



(19) 대한민국특허청(KR)  
(12) 공개특허공보(A)

(11) 공개번호 10-2015-0126481

(43) 공개일자 2015년11월12일

(51) 국제특허분류(Int. Cl.)

*B62D 41/00* (2006.01) *G06F 19/00* (2011.01)

(21) 출원번호 10-2014-0053388

(22) 출원일자 2014년05월02일

심사청구일자 2014년05월02일

(71) 출원인

국민대학교산학협력단

서울특별시 성북구 정릉로 77 (정릉동, 국민대학교)

(72) 발명자

최은미

서울특별시 강남구 삼성로63길 37, 102동 202호  
(대치동, 한티아파트)

조원희

서울특별시 성북구 솔샘로25길 28 풍림아파트 12  
1동 602호 (정릉동)

(74) 대리인

특허법인 제나

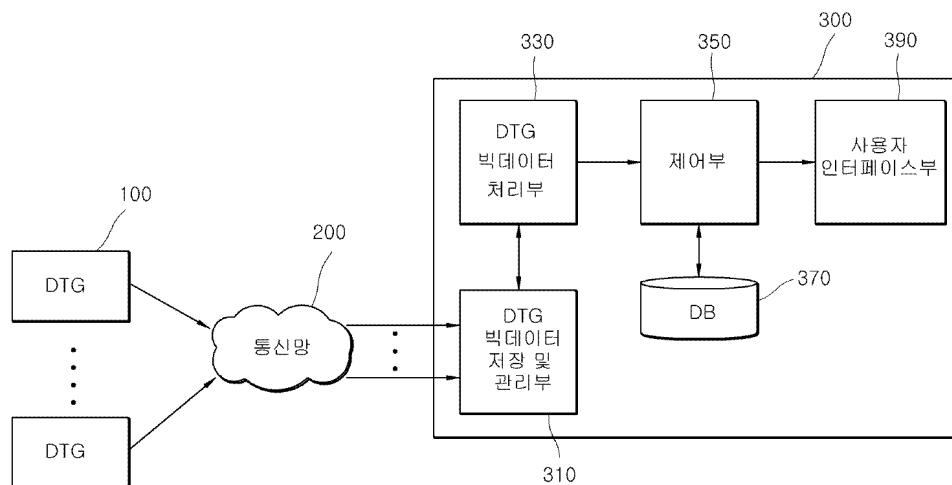
전체 청구항 수 : 총 19 항

(54) 발명의 명칭 DTG 빅데이터 처리 및 분석 시스템

### (57) 요약

DTG(Digital Tachograph) 단말에 의해 수집된 운행기록 데이터를 이용하여 빅데이터 처리 및 분석을 수행하는 시스템으로서, 상기 DTG 단말에 의해 수집된 로우 데이터(raw data)의 운행기록 데이터를 저장하는 DTG 빅데이터 저장부; 및 상기 DTG 빅데이터 저장부에 저장된 로우 데이터의 운행기록 데이터를 정제(refine)하고, 정제된 운행기록 데이터에 근거하여 통계 데이터를 획득하고, 상기 정제된 운행기록 데이터 및 상기 획득된 통계 데이터 중 적어도 하나에 근거하여 마이닝 분석을 수행하는 DTG 빅데이터 처리부를 포함하는 DTG 빅데이터 처리 및 분석 시스템이 제공된다.

### 대표도



이 발명을 지원한 국가연구개발사업

과제고유번호 1711002552

부처명 미래창조과학부

연구관리전문기관 정보통신산업진흥원

연구사업명 정보통신산업진흥원사업 IT/SW 창의연구과정 (기술개발형)

연구과제명 디지털 운행기록계 빅데이터 분석 기술 연구

기 여 율 1/1

주관기관 국민대학교 산학협력단

연구기간 2013.09.01 ~ 2014.08.31

---

## 명세서

### 청구범위

#### 청구항 1

DTG(Digital Tachograph) 단말에 의해 수집된 운행기록 데이터를 이용하여 빅데이터 처리 및 분석을 수행하는 시스템으로서,

상기 DTG 단말에 의해 수집된 로우 데이터(raw data)의 운행기록 데이터를 저장하는 DTG 빅데이터 저장부; 및

상기 DTG 빅데이터 저장부에 저장된 로우 데이터의 운행기록 데이터를 정제(refine)하고, 정제된 운행기록 데이터에 근거하여 통계 데이터를 획득하고, 상기 정제된 운행기록 데이터 및 상기 획득된 통계 데이터 중 적어도 하나에 근거하여 마이닝 분석을 수행하는 DTG 빅데이터 처리부;

를 포함하는 DTG 빅데이터 처리 및 분석 시스템.

#### 청구항 2

제1 항에 있어서,

상기 운행기록 데이터는, 식별 데이터와 센싱 데이터를 포함하고,

상기 식별 데이터는, 상기 DTG 단말이 장착된 차량의 차대번호, 유형, 차량등록번호, 운송 사업자 등록번호 및 운전자 식별코드 필드에 대한 레코드를 포함하고,

상기 센싱 데이터는, 상기 DTG 단말이 장착된 차량의 주행거리, 주행시간, 데이터 획득주기, 데이터 획득일시, 속도, 분당 엔진 회전수, 브레이크 신호, 위치, 방위각 및 가속도 필드에 대한 레코드를 포함하는, DTG 빅데이터 처리 및 분석 시스템.

#### 청구항 3

제2 항에 있어서,

상기 DTG 빅데이터 처리부는, 상기 센싱 데이터의 각 필드 별로 레코드가 제한 범위 내의 값을 갖는지 여부를 판단하고, 상기 레코드가 상기 제한 범위 외의 값을 갖는 경우 상기 레코드를 제거하거나 상기 레코드를 상기 제한 범위 내의 값에 상응하도록 보정하여 상기 로우 데이터의 운행기록 데이터를 정제하는, DTG 빅데이터 처리 및 분석 시스템.

#### 청구항 4

제3 항에 있어서,

상기 DTG 빅데이터 처리부는, 상기 센싱 데이터의 각 필드 별로 연속하는 레코드들을 편차 비교하여 이상치(outlier)를 검출하고, 이상치로 검출된 레코드를 제거하여 기 정제된 로우 데이터의 운행기록 데이터를 다시 정제하는, DTG 빅데이터 처리 및 분석 시스템.

#### 청구항 5

제3 항에 있어서,

상기 DTG 빅데이터 처리부는, 상기 센싱 데이터에서 서로 상관관계에 있는 필드들의 레코드를 비교하여 이상치를 검출하고, 이상치로 검출된 레코드를 제거하여 정제된 로우 데이터의 운행기록 데이터를 다시 정제하는, DTG

빅데이터 처리 및 분석 시스템.

#### 청구항 6

제1 항에 있어서,

상기 획득된 통계 데이터는, 운행 통계 데이터와 성향 통계 데이터를 포함하고,

상기 운행 통계 데이터는, 일주행거리, 총 주행 시간, 평균 속도, 최고 속도, 및 최고 분당 엔진 회전수 필드에 대한 레코드를 포함하고,

상기 성향 통계 데이터는, 과속 횟수, 위험과속 횟수, 장기고속 횟수, 급가속 횟수, 급감속 횟수, 급출발 횟수, 급정지 횟수, 앞지르기 횟수, 진로변경 횟수, 급회전 횟수, 공회전 횟수, 및 고 분당 엔진 회전수 횟수 필드에 대한 레코드를 포함하는, DTG 빅데이터 처리 및 분석 시스템.

#### 청구항 7

제6 항에 있어서,

상기 DTG 빅데이터 처리부는, 상기 정제된 운행기록 데이터를 통계적으로 계산하여 상기 통계 데이터를 생성하고, 생성된 통계 데이터에서 필드 별 레코드가 미리 설정된 시간 범위 내에서 반복 계산된 것인지 여부를 판단하고, 반복 계산된 것으로 판단된 레코드를 보정하여 상기 통계 데이터를 획득하는, DTG 빅데이터 처리 및 분석 시스템.

#### 청구항 8

제6 항에 있어서,

상기 DTG 빅데이터 처리부는, 상기 성향 통계 데이터를 기초로 상기 DTG 단말이 장착된 차량의 운전자의 운전 성향을 유형 별로 분류하고, 상기 분류된 유형 각각에 대해 상기 운전자의 운전 성향 정도를 지수화하기 위한 마이닝 분석을 수행하는, DTG 빅데이터 처리 및 분석 시스템.

#### 청구항 9

제8 항에 있어서,

상기 DTG 빅데이터 처리부는, 상기 분류된 유형의 성향 정도를 지수화하여 나타내는 지수정보를 생성하고, 상기 성향 통계 데이터의 필드 중 적어도 하나와 상기 지수정보의 상관관계를 분석하기 위한 마이닝 분석을 수행하는, DTG 빅데이터 처리 및 분석 시스템.

#### 청구항 10

제6 항에 있어서,

상기 DTG 빅데이터 처리부는, 상기 정제된 운행기록 데이터 및 상기 획득된 통계 데이터에 근거하여, 특정 주기를 기준으로 상기 DTG 단말이 장착된 차량의 운전자의 운전성향을 분석하기 위한 마이닝 분석을 수행하는, DTG 빅데이터 처리 및 분석 시스템.

#### 청구항 11

제6 항에 있어서,

상기 획득된 통계 데이터는, 사고 이력 및 정비 이력 필드에 대한 레코드를 구비하는 이력 데이터를 더 포함하고,

상기 DTG 빅데이터 처리부는, 상기 이력 데이터, 상기 정제된 운행기록 데이터 및 상기 획득된 통계 데이터에 근거하여, 상기 DTG 단말이 장착된 차량의 사고위험도 및 정비시기를 예측하기 위한 마이닝 분석을 수행하는, DTG 빅데이터 처리 및 분석 시스템.

#### 청구항 12

제11 항에 있어서,

상기 DTG 빅데이터 처리부는, 상기 사고위험도의 정도 또는 상기 정비시기 도래여부를 지수화하여 나타내는 지수정보를 생성하는, DTG 빅데이터 처리 및 분석 시스템.

#### 청구항 13

제6 항에 있어서,

상기 DTG 빅데이터 처리부는, 상기 정제된 운행기록 데이터 및 상기 획득된 통계 데이터에 근거하여, 사고 위험성이 높거나 대기 오염도를 증가시키는 공간을 분별하기 위한 마이닝 분석을 수행하는, DTG 빅데이터 처리 및 분석 시스템.

#### 청구항 14

제6 항에 있어서,

상기 DTG 빅데이터 처리부는, 상기 정제된 운행기록 데이터 및 상기 획득된 통계 데이터에 근거하여, 행정 구역별로 운전자들의 운전 성향을 분석하기 위한 마이닝 분석을 수행하는, DTG 빅데이터 처리 및 분석 시스템.

#### 청구항 15

제8 항에 있어서,

상기 DTG 빅데이터 처리부는, 상기 분류된 유형의 성향 정도를 지수화하여 나타내는 지수정보를 생성하고, 생성된 지수정보, 상기 정제된 운행기록 데이터 및 상기 획득된 통계 데이터에 근거하여 운송 사업자 별로 상기 지수정보의 분포상태를 확인하기 위한 마이닝 분석을 수행하는, DTG 빅데이터 처리 및 분석 시스템.

#### 청구항 16

제8 항에 있어서,

상기 DTG 빅데이터 처리부는, 상기 분류된 유형의 성향 정도를 지수화하여 나타내는 지수정보를 생성하고, 생성된 지수정보, 상기 정제된 운행기록 데이터 및 상기 획득된 통계 데이터에 근거하여 운전자가 상기 생성된 지수정보의 인지 전과 인지 후의 운전 성향 변화를 분석하기 위한 마이닝 분석을 수행하는, DTG 빅데이터 처리 및 분석 시스템.

#### 청구항 17

제1 항에 있어서,

상기 DTG 빅데이터 저장부는, 상기 운행기록에 관한 로우 데이터를 저장하도록 분산 파일 시스템으로 구성되는,

DTG 빅데이터 처리 및 분석 시스템.

## 청구항 18

제1 항에 있어서,

상기 복수의 통계 데이터 및 상기 결과정보를 이용하여 분석정보를 생성하는 제어부; 및 상기 결과정보 및 상기 분석정보를 사용자에게 제공하는 사용자 인터페이스부를 더 포함하되,

상기 제어부는, 상기 DTG 빅데이터 처리부에 의해 처리된 데이터에 근거하여, 운전자 위험 운전 성향에 관한 분석정보, 연비 지수 및 환경개선 지수 중 적어도 하나에 관한 지수분석정보, 시기별 및 회사별 통계분석정보 중 적어도 하나를 생성하는, DTG 빅데이터 처리 및 분석 시스템.

## 청구항 19

제18 항에 있어서,

상기 사용자 인터페이스부는, 상기 운행기록 데이터에 관한 분석결과를 시각화(visualization) 처리하는 시각화 처리부를 포함하되,

상기 시각화 처리부는, 운행기록 데이터에 포함된 단일의 운행기록 필드 별로의 군집분석에 관한 시각화, 상기 단일 운행기록 필드 별로의 통계분석에 관한 시각화, 상기 운행기록 필드 별로의 스케일 재분류에 관한 시각화, 상기 운행기록 데이터 중 상관관계에 있는 어느 2개의 운행기록 필드 간의 교차분석에 관한 시각화, 상기 운행기록 데이터에 관한 시계열적 로우 데이터 분석에 관한 시각화, 및 단기 또는 장기 마이닝 분석에 관한 시각화 중 적어도 하나의 처리를 수행하는, DTG 빅데이터 처리 및 분석 시스템.

## 발명의 설명

### 기술 분야

[0001] 본 발명은 DTG 빅데이터 처리 및 분석 시스템에 관한 것으로서, 보다 상세하게는 체계적인 데이터 정제 및 데이터 마이닝 절차를 통해 디지털 운행기록에 관한 빅데이터로부터 유용한 정보를 획득할 수 있도록 하는 DTG 빅데이터 처리 및 분석 시스템에 관한 것이다.

### 배경 기술

[0002] 빅데이터는 일반적인 데이터베이스 시스템으로는 수집, 저장 및 분석하기 어려운 방대한 양의 데이터를 의미한다. 빅데이터의 예로는, 소셜 미디어, 소셜 네트워크 서비스에서 발생하는 소셜 데이터, 네트워크 트래픽 로그 및 웹 서버나 응용 프로그램의 웹 로그, 센싱 장비에 의해 획득되는 센싱 데이터와 로그 등을 들 수 있으며, 상기 네트워크 트래픽 로그 등과 같은 시계열 데이터(time-series data)의 경우 시간이 지날수록 데이터가 축적되어 다루어야 할 데이터의 크기가 폭발적으로 증가한다.

[0003] 상기 시계열 데이터의 일 예로는 디지털 운행기록장치(Digital Tachograph 또는 Digital Tachometer, 이하, DTG로 통칭함)에 의해 획득된 운행기록 데이터를 들 수 있다. DTG는 차량의 위치, 운전자의 조작에 의한 차량 시동, 급발진, 속도 등의 운행기록 데이터를 수집하기 위한 장치이며, 이에 의해 획득된 운행기록 데이터는 관리자, 예컨대 운송 사업자, 교통안전공단 등에 제공되어 운행 이력관리, 사고 발생시 사고 원인 규명, 사고 예방 등을 위해 사용된다.

[0004] 최근 정부 시책에 따라 택시와 버스 등 사업용 차량에 대하여 DTG의 장착이 의무화되는 등 향후 DTG의 설치 및 보급이 확대될 예정이며, 시계열 데이터의 특성상 처리 및 분석이 필요한 운행기록 데이터의 양이 급격히 증가하게 될 것인 바, 방대한 양의 운행기록 데이터를 체계적으로 분석하고 분석을 통해 획득한 유용한 정보를 관리자 또는 사용자가 효율적으로 이용할 수 있게 하기 위한 DTG 빅데이터 처리 및 분석 시스템의 구축이 필요하다.

## 발명의 내용

### 해결하려는 과제

[0005] 본 발명의 기술적 사상이 이루고자 하는 기술적 과제는, 디지털 운행기록에 관한 빅데이터를 체계적으로 분석할 수 있고, 분석을 통해 획득한 결과정보를 관리자 또는 사용자가 효율적으로 이용할 수 있도록 하는 DTG 빅데이터 처리 및 분석 시스템을 제공하는 것이다.

### 과제의 해결 수단

[0006] 본 발명의 일 측면에 따르면, DTG(Digital Tachograph) 단말에 의해 수집된 운행기록 데이터를 이용하여 빅데이터 처리 및 분석을 수행하는 시스템으로서, 상기 DTG 단말에 의해 수집된 로우 데이터(raw data)의 운행기록 데이터를 저장하는 DTG 빅데이터 저장부; 및 상기 DTG 빅데이터 저장부에 저장된 로우 데이터의 운행기록 데이터를 정제(refine)하고, 정제된 운행기록 데이터에 근거하여 통계 데이터를 획득하고, 상기 정제된 운행기록 데이터 및 상기 획득된 통계 데이터 중 적어도 하나에 근거하여 마이닝 분석을 수행하는 DTG 빅데이터 처리부를 포함하는 DTG 빅데이터 처리 및 분석 시스템이 제공된다.

[0007] 일 실시예에서, 상기 운행기록 데이터는, 식별 데이터와 센싱 데이터를 포함하고,

[0008] 상기 식별 데이터는, 상기 DTG 단말에 장착된 차량의 차대번호, 유형, 차량등록번호, 운송 사업자 등록번호 및 운전자 식별코드 필드에 대한 레코드를 포함하고,

[0009] 상기 센싱 데이터는, 상기 DTG 단말에 장착된 차량의 주행거리, 주행시간, 데이터 획득주기, 데이터 획득일시, 속도, 분당 엔진 회전수, 브레이크 신호, 위치, 방위각 및 가속도 필드에 대한 레코드를 포함할 수 있다.

[0010] 일 실시예에서, 상기 DTG 빅데이터 처리부는, 상기 센싱 데이터의 각 필드 별로 레코드가 제한 범위 내의 값을 갖는지 여부를 판단하고, 상기 레코드가 상기 제한 범위 외의 값을 갖는 경우 상기 레코드를 제거하거나 상기 레코드를 상기 제한 범위 내의 값에 상응하도록 보정하여 상기 로우 데이터의 운행기록 데이터를 정제할 수 있다.

[0011] 일 실시예에서, 상기 DTG 빅데이터 처리부는, 상기 센싱 데이터의 각 필드 별로 연속하는 레코드들을 편차 비교하여 이상치(outlier)를 검출하고, 이상치로 검출된 레코드를 제거하여 기 정제된 로우 데이터의 운행기록 데이터를 다시 정제할 수 있다.

[0012] 일 실시예에서, 상기 DTG 빅데이터 처리부는, 상기 센싱 데이터에서 서로 상관관계에 있는 필드들의 레코드를 비교하여 이상치를 검출하고, 이상치로 검출된 레코드를 제거하여 정제된 로우 데이터의 운행기록 데이터를 다시 정제할 수 있다.

[0013] 일 실시예에서, 상기 획득된 통계 데이터는, 운행 통계 데이터와 성향 통계 데이터를 포함하고,

[0014] 상기 운행 통계 데이터는, 일주행거리, 총 주행 시간, 평균 속도, 최고 속도, 및 최고 분당 엔진 회전수 필드에 대한 레코드를 포함하고,

[0015] 상기 성향 통계 데이터는, 과속 횟수, 위험과속 횟수, 장기과속 횟수, 급가속 횟수, 급감속 횟수, 급출발 횟수, 급정지 횟수, 앞지르기 횟수, 진로변경 횟수, 급회전 횟수, 공회전 횟수, 및 고 분당 엔진 회전수 횟수 필드에 대한 레코드를 포함할 수 있다.

[0016] 일 실시예에서, 상기 DTG 빅데이터 처리부는, 상기 정제된 운행기록 데이터를 통계적으로 계산하여 상기 통계 데이터를 생성하고, 생성된 통계 데이터에서 필드 별 레코드가 미리 설정된 시간 범위 내에서 반복 계산된 것인

지 여부를 판단하고, 반복 계산된 것으로 판단된 레코드를 보정하여 상기 통계 데이터를 획득할 수 있다.

- [0017] 일 실시예에서, 상기 DTG 빅데이터 처리부는, 상기 성향 통계 데이터를 기초로 상기 DTG 단말이 장착된 차량의 운전자의 운전 성향을 유형 별로 분류하고, 상기 분류된 유형 각각에 대해 상기 운전자의 운전 성향 정도를 지수화하기 위한 마이닝 분석을 수행할 수 있다.
- [0018] 일 실시예에서, 상기 DTG 빅데이터 처리부는, 상기 분류된 유형의 성향 정도를 지수화하여 나타내는 지수정보를 생성하고, 상기 성향 통계 데이터의 필드 중 적어도 하나와 상기 지수정보의 상관관계를 분석하기 위한 마이닝 분석을 수행할 수 있다.
- [0019] 일 실시예에서, 상기 DTG 빅데이터 처리부는, 상기 정제된 운행기록 데이터 및 상기 획득된 통계 데이터에 근거하여, 특정 주기를 기준으로 상기 DTG 단말이 장착된 차량의 운전자의 운전성향을 분석하기 위한 마이닝 분석을 수행할 수 있다.
- [0020] 일 실시예에서, 상기 획득된 통계 데이터는, 사고 이력 및 정비 이력 필드에 대한 레코드를 구비하는 이력 데이터를 더 포함하고,
- [0021] 상기 DTG 빅데이터 처리부는, 상기 이력 데이터, 상기 정제된 운행기록 데이터 및 상기 획득된 통계 데이터에 근거하여, 상기 DTG 단말이 장착된 차량의 사고위험도 및 정비시기를 예측하기 위한 마이닝 분석을 수행할 수 있다.
- [0022] 일 실시예에서, 상기 DTG 빅데이터 처리부는, 상기 사고위험도의 정도 또는 상기 정비시기 도래여부를 지수화하여 나타내는 지수정보를 생성할 수 있다.
- [0023] 일 실시예에서, 상기 DTG 빅데이터 처리부는, 상기 정제된 운행기록 데이터 및 상기 획득된 통계 데이터에 근거하여, 사고 위험성이 높거나 대기 오염도를 증가시키는 공간을 분별하기 위한 마이닝 분석을 수행할 수 있다.
- [0024] 일 실시예에서, 상기 DTG 빅데이터 처리부는, 상기 정제된 운행기록 데이터 및 상기 획득된 통계 데이터에 근거하여, 행정 구역별로 운전자들의 운전 성향을 분석하기 위한 마이닝 분석을 수행할 수 있다.
- [0025] 일 실시예에서, 상기 DTG 빅데이터 처리부는, 상기 분류된 유형의 성향 정도를 지수화하여 나타내는 지수정보를 생성하고, 생성된 지수정보, 상기 정제된 운행기록 데이터 및 상기 획득된 통계 데이터에 근거하여 운송 사업자 별로 상기 지수정보의 분포상태를 확인하기 위한 마이닝 분석을 수행할 수 있다.
- [0026] 일 실시예에서, 상기 DTG 빅데이터 처리부는, 상기 분류된 유형의 성향 정도를 지수화하여 나타내는 지수정보를 생성하고, 생성된 지수정보, 상기 정제된 운행기록 데이터 및 상기 획득된 통계 데이터에 근거하여 운전자가 상기 생성된 지수정보의 인지 전과 인지 후의 운전 성향 변화를 분석하기 위한 마이닝 분석을 수행할 수 있다.
- [0027] 일 실시예에서, 상기 DTG 빅데이터 저장부는, 상기 운행기록에 관한 로우 데이터를 저장하도록 분산 파일 시스템으로 구성될 수 있다.
- [0028] 일 실시예에서, 상기 복수의 통계 데이터 및 상기 결과정보를 이용하여 분석정보를 생성하는 제어부; 및 상기 결과정보 및 상기 분석정보를 사용자에게 제공하는 사용자 인터페이스부를 더 포함하되,
- [0029] 상기 제어부는, 상기 DTG 빅데이터 처리부에 의해 처리된 데이터에 근거하여, 운전자 위험 운전 성향에 관한 분석정보, 연비 지수 및 환경개선 지수 중 적어도 하나에 관한 지수분석정보, 시기별 및 회사별 통계분석정보 중 적어도 하나를 생성할 수 있다.
- [0030] 일 실시예에서, 상기 사용자 인터페이스부는, 상기 운행기록 데이터에 관한 분석결과를 시각화(visualization) 처리하는 시각화 처리부를 포함하되,
- [0031] 상기 시각화 처리부는, 상기 운행기록 데이터에 포함된 단일의 운행기록 필드 별로의 군집분석에 관한 시각화, 상기 단일 운행기록 필드 별로의 통계분석에 관한 시각화, 상기 운행기록 필드 별로의 스케일 재분류에 관한 시각화, 상기 운행기록 데이터 중 상관관계에 있는 어느 2개의 운행기록 필드 간의 교차분석에 관한 시각화, 상기



운행기록 데이터에 관한 시계열적 로우 데이터 분석에 관한 시각화, 및 단기 또는 장기 마이닝 분석에 관한 시각화 중 적어도 하나의 처리를 수행할 수 있다.

### 발명의 효과

- [0032] 본 발명의 기술적 사상에 의한 DTG 빅데이터 처리 및 분석 시스템은, 수집된 운행기록 데이터와 같은 빅데이터를 분석함에 있어서, 로우 데이터(raw data)에 대해 2단계의 정제를 수행하고, 정제된 로우 데이터로부터 통계 데이터를 추출하고, 단기적 관점과 장기적 관점의 복합적인 데이터 마이닝을 수행함으로써, 분석 오류를 줄임과 동시에 분석 성능을 향상시킨다.
- [0033] 또한, 본 발명의 기술적 사상에 의한 DTG 빅데이터 처리 및 분석 시스템은, 수집된 운행기록 데이터와 같은 빅데이터를 분석하여 추출한 결과정보를 사용자에게 제공함으로써, 차량 운전자의 안전 운전 유도, 차량 연비 향상 유도, 사전 정비 유도 및 사고 방지를 가능하게 하고, 운송 사업자의 관리 효율성을 향상시키며, 공공기관에서 사고위험 및 환경악화를 조래하는 행정구역, 도로 등을 효율적으로 개선할 수 있도록 한다.
- [0034] 또한, 본 발명의 기술적 사상에 의한 DTG 빅데이터 처리 및 분석 시스템은, 추출된 정보들을 시각화 처리하여 사용자에게 제공함으로써, 빅데이터로부터 추출된 정보를 관리자 또는 사용자가 직관적으로 인지할 수 있도록 하고 효과적으로 이용할 수 있도록 한다.

### 도면의 간단한 설명

- [0035] 본 발명의 상세한 설명에서 인용되는 도면을 보다 충분히 이해하기 위하여 각 도면의 간단한 설명이 제공된다.
- 도 1은 본 발명의 기술적 사상에 의한 일 실시예에 따른 DTG 빅데이터 처리 및 분석 시스템을 설명하기 위한 도면이다.
- 도 2는 본 발명의 기술적 사상에 의한 일 실시예에 따른 DTG 빅데이터 처리 및 분석 시스템에서 저장 및 관리부를 보다 상세히 설명하기 위한 도면이다.
- 도 3은 본 발명의 기술적 사상에 의한 일 실시예에 따른 DTG 빅데이터 처리 및 분석 시스템의 처리부에서 수행되는 빅데이터 처리 및 분석 방법을 설명하기 위한 순서도이며, 도 4 내지 도 10은 도 3의 빅데이터 처리 및 분석 방법의 각 단계들을 보다 상세히 설명하기 위한 도면들이다.
- 도 11은 본 발명의 기술적 사상에 의한 일 실시예에 따른 DTG 빅데이터 처리 및 분석 시스템에서 제어부를 보다 상세히 설명하기 위한 도면이다.
- 도 12는 본 발명의 기술적 사상에 의한 일 실시예에 따른 DTG 빅데이터 처리 및 분석 시스템에서 사용자 인터페이스부를 보다 상세히 설명하기 위한 도면이고, 도 13a 내지 도 13f는 사용자 인터페이스부에서 제공하는 데이터 시각화의 구현 예들을 설명하기 위한 도면이다.
- 도 14는 본 발명의 기술적 사상에 의한 일 실시예에 따른 DTG 빅데이터 처리 및 분석 시스템에서 운행기록 데이터의 저장 포맷을 설명하기 위한 도면이다.

### 발명을 실시하기 위한 구체적인 내용

- [0036] 이하, 첨부 도면을 참조하여 본 발명의 실시예들을 상세히 설명한다. 도면상의 동일한 구성요소에 대해서는 동일한 참조부호를 사용하고, 이들에 대한 중복된 설명은 생략한다.
- [0037] 본 발명의 실시예들은 당해 기술 분야에서 통상의 지식을 가진 자에게 본 발명을 더욱 완전하게 설명하기 위하여 제공되는 것으로, 아래의 실시예들은 여러 가지 다른 형태로 변형될 수 있으며, 본 발명의 범위가 아래의 실시예들로 한정되는 것은 아니다. 오히려, 이들 실시예는 본 개시를 더욱 충실하고 완전하게 하며 당업자에게 본 발명의 사상을 완전하게 전달하기 위하여 제공되는 것이다.
- [0038] 본 발명의 실시예들을 설명함에 있어서, 관련된 공지 기능 또는 구성에 대한 구체적인 설명이 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명을 생략할 것이다. 또한, 본 명세서의 설명 과정에서 이용되는 숫자(예를 들어, 제1, 제2 등)는 하나의 구성요소를 다른 구성요소와 구분하기 위한 식별기호에

불과하다. 또한, 본 명세서에서, 일 구성요소가 다른 구성요소와 "연결된다" 거나 "접속된다" 등으로 언급된 때에는, 상기 일 구성요소가 상기 다른 구성요소와 직접 연결되거나 또는 직접 접속될 수도 있지만, 특별히 반대되는 기재가 존재하지 않는 이상, 중간에 또 다른 구성요소를 매개하여 연결되거나 또는 접속될 수도 있다고 이해되어야 할 것이다.

[0039] 그리고, 본 명세서에서 사용되는 용어(terminology)들은 본 발명의 실시예들을 적절히 표현하기 위해 사용된 용어들이어서, 이는 사용자, 운용자의 의도 또는 본 발명이 속하는 분야의 관례 등에 따라 달라질 수 있다. 따라서, 본 용어들에 대한 정의는 본 명세서 전반에 걸친 내용을 토대로 내려져야 할 것이다. 각 도면에 제시된 동일한 참조부호는 동일한 부재를 나타낸다.

[0040] 도 1은 본 발명의 기술적 사상에 의한 일 실시예에 따른 DTG 빅데이터 처리 및 분석 시스템을 설명하기 위한 도면이다.

[0041] 도 1에서는 DTG 단말을 통해 직접 획득된 운행기록 데이터로 구성되는 빅데이터를 처리 및 분석하는 시스템을 예시하였으나, 본 발명의 기술적 사상은 이에 한정되지 않는다. 예를 들어, 본 발명의 기술적 사상에 따른 DTG 빅데이터 처리 및 분석 시스템은, DTG 단말을 통해 직접 획득되는 데이터 이외에도, 차량 등의 운송수단에 설치된 각종 센서로부터 획득된 운행기록에 관한 빅데이터도 함께 처리할 수 있을 것이다. 다만, 이하 본 명세서에서는 설명의 편의상 DTG 단말을 통해 직접 획득된 DTG 빅데이터를 처리 및 분석하는 경우를 중심으로 설명하기로 한다.

[0042] 도 1을 참조하면, DTG 빅데이터 처리 및 분석 시스템(300)이 통신망(200)을 통해서 복수의 DTG 단말(100)과 연결된다. 여기서, DTG 단말(100)은 차량 등 운송수단에 탑재되어 해당 운송수단에서의 운행기록 데이터를 수집 및 획득하며, 획득된 운행기록 데이터를 통신망(200)을 통해서 DTG 빅데이터 처리 및 분석 시스템(300)으로 전송할 수 있다. 다만, 도 1에서는 운행기록 데이터가 통신망(200)을 통해 DTG 빅데이터 처리 및 분석 시스템(300)으로 전송되는 경우를 예시하였지만, 반드시 이에 한정되는 것이 아님은 물론이다.

[0043] DTG 단말(100)에 의해 획득되는 운행기록 데이터는 도 14에 도시된 바와 같은 저장 포맷을 가질 수 있다. 다만, 사용되는 DTG 저장 포맷에 따라 기록될 수 있는 운행기록 데이터는 다양하게 변형 가능함은 물론이다. 예를 들어, 냉각수 온도, 배터리 전압, 에어 게이지의 전압, 연료 게이지의 전압, 엔진 온도, 연료 잔량 등이 운행 기록 데이터에 포함될 수 있다. 또한, 데이터 획득 주기, 사고이력, 정비이력 등이 운행기록 데이터에 포함될 수 있다.

[0044] DTG 빅데이터 처리 및 분석 시스템(300)은, DTG 빅데이터 저장 및 관리부(310, 이하 저장 및 관리부), DTG 빅데이터 처리부(330, 이하 처리부), 제어부(350), 데이터베이스(370) 및 사용자 인터페이스부(390)를 포함할 수 있으며, 각 구성부에 대해서는 이하 도 2 내지 도 13을 참조하여 차례로 설명하기로 한다.

#### [0045] [DTG 빅데이터 저장 및 관리부의 기능 및 역할 - 도 2]

[0046] 도 2는 DTG 빅데이터 처리 및 분석 시스템(300)에서 저장 및 관리부(310)를 보다 상세히 설명하기 위한 도면이다.

[0047] 도 2를 참조하면, 저장 및 관리부(310)는 관리부(311) 및 데이터 저장수단(313)을 포함할 수 있다.

[0048] 관리부(311)는 통신망(200, 도 1 참조)을 통해 복수의 DTG 단말(100, 도 1 참조)로부터 제공되는 로우 데이터(raw data)의 운행기록 데이터를 수신할 수 있다. 관리부(311)는 데이터 저장수단(313)으로의 상기 로우 데이터의 운행기록 데이터 저장 및 데이터 저장수단(313)으로부터의 상기 로우 데이터의 운행기록 데이터 출력을 제어할 수 있다.

[0049] 데이터 저장수단(313)은 상기 로우 데이터의 운행기록 데이터를 저장할 수 있다. 데이터 저장수단(313)은, 예컨대 물리적으로 떨어져 있는 여러 개의 스토리지들로 구성될 수 있으며, 각각의 스토리지들은 동일한 상기 로우 데이터의 운행기록 데이터를 상호 공유 및 저장할 수 있다.

[0050] 즉, 저장 및 관리부(310)는, 관리부(310) 및 데이터 저장수단(313)을 포함하는 하둡 분산 파일 시스템(hadoop

distributed file system)과 같은 분산 파일 시스템으로 구성될 수 있다.

**[DTG 빅데이터 처리부의 기능 및 역할 - 도 3 내지 도 10]**

도 3은 DTG 빅데이터 처리 및 분석 시스템(300)의 처리부(330)에서 수행되는 빅데이터 처리 및 분석 방법을 설명하기 위한 순서도이며, 도 4 내지 도 10은 도 3의 빅데이터 처리 및 분석 방법의 각 단계들을 보다 상세히 설명하기 위한 도면들이다. 도 3에 도시된 빅데이터 처리 및 분석 방법의 각 단계들은 해당 알고리즘을 구현하는 적어도 하나의 소프트웨어를 통해 처리부(330)에서 수행될 수 있다. 이하에서는, 도 3에 도시된 빅데이터 처리 및 분석 방법의 각 단계들을 설명함에 있어서, 도 4 내지 도 10을 함께 참조하여 설명한다.

**기본 정제 프로세스(S3310)**

도 3 및 도 4를 참조하면, 처리부(330)는 저장 및 관리부(310)로부터 제공되는 운행기록 데이터를 기본 정제할 수 있다(S3310). 여기서, 저장 및 관리부(310)로부터 제공되는 운행기록 데이터는 도 14에 도시된 바와 같은 저장 포맷을 갖는 정제되지 않은 로우 데이터일 수 있으며, 식별 데이터와 센싱 데이터를 포함할 수 있다.

여기서, 상기 식별 데이터는, DTG 단말(100, 도 1 참조)이 장착된 차량 및 해당 차량의 운전자, 해당 장착된 차량이 속한 운송 사업자 등을 식별하기 위한 데이터로, 예컨대 차량의 차대번호, 운전자 식별코드, 운송 사업자 등록번호 등의 필드에 대한 레코드를 포함할 수 있다. 또한 상기 센싱 데이터는, 다양한 센서에 의해 획득되어 DTG 단말(100)에 의해 수집된 데이터로, 예컨대 주행거리, 차량속도, 브레이크 신호, 위치 등의 필드에 대한 레코드를 포함할 수 있다(도 14의 DTG 저장 포맷 참조).

기본 정제 프로세스(S3310)를 상세히 설명하면, 먼저 처리부(330)는 상기 로우 데이터의 운행기록 데이터(보다 구체적으로는, 운행기록 데이터에 포함된 상기 센싱 데이터를 의미함)에서 각 필드 별로 레코드가 정상적인 센싱 범위(즉, 소정의 제한 범위) 내의 값을 갖는지 여부를 판단하여 이상치(outlier)를 검출할 수 있다(S3311).

예를 들면, 차량속도 필드의 경우, 처리부(330)는 각 레코드가 차량 속도 센서가 정상적으로 동작할 때의 센싱 범위(예컨대, 0~250km/h) 내의 값을 갖는지 여부를 판단하여, 상기 센싱 범위 외의 값을 갖는 레코드를 이상치로 검출할 수 있다.

다른 예를 들면, 분당 엔진 회전수 필드의 경우, 처리부(330)는 각 레코드가 분당 엔진 회전수 센서의 센싱 범위(예컨대, 0~6000rpm) 내의 값을 갖는지 여부를 판단하여, 상기 센싱 범위 외의 값을 갖는 레코드를 이상치로 검출할 수 있다.

또 다른 예를 들면, GPS 위치 필드의 경우, 처리부(330)는 각 레코드가 각각 대한민국 영토 범위 내의 경도와 위도에 대응하는 x 및 y 좌표(예컨대,  $125.8 < x < 131.0$ ,  $33.0 < y < 39.0$ ) 내의 값을 갖는지 여부를 판단하여, 상기 범위 외의 값을 갖는 레코드를 이상치로 검출할 수 있다.

한편, 브레이크 신호 센서의 센싱 값이 0과 1인 브레이크 신호 필드의 경우는 상기 0과 1이 이상치를 검출하기 위한 기준치(reference value)로 설정될 수 있고, 처리부(330)는 상기 브레이크 신호 필드의 레코드들 중에서 상기 기준치 이외의 값을 갖는 레코드를 이상치로 검출할 수 있다.

상술한 바에 따라, 처리부(330)는 각 필드 별로 이상치로 검출된 레코드를 제거하거나 보정할 수 있다(S3313). 상술한 예들 중 상기 차량속도 필드의 예를 들면, 처리부(330)는 상기 차량 속도 필드에서 이상치로 검출된 레코드들 중 상기 제한 범위의 상한 또는 하한과 기준 편차(예컨대, 속도 센싱 장치의 센싱 오차 범위)보다 큰 편차를 갖는 레코드는 제거할 수 있고, 또는 상기 기준 편차 이내의 편차를 갖는 레코드는 상기 제한 범위의 상한 또는 하한에 상응하는 값으로 보정할 수 있다.

**비교 정제 프로세스(S3330)**

도 3 및 도 5를 참조하면, 기본 정제 처리된 운행기록 데이터를 다시 비교 정제(S3330)할 수 있다.

일 실시예에서, 처리부(330)는 상기 기본 정제 처리된 운행기록 데이터에서 각 필드 별로 연속하는 레코드들 간

의 편차 비교를 통해 이상치를 검출할 수 있다(S3331).

[0065] 예를 들면, 차량속도 필드에서 연속하는 레코드들이 현격한 차이를 보이는 경우(앞의 레코드가 0km/h, 연이은 뒤의 레코드가 100km/h), 앞의 레코드 및/또는 뒤의 레코드를 이상치로 검출할 수 있다. 유사하게, 가속도 필드, 분당 엔진 회전수 필드, 방위각 필드 등의 경우도 연속하는 레코드들의 편차 비교를 통해 이상치를 검출할 수 있다.

[0066] 다른 실시예에서, 처리부(330)는 상기 기본 정제 처리된 운행기록 데이터에서 서로 상관관계(correlation)에 있는 필드들의 레코드를 비교하여 이상치를 검출할 수도 있다(S3333).

[0067] 예를 들어, 도 6a 및 도 6b를 참조하면, 상기 차량속도 필드와 상기 분당 엔진 회전수 필드의 경우 도 6a에서와 같이 상호 비례관계를 형성(피어슨 상관계수가 0.89826)하는 것이 정상적인데, 비정상적으로 도 6b에서와 같이 상호 반비례관계를 형성(피어슨 상관계수가 -0.79212)할 수 있으므로, 처리부(330)는 서로 상관관계에 있는 필드들의 레코드 비교를 통해 비정상적인 관계를 형성하는 레코드들을 이상치로 검출할 수 있다.

[0068] 이에 따라, 처리부(330)는 편차 비교(S3331) 및/또는 상관관계 비교(S3333)를 통해 이상치로 검출된 레코드들을 제거할 수 있다(S3335).

[0069] 이와 같은, 2단계의 운행기록 데이터 정제를 통해 DTG 빅데이터 처리 및 분석 시스템(300)의 데이터 분석 신뢰성이 향상될 수 있다.

#### [0070] 통계 데이터 추출 프로세스(S3350)

[0071] 도 3 및 도 7을 참조하면, 기본 정제 및 비교 정제 처리된 운행기록 데이터에 근거하여 통계 데이터를 획득할 수 있다(S3350).

[0072] 여기서, 상기 통계 데이터는 정제된 운행기록 데이터를 통계 처리하여 추출한 것으로, 아래의 표 1과 같은 저장 포맷을 가질 수 있다.

표 1

항목	자료형	항목	자료형	항목	자료형
자동차번호	String	급감속건수	Integer	100km 환산 위험과속건수	Float
년월일시	Integer	급출발건수	Integer	100km 환산 장기과속건수	Float
일주행거리	Integer	급정지건수	Integer	100km 환산 급가속건수	Float
총주행시간	Integer	앞지르기건수	Integer	100km 환산 급감속건수	Float
평균속도	Integer	진로변경건수	Integer	100km 환산 급출발건수	Float
최고속도	Integer	급회전건수	Integer	100km 환산 급정지건수	Float
최고RPM	Integer	고RPM건수	Integer	100km 환산 앞지르기건수	Float
과속건수	Integer	공회전건수	Integer	100km 환산 진로변경건수	Float
위험과속건수	Integer	일일총운행시간	Float	100km 환산 급회전건수	Float
장기과속건수	Integer	연속운행시간	Float		
급가속건수	Integer	100km 환산 과속건수	Float		

[0073]

[0074] 여기서, 상기 통계 데이터의 필드 중, DTG 단말(100)이 장착된 차량의 운행과 관련된 필드, 예컨대 일주행거리, 총 주행 시간, 평균 속도, 최고 속도등 필드 및 이에 대한 레코드는 운행 통계 데이터로 분류될 수 있다.

[0075] 또한, DTG 단말(100)이 장착된 차량 운전자의 운전 성향과 관련된 필드, 예컨대 과속 횟수, 위험과속 횟수, 장기과속 횟수, 급가속 횟수, 급감속 횟수, 환산 과속건수, 환산 위험과속건수 필드 및 이에 대한 레코드는 성향 통계 데이터로 분류될 수 있다.

[0076] 도 7을 참조하여 통계 데이터 획득 프로세스(S3350)를 더 상세히 설명하면, 먼저 처리부(330)는 정제 처리된 운

행기록 데이터를 통계적으로 계산하여 통계 데이터를 생성할 수 있다(S3351).

- [0077] 예를 들면, 처리부(330)는 상기 정제 처리된 운행기록 데이터에서 차량속도 필드의 레코드들을 통계적으로 계산(예컨대, 평균치 계산)하여 평균속도 등의 통계 데이터 필드를 생성할 수 있다.
- [0078] 다른 예를 들면, 처리부(330)는 상기 차량속도 필드의 레코드들을 통계적으로 계산(예컨대, GPS 위치 정보를 기초로 확인된 도로의 제한 속도를 제한치로 하여 전 구간 또는 일부 구간에서 상기 제한치 초과 건수의 계산)하여 과속 횟수, 위험과속 횟수, 장기과속 횟수 등의 통계 데이터 필드를 생성할 수 있다. 유사하게, 처리부(330)는 가속도 필드 레코드들을 통계적으로 계산(예컨대, 임의의 값을 기준으로 전 구간 또는 일부 구간에서 상기 제한치 초과 건수의 계산)하여 급가속 횟수, 급감속 횟수 등의 통계 데이터 필드를 생성할 수 있다.
- [0079] 또 다른 예를 들면, 처리부(330)는 차량의 주행거리, 주행시간 필드 레코드들을 통계적으로 계산(예컨대, 누적치 계산, 일부 구간 누적치 계산)하여 일 주행 거리, 상기 차량의 일 주행 시간 등의 통계 데이터 필드를 생성할 수 있다.
- [0080] 또 다른 예를 들면, 처리부(330)는 차량속도, 방위각, 분당 엔진 회전수 등의 필드 레코드들을 복합적으로 통계적으로 계산(예컨대, 차량속도 필드 레코드를 기준으로 방위각 필드 레코드 또는 분당 엔진 회전수 필드 레코드들의 변화를 계산)하여 앞지르기횟수, 진로변경횟수, 급회전횟수 등의 통계 데이터 필드를 생성할 수 있다.
- [0081] 이어서, 처리부(330)는 생성된 복수의 통계 데이터에서 필드 레코드가 기 설정된 시간 범위 내에서 반복 계산된 것인지를 여부를 판단하여 반복 계산된 것으로 판단된 레코드를 보정할 수 있다(S3353).
- [0082] 또한 도 8을 참조하여 상술한 예들 중 과속횟수 필드의 레코드 생성 예를 살펴볼 때, 처리부(330)는 데이터 획득 주기가 1초인 경우 기 설정된 시간 범위(이하, 윈도우), 예컨대 15초 이내에서 과속횟수를 최대 15회로 계산할 수 있다. 이 때 운전자가 연속적으로 과속하고 있는 경우라면, 실질적으로 상기 과속횟수 필드의 레코드가 1로 생성되어야 하나 15로 생성될 수 있으며, 이는 데이터 분석 신뢰성을 저하시킬 수 있다.
- [0083] 따라서, 처리부(330)는 과속횟수 필드의 레코드가 상기와 같이 상기 윈도우 내에서 반복 계산된 것인지를 판단하여 이상치를 검출하고, 상기 과속횟수 필드의 레코드를 1로 보정할 수 있다.
- [0084] 이와 같은, 통계 데이터의 보정을 통해 DTG 빅데이터 처리 및 분석 시스템(300)의 데이터 분석 신뢰성이 더 향상될 수 있다.
- [0085] **단기적 관점의 데이터 마이닝 프로세스(S3370)**
- [0086] 도 3 및 도 9를 참조하면, 처리부(330)는 정제된 운행기록 데이터 및 획득된 통계 데이터 중 적어도 하나에 근거하여 단기적 관점의 데이터 마이닝 분석을 수행하여 특정 결과정보를 생성할 수 있다(S3370).
- [0087] 여기서, 단기적 관점의 데이터 마이닝 분석은 시계열적으로 단기 관점의 분석, 즉 실시간성 데이터 분석 개념은 물론, 개별적인 DTG 단말(100)에 대한 미시적 관점의 데이터 분석이 케이스를 포함할 수 있다. 단기적 관점의 데이터 마이닝을 위한 기법으로는, 예컨대 사고이력, 정비이력 필드 등과 같이 사고/정비 예측과 관련된 직접적인 정보가 있는 경우에는 지도학습형(supervised learning type) 분석 기법인 회귀분석 기법, 의사결정나무 기법, 신경망 기법 등이 이용될 수 있으며, 예컨대 운전자 운전유형 등과 같이 직접적으로 관련된 정보가 없는 경우에는 비지도학습형(unsupervised learning type) 분석 기법인 군집 분석 기법, 연관성 분석 기법 등이 이용될 수 있다.
- [0088] 도 9를 참조하여 단기적 관점의 데이터 마이닝 프로세스(S3370)를 더 상세히 설명하면, 처리부(330)는 획득된 통계 데이터에서 성향과 관련된 필드, 즉 성향 통계 데이터를 기초로 차량 운전자의 운전 성향을 유형 별로 분류할 수 있다(S3371).
- [0089] 일 실시예에서, 처리부(330)는 과속횟수, 위험과속횟수, 장기과속횟수, 급가속횟수, 급감속횟수, 급출발횟수, 앞지르기횟수, 진로변경횟수, 공회전횟수 등에 대한 필드들을 포함하는 성향 통계 데이터에 근거하여 해당 운전자의 운전 유형을 위험 운전자, 조급형 운전자, 에너지 과소비형 운전자 유형 등으로 분류할 수 있다. 이 때, 데이터 마이닝 기법으로는 앞서 예시한 지도학습형 분석 기법들이 이용될 수 있다.
- [0090] 다른 실시예에서, 처리부(330)는 분류된 유형 별로 대응되는 특징을 지수화하여 나타내는 지수정보를 생성할 수



있다(S3373).

- [0091] 예를 들면, 처리부(330)는 상기 위험 운전자 유형에 대하여 과속 횟수, 위험과속 횟수, 장기과속 횟수 등에 대한 필드 레코드를 기초로 위험성의 정도를 지수화하여 나타내는 지수정보를 생성할 수 있다.
- [0092] 다른 예를 들면, 처리부(330)는 에너지 과소비형 운전자 유형에 대하여 분당 엔진 회전수, 급가속 횟수, 공회전 횟수 등에 대한 필드 레코드를 기초로 연비와 이산화탄소 배출량을 추정할 수 있고, 각각의 추정 결과를 지수화하여 연비 지수정보와 환경 지수정보를 생성할 수 있다.
- [0093] 상기 위험운전 지수정보, 상기 연비 지수정보 및 상기 환경 지수정보는, 제어부(350, 도 1 참조) 및 사용자 인터페이스부(370, 도 1 참조)를 통해 운전자에게 제공되어, 운전자가 운전습관을 개선할 수 있도록 한다. 또한, 상기 위험운전 지수정보는, 제어부(350, 도 1 참조) 및 사용자 인터페이스부(370, 도 1 참조)를 통해 운전자가 속한 운송 사업자 등으로 제공되어, 운송 사업자가 상기 운전자를 관리, 지도 및 교육할 수 있도록 한다.
- [0094] 한편, 처리부(330)는 분류된 유형과 무관하게 운행기록 데이터에 포함되는 사고이력, 정비이력 필드를 기초로 사고위험성, 고장위험성을 추정할 수 있고, 사고위험성과 고장위험성 결과를 지수화하여 사고위험 지수정보와 정비대상 지수정보를 생성할 수도 있다.
- [0095] 상기 사고위험 지수정보와 상기 고장위험 지수정보도, 상기 위험운전 지수정보 등과 유사하게 운전자, 운송 사업자 등으로 제공되어, 상기 운전자, 운송 사업자 등이 사고를 방지할 수 있도록 하고, 적정시기에 차량을 정비할 수 있도록 한다.
- [0096] 또한 일 실시예에서, 처리부(330)는 특정한 지수정보와 성향 통계 데이터의 특정 필드사이 또는 특정한 지수정보와 정제된 운행기록 데이터의 특정 필드 사이의 상관관계를 분석하기 위한 마이닝을 수행할 수 있고, 이를 통해 이들 사이의 상호 영향 분석결과를 지시하는 결과정보를 추출할 수 있다(S3375).
- [0097] 예를 들면, 처리부(330)는 상기 위험운전 지수정보를 차량종류 또는 차량연식에 관한 필드 사이의 상관관계를 분석하여 상호 영향을 분석하고, 이들의 상호 영향 분석결과를 지시하는 결과정보를 추출할 수 있다.
- [0098] **장기적 관점의 데이터 마이닝 프로세스(S3390)**
- [0099] 도 3 및 도 10을 참조하면, 처리부(330)는 정제된 운행기록 데이터, 획득된 통계 데이터, 생성된 지수정보 등을 이용하여 장기적 관점의 데이터 마이닝 분석을 수행하여 특정 결과정보를 생성할 수 있다(S3390).
- [0100] 여기서, 장기적 관점의 데이터 마이닝은 시계열적으로 장기 관점의 데이터분석 개념은 물론 다양한 DTG 단말(100)에 대한 거시적 관점의 데이터 분석 개념을 포괄하는 개념이다.
- [0101] 일 실시예에서, 처리부(330)는 정제된 운행기록 데이터, 획득된 통계 데이터, 지수정보에 근거하여, 특정 주기에서의 반복되는 성향을 분석하고, 분석결과를 나타내는 결과정보를 추출할 수 있다(도 10의 S3391).
- [0102] 예를 들면, 처리부(330)는 획득된 통계 데이터의 성향 통계 데이터의 필드들에서 요일별, 시간대별, 계절별, 또는 이벤트별(명절, 연휴, 사고 등의 비정규적인 사건)로 반복되는 패턴을 분석하여, 특정 주기에서의 운전자의 운전성향, 운전습관 등을 나타내는 결과정보를 추출할 수 있다. 상기 결과정보는, 특정 운전자의 특정 주기에서의 운전성향, 운전습관뿐만 아니라, 모든 운전자들의 공통된 운전성향, 운전습관 등을 나타내는 정보도 포함할 수 있다.
- [0103] 한편, 처리부(330)는 위험운전 지수정보, 연비 지수정보 등을 DTG 빅데이터 처리 및 분석 시스템(300)에 기저장되어 있거나 외부 기기로부터 제공되는 날씨정보, 교통정보 등의 기타 정보와 상관 분석하여, 이들 사이의 영향 분석결과를 지시하는 결과정보를 추출할 수도 있을 것이다.
- [0104] 다른 실시예에서, 처리부(330)는 정제된 운행기록 데이터, 획득된 통계 데이터, 지수정보에 근거하여, 특정 필드의 변화를 예측하기 위한 학습 모델을 생성하고, 상기 특정 필드의 변화에 대한 예측 결과를 나타내는 결과정보를 추출할 수 있다(도 10의 S3393).
- [0105] 예를 들면, 처리부(330)는 사고이력 필드를 기초로 사고 발생 시기 등을 예측하기 위한 학습 모델을 만들고, 사

고위험 지수정보와 함께 예측되는 사고 발생 시기와 위험성 등을 나타내는 결과정보를 추출할 수 있다.

[0106] 다른 예를 들면, 처리부(330)는 정비이력 필드를 기초로 고장 발생 시기 등을 예측하기 위한 학습 모델을 만들고, 정비대상 지수정보와 함께 예측되는 정비시기와 고장 위험성 등을 지시하는 결과정보를 추출할 수 있다.

[0107] 한편, 처리부(330)는, 다양한 DTG 단말들 각각에 대한 사고 발생시기와 위험성 또는 고장 발생 시기와 위험성 등을 지시하는 결과정보로부터 공통된 특징, 예컨대 특정 차종(혹은 특정 연식의 차량)의 공통된 사고 발생 시기와 공통된고장 발생 시기 등을 추출하여, 이들을 통해 특정 차종에 대한 사고위험성과 고장위험성 등을 지시하는 결과정보도 생성할 수 있음은 물론이다.

[0108] 또 다른 실시예에서, 처리부(330)는 정제된 운행기록 데이터, 획득된 통계 데이터, 지수정보에 근거하여, 공간 데이터 관련 특징을 분석하고, 분석결과를 나타내는 결과정보를 추출할 수 있다(S3395).

[0109] 예를 들면, 처리부(330)는 다양한 DTG 단말들에 대한 GPS 위치 필드 등과 상기 다양한 DTG 단말들에 대한 위험 운전 지수정보, 연비 지수정보, 환경 지수정보 등을 기초로, 고위험도로 상승급정지 도로, 이산화탄소 배출량 과다 유발 도로 등을 나타내는 결과정보를 추출할 수 있다.

[0110] 한편, 처리부(330)는 다양한 DTG 단말들의 GPS 위치 필드, 위험운전 지수정보, 연비 지수정보, 환경 지수정보 등을, DTG 빅데이터 처리 및 분석 시스템(300, 도 1)에 기저장되어 있거나 외부 기기로부터 제공되는 전자지도 맵 데이터와 매칭시켜 행정 구역별로 운전자들의 성향과 관련된 지수 분포를 나타내는 결과정보를 추출할 수도 있다.

[0111] 또 다른 실시예에서, 처리부(330)는 정제된 운행기록 데이터, 획득된 통계 데이터, 지수정보에 근거하여, 특정 필드를 기준으로 하는 지수정보의 분포상태를 나타내는 결과정보를 추출할 수 있다(S3397).

[0112] 예를 들면, 처리부(330)는 운송 사업자 등록번호 필드를 기준으로, 운전자들의 위험운전 지수, 연비 지수 정보, 환경 지수정보 등의 분포상태를 분석하여, 업체별 운전자 유형, 업체별 지수정보 분포상태를 나타내는 결과정보를 추출할 수 있다.

[0113] 또 다른 실시예에서, 처리부(330)는 정제된 운행기록 데이터, 획득된 통계 데이터, 지수정보에 근거하여, 사용자의 결과정보 인지 전과 인지 후의 변화를 나타내는 결과정보를 추출할 수 있다(S3399).

[0114] 예를 들면, 처리부(330)는, 제어부(350)가 위험운전 지수정보, 연비 지수정보, 환경 지수정보, 결과정보 등을 복합적으로 해석하여 생성한 분석정보, 알람 메시지 등을 해당 운전자에게 제공한 일자, 또는 상기 운전자가 속한 운송 사업자에 제공되어 해당 운전자를 대상으로 하는 교육이 실시된 일자 등의 정보를 상기 운전자, 상기 운송 사업자 등으로부터 제공받을 수 있다. 이에 따라, 처리부(330)는 상기 운전자의 특정 결과정보 인지 전과 인지 후, 또는 상기 운전자의 교육 전과 후에 대한 운행기록 데이터, 획득된 통계 데이터, 지수정보의 변화를 확인하여 운전자의 운전 성향 변화 등을 분석할 수 있고, 분석결과를 나타내는 결과정보를 추출할 수 있다.

[0115] 이와 같은, 복합적이고 효율적인 데이터 마이닝을 통해 DTG 빅데이터 처리 및 분석 시스템(300)의 분석 성능이 향상될 수 있다.

#### [제어부의 기능 및 역할 - 도 11]

[0117] 도 11은 DTG 빅데이터 처리 및 분석 시스템(300)에서 제어부(350)를 보다 상세히 설명하기 위한 도면이다. 도 11을 참조하면, 제어부(350)는 해석부(351) 및 관리부(353)를 포함할 수 있다.

[0118] 해석부(351)는, 처리부(330)에서 생성된 지수정보들과 결과 정보들 및 복수의 통계 데이터를 복합적으로 해석하여 해석결과를 반영한 다양한 분석정보 및 알람 메시지를 생성할 수 있다.

[0119] 예컨대, 해석부(351)는 차량 운전자 개개인의 운전 성향에 관한 분석정보, 연비 지수 및 환경개선 지수 중 적어도 하나에 관한 지수분석정보, 시기별 및 회사별 통계분석정보 등을 생성할 수 있다. 해석부(351)는 사고위험 지수정보 등을 운전자와 운송 사업자 측에 제공하기 위한 알람 메시지를 생성할 수 있다.

- [0120] 또한, 해석부(351)는 행정구역별 운전자들의 운전성향과 특정 도로의 사고 위험, 환경 영향 등을 해석한 결과를 반영하는 분석정보 및 알람 메시지를 생성할 수 있다.
- [0121] 관리부(353)는 상기 분석정보들 및 상기 알람 메시지 등과, 처리부(330)로부터 제공되는 통계 데이터, 지수정보들 및 결과정보들을 데이터베이스(370)에 저장하여 관리할 수 있다.
- [0122] 관리부(353)는 기 설정된 주기에 따라 또는 상기 운전자, 상기 운송 사업자, 교통안전공단 등의 요청이 있는 경우, 사용자 인터페이스부(390)를 통해 상기 분석정보 등을 상기 운전자, 상기 운송 사업자, 상기 교통안전공단 등에 제공할 수 있다.
- [0123] 이에 따라, 차량 운전자의 안전 운전 유도, 차량 연비 향상 유도, 사전 정비 유도 및 사고 방지가 가능해지고, 운송 사업자의 관리 효율성이 향상될 수 있으며, 교통안전공단과 같은 공공기관에서 사고위험 및 환경악화를 초래하는 행정구역, 도로 등을 효율적으로 개선할 수 있도록 한다.
- [0124] **[사용자 인터페이스부의 기능 및 역할 - 도 12, 도 13a 내지 도 13f]**
- [0125] 도 12는 사용자 인터페이스부(390)를 보다 상세히 설명하기 위한 도면이고, 도 13a 내지 도 13f는 사용자 인터페이스부(390)에서 제공하는 시각화(visualization) 구현 예들을 나타낸 도면이다.
- [0126] 도 12를 참조하면, 사용자 인터페이스부(390)는 시각화 처리부(391)를 포함할 수 있다.
- [0127] 시각화 처리부(391)는, 제어부(350)로부터 제공되는 통계 데이터, 결과정보 등을 시각화하여 운전자, 운송 사업자 등의 사용자가 직접 인지할 수 있도록 한다. 예컨대, 시각화 처리부(391)는 운행기록 데이터에 포함된 단일의 운행기록 필드 별로의 군집분석에 관한 시각화, 단일 운행기록 필드 별로의 통계분석에 관한 시각화, 단일 운행기록 필드 별로의 스케일 재분류에 관한 시각화, 운행기록 데이터 중 상관관계에 있는 어느 2개의 운행기록 필드 간의 교차분석에 관한 시각화, 운행기록 데이터에 관한 시계열적 로우 데이터 분석에 관한 시각화 등을 수행할 수 있다.
- [0128] 또한, 시각화 처리부(391)는 앞서 설명된 단기적 관점의 데이터 마이닝 프로세스(S3370) 및 장기적 관점의 데이터 마이닝 프로세스(S3390)의 다양한 마이닝 분석과 관련된 시각화 처리, 즉 마이닝 분석을 수행하여 결과 정보를 도출하기 위한 시각화 처리를 수행할 수도 있다. 이와 같은, 시각화 처리부(391)의 시각화 처리를 통해, 사용자가 직접 단기적 관점의 데이터 마이닝 프로세스(S3370) 및 장기적 관점의 데이터 마이닝 프로세스(S3390)의 다양한 마이닝 분석을 수행할 수 있음은 물론이다.
- [0129] 도 13a 내지 도 13f를 참조하여 시각화의 예들을 더 상세히 설명하면, 시각화 처리부(391)는 도 13a에 도시된 바와 같이 통계 데이터에서 차량들의 과속횟수, 급가속횟수 등의 필드 레코드들을 군집분석이 가능하도록 시각화할 수 있다. 또한, 시각화 처리부(391)는 도 13b에 도시된 바와 같이 운행기록 데이터에서 차량의 속도, 분당 엔진 회전수 등의 필드 레코드들을 교차분석이 가능하도록 시각화할 수 있다.
- [0130] 시각화 처리부(391)는 도 13c 및 도 13d에 도시된 바와 같이 통계 데이터의 평균속도와 최고속도 필드를 각각 히스토그램으로 표현할 수 있으며, 도 13e에 도시된 바와 같이 통계 데이터에서 차량들의 과속건수 이상치를 쉽게 확인할 수 있도록 산포도와 같은 형태로 시각화할 수 있고, 또한 도 13f에 도시된 바와 같이 운행기록 데이터에서 시계열적인 차량속도 필드 레코드의 이상치를 쉽게 확인할 수 있도록 시각화할 수 있다.
- [0131] 이와 같이, DTG 빅데이터 처리 및 분석 시스템(300)은, 추출된 정보들을 사용자 인터페이스부(390)를 통해 시각화 처리하여 사용자에게 제공함으로써, 사용자가 빅데이터로부터 추출된 정보들을 직관적으로 인지할 수 있도록 하고 효과적으로 이용할 수 있도록 하며, 또한 직접 마이닝 분석도 수행할 수 있도록 한다.
- [0132] 이상, 본 발명을 바람직한 실시예를 들어 상세하게 설명하였으나, 본 발명은 상기 실시예에 한정되지 않고, 본 발명의 기술적 사상 및 범위 내에서 당 분야에서 통상의 지식을 가진 자에 의하여 여러가지 변형 및 변경이 가능하다.



# 부호의 설명

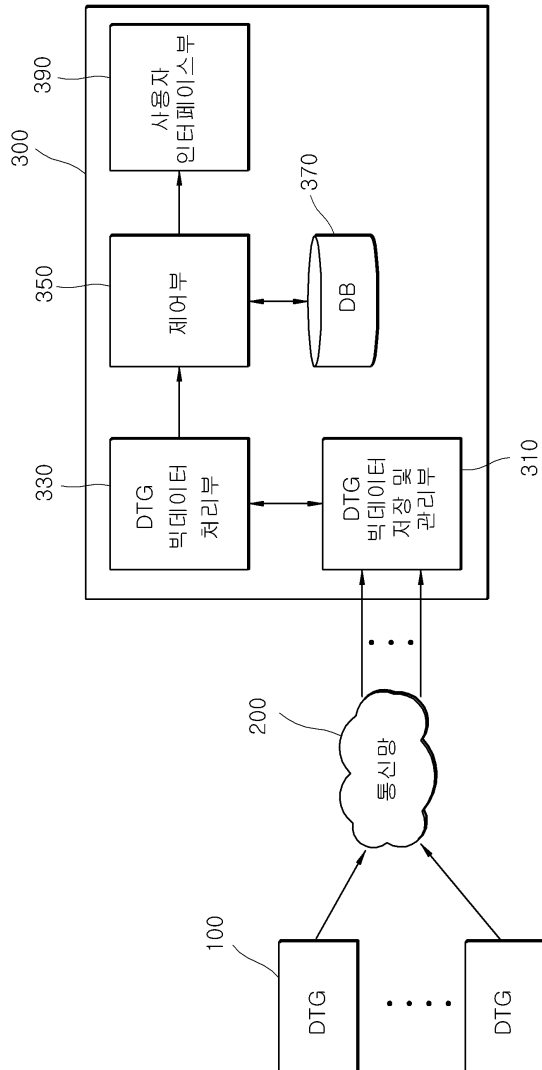
100: DTG 단말

200: 통신망

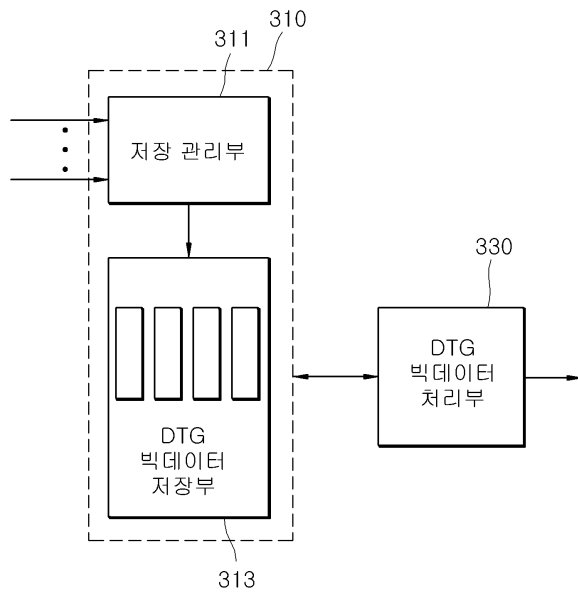
300: DTG 빅데이터 처리 및 분석 시스템

## 도면

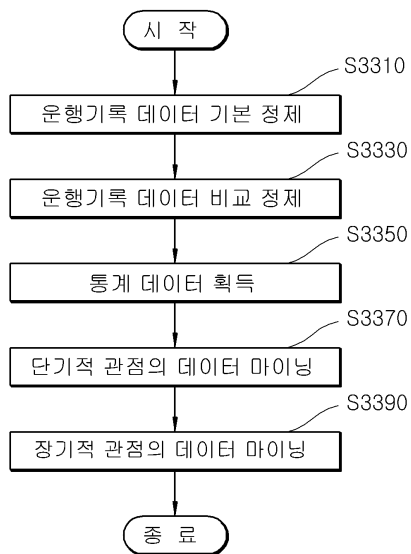
### 도면1



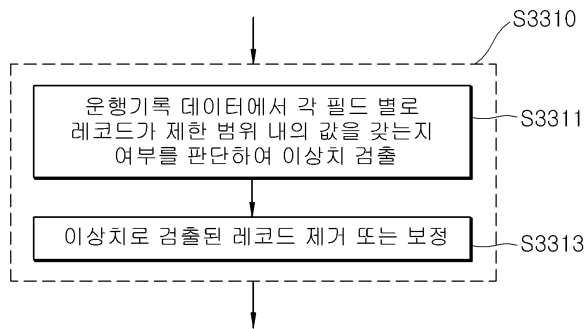
도면2



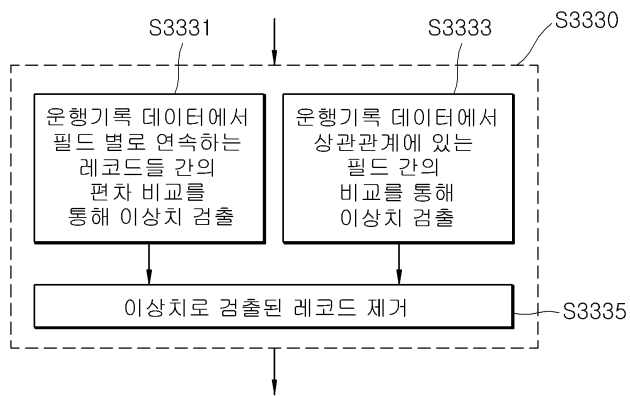
도면3



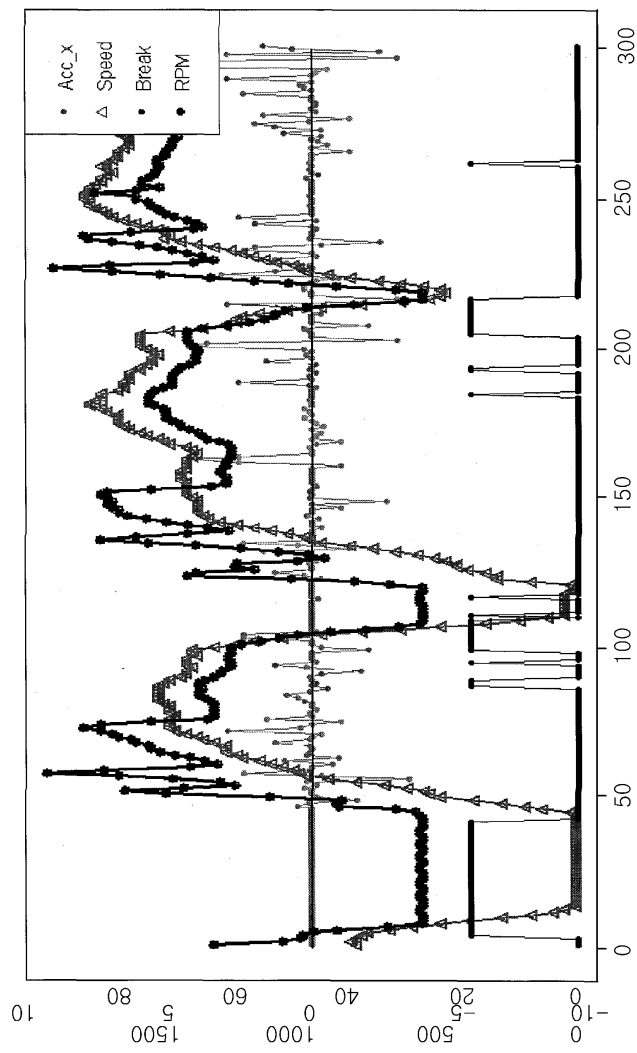
도면4



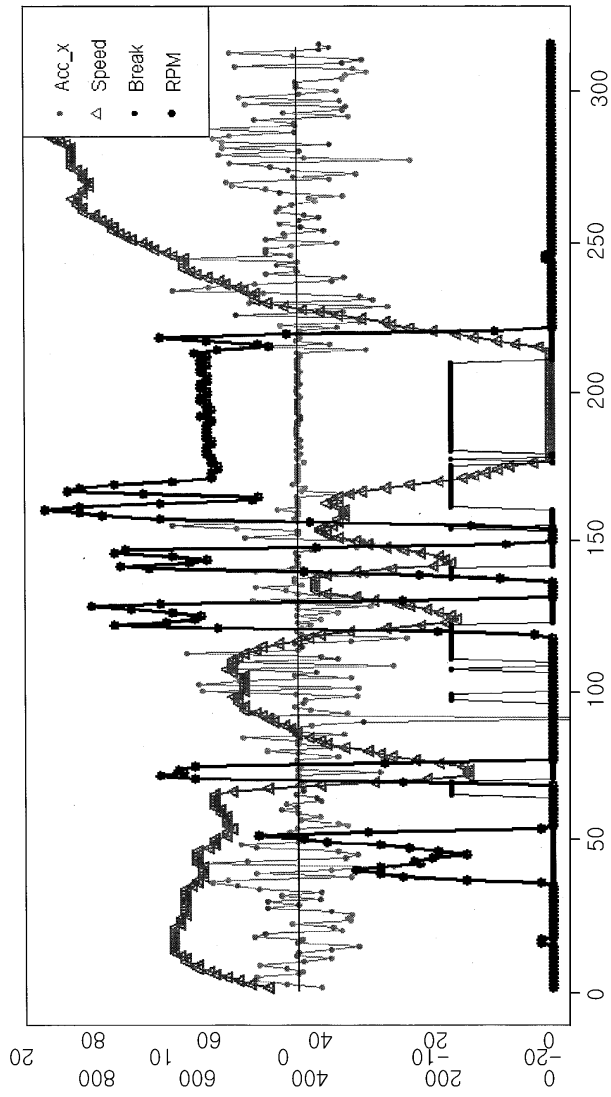
도면5



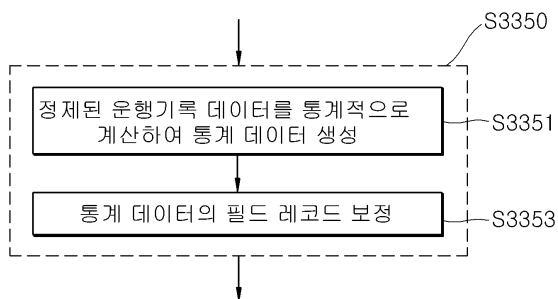
도면6a



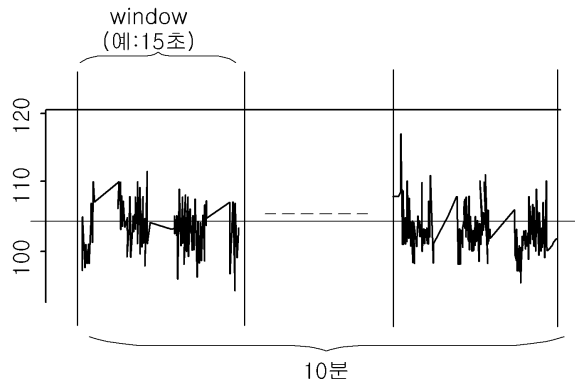
도면6b



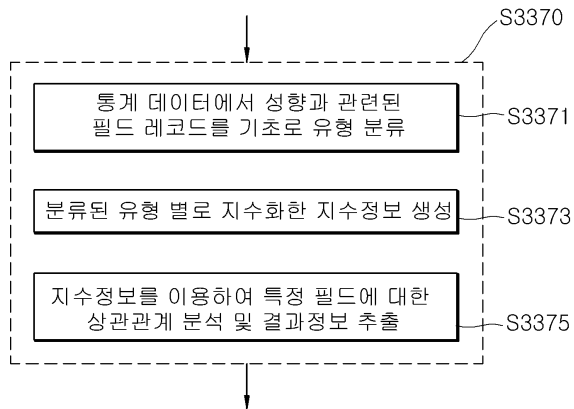
도면7



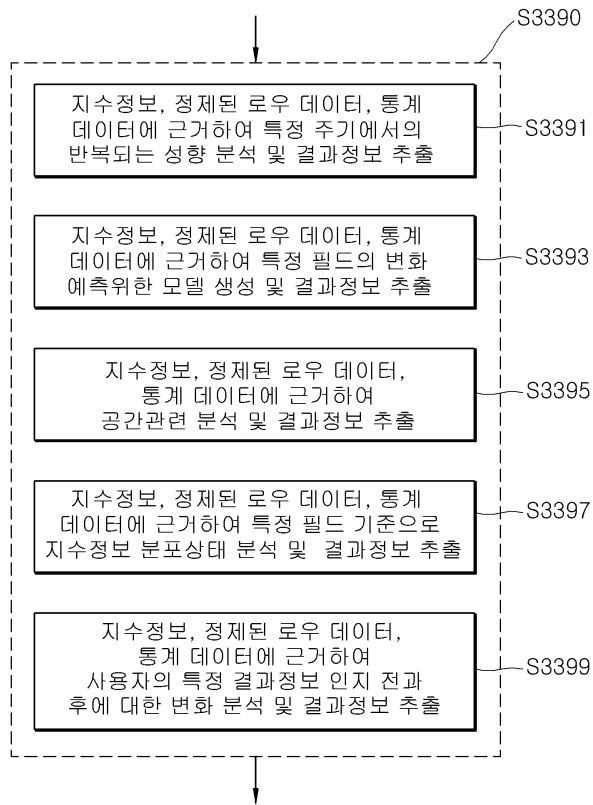
도면8



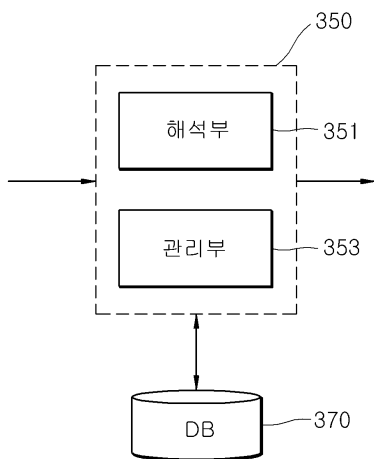
도면9



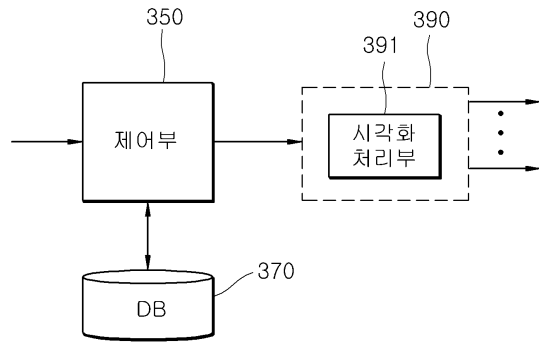
도면10



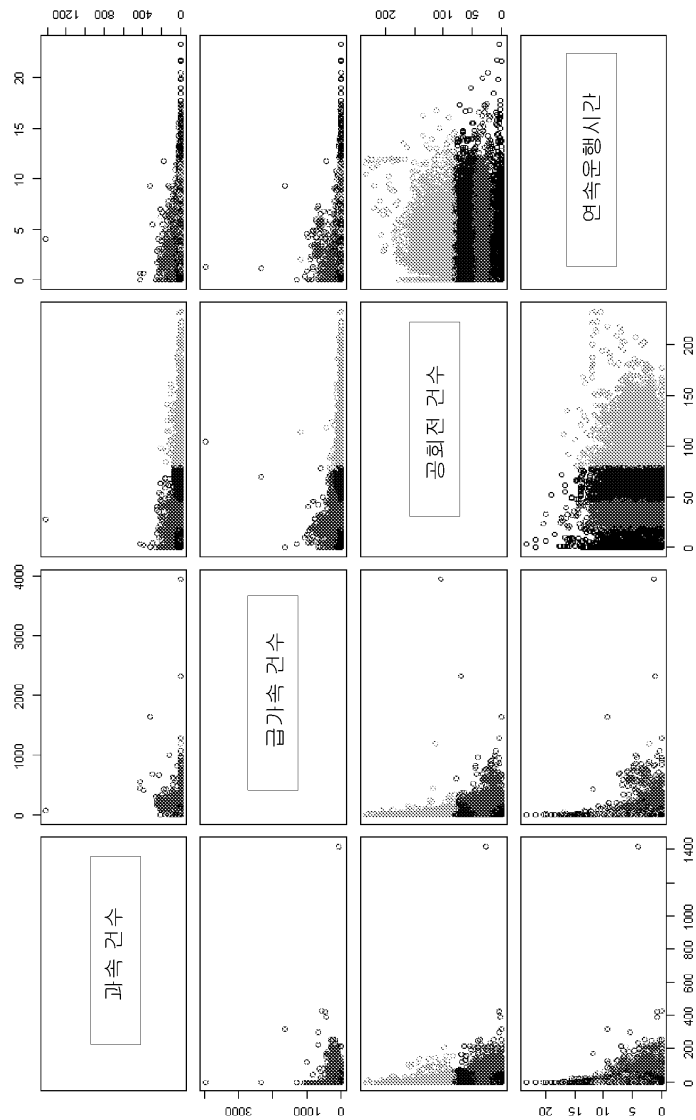
도면11



도면12

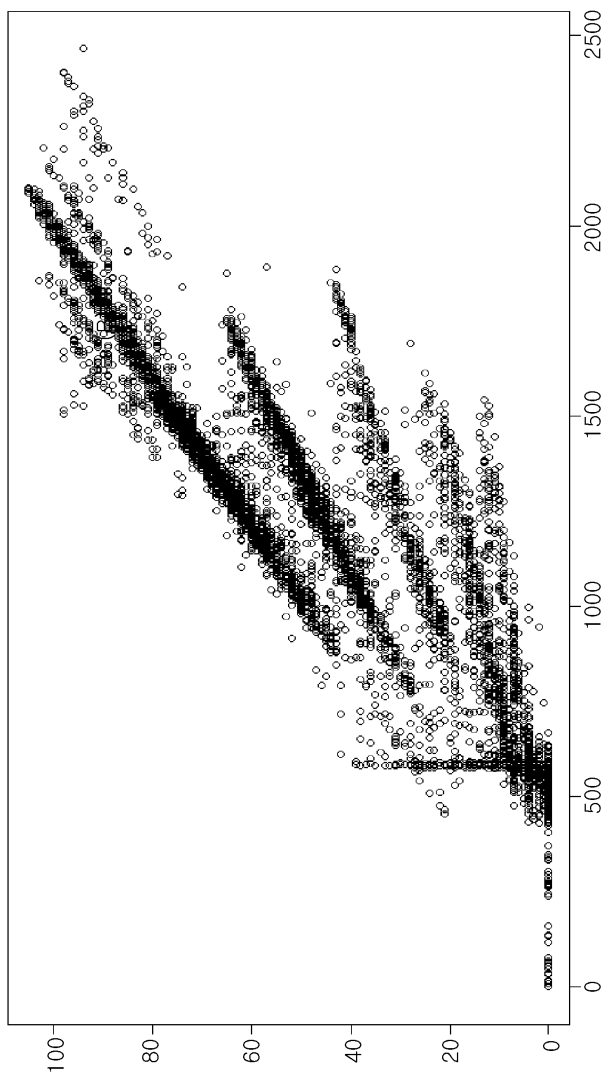


도면13a

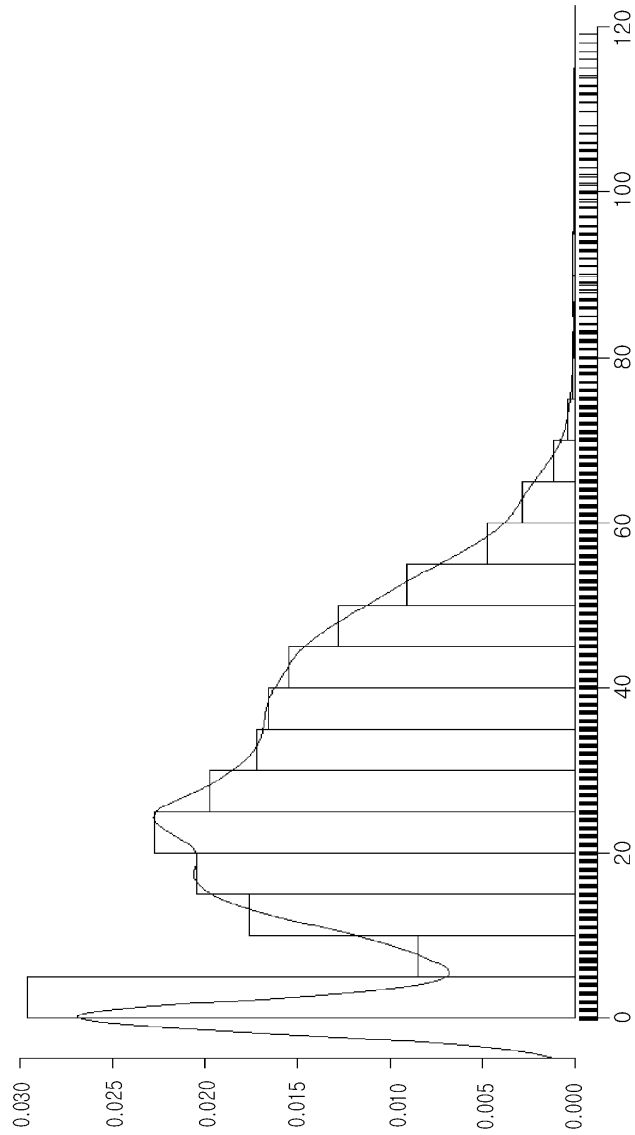




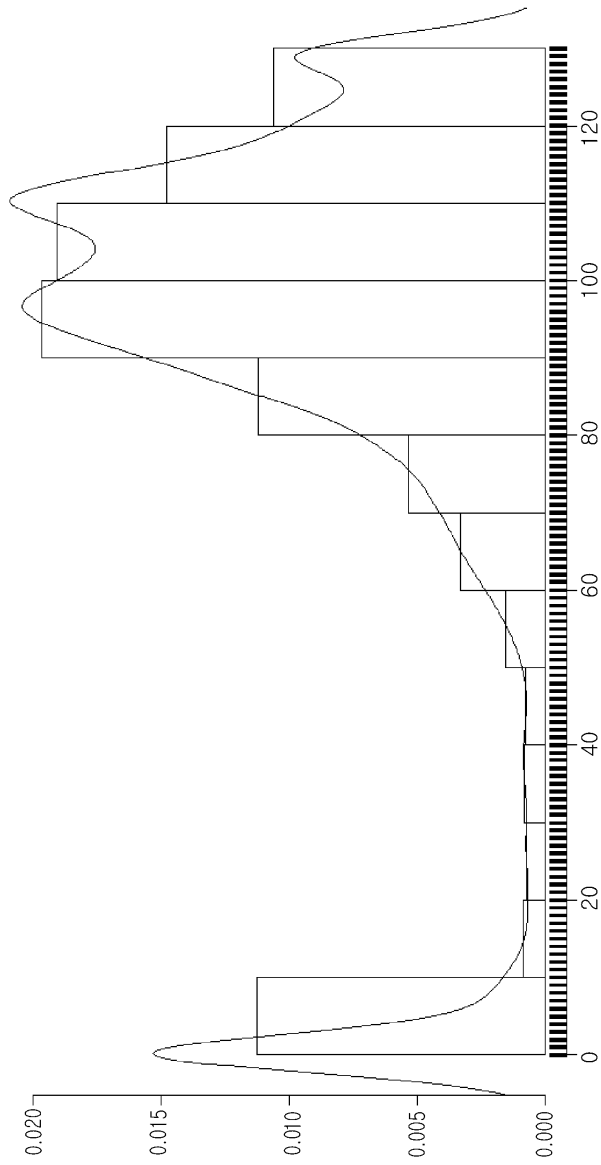
도면13b



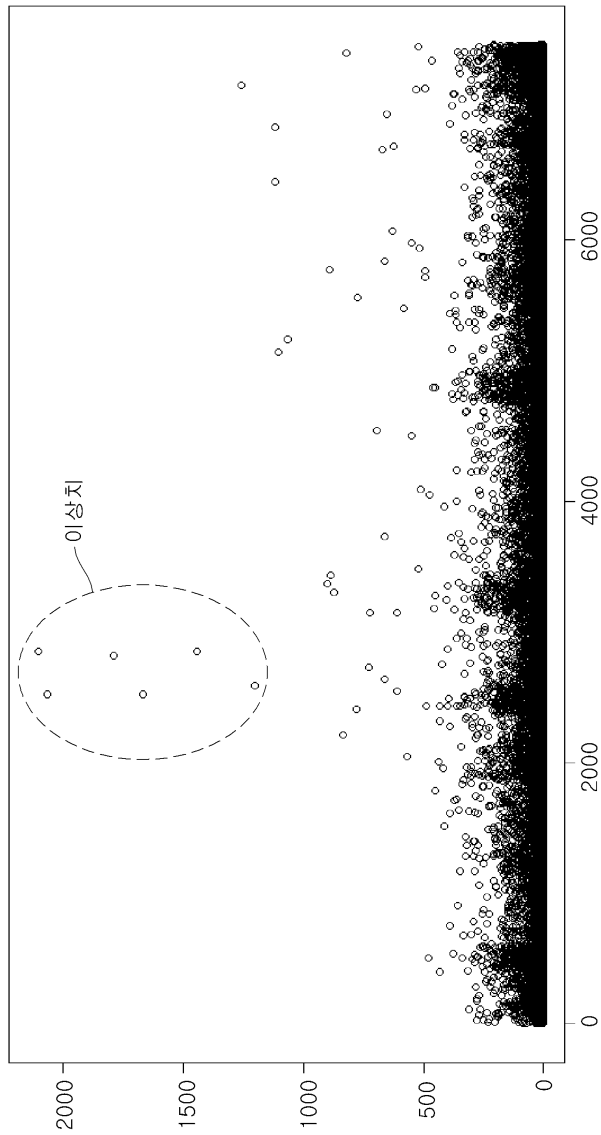
도면13c



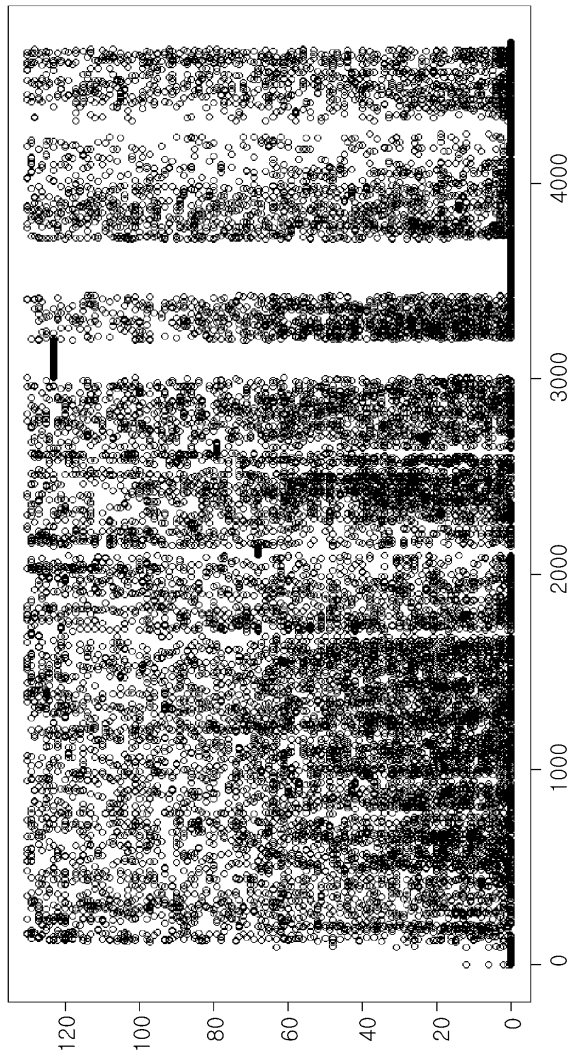
도면13d



도면13e



도면13f



도면14

항목		자릿수	표기방법	표기사항
운행기록장치 모델명		20	오른쪽으로 정렬하고 빈칸은 '#'으로 표기	최초 사용 시 등록
자대번호		17	영문(대문자)아라비아숫자 전부 표기	상등
자동차 유형		2	11: 시내버스, 12: 농어촌버스, 13: 마을버스 14: 시외버스, 15: 고속버스, 16: 전세버스 17: 특수여객자동차, 21: 일반택시, 22: 개인택시 31: 일반화물자동차, 32: 개별화물자동차 41: 비사업용 자동차	"
자동차 등록번호		12	자동차등록번호 전부 표기 (한글 하나에 두 자리, 빈칸은 '#'으로 표기)	"
운송사업자 등록번호		10	사업자등록번호 전부 표기	"
운전자코드		18	운전자의 자격증번호 (빈칸은 '#'으로 표기하고, 중간자 '-'는 생략)	자동차 운송사업자 설정
주행 거리 (km)	일일주행거리	4	00시부터 24시까지 주행한 거리 (범위: 0000 ~ 9999)	실시간
	누적주행거리	7	최초등록일로부터 누적한 거리 (범위: 000000 ~ 9999999)	"
정보발생 일시		14	YYMMDDhhmmss (연/월/일/시/분/0.01초)	"
차량속도(km/h)		3	범위: 000 ~ 255	"
분당 엔진회전수(RPM)		4	범위: 0000 ~ 9999	"
브레이크 신호		1	범위: 0(off) 또는 1(on)	"
차량위치 (GPS X, Y 좌표)	X	9	10진수로 표기	"
	Y	9	(예: 127.123456*1000000 = 127123456)	
위성항법장치(GPS) 방위각		3	범위: 0 ~ 360 (0 ~ 360°에서 1°를 1로 표현)	"
가속도 (m/sec <sup>2</sup> )	ΔVx	6	범위: -100.0 ~ +100.0	"
	ΔVy	6		
기기 및 통신상태코드 (백업 수집 주기 내)		2	00: 운행기록장치 정상 11: 위치추적장치(GPS수신기) 이상 12: 속도센서 이상 13: RPM 센서 이상 14: 브레이크 신호감지 센서 이상 21: 센서 입력부 장치 이상 22: 센서 출력부 장치 이상 31: 데이터 출력부 장치 이상 32: 통신 장치 이상 41: 운행거리 산정 이상 99: 전원공급 이상	"