리할 수 있다. 또한,

이로써, 사용자의 분석 요구사항에 해당되는 데이터 처리 성능을 기대할 수 있다.

- [0047] 보다 구체적으로, 일실시예에 따른 지능적 전처리 작업 관리부(120)는 전처리의 필요성을 판단하기 위해 메타데이터에 대한 통계 분석을 수행할 수 있다.
- [0048] 일실시예에 따른 메타데이터는 대용량 데이터의 통계적 정보를 포함하고 있으며, 예를 들어 데이터의 컬럼 정보, 라인 수, 컬럼들의 값 분포 등의 분석 결과가 메타데이터에 포함될 수 있다.
- [0049] 일실시예에 따른 지능적 전처리 작업 관리부(120)는 컬럼들의 값 분포를 통해 데이터의 연관성이 있는지 여부를 판단할 수 있다. 즉, 일실시예에 따른 지능적 전처리 작업 관리부(120)는 저장된 메타데이터에 대해 연관 분석과 클러스터링 분석을 통해 컬럼들의 값 분포를 생성함으로써, 통계 분석을 처리할 수 있다.
- [0050] 일례로, 지능적 전처리 작업 관리부(120)는 데이터가 양의 상관관계 또는 음의 상관관계를 일정 수치 이상 가질 경우 연관성이 있는 것으로 연관 분석을 처리할 수 있다.
- [0051] 일실시예에 따른 지능적 전처리 작업 관리부(120)는 연관 분석의 결과에 따라 연관성이 있는 컬럼에 대해 상기 클러스터링 분석을 처리할 수 있다.
- [0052] 이때, 일실시예에 따른 지능적 전처리 작업 관리부(120)는 클러스터 수를 증가시키며 평균 클러스터 간 거리($D_{Average}$)의 계산할 수 있다.
- [0053] 또한, 일실시예에 따른 지능적 전처리 작업 관리부(120)는 계산된 평균 클러스터 간 거리($D_{Average}$)를 균등 분포시 클러스터 간 거리(D_{Max})에 대비하여 데이터가 분포된 정도를 판단할 수 있다.
- [0054] 일실시예에 따른 지능적 전처리 작업 관리부(120)는 인접한 두 클러스터의 중심점 간 차이 값들을 누적한 값과 클러스터의 개수간 비율에 기초해서 상기 평균 클러스터 간 거리($D_{Average}$)를 계산할 수 있다.
- [0055] 보다 구체적으로, 일실시예에 따른 지능적 전처리 작업 관리부(120)는 [수학식 1]을 통해 연관 분석의 결과 연관성이 있는 컬럼에 대해 클러스터링 분석을 수행하고 클러스터 수를 늘려가며 평균 클러스터 간 거리 계산으로 데이터가 균일한 분포를 가지는지 판단할 수 있다.

[0057] [수학식 1]

$$D_{Average} = \frac{\sum_2^m (C_n - C_{n-1})}{m - 1}$$

[0060] [수학식 1]에서, $D_{Average}$ 는 평균 클러스터 간 거리, C_n 은 n번째 클러스터의 중심점, m 은 클러스터 개수를 나타낸다.

[0061] 한편, 일실시예에 따른 지능적 전처리 작업 관리부(120)는 클러스터가 1개일 때의 최대 지름값과 전체의 클러스터 개수의 비율에 기초하여 상기 균등 분포 시 클러스터 간 거리(D_{Max})를 계산할 수 있다.

[0062] 보다 구체적으로, 일실시예에 따른 지능적 전처리 작업 관리부(120)는 아래 [수학식 2]를 통해 균등 분포시 클러스터 간 거리(D_{Max})를 계산할 수 있다.

[0064] [수학식 2]

$$D_{Max} = \frac{D_0}{m - 1}$$

- [0067] [수학식 2]에서, D_{Max} 는 클러스터 개수가 m 개일 때, 균등 분포 시 클러스터 간 거리, D_0 는 클러스터 개수가 1개일 때의 최대 지름을 나타낼 수 있다.
- [0068] 일실시예에 따른 지능적 전처리 작업 관리부(120)는 데이터가 분포된 정도에 기초하여 상기 데이터에 대한 전처리 우선 순위를 결정할 수 있다.
- [0069] 일례로, 지능적 전처리 작업 관리부(120)는 연관성이 있는 데이터에 대해서만 분포된 정도를 판단하고 이를 근거로 전처리 우선 순위를 결정할 수 있다.
- [0070] 지능적 전처리 작업 관리부(120)는 아래 [수학식 3]과 같이 계산된 평균 클러스터 간 거리($D_{Average}$)와 균등 분포 시 클러스터 간 거리(D_{Max})의 차이에 대한 절대값을, 균일 분포 임계값($T_{cluster}$)에 대비하여 데이터가 분포된 정도를 판단할 수 있다.
- [0072] [수학식 3]
- [0073] $|D_{Max} - D_{Average}| \leq T_{Cluster}$
- [0075] [수학식 3]에서, $T_{Cluster}$ 는 균일한 분포 여부에 대한 임계값을 나타낼 수 있다.
- [0076] 이후, 일실시예에 따른 지능적 전처리 작업 관리부(120)는 통계 분석의 처리 결과에 따라 저장된 데이터에 대한 전처리 우선 순위를 결정하고, 결정된 전처리 우선 순위에 따라 데이터에 대한 전처리를 수행할 수 있다.
- [0077] 일실시예에 따른 분산 데이터처리 작업 관리부(130)는 전처리가 처리된 데이터 중에서, 입력되는 분석 요구사항에 상응하는 데이터들의 분산 데이터처리를 수행할 수 있다.
- [0078] 일실시예에 따른 분산 데이터처리 작업 관리부(130)는 모든 처리 작업에 대한 워크플로우를 생성하고, 작업 스케줄링을 수행하여, 해당 작업을 분산 빅데이터 처리 노드를 통해 스트림 처리, 마이크로 배치 처리, 배치 처리에 각각 할당할 수 있다.
- [0079] 도 2는 평균 클러스터 간 거리 측정 결과를 나타내는 실시예(200)이다.
- [0080] 도 2에서 보는 바와 같이, 일실시예에 따른 지능적 데이터 전처리 시스템은 각 클러스터 간 거리를 측정하고 이에 대한 평균값으로 평균 클러스터 간 거리를 측정할 수 있다.
- [0081] 일실시예에 따른 지능적 데이터 전처리 시스템은 [수학식 1]을 통해 각 클러스터 간 거리의 평균값을 산출할 수 있고, [수학식 2]를 통해 데이터를 클러스터 개수 1개로 클러스터링을 수행했을 때, 나타나는 클러스터의 지름을 산출할 수 있다.
- [0082] 또한, 일실시예에 따른 지능적 데이터 전처리 시스템은 균등 분포 시 클러스터 간 거리에서 평균 클러스터 간 거리의 차에 대한 임계값 비교를 수행하기 위해, [수학식 3]을 통해 균일한 분포 여부의 임계값을 비교할 수 있다.
- [0083] 이를 통해 일실시예에 따른 지능적 데이터 전처리 시스템은 균등하게 데이터가 분포된 정도에 따라 순위를 정하고 데이터 전처리 후보를 선정할 수 있다.
- [0084] 도 3은 전처리 여부에 따른 성능 비교를 설명하는 도면(300)이다.
- [0085] 도 3의 경우 전처리를 미리 수행해둔 경우의 데이터와 전처리를 수행하지 않은 데이터의 전체 처리 시간을 비교한 도면(300)이다.
- [0086] 전처리의 종류, 데이터의 크기와 종류 등에 따라 차이가 발생할 수 있으나, 도 3를 통해 본 발명에서 제시하는 지능적으로 전처리를 미리 수행하는 방법에 대한 성능 개선 가능성을 확인할 수 있다. 이 외에도 분석 결과에 따라 전처리 우선순위를 정하고 순서대로 전처리를 미리 수행하기 때문에 데이터 처리 단계에서 우선순위가 낮은 경우 전처리가 수행되어 있지 않아 도 3의 전처리가 되어 있지 않은 경우와 같이 처리 시간이 증가될 수 있다.

- [0087] 하지만 도 2와 같은 지능적 전처리 우선순위 방법에 따라 전처리가 미리 되어 있는 가능성을 높이고 전체 처리 수행 시간을 단축시킬 수 있다.
- [0088] 따라서 일실시예에 따른 지능적 데이터 전처리 시스템은 수행한 데이터 전처리로 실제 데이터 분석 요구 시 사용자가 요구하는 대용량 데이터에 대해 미리 준비하여 분석 성능을 높이고 사용자의 성능 요구사항을 충족시킬 수 있다.
- [0089] 도 4는 다른 일실시예에 따른 지능적 데이터 전처리 시스템(400)을 설명하는 도면이다.
- [0090] 일실시예에 따른 지능적 데이터 전처리 시스템(400)은 실시간 수집되는 스트림 데이터와 저장되어 있는 대용량의 데이터를 처리하는데 필요한 전체 처리 시간의 지연 시간을 최소화할 수 있다.
- [0091] 이를 위해, 일실시예에 따른 지능적 데이터 전처리 시스템(400)은 데이터베이스(410), 전처리 우선 순위 산출부(420), 및 전처리부(430)를 포함할 수 있다.
- [0092] 일실시예에 따른 데이터베이스(410)는 데이터와 상기 데이터에 상응하는 메타데이터를 기록하고 유지할 수 있다.
- [0093] 일실시예에 따른 전처리 우선 순위 산출부(420)는 메타데이터에 기초하여, 상기 데이터의 연관성과 분포된 정도를 판단하여, 상기 데이터에 대한 전처리 우선 순위를 산출할 수 있다.
- [0094] 일례로, 전처리 우선 순위 산출부(420)는 데이터가 양의 상관관계 또는 음의 상관관계를 일정 수치 이상 가질 경우 연관성이 있는 것으로 판단할 수 있다.
- [0095] 또한, 전처리 우선 순위 산출부(420)는 연관성이 있는 데이터에 대해서만 분포된 정도를 판단하되, 메타데이터에 포함된 클러스터 수를 증가시키며 평균 클러스터 간 거리($D_{Average}$)의 계산할 수 있다. 또한, 전처리 우선 순위 산출부(420)는 계산된 평균 클러스터 간 거리($D_{Average}$)와 균등 분포시 클러스터 간 거리(D_{Max})의 차이에 대한 절대값을 산출할 수 있다. 또한, 전처리 우선 순위 산출부(420)는 상기 산출된 절대값을 균일 분포 임계값($T_{cluster}$)에 대비하여 상기 데이터가 분포된 정도를 판단할 수 있다.
- [0096] 일실시예에 따른 전처리부(430)는 산출된 전처리 우선 순위에 기초하여 상기 데이터에 대한 전처리를 수행할 수 있다.
- [0097] 도 5는 일실시예에 따른 지능적 데이터 전처리 시스템의 동작 방법을 설명하는 도면이다.
- [0098] 일실시예에 따른 지능적 데이터 전처리 시스템의 동작 방법은 실시간으로 데이터를 수집할 수 있다(단계 501). 또한, 실시간 데이터 수집과는 별개로 신규 전처리 요청 목록을 수신할 수 있다(단계 502).
- [0099] 다음으로, 일실시예에 따른 지능적 데이터 전처리 시스템의 동작 방법은 실시간으로 수집된 데이터에 대해서 메타데이터를 생성하고, 데이터와 메타데이터를 연관지어 데이터베이스에 저장할 수 있다(단계 503).
- [0100] 일실시예에 따른 지능적 데이터 전처리 시스템의 동작 방법은 저장된 메타데이터에 대한 통계 분석을 처리할 수 있다(단계 504).
- [0101] 통계 분석을 위해, 지능적 데이터 전처리 시스템의 동작 방법은 저장된 메타데이터에 대해 연관 분석과 클러스터링 분석을 통해 컬럼들의 값 분포를 생성하여, 통계 분석을 처리할 수 있다.
- [0102] 일실시예에 따른 지능적 데이터 전처리 시스템의 동작 방법은 통계 분석된 결과에 신규 전처리 요청 목록을 반영하여 데이터에 대한 전처리 필요성을 판단할 수 있다(단계 505).
- [0103] 단계 505에서의 판단 결과, 대용량 데이터에 대한 전처리가 필요한 경우라면 일실시예에 따른 지능적 데이터 전처리 시스템의 동작 방법은 대용량 데이터 전처리를 수행할 수 있다(단계 506).
- [0104] 일례로, 일실시예에 따른 지능적 데이터 전처리 시스템의 동작 방법은 통계 분석의 처리 결과에 따라 상기 저장된 데이터에 대한 전처리 우선 순위를 결정하고, 결정된 전처리 우선 순위에 따라 전처리를 수행할 수 있다.
- [0105] 일실시예에 따른 지능적 데이터 전처리 시스템의 동작 방법은 통계 분석을 처리하기 위해, 데이터가 양의 상관관계 또는 음의 상관관계를 일정 수치 이상 가질 경우 연관성이 있는 것으로 상기 연관 분석을 처리할 수 있다.
- [0106] 또한, 일실시예에 따른 지능적 데이터 전처리 시스템의 동작 방법은 통계 분석을 처리하기 위해, 연관 분석의 결과에 따라 연관성이 있는 컬럼에 대해 클러스터링 분석을 처리하되, 클러스터 수를 증가시키며 평균 클러스터

간 거리($D_{Average}$)의 계산할 수 있다.

- [0107] 또한, 일실시예에 따른 지능적 데이터 전처리 시스템의 동작 방법은 계산된 평균 클러스터 간 거리($D_{Average}$)를 균등 분포시 클러스터 간 거리(D_{Max})에 대비하여 상기 데이터가 분포된 정도를 판단할 수 있다.
- [0108] 일실시예에 따른 지능적 데이터 전처리 시스템의 동작 방법은 통계 분석을 처리하기 위해, 인접한 두 클러스터의 중심점 간 차이 값들을 누적한 값과 클러스터의 개수간 비율에 기초해서 평균 클러스터 간 거리($D_{Average}$)를 계산할 수 있다.
- [0109] 또한, 일실시예에 따른 지능적 데이터 전처리 시스템의 동작 방법은 클러스터가 1개일 때의 최대 지름값과 전체의 클러스터 개수의 비율에 기초하여 상기 균등 분포시 클러스터 간 거리(D_{Max})를 계산할 수 있다.
- [0110] 만약, 단계 505에서의 판단 결과, 대용량 데이터에 대한 전처리가 필요하지 않는 경우라면 일실시예에 따른 지능적 데이터 전처리 시스템의 동작 방법은 대용량 데이터 전처리를 수행하지 않고, 프로세스를 종료할 수 있다.
- [0111] 이후, 지능적 데이터 전처리 시스템의 동작 방법은 전처리가 처리된 데이터 중에서, 입력되는 분석 요구사항에 상응하는 데이터들의 분산 데이터처리를 수행할 수 있다.
- [0112] 예를 들면, 지능적 데이터 전처리 시스템의 동작 방법은 전처리가 처리된 데이터 중에서, 입력되는 분석 요구사항에 상응하는 데이터들의 분산 데이터처리를 수행할 수 있다. 이를 위해, 지능적 데이터 전처리 시스템의 동작 방법은 처리 작업에 대한 워크플로우를 생성하고, 생성된 워크플로우에 따라 작업들을 스케줄링할 수 있다. 또한, 스케줄링 하는 작업들을 스트림 처리, 마이크로 배치 처리, 배치 처리 중에서 어느 하나의 처리에 할당할 수 있다.
- [0113] 결국, 본 발명을 이용하면 빅데이터 분석을 요구하는 산업 전반에서 다양한 사용자 및 시스템 요소들을 고려한 최적의 워크플로우를 생성하고 사용자 요구사항을 충족시켜 서비스 품질을 높일 수 있다. 또한, 통계적 분석 방법을 이용한 프로액티브(Proactive) 방식의 데이터 전처리를 수행할 수 있다. 뿐만 아니라, 사용자 요구사항에 따른 지능적 데이터 전처리를 수행하여 실시간 데이터 처리의 지연시간을 최소화 할 수 있다.
- [0115] 이상에서 설명된 장치는 하드웨어 구성요소, 소프트웨어 구성요소, 및/또는 하드웨어 구성요소 및 소프트웨어 구성요소의 조합으로 구현될 수 있다. 예를 들어, 실시예들에서 설명된 장치 및 구성요소는, 예를 들어, 프로세서, 콘트롤러, ALU(arithmetic logic unit), 디지털 신호 프로세서(digital signal processor), 마이크로컴퓨터, FPA(field programmable array), PLU(programmable logic unit), 마이크로프로세서, 또는 명령(instruction)을 실행하고 응답할 수 있는 다른 어떠한 장치와 같이, 하나 이상의 범용 컴퓨터 또는 특수 목적 컴퓨터를 이용하여 구현될 수 있다. 처리 장치는 운영 체제(OS) 및 상기 운영 체제 상에서 수행되는 하나 이상의 소프트웨어 애플리케이션을 수행할 수 있다. 또한, 처리 장치는 소프트웨어의 실행에 응답하여, 데이터를 접근, 저장, 조작, 처리 및 생성할 수도 있다. 이해의 편의를 위하여, 처리 장치는 하나가 사용되는 것으로 설명된 경우도 있지만, 해당 기술분야에서 통상의 지식을 가진 자는, 처리 장치가 복수 개의 처리 요소(processing element) 및/또는 복수 유형의 처리 요소를 포함할 수 있음을 알 수 있다. 예를 들어, 처리 장치는 복수 개의 프로세서 또는 하나의 프로세서 및 하나의 콘트롤러를 포함할 수 있다. 또한, 병렬 프로세서(parallel processor)와 같은, 다른 처리 구성(processing configuration)도 가능하다.
- [0116] 소프트웨어는 컴퓨터 프로그램(computer program), 코드(code), 명령(instruction), 또는 이들 중 하나 이상의 조합을 포함할 수 있으며, 원하는 대로 동작하도록 처리 장치를 구성하거나 독립적으로 또는 결합적으로(collectively) 처리 장치를 명령할 수 있다. 소프트웨어 및/또는 데이터는, 처리 장치에 의하여 해석되거나 처리 장치에 명령 또는 데이터를 제공하기 위하여, 어떤 유형의 기계, 구성요소(component), 물리적 장치, 가상 장치(virtual equipment), 컴퓨터 저장 매체 또는 장치, 또는 전송되는 신호 파(signal wave)에 영구적으로, 또는 일시적으로 구체화(embodiment)될 수 있다. 소프트웨어는 네트워크로 연결된 컴퓨터 시스템 상에 분산되어서, 분산된 방법으로 저장되거나 실행될 수도 있다. 소프트웨어 및 데이터는 하나 이상의 컴퓨터 판독 가능 기록 매체에 저장될 수 있다.
- [0117] 실시예에 따른 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 상기 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 상기 매체에 기록되는 프로그램 명령은 실시예를 위하여 특별히

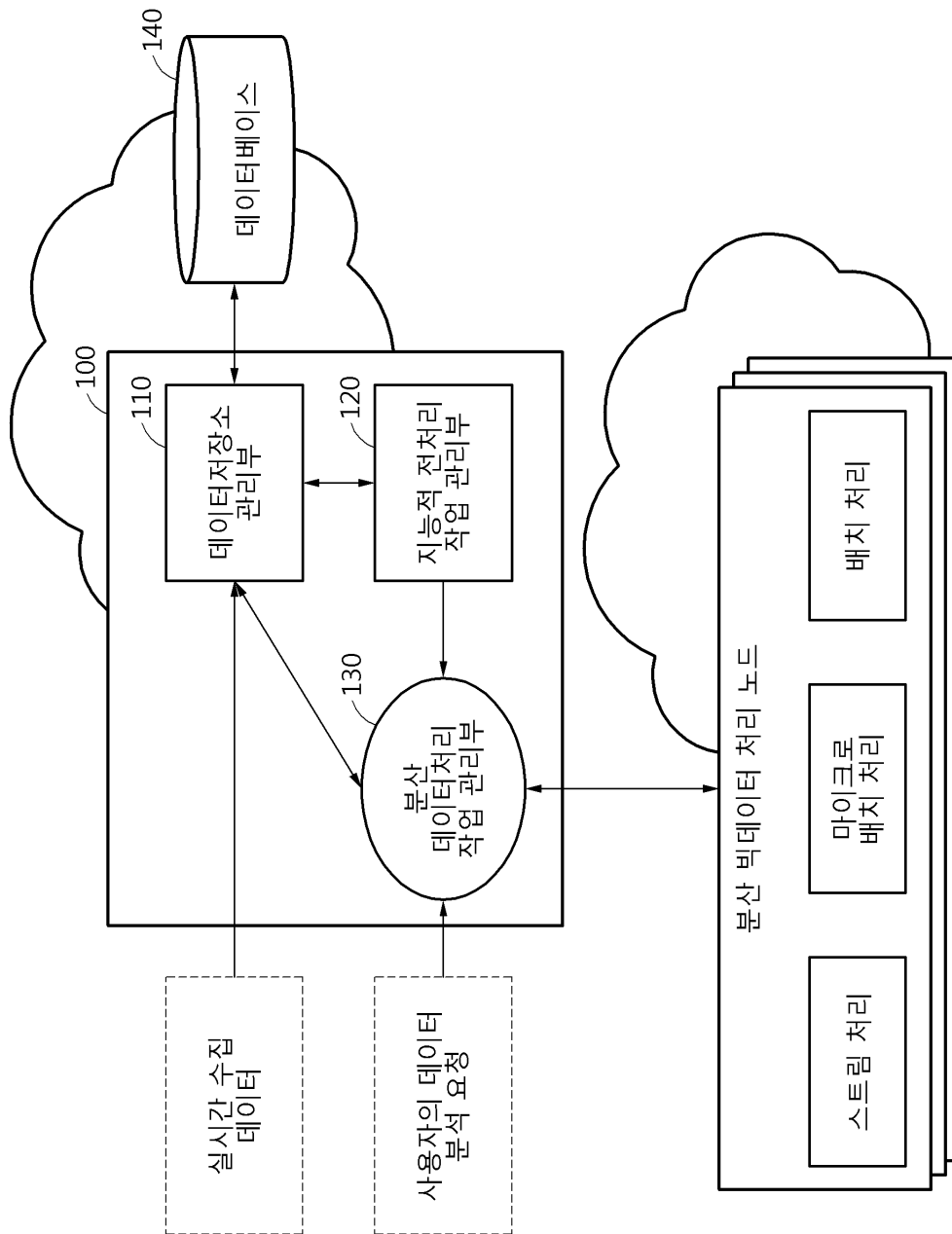
설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(magnetic media), CD-ROM, DVD와 같은 광기록 매체(optical media), 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical media), 및 롬(ROM), 램(RAM), 플래시 메모리 등과 같은 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다. 프로그램 명령의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함한다. 상기된 하드웨어 장치는 실시예의 동작을 수행하기 위해 하나 이상의 소프트웨어 모듈로서 작동하도록 구성될 수 있으며, 그 역도 마찬가지이다.

[0118] 이상과 같이 실시예들이 비록 한정된 도면에 의해 설명되었으나, 해당 기술분야에서 통상의 지식을 가진 자라면 상기의 기재로부터 다양한 수정 및 변형이 가능하다. 예를 들어, 설명된 기술들이 설명된 방법과 다른 순서로 수행되거나, 및/또는 설명된 시스템, 구조, 장치, 회로 등의 구성요소들이 설명된 방법과 다른 형태로 결합 또는 조합되거나, 다른 구성요소 또는 균등물에 의하여 대치되거나 치환되더라도 적절한 결과가 달성될 수 있다.

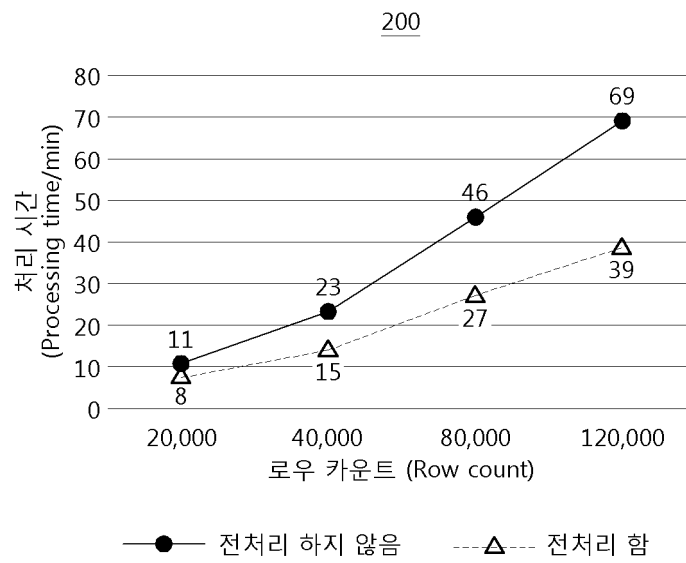
[0119] 그러므로, 다른 구현들, 다른 실시예들 및 특허청구범위와 균등한 것들도 후술하는 특허청구범위의 범위에 속한다.

도면

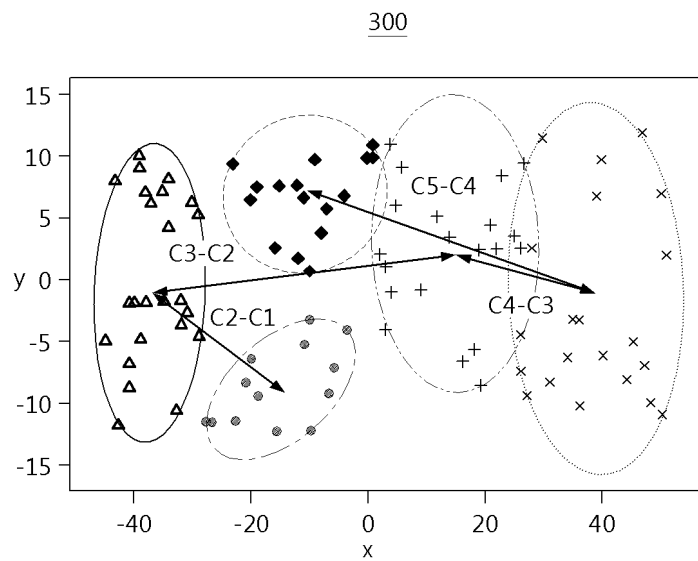
도면1



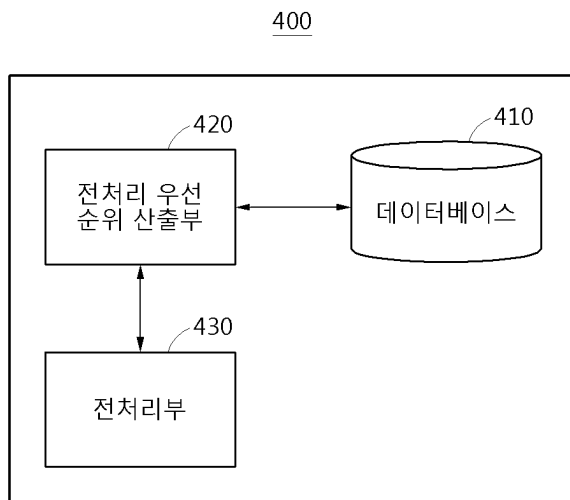
도면2



도면3



도면4



도면5

