



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2015년05월04일

(11) 등록번호 10-1516679

(24) 등록일자 2015년04월24일

(51) 국제특허분류(Int. Cl.)

G06F 17/00 (2006.01)

(21) 출원번호 10-2014-0159809

(22) 출원일자 2014년11월17일

심사청구일자 2014년11월17일

(56) 선행기술조사문헌

KR1020140067065 A*

*는 심사관에 의하여 인용된 문헌

(73) 특허권자

주식회사 넷스루

서울특별시 강남구 선릉로90길 44, 정인빌딩 8층, 9층, 10층(대치동)

(72) 발명자

서동열

경기도 성남시 분당구 내정로 152, 128동 903호 (수내동, 파크타운롯데아파트)

조보연

경기도 용인시 수지구 정든로 22, 902동 510호 (죽전동, 도담마을죽전파크빌)

최현준

서울특별시 관악구 인현13길 6-6, 2층

(74) 대리인

이철희

전체 청구항 수 : 총 8 항

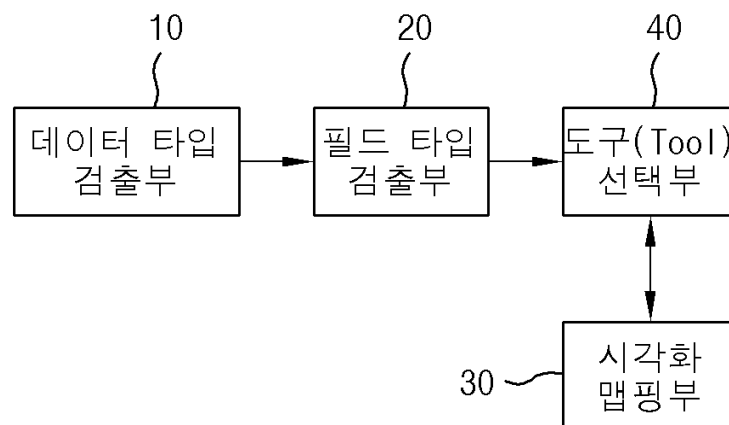
심사관 : 이석형

(54) 발명의 명칭 차트 추천 장치 및 방법

(57) 요약

본 발명은 데이터 파일 또는 시트의 선택된 영역의 필드를 분석하여 차트와 같은 시각화 방법을 선택하여 추천하는 차트 추천 장치 및 방법을 개시하며, 차트 추천 장치는 선택된 블록의 데이터 단위들에 적용되는 필드 별로 데이터 타입을 검출하는 데이터 타입 검출부, 필드 별로 필드 타입을 검출 및 결정하는 필드 타입 검출부, 시각화 코드에 대응하여 맵핑 룰을 갖는 하나 이상의 시각화 도구를 제공하는 시각화 맵핑부, 및 시각화 코드를 정의하고 시각화 코드에 해당하는 시각화 도구를 제공하는 도구 선택부를 포함한다.

대표도 - 도1



이 발명을 지원한 국가연구개발사업

과제고유번호 10047040

부처명 미래창조과학부

연구관리전문기관 한국산업기술평가관리원

연구사업명 산업융합원천기술개발사업

연구과제명 빅데이터 검색 분석 결과 시각화를 위한 간편화 UI 개발 지원 도구

기 여 율 1/1

주관기관 (주)넷스루

연구기간 2013.09.01 ~ 2014.12.31

명세서

청구범위

청구항 1

삭제

청구항 2

분석을 위하여 선택된 블록의 데이터 단위들에 적용되는 필드 별로 데이터 타입을 검출하며, 상기 데이터 타입은 제1 속성값들에 의하여 복수로 구분되는 데이터 타입 검출부;

상기 필드 별로 상기 제1 속성값과 필드값들 중 어느 하나를 기준으로 필드 타입을 검출 및 결정하며, 상기 필드 타입은 제2 속성값들에 의하여 복수로 구분되는 필드 타입 검출부;

시각화 코드에 대응하여 맵핑 룰을 갖는 하나 이상의 시각화 도구를 제공하는 시각화 맵핑부; 및

상기 필드 타입 검출부에서 결정된 상기 필드 별 상기 필드 타입을 참조하여 상기 시각화 코드를 정의하고 상기 시각화 코드에 해당하는 하나 이상의 상기 시각화 도구를 상기 시각화 맵핑부에서 구하며, 상기 맵핑 룰에 의해 표현되는 하나 이상의 상기 시각화 도구를 제공하는 도구 선택부;를 포함하며,

상기 필드 타입 검출부는,

상기 제2 속성값에 일대일 매칭되는 상기 제1 속성값을 갖는 필드에 대하여 상기 제1 속성값에 매칭되는 상기 제2 속성값으로 상기 필드 타입을 검출하고,

상기 제2 속성값에 일대일 매칭되지 않는 상기 제1 속성값을 갖는 필드에 대하여 필드값들을 분석하고 상기 제2 속성값을 판단함으로써 상기 필드 타입을 검출하는 차트 추천 장치.

청구항 3

제2 항에 있어서,

상기 데이터 타입 검출부는, 상기 필드의 상기 데이터 타입을 데이트-타임, 스트링 및 정수 및 실수를 포함하는 상기 제1 속성값들로 구분하고,

상기 필드 타입 검출부는 상기 필드의 타입을 데이터-타임, 범주형 및 숫자형을 포함하는 상기 제2 속성값들로 구분하며, 상기 데이터 타입이 상기 정수 또는 상기 실수로 구분된 상기 필드에 대하여 범주형 및 숫자형으로 구분하는 차트 추천 장치.

청구항 4

제2 항에 있어서,

상기 블록은 현재 실행 중인 어플리케이션 프로그램의 시트에서 선택하여 제공되는 차트 추천 장치.

청구항 5

제2 항에 있어서,

상기 블록은 데이터 파일에 포함된 전체 또는 일부 데이터 단위들을 분석하여 제공되는 차트 추천 장치.

청구항 6

삭제

청구항 7

분석을 위하여 어플리케이션 프로그램 또는 데이터 파일에서 선택된 블록의 데이터 단위들에 적용되는 필드 별로 데이터 타입을 검출하는 단계;

제1 속성값과 필드값들 중 어느 하나를 기준으로 상기 데이터 타입이 복수로 구분된 상기 블록의 상기 데이터 단위들에 대하여 상기 필드 별로 필드 타입을 검출 및 결정하는 단계;

시각화 코드에 대응하여 맵핑 룰을 갖는 하나 이상의 시각화 도구를 제공하는 단계;

제2 속성값들에 의하여 복수로 구분되는 상기 필드별 상기 필드 타입을 참조하여 상기 시각화 코드를 정의하는 단계;

정의된 상기 시각화 코드에 해당하는 하나 이상의 시각화 도구를 구하는 단계; 및

구해진 하나 이상의 상기 시각화 도구에 대응하는 맵핑 룰에 따르는 시각화 도구를 제공하는 단계;를 포함하며, 상기 제1 속성값들에 의하여 상기 데이터 타입을 복수로 구분하여 상기 데이터 타입을 검출하는 단계는 데이터 타입 검출부에서 수행되고,

상기 제2 속성값들에 의하여 상기 필드 타입을 복수로 구분하여 상기 필드 타입을 검출 및 결정하는 단계는 필드 타입 검출부에서 수행되며,

상기 시각화 코드에 대응하는 하나 이상의 시각화를 위한 시각화 도구를 제공하는 단계는 시각화 맵핑부에서 수행되고,

상기 시각화 코드의 정의, 정의된 상기 시각화 코드에 해당하는 하나 이상의 상기 시각화 도구를 구하는 것 및 상기 시각화 도구에 대응하는 상기 맵핑 룰에 따르는 상기 시각화 도구는 도구 선택부에서 수행됨을 특징으로 하는 차트 추천 방법.

청구항 8

제7 항에 있어서,

상기 시각화 코드는 상기 블록에 포함되는 전체 필드에 대응하는 하나 이상의 상기 제2 속성값을 조합하여 결정되며,

상기 제2 속성값은,

정의된 제2 속성값에 일대일 매칭되는 상기 제1 속성값을 갖는 필드에 대하여 상기 제1 속성값에 매칭되는 값으로 결정되고,

정의된 제2 속성값에 일대일 매칭되지 않는 상기 제1 속성값을 갖는 필드에 대하여 필드값들을 분석하고 상기 제2 속성값을 판단하여 결정되는 차트 추천 방법.

청구항 9

제7 항에 있어서,

상기 블록은 현재 실행 중인 어플리케이션 프로그램의 시트에서 선택하여 제공되는 차트 추천 방법.

청구항 10

제7 항에 있어서,

상기 블록은 데이터 파일에 포함된 전체 또는 일부 데이터 단위들을 분석하여 제공되는 차트 추천 방법.

발명의 설명

기술 분야

[0001] 본 발명은 차트 추천 장치 및 방법에 관한 것으로서, 보다 상세하게는 데이터 파일 또는 시트의 선택된 영역의 필드를 분석하여 차트와 같은 시각화 방법을 선택하여 추천하는 차트 추천 장치 및 방법에 관한 것이다.

배경 기술

[0002] 데이터베이스나 스프레드 시트를 활용하는 다양한 프로그램이나 어플리케이션이 많이 활용되고 있으며, 데이터베이스와 스프레드 시트에 포함된 정보들은 차트를 이용하여 시각적으로 표현될 수 있다.

[0003] 차트는 막대, 도넛, 라인, 히스토그램 등 기초적인 기능부터 고급 기능을 구현하기 위한 것까지 다양한 형태로 표현될 수 있다.

[0004] 그러나, 일반 사용자들은 어떤 차트를 사용해야 확보된 정보를 이용하여 원하는 목적에 맞는 효과적인 시각 효과를 얻을 수 있을지 잘 알지 못한다. 즉, 정보를 시각화하여 표현하는 다양한 방법이 있음에도, 사용자는 시각화 도구의 종류, 표현 방법, 사용 방법 등을 잘 몰라서 차트를 적절하게 이용하지 못한다.

[0005] 이는 차트로 구현되는 그래프나 다이어그램의 종류가 많을 뿐만 아니라 데이터의 성격에 따라 효과적으로 사용할 수 있는 그래프가 각각 다르며, 각각의 그래프나 다이어그램에 따른 설정 항목과 표현 방법이 달라서 쉽게 사용하기 힘들기 때문이다.

[0006] 따라서, 일반 사용자들이 확보된 정보를 이용하여 효과적인 차트를 구현할 수 있는 방법의 제시가 필요하다.

발명의 내용

해결하려는 과제

[0007] 본 발명의 목적은 사용자가 손쉽게 정보를 시각화할 수 있도록 정보를 분석하고 정보의 특성을 판단하며 정보의 특성에 맞게 시각화 도구를 추천할 수 있는 차트 추천 장치 및 방법을 제공함에 있다.

[0008] 본 발명의 다른 목적은 프로그램이나 어플리케이션 또는 클라이언트-서버 환경에서 사용자가 선택한 정보를 차트를 이용하여 손쉽게 시각화할 수 있는 차트 추천 장치 및 방법을 제공함에 있다.

과제의 해결 수단

[0009] 본 발명의 차트 추천 장치는, 분석을 위하여 선택된 블록의 데이터 단위들에 적용되는 필드 별로 데이터 타입을 검출하며, 상기 데이터 타입은 제1 속성값들에 의하여 복수로 구분되는 데이터 타입 검출부; 상기 필드 별로 필드 타입을 검출 및 결정하며, 상기 필드 타입은 제2 속성값들에 의하여 복수로 구분되는 필드 타입 검출부; 시각화 코드에 대응하여 맵핑 룰을 갖는 하나 이상의 시각화 도구를 제공하는 시각화 맵핑부; 및 상기 필드 타입 검출부에서 결정된 상기 필드 별 상기 필드 타입을 참조하여 상기 시각화 코드를 정의하고 상기 시각화 코드에 해당하는 하나 이상의 상기 시각화 도구를 상기 시각화 맵핑부에서 구하며, 상기 맵핑 룰에 의해 표현되는 하나 이상의 상기 시각화 도구를 제공하는 도구 선택부;를 포함함을 특징으로 한다.

[0010] 또한, 본 발명의 차트 추천 방법은, 분석을 위하여 어플리케이션 프로그램 또는 데이터 파일에서 선택된 블록의 데이터 단위들에 적용되는 필드 별로 데이터 타입을 검출하는 단계; 제1 속성값들에 의하여 상기 데이터 타입이 복수로 구분된 상기 블록의 상기 데이터 단위들에 대하여 상기 필드 별로 필드 타입을 검출 및 결정하는 단계; 시각화 코드에 대응하여 맵핑 룰을 갖는 하나 이상의 시각화 도구를 제공하는 단계; 제2 속성값들에 의하여 복수로 구분되는 상기 필드별 상기 필드 타입을 참조하여 상기 시각화 코드를 정의하는 단계; 정의된 상기 시각화 코드에 해당하는 하나 이상의 시각화 도구를 구하는 단계; 및 구해진 하나 이상의 상기 시각화 도구에 대응하는 맵핑 룰에 따르는 시각화 도구를 제공하는 단계;를 포함함을 특징으로 한다.

발명의 효과

[0011] 본 발명에 의하면 사용자가 분석을 위하여 선택한 정보의 특성에 맞는 차트들이 추천될 수 있고, 사용자는 추천된 차트들 중 원하는 목적에 맞는 차트를 선택할 수 있다. 그러므로, 사용자가 편리하게 차트를 이용할 수 있는 이점이 있다.

[0012] 또한, 본 발명에 의하면 프로그램이나 어플리케이션 또는 클라이언트-서버 환경에서 차트를 추천하고 선택할 수 있는 기능이 제공될 수 있어서 사용자는 다양한 환경에서 선택한 정보에 맞는 차트를 손쉽게 추천받고 이용할 수 있는 이점이 있다.

도면의 간단한 설명

[0013] 도 1은 본 발명의 차트 추천 장치의 바람직한 실시예를 나타내는 블록도.

도 2는 본 발명의 차트 추천 방법의 바람직한 실시예를 나타내는 흐름도.

도 3은 어플리케이션 환경에서 구현되는 차트 추천 장치를 예시한 블록도.

도 4는 클라이언트-서버 환경에서 구현되는 차트 추천 장치를 예시한 블록도.

도 5는 차트의 일 예를 표현한 그래프.

도 6은 차트의 다른 예를 표현한 그래프.

도 7은 데이터 파일의 분석을 위하여 이용될 수 있는 인터페이스 윈도우를 예시한 도면.

도 8은 프로그램 상에서 이용될 수 있는 인터페이스 윈도우를 예시한 도면.

발명을 실시하기 위한 구체적인 내용

[0014] 이하, 첨부된 도면을 참조하여 본 발명의 바람직한 실시예를 상세하게 설명한다. 본 명세서 및 특허청구범위에 사용된 용어는 통상적이거나 사전적 의미로 한정되어 해석되지 아니하며, 본 발명의 기술적 사항에 부합하는 의미와 개념으로 해석되어야 한다.

[0015] 본 명세서에 기재된 실시예와 도면에 도시된 구성은 본 발명의 바람직한 실시예이며, 본 발명의 기술적 사상을 모두 대변하는 것이 아니므로, 본 출원 시점에서 이들을 대체할 수 있는 다양한 균등물과 변형예들이 있을 수 있다.

[0016] 본 발명은 데이터베이스, 파일 시스템이나 스프레드 시트를 활용하는 다양한 프로그램이나 어플리케이션에 포함된 정보들을 차트를 이용하여 시각적으로 표현하기 위한 기술을 개시하며, 사용자가 선택한 정보를 시각화할 수 있도록 정보를 분석하고 정보의 특성을 판단하며 특성에 맞게 자동으로 시각화 도구를 추천하는 기술을 개시한다.

[0017] 사용자는 차트로 표현하기 위한 정보를 선택할 수 있으며, 정보는 데이터베이스, 파일 시스템이나 스프레드 시트와 같이 데이터 단위들을 포함하는 것을 총칭하는 것으로 이해될 수 있으며, 각 데이터 단위들은 하나 이상의 필드에 대응하는 값을 가질 수 있다. 여기에서, 데이터 단위는 데이터 레코드(Record)로 이해될 수 있다.

[0018] 이하, 본 발명의 실시예의 설명의 편의를 위하여 하기 <표 1>을 예시할 수 있다. <표 1>은 사용자가 선택한 정보의 일 예이고, 요일 별 도로 이용 상황을 예시한 것이며, 필드로 설정된 '요일'은 월, 화, 수, 목, 금, 토, 일로 구분될 수 있고, 필드로 설정된 '도로 명'은 제1 도로 내지 제8 도로 및 기타로 구분될 수 있으며, 필드로 설정된 '카운트(Count)'는 이용한 수를 기재한 것이다. 각 데이터 단위는 '요일', '도로 명' 및 '카운트'에 대응하는 값을 갖는다.

표 1

요일	도로명	Count
월	제2 도로	4532
월	제1 도로	0
화	제8 도로	4742
수	제7 도로	5512
...		
토	제2 도로	2233
토	기타	5616
일	제5 도로	1121
일	제4 도로	1963

[0020] 도 1을 참조하면, 본 발명의 차트 추천 장치의 실시예는 데이터 타입 검출부(10), 필드 타입 검출부(20), 시각화 맵핑부(30) 및 도구(Tool) 선택부(40)를 포함한다.

[0021] 데이터 타입 검출부(10)는 분석을 위하여 <표 1>과 같이 선택된 블록의 데이터 단위들에 적용되는 필드 별로 데이터 타입을 검출하며, 데이터 타입은 제1 속성값들에 의하여 복수로 구분될 수 있다.

[0022] 여기에서, 데이터 타입은 원시 데이터의 유형을 구분하기 위한 것이며, 일 예로, '정수(Integer)', '실수(Real)', '데이트-타임(Date-Time)', '문자열(String)'로 구분할 수 있다. '정수(Integer)', '실수(Real)', '데이트-타임(Date-Time)', '문자열(String)'은 제1 속성값들을 예시한 것이다.

[0023] 필드 타입 검출부(20)는 필드 별로 필드 타입을 검출 및 결정하며, 필드 타입은 제2 속성값들에 의하여 복수로

구분될 수 있다.

[0024] 필드 타입은 필드의 유형을 판단하기 위한 것이며, 일 예로, '범주형(Categorical)' 데이터, '숫자형(Numerical)' 데이터, '데이트-타임형(Date-Time)' 데이터로 구분할 수 있다. '범주형(Categorical)' 데이터, '숫자형(Numerical)' 데이터, '데이트-타임(Date-Time)형' 데이터는 제2 속성값들을 예시한 것이다.

[0025] 시각화 맵핑부(30)는 시각화 코드에 대응하여 맵핑 룰을 갖는 하나 이상의 시각화 도구를 제공한다.

[0026] 시각화 맵핑부(30)는 하기 <표 2>와 같이 시각화 도구들에 대응하는 맵핑 룰을 시각화 코드에 대응하여 제공할 수 있다.

표 2

[0027]

시각화 도구	C	N	CC	CN	NN	TN	CCN	TCN	CNN
영역 (AREA)	X:cat. Y:누적 cnt								
싱글 바 (Single Bar)	X:cat. Y:cnt	X:Rangr Y:cnt		X:C Y:N		X:T Y:N			
멀티 바 (Multi-Bar)			X:C1-C Y:cnt				X:C1-C2 Y:N	X:T Y:N Bar: C	
스택 바 (Stacked Bar)			X:C1 Y:cnt Seg : C2				X:C1 Y:N Seg : C2		
...

[0028] 상기 <표 2>에서, 시각화 도구는 '영역', '싱글 바', '멀티 바', '스택 바' 등으로 예시될 수 있으며, 그외 싱글 라인, 멀티 라인, 파이, 싱글 도넛, 멀티 도넛, 스캐터 플롯(Scatter Plot), 버블(Bubble), 싱글 박스 플롯(Single Box Plot), 멀티 박스 플롯(Multi-box Plot) 등 다양한 도구들이 적용될 수 있다.

[0029] 그리고, 시각화 코드는 'C', 'N', 'CC', 'CN', 'NN', 'TN', 'CCN', 'TCN', 'CNN' 등 다양하게 구분될 수 있다. 여기에서, 'C'는 '범주형(Categorical)' 데이터를 지칭하는 필드 코드이며, 'N'은 '숫자형(Numerical)' 데이터를 지칭하는 필드 코드이고, 'T'는 '데이트-타임(Date-Time)형' 데이터를 지칭하는 필드 코드이다. 즉, 시각화 코드는 하나 이상의 필드 코드가 조합되어 형성될 수 있다. 시각화 코드에서, 'C'는 '범주형' 데이터를 싱글로 포함하는 경우를 의미하고, 'N'은 '숫자형' 데이터를 싱글로 포함하는 경우를 의미하며, 'CC'는 '범주형' 데이터를 두 개 포함하는 것을 의미하고, 'CN'은 '범주형' 데이터와 '숫자형' 데이터를 포함하는 경우를 의미하며, 'NN'은 '숫자형' 데이터를 두 개 포함하는 것을 의미하고, 'TN'은 '데이트-타임형' 데이터와 '숫자형' 데이터를 포함하는 것을 의미하며, 'CCN'은 '범주형' 데이터 두 개와 '숫자형' 데이터 하나를 포함하는 것을 의미하고, 'TCN'은 '데이트-타임형' 데이터와 '숫자형' 데이터 및 '범주형' 데이터를 포함하는 것을 의미하며, 'CNN'은 '범주형' 데이터와 두 개의 '숫자형' 데이터를 포함하는 것을 의미한다. 상술한 시각화 코드는 본 발명의 실시예를 위하여 예시된 것이며 보다 다양하게 예시될 수 있다.

[0030] 하나의 시각화 코드에 하나 또는 복수 개의 시각화 도구가 대응될 수 있으며, 시각화 도구는 각 시각화 코드에 따르는 맵핑 룰이 규정될 수 있다.

[0031] 구체적인 일례로, 'CCN'의 경우, '멀티 바'와 '스택 바'가 시각화 도구로 제공될 수 있으며, '멀티 바'는 시각화 코드 'CCN'에 대응하여 "X:C1-C2, Y:N"이라는 맵핑 룰을 가지며, '스택 바'는 "X:C1, Y: N, Seg : C2" 라는 맵핑 룰을 가질 수 있다. "X:C1-C2, Y:N"은 X축에 두 개의 '범주형' 데이터(C1, C2)를 표현하고 Y 축에 하나의 '숫자형' 데이터를 표현하는 것에 대응하는 맵핑 룰을 표현한 것으로 이해될 수 있으며, 이 경우는 후술하는 도 5와 같이 표현될 수 있다. "X:C1, Y:N, Seg :C2"는 X축에 하나의 '범주형' 데이터(C1)를 표현하고 Y축에 하나의 '숫자형' 데이터(N)를 표현하며, Y축의 '숫자형 데이터'를 하나의 '범주형' 데이터(C2)로 구분하여 표현하는 것으로 이해될 수 있으며, 이 경우는 후술하는 도 6과 같이 표현될 수 있다.

[0032] 한편, 도구 선택부(40)는 필드 타입 검출부(20)에서 결정된 필드 별 필드 타입을 참조하여 시각화 코드를 정의

하고 시각화 코드에 해당하는 하나 이상의 시각화 도구를 시각화 맵핑부(30)에서 구하며, 구해진 하나 이상의 시각화 도구에 대응하는 맵핑 룰에 의한 시각화 도구를 제공한다.

- [0033] 상술한 바와 같이 구성된 본 발명의 실시예는 도 2와 같은 방법에 의하여 차트를 추천할 수 있다.
- [0034] 한편, 도 3은 어플리케이션 환경에서 구현되는 차트 추천 장치를 예시한 블록도이며, 컴퓨터나 서버에 탑재되어 실행되는 프로그램 또는 어플리케이션(100)에 대응하여 차트 제공 장치(200)가 연동하도록 구성될 수 있다. 이 경우, 사용자가 분석을 위하여 프로그램 또는 어플리케이션(100)으로 예시될 수 있는 스프레드 시트의 전체 또는 일부를 블록으로 선택한 후 프로그램 또는 어플리케이션(100)에서 차트 제공 기능을 위하여 제공되는 아이콘(또는 버튼)을 클릭하거나 명령어를 입력하면, 그에 대응하여 차트 제공 장치(200)는 도 2를 참조하여 설명되는 동작을 수행할 수 있다.
- [0035] 그리고, 도 4는 클라이언트-서버 환경에서 구현되는 차트 제공 장치를 예시한 블록도이고, 하나의 서버(170)에 다수의 클라이언트(150)가 연결된 것을 예시하고, 서버(170)에 차트 제공 장치(200)가 연동하도록 구성될 수 있다. 여기에서, 서버(170)와 다수의 클라이언트(150)는 인터넷 프로토콜 기반의 네트워크를 통하여 통신 가능하도록 접속될 수 있으며, 차트 제공 장치(200)는 서버(170)에 내장되거나 또는 별도의 서버(도시되지 않음)에 탑재되어 서버(170)와 연동되도록 구성될 수 있다. 이 경우, 사용자가 클라이언트(150)를 통하여 분석을 위한 데이터 파일을 선택하거나 현재 활성화된 스프레드 시트의 전체 또는 일부를 블록으로 선택한 후 차트 제공 기능을 위하여 제공되는 아이콘(또는 버튼)을 클릭하거나 명령어를 입력하면, 차트 제공 요청이 서버(170)를 차트 제공 장치(200)에 전달되고, 상기한 차트 제공 요청에 대응하여 차트 제공 장치(200)는 도 2를 참조하여 설명되는 동작을 수행할 수 있다.
- [0036] 상기와 같은 사용자의 요청에 대응하여, 차트 제공 장치(200)로 구성될 수 있는 도 1의 실시예는 도 2와 같이 동작될 수 있다.
- [0037] 사용자가 분석을 위하여 블록을 선택한 후 차트 제공을 요청하면, 데이터 타입 검출부(10)에 의하여 선택된 블록의 데이터 단위들에 적용되는 필드 별로 데이터 타입을 검출하는 단계(S10)가 수행된다.
- [0038] 데이터 타입 검출부(10)는 원시 데이터의 유형 즉 데이터 타입을 필드 별로 구분한다.
- [0039] <표 1>에서 필드로 설정된 '요일', '도로 명', '카운트'는 각각 문자열, 문자열, 정수이며, 데이터 타입 검출부(10)는 데이터 타입에 대한 제1 속성값들을 각 필드에 대응하여 문자열, 문자열 및 정수로 검출할 수 있다.
- [0040] 그 후, 필드 타입 검출부(20)에 의하여 필드들에 대한 필드 타입에 대한 검출이 수행하며, 필드 타입 검출부(20)는 선택된 블록의 데이터 단위들에 적용되는 필드를 하나씩 선택하여 필드 타입을 검출할 수 있다(S12).
- [0041] 필드 타입 검출부(20)는 제2 속성값에 일대일 매칭되는 제1 속성값을 갖는 필드에 대하여 제1 속성값에 매칭되는 제2 속성값으로 필드 타입을 검출한다.
- [0042] 그리고, 필드 타입 검출부(20)는 제2 속성값에 일대일 매칭되지 않는 제1 속성값을 갖는 필드에 대하여 필드값들을 분석하고 제2 속성값을 판단함으로써 필드 타입을 검출한다.
- [0043] 보다 구체적으로 예시하면, 필드 타입 검출부(20)는 제1 속성값이 '데이트-타임'인 경우 제2 속성값을 '데이트-타임형'으로 검출하고, 제1 속성값이 문자열인 경우 제2 속성값을 '범주형'으로 검출한다. 즉, 제1 속성값의 '데이트-타임'과 제2 속성값의 '데이터 타임형'이 일대일 매칭되도록 설정되고, 제1 속성값의 문자열과 제2 속성값의 '범주형'이 일대일 매칭되도록 설정된다.
- [0044] 필드 타입 검출부(20)는 제1 속성값이 정수형 또는 실수형인 경우 이들에 일대일 매칭되는 제2 속성값이 없기 때문에 필드값들을 분석하여 '범주형' 또는 '숫자형'으로 제2 속성값을 판단할 수 있다.
- [0045] 이를 위하여, 필드 타입 검출부(20)는 필드 타입을 검출할 필드(S12)를 선택한 후 순차적으로 현재 선택된 필드의 데이터 타입 즉 제1 속성값이 '타이트-타임'인가 판단하는 단계(S20), 현재 선택된 필드의 데이터 타입 즉 제1 속성값이 '문자열'인가 판단하는 단계(S22) 및 현재 선택된 필드의 전체 필드값에서 '구분값' 비율이 50% 이상인가 판단하는 단계(S24)를 순차적으로 수행한다.
- [0046] 필드 타입 검출부(20)는 현재 선택된 필드의 데이터 타입 즉 제1 속성값이 '데이트-타임'으로 판단되면(S20) 현재 선택된 필드의 필드 타입 즉 제2 속성값을 '데이트-타임형'으로 결정한다(S26). 그리고, 필드 타입 검출부(20)는 현재 선택된 필드의 데이터 타입 즉 제1 속성값이 '문자열'로 판단되면(S22) 현재 선택된 필드의 필드 타입 즉 제2 속성값을 '범주형'으로 결정한다(S28).

- [0047] 단계 S20와 단계 S22를 수행한 후 필드 타입 검출부(20)는 현재 선택된 필드가 제2 속성값에 일대일 매칭되지 않는 제1 속성값을 갖는 것으로 판단하고 현재 선택된 필드의 전체 필드값에서 '구분값' 비율이 50% 이상인가 판단한다(S24). 필드 타입 검출부(20)는 '구분값' 비율이 50% 미만인 경우 현재 선택된 필드의 필드 타입 즉 제2 속성값을 '범주형'으로 결정하고(S30), '구분값' 비율이 50% 이상인 경우 현재 선택된 필드의 필드 타입 즉 제2 속성값을 '숫자형'으로 결정한다(S32).
- [0048] 여기에서, '구분값'은 선택된 필드의 제1 속성값이 실수형이나 정수형인 경우 제2 속성값을 판단하기 위하여 구하는 값이다. '구분값'은 선택된 블록의 데이터 단위들에 적용되는 필드의 값들이 분류 가능한 개수로 구분되는지 아니면 일정 수준 이상의 개수로 분산되어서 분류가 의미없는지 판단하기 위하여 현재 선택된 필드의 필드값들에 대한 카테고리 수를 카운트 한 값이다. 본 발명의 실시예는 구체적인 예로 50% 이상인 경우를 예시한다. 보다 구체적으로, 선택된 블록의 데이터 단위들이 1000개인 경우, 현재 선택된 필드의 필드값들이 500개 미만으로 분류되면 '범주형'으로 제2 속성값이 판단될 수 있고, 500개 이상으로 분류되면 '숫자형'으로 제2 속성값이 판단될 수 있다.
- [0049] 한편, 상술한 과정에 의하여 현재 선택된 필드의 필드 타입이 판단될 수 있으며, 필드 타입 검출부(20)는 현재 선택된 필드의 필드 타입 즉 제2 속성값을 단계 S26~ 단계 S32에서 판단된 것으로 결정한다.
- [0050] 그 후, 필드 타입 검출부(20)는 나머지 필드에 대해서 단계 S12 내지 단계 S40를 반복적으로 수행하여 필드 타입을 결정한다.
- [0051] 한편, 도구 선택부(40)는 필드 타입 검출부(20)에서 결정된 필드 별 필드 타입을 참조하여 시각화 코드를 정의한다.
- [0052] <표 1>의 경우, 필드 타입 검출부(20)는 '요일'과 '도로 명'의 필드 타입을 '범주형'으로 결정하고 '카운트'의 필드 타입을 '숫자형'으로 결정할 수 있다.
- [0053] 상기한 필드 타입 검출부(20)의 필드 타입 결정에 대응하여, 도구 선택부(40)는 두 개의 범주형의 필드 코드 'C'와 하나의 숫자형의 필드 코드 'N'을 조합하여 시각화 코드를 'CCN'으로 정의할 수 있다.
- [0054] 도구 선택부(40)는 시각화 코드에 해당하는 하나 이상의 시각화 도구를 시각화 맵핑부(30)에서 구하며, <표 1>의 경우에 <표 2>를 참조하여 시각화 도구를 구하면 '멀티 바'와 '스택 바'가 시각화 도구로 구해질 수 있다.
- [0055] 도구 선택부(40)는 '멀티 바'와 '스택 바'의 맵핑 룰에 따라서 도 5 및 도 6과 같은 시각화 도구를 제공할 수 있다.
- [0056] 도 5를 참조하면, '멀티 바'는 시각화 코드 'CCN'에 대응하여 "X:C1-C2, Y:N"이라는 맵핑 룰을 가지며, X축에 두 개의 '범주형' 데이터인 '도로 명'과 '요일'을 표현하고 Y 축에 '숫자형' 데이터인 '카운트'를 표현하는 것으로 구현될 수 있다.
- [0057] 도 6을 참조하면, '스택 바'는 시각화 코드 'CCN'에 대응하여 "X:C1, Y: N, Seg : C2" 라는 맵핑 룰을 가지며, X축에 범주형 데이터인 '요일'을 표현하고 Y축에 '숫자형' 데이터인 '카운트'를 표현하며 '범주형' 데이터인 '도로 명'은 '숫자형' 데이터인 '카운트'에 구분되도록 표시되는 것으로 구현될 수 있다.
- [0058] 한편, 본 발명의 실시예는 도 4와 같이 클라이언트-서버 환경에서 데이터 파일을 전송하여 분석하는 경우, 도 7과 같이 데이터 파일을 미리 분석한 후 분석할 필드를 선택하도록 인터페이스 윈도우가 서버 또는 차트 제공 장치에 의하여 제공될 수 있다.
- [0059] 이 경우, 사용자에게 의하여 분석할 필드가 선택될 수 있으며, 선택된 필드에 대한 썬머리(Summary)와 뷰(View)가 제공될 수 있다. 여기에서, 썬머리(Summary)는 선택된 필드들에 대한 분석 결과와 분석 결과에 대응하는 정보를 표시하는 추가적인 윈도우를 제공하도록 구현될 수 있다. 그리고, 뷰(View)는 선택된 필드에 대하여 제공되는 시각화 도구를 미리보기하기 위한 윈도우를 제공하도록 구현될 수 있다.
- [0060] 따라서, 사용자가 분석을 위하여 정보를 선택하면, 상술한 본 발명의 실시예는 선택한 정보의 특성에 맞는 차트들이 추천될 수 있고, 사용자는 추천된 차트들 중 원하는 목적에 맞는 차트를 선택할 수 있다.
- [0061] 또한, 본 발명의 실시예는 프로그램이나 어플리케이션 또는 클라이언트-서버 환경에서 차트를 추천하고 선택할 수 있는 기능이 제공될 수 있어서 사용자는 다양한 환경에서 선택한 정보에 맞는 차트를 손쉽게 추천받고 이용할 수 있다.

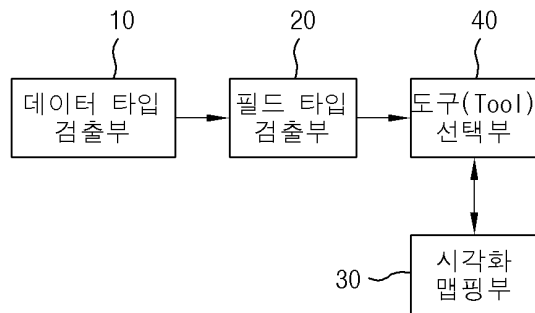
부호의 설명

[0062]

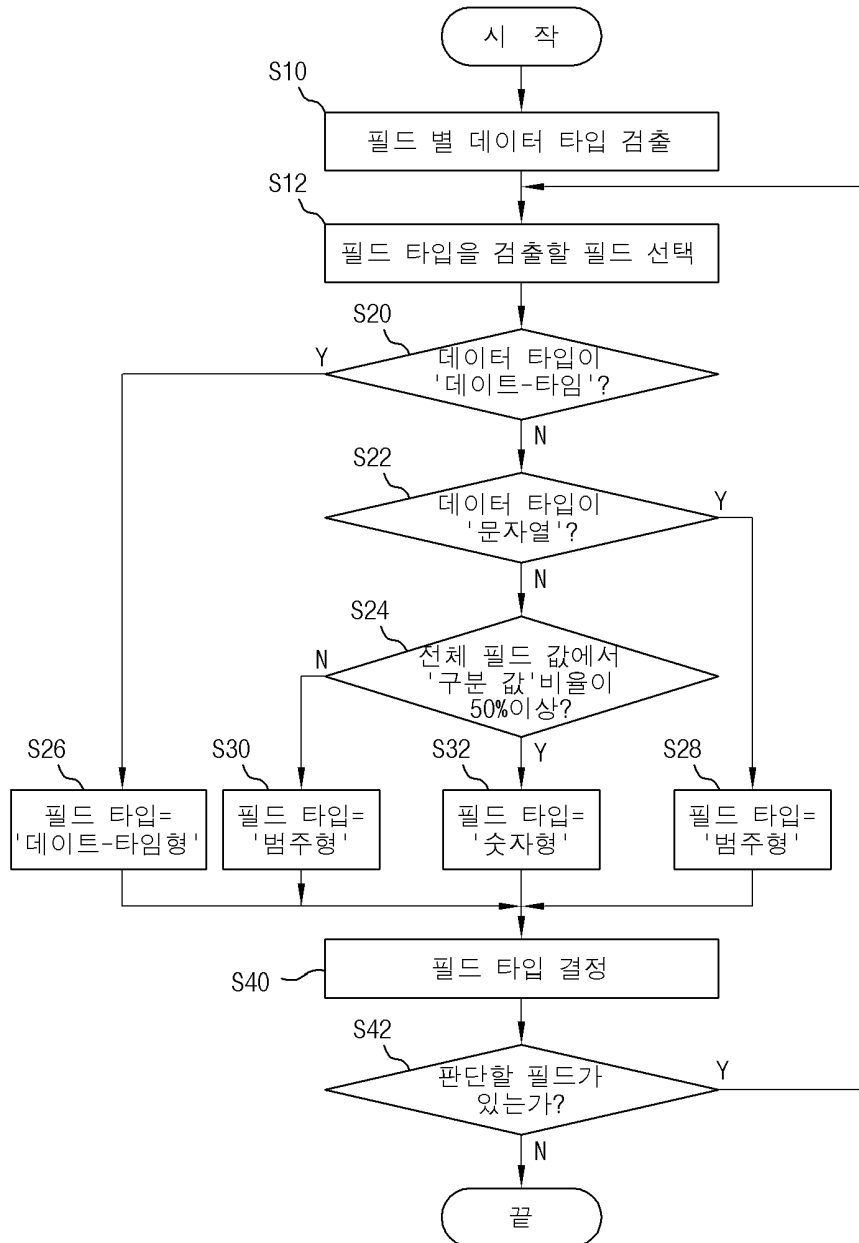
10 : 데이터 타입 검출부 20 : 필드 타입 검출부
30 : 시각화 맵핑부 40 : 도구 선택부

도면

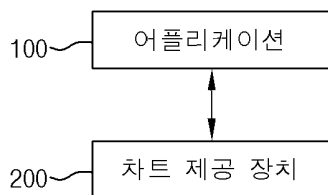
도면1



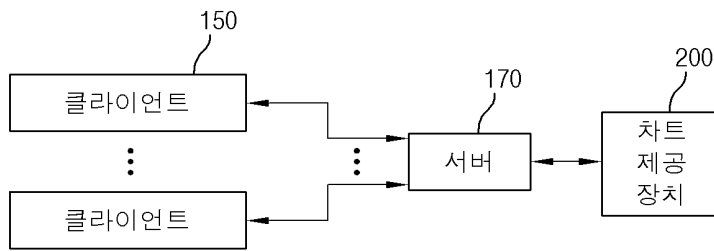
도면2



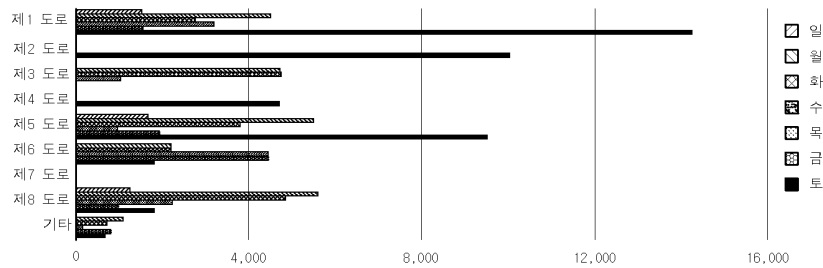
도면3



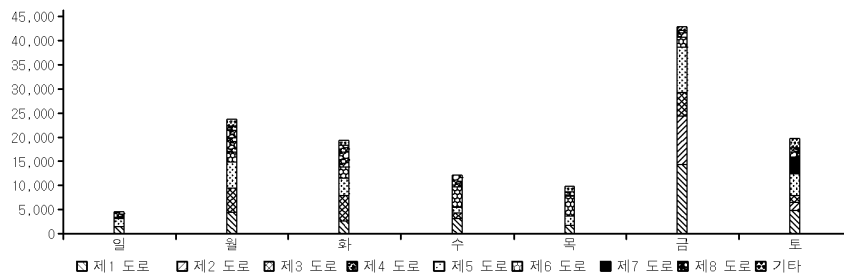
도면4



도면5



도면6



도면7

Category

Categorical

Numerical

Categorical

Categorical

Categorical

Type

String

Integer

Integer

String

String

Name

ItemID

Counts

DivCode

DivStr

isSell

A	1	0001	Clothes	Yes
B	100	0002	Shoes	No
C	10000			
.....				

Summary

View

도면8

	A	A	B	C
1	No	Movie Title		
2	1			
3	2			
4	3			
5	4			
6	5			
7	6			
8	7			
9	8			
10	9			
11	10			
12	11			
13	12			
14	13			
15	14			
16	15			
17	16			
18	17			
19	18			
20	19			
21	20			
22	21			
23	22			
24	23			
25	24			
26	25			
27	26			
28	27			
29	28			
30	29			