



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2024년02월07일
(11) 등록번호 10-2634331
(24) 등록일자 2024년02월01일

- (51) 국제특허분류(Int. Cl.)
G06F 7/76 (2006.01) G06F 17/13 (2006.01)
G06F 7/24 (2006.01) H04W 4/029 (2018.01)
H04W 4/16 (2009.01) H04W 8/22 (2009.01)
- (52) CPC특허분류
G06F 7/76 (2021.08)
G06F 17/13 (2013.01)
- (21) 출원번호 10-2021-0179684
- (22) 출원일자 2021년12월15일
심사청구일자 2021년12월15일
- (65) 공개번호 10-2023-0091226
- (43) 공개일자 2023년06월23일
- (56) 선행기술조사문헌
JP2016048529 A
JP2017143472 A
JP2015056018 A
JP2015049681 A

(73) 특허권자
재단법인 서울연구원
서울 서초구 서초동 391
에스케이텔레콤 주식회사
서울특별시 중구 을지로 65 (을지로2가)
에스케이 주식회사
서울특별시 종로구 종로 26 (서린동)

(72) 발명자
최수범

최준영

(뒷면에 계속)

(74) 대리인
손태원

전체 청구항 수 : 총 8 항

심사관 : 지정훈

(54) 발명의 명칭 통신데이터 위치정보에 미분기법을 적용한 효과적인 체류정보 추출 방법 및 장치

(57) 요약

본 발명은 통신데이터 위치정보에 미분기법을 적용한 효과적인 체류정보 추출 방법 및 장치에 관한 것으로서, (A) 콜로그가 발생된 원본 통신데이터를 불러오는 단계; (B) 원본 통신데이터로부터 단말기ID 정보와 시간순으로 정렬하는 단계; (C) 정렬된 위치정보와 단말기ID 정보를 각각 미분하는 단계; (D) 미분된 위치정보와 단말기ID

(뒷면에 계속)

대표도 - 도1



정보를 각각 이진 변환함으로써 이진 미분 열을 생성하는 단계; (E) 생성된 이진 미분 열로부터 시간 순서대로 단말기 측 같은 위치가 연속해서 나오는 행을 탐색하는 단계; (F) 같은 위치에서 연속으로 발생하는 콜로그의 첫 행(체류 시작)과 마지막 행(체류 종료)을 남기고 제거하는 단계; (G) 콜로그의 마지막 행을 열로 변환 생성하여 데이터를 축약하는 단계; (H) 축약된 데이터로부터 체류시간을 추출하고, 체류시간을 기준으로 이동 중인지 체류 중인지를 구분하는 단계; 를 포함하는 것을 특징으로 한다. 본 발명에 따르면, 단말기(암호화된 ID) 정보와 위치(위도, 경도) 정보 및 통신시간 정보 등의 콜로그 데이터를 단말기와 시간순으로 정렬한 후 미분기법을 적용함으로써 같은 위치에서 중복 발생하는 기록과 위치가 변하는 기록을 구분하여 체류 위치별 시작과 종료 시각으로 데이터를 축약할 수 있으며, 이와 같은 축약된 정보를 토대로 하나의 단말기로부터 위치별 체류정보를 매우 효과적으로 추출 및 빠른 속도로 처리할 수 있다.

(52) CPC특허분류

- G06F 7/24 (2013.01)
- H04W 4/029 (2020.05)
- H04W 4/16 (2013.01)
- H04W 8/22 (2013.01)

김민원



(72) 발명자

강현구



명세서

청구범위

청구항 1

- (A) 콜로그(단말기ID 정보, 위치정보 및 통신시간 정보 포함)가 발생된 원본 통신데이터를 불러오는 단계;
- (B) 원본 통신데이터로부터 단말기ID 정보와 시간순으로 정렬하는 단계;
- (C) 정렬된 위치정보와 단말기ID 정보를 각각 미분하는 단계;
- (D) 미분된 위치정보와 단말기ID 정보를 각각 이진 변환함으로써 이진 미분 열을 생성하는 단계;
- (E) 생성된 이진 미분 열로부터 시간 순서대로 단말기 측 같은 위치가 연속해서 나오는 행을 탐색하는 단계;
- (F) 같은 위치에서 연속으로 발생하는 콜로그의 첫 행(체류 시작)과 마지막 행(체류 종료)을 남기고 제거하는 단계;
- (G) 콜로그의 마지막 행을 열로 변환 생성하여 데이터를 축약하는 단계;
- (H) 축약된 데이터로부터 체류시간을 추출하고, 체류시간을 기준으로 이동 중인지 체류 중인지를 구분하는 단계; 를 포함하는 것을 특징으로 하는 통신데이터 위치정보에 미분기법을 적용한 효과적인 체류정보 추출방법.

청구항 2

- (1) 일정 지역 및 일정 기간 동안에 발생한 콜로그 데이터를 불러오는 단계;
- (2) 불러온 콜로그 데이터로부터 단말기ID를 1차 조건으로 하고 통신시간을 2차 조건으로 하여 정렬하는 단계;
- (3) 정렬된 콜로그 데이터로부터 단말기ID, 위치정보 1(위도 또는 X), 위치정보 2(경도 또는 Y) 각각에 대해 열을 미분함으로써 시간에 따른 변화를 계산하여 저장하는 단계;
- (4) 미분된 단말기ID, 위치정보 1(위도 또는 X), 위치정보 2(경도 또는 Y) 각각에 대해 '0' 또는 '1'의 이진 변환을 수행하는 단계;
- (5) 이진 미분 단말기ID, 이진 미분 위치정보 1, 이진 미분 위치정보 2 중에서 어느 하나라도 '1'이면 '1', 모두 '0'이면 '0'으로 하여 행을 추가함으로써 단말기 또는 위치의 변화를 표시하는 단계;
- (6) 상기 (5)단계에서 계산된 이진 미분 위치정보 열(j)을 1행씩 위로 시프트시키고 빈칸은 '1'로 채워 다음 열(k)을 생성하는 단계;
- (7) 상기 (5)단계에서 계산된 이진 미분 위치정보 열(j) 측 행과 다음 열(k) 측 행을 더한 값으로 차기 열(l)을 생성하는 단계;
- (8) 상기 (7)단계에서 생성된 차기 열(l)에서 '1'로 구분된 행들만 추출하되, 같은 단말기 및 같은 위치의 행 중에서 첫 행은 체류시작시간으로 하고, 마지막 행은 체류종료시간으로 하여 같은 단말기 및 같은 위치로 짝(체류시작과 체류종료)을 이루는 콜로그 정보를 한 개의 행으로 축약함으로써 (m)열을 생성하는 단계;
- (9) 상기 (8)단계에서 생성된 (m)열의 체류종료시간과 체류시작시간인 (b)열의 차를 계산하여 체류시간 (n)열을 생성하는 단계; 를 포함하는 것을 특징으로 하는 통신데이터 위치정보에 미분기법을 적용한 효과적인 체류정보 추출방법.

청구항 3

제 2항에 있어서,

상기 (5)단계에서 계산된 이진 미분 위치정보 열(j) 측 행과 다음 열(k) 측 행을 더한 값으로 생성된 차기 열

(1)에 있어,

같은 단말기 및 같은 위치에서 3건 이상 콜로그 발생하고 이전 행과 다음 행에서 위치변화가 없는 경우에는 '0'으로 기록하고,

단말기가 변경되거나 같은 단말기 및 같은 위치에서 2건 이상 콜로그 발생하고 이전 행 또는 다음 행에서 위치가 변화한 경우의 첫 행(체류 시작)과 마지막 행(체류 종료)은 '1'로 기록하며,

같은 단말기 및 같은 위치에서 발생한 콜로그가 1건인 경우에는 '2'로 기록하여 구분하는 것을 특징으로 하는 통신데이터 위치정보에 미분기법을 적용한 효과적인 체류정보 추출방법.

청구항 4

제 2항에 있어서,

상기 축약 생성된 (m)열에서는,

상기 (7)단계에서 생성된 차기 열(1)에서 '0'으로 구분된 행은 제거하고, '2'로 구분된 행은 체류종료시간을 '공란'으로 표시하며, '1'로 구분된 행 중에서 체류종료시간에 해당하는 행은 제거하는 것을 특징으로 하는 통신데이터 위치정보에 미분기법을 적용한 효과적인 체류정보 추출방법.

청구항 5

제 1항에 있어서,

상기 (A)단계와 (B)단계 사이에

원본 통신데이터로부터 시간 변수에서 초 단위를 제거하는 단계; 가 추가되는 것을 특징으로 하는 통신데이터 위치정보에 미분기법을 적용한 효과적인 체류정보 추출방법.

청구항 6

제 2항에 있어서,

상기 (1)단계와 (2)단계 사이에 불러온 콜로그 데이터로부터 시간 변수에서 초 단위를 제거하는 단계; 가 추가되는 것을 특징으로 하는 통신데이터 위치정보에 미분기법을 적용한 효과적인 체류정보 추출방법.

청구항 7

콜로그가 발생된 원본 통신데이터를 불러오되, 원본 통신데이터로부터 정렬된 단말기ID 정보와 위치정보를 미분하여 같은 위치가 연속되는지 아닌지를 판단하는 데이터 판단부;

같은 위치가 연속해서 나오는 행이 발생되면 같은 위치가 연속으로 발생하는 콜로그의 첫 행(체류 시작)과 마지막 행(체류 종료)을 남기고 제거하는 데이터 제거부;

콜로그의 마지막 행을 열로 변환 생성하여 데이터를 축약하는 데이터 축약부;

축약된 데이터로부터 체류시간을 추출하고, 체류시간을 기준으로 이동 중인지 체류 중인지를 구분하는 체류정보 추출부; 를 포함하는 것을 특징으로 하는 통신데이터 위치정보에 미분기법을 적용한 효과적인 체류정보 추출장치.

청구항 8

제 7항에 있어서,

상기 데이터 판단부는,

콜로그가 발생된 원본 통신데이터를 불러와서 원본 통신데이터로부터 시간 변수에서 초 단위를 제거하는 시간변수 처리부;

원본 통신데이터로부터 단말기ID 정보와 시간순으로 정렬하는 데이터 정렬부;

원본 통신데이터로부터 정렬된 위치정보와 단말기 ID 정보를 각각 미분하는 데이터 미분처리부;

미분된 데이터를 이진화 처리하여 이진 코드로 변환하는 이진 변환부;

이진 변환 후, 단말기 측 같은 위치가 연속해서 나오는 행이 있는지를 탐색하는 데이터 탐색부; 로 이루어지는 것을 특징으로 하는 통신데이터 위치정보에 미분기법을 적용한 효과적인 체류정보 추출장치.

발명의 설명

기술 분야

[0001] 본 발명은 통신데이터 위치정보에 미분기법을 적용한 효과적인 체류정보 추출 방법 및 장치에 관한 것으로서, 더욱 상세하게는 단말기(암호화된 ID) 정보와 위치(위도, 경도) 정보 및 통신시간 정보 등의 콜로그(Call log) 데이터를 단말기와 시간순으로 정렬한 후 미분기법을 적용함으로써 같은 위치에서 중복 발생하는 기록과 위치가 변하는 기록을 구분하여 체류 위치별 시작과 종료 시각으로 데이터를 축약할 수 있도록 하며, 이를 통해 축약된 정보를 토대로 위치별 체류정보를 매우 효과적으로 추출할 수 있도록 한 통신데이터 위치정보에 미분기법을 적용한 효과적인 체류정보 추출 방법 및 장치에 관한 것이다.

배경 기술

[0002] 이동통신 단말기와 통신기지국 간에는 상호간에 신호정보를 주고받는 것으로서, 단말기의 미 사용시 또는 단말기의 사용시 기록되는 신호정보를 갖게 되며, 통신데이터가 발생된다.

[0003] 이러한 통신데이터는 통신기지국에서 생성되는 콜로그(Call log)를 기반으로 해당 사용자에게 대한 단말기별 위치를 추정할 수 있는 정보가 담겨져 기록되는데, 상기 콜로그(Call log)는 단말기(암호화된 ID) 정보와 위치(위도, 경도) 정보 및 접속(통신)시간 정보 등을 포함한다.

[0004] 하지만, 이와 같은 콜로그를 기반으로 하는 통신데이터는 하루에도 단말기별 수백건 이상의 데이터가 쌓여 이를 저장하는데도 많은 자원이 필요하고, 이를 분석하기에도 어려움이 많은 실정에 있으며, 방대한 데이터량에 의해 처리속도 또한 빠르지 않다.

[0005] 이에 따라, 통신데이터에 대해 데이터를 축약하기 위한 다양한 알고리즘들이 개발되고 있으나, 아직까지 효율성 및 처리속도 등에 있어 미흡한 부분이 많은 등 여전히 개선이 요구되는 실정에 있다.

[0006] 한편, 종래 사용되고 있는 통신데이터 축약기술을 살펴보면, 축적된 통신데이터에서 단말기별 데이터들을 시간대별로 축약하기 위해서 단말기 정보들을 그룹화하고 일정 시간 간격으로 콜로그 데이터가 발생하였으면 위치를 추가하는 방식으로 정리하고 있으며, 이로부터 체류정보 등을 추출하고 있다.

[0007] 부연하여, 종래에는 일정 지역, 일정 기간 동안에 발생한 콜로그 데이터를 불러온 후 동일 단말기별 콜로그들을 추출하고, 같은 위치에서 연속하여 발생한 콜로그들로부터 체류시간을 추출하며, 이러한 과정을 전체 데이터에 존재하는 단말기 개수만큼 콜로그 추출을 반복하고 있는 수준으로서, 해당 단말기의 사용자에게 대한 체류정보를 추출하는데 따른 처리 속도 및 시간이 오래 소요되는 문제점 및 효율성이 떨어지는 문제점이 있다.

[0008] 한편, 종래 선행기술문헌을 살펴보았을 때, 국내 등록특허공보 제10-2297845호에 있어, "모바일 단말의 위치정보를 이용하여 사용자의 이동을 파악하는 방법에 있어서, 모바일 단말을 이용하여 수집한 위치정보로부터 사용자의 이동경로를 추출하는 단계; 및 상기 이동경로를 포함하는 대상지역에 복수의 그리드(grid)를 적용한 맵을 기준으로, 상기 이동경로를 교통수단의 주행경로와 비교함으로써 상기 사용자가 이용한 교통수단을 파악하는 단계;를 포함하며, 상기 복수의 그리드 각각은 복수의 단위구역(unit zone)들로 구성되며, 서로 다른 그리드에 포함되는 단위구역의 크기는 상이한 것을 특징으로 하는 모바일 단말의 위치정보를 이용하여 사용자의 이동을 파악하기 위한 방법, 이를 수행하기 위한 장치 및 시스템"을 개시하고 있다.

[0009] 하지만, 상술한 국내 등록특허공보 제10-2297845호에서는 GPS 데이터, Wi-Fi 데이터, 셀룰러 데이터 및 GPS를

통해 위치정보를 수집하고 사용자가 이용한 교통수단을 파악하는데 특화되어 있는 기술로서, 이하에서 제안하는 본 발명의 기술과는 처리기법 및 처리과정에서 차이를 갖는다 할 수 있다.

선행기술문헌

특허문헌

[0010] (특허문헌 0001) 대한민국 등록특허공보 제10-2297845호

발명의 내용

해결하려는 과제

[0011] 본 발명은 상기의 종래 문제점들을 개선 및 이를 감안하여 안출된 것으로서, 단말기(암호화된 ID) 정보와 위치(위도, 경도) 정보 및 통신시간 정보 등의 콜로그(Call log) 데이터를 단말기와 시간순으로 정렬한 후 미분기법을 적용함으로써 같은 위치에서 중복 발생하는 기록과 위치가 변하는 기록을 구분하여 체류 위치별 시작과 종료 시각으로 데이터를 축약할 수 있도록 하며, 이를 통해 축약된 정보를 토대로 위치별 체류정보를 매우 효과적으로 추출할 수 있도록 한 통신데이터 위치정보에 미분기법을 적용한 효과적인 체류정보 추출 방법 및 장치를 제공하는데 그 목적이 있다.

[0012] 본 발명은 기존에 비해 통신데이터를 보다 빠르고 효율적으로 처리 가능함은 물론 실시간 처리가 가능하게 하고, 방대한 통신데이터의 저장을 위한 자원을 줄일 수 있도록 한 통신데이터 위치정보에 미분기법을 적용한 효과적인 체류정보 추출 방법 및 장치를 제공하는데 그 목적이 있다.

[0013] 본 발명은 감염병 확진자 등의 발생시 해당자에 대한 이동경로 등 동선을 파악함과 더불어 접촉자를 선별하는데 활용할 수 있도록 한 통신데이터 위치정보에 미분기법을 적용한 효과적인 체류정보 추출 방법 및 장치를 제공하는데 그 목적이 있다.

과제의 해결 수단

[0014] 상기의 목적을 달성하기 위한 통신데이터 위치정보에 미분기법을 적용한 효과적인 체류정보 추출방법은, (A) 콜로그(단말기ID 정보, 위치정보 및 통신시간 정보 포함)가 발생된 원본 통신데이터를 불러오는 단계; (B) 원본 통신데이터로부터 단말기ID 정보와 시간순으로 정렬하는 단계; (C) 정렬된 위치정보와 단말기ID 정보를 각각 미분하는 단계; (D) 미분된 위치정보와 단말기ID 정보를 각각 이진 변환함으로써 이진 미분 열을 생성하는 단계; (E) 생성된 이진 미분 열로부터 시간 순서대로 단말기 측 같은 위치가 연속해서 나오는 행을 탐색하는 단계; (F) 같은 위치에서 연속으로 발생하는 콜로그의 첫 행(체류 시작)과 마지막 행(체류 종료)을 남기고 제거하는 단계; (G) 콜로그의 마지막 행을 열로 변환 생성하여 데이터를 축약하는 단계; (H) 축약된 데이터로부터 체류시간을 추출하고, 체류시간을 기준으로 이동 중인지 체류 중인지를 구분하는 단계;를 포함하는 것을 특징으로 한다.

[0015] 또한, 상기의 목적을 달성하기 위한 통신데이터 위치정보에 미분기법을 적용한 효과적인 체류정보 추출방법은, (1) 일정 지역 및 일정 기간 동안에 발생한 콜로그 데이터를 불러오는 단계; (2) 불러온 콜로그 데이터로부터 단말기ID를 1차 조건으로 하고 통신시간을 2차 조건으로 하여 정렬하는 단계; (3) 정렬된 콜로그 데이터로부터 단말기ID, 위치정보 1(위도 또는 X), 위치정보 2(경도 또는 Y) 각각에 대해 열을 미분함으로써 시간에 따른 변화를 계산하여 저장하는 단계; (4) 미분된 단말기ID, 위치정보 1(위도 또는 X), 위치정보 2(경도 또는 Y) 각각에 대해 '0' 또는 '1'의 이진 변환을 수행하는 단계; (5) 이진 미분 단말기ID, 이진 미분 위치정보 1, 이진 미분 위치정보 2 중에서 어느 하나라도 '1'이면 '1', 모두 '0'이면 '0'으로 하여 행을 추가함으로써 단말기 또는 위치의 변화를 표시하는 단계; (6) 상기 (5)단계에서 계산된 이진 미분 위치정보 열(j)을 1행씩 위로 시프트시키고 빈칸은 '1'로 채워 다음 열(k)을 생성하는 단계; (7) 상기 (5)단계에서 계산된 이진 미분 위치정보 열(j) 측 행과 다음 열(k) 측 행을 더한 값으로 차기 열(l)을 생성하는 단계; (8) 상기 (7)단계에서 생성된 차기 열(l)에서 '1'로 구분된 행들만 추출하되, 같은 단말기 및 같은 위치의 행 중에서 첫 행은 체류시작시간으로 하고, 마지막 행은 체류종료시간으로 하여 같은 단말기 및 같은 위치로 짝(체류시작과 체류종료)을 이루는 콜로그 정보를 한 개의 행으로 축약함으로써 (m)열을 생성하는 단계; (9) 상기 (8)단계에서 생성된 (m)열의 체류종료시간과 체류시작시간인 (b)열의 차를 계산하여 체류시간 (n)열을 생성하는 단계;를 포함하는 것을 특징으로

한다.

- [0016] 여기에서, 상기 (5)단계에서 계산된 이진 미분 위치정보 열(j) 측 행과 다음 열(k) 측 행을 더한 값으로 생성된 차기 열(1)에 있어, 같은 단말기 및 같은 위치에서 3건 이상 콜로그 발생하고 이전 행과 다음 행에서 위치변화가 없는 경우에는 '0'으로 기록하고, 단말기가 변경되거나 같은 단말기 및 같은 위치에서 2건 이상 콜로그 발생하고 이전 행 또는 다음 행에서 위치가 변화한 경우의 첫 행(체류 시작)과 마지막 행(체류 종료)은 '1'로 기록하며, 같은 단말기 및 같은 위치에서 발생한 콜로그가 1건인 경우에는 '2'로 기록하여 구분하도록 구성할 수 있다.
- [0017] 여기에서, 상기 축약 생성된 (m)열에서는, 상기 (7)단계에서 생성된 차기 열(1)에서 '0'으로 구분된 행은 제거하고, '2'로 구분된 행은 체류종료시간을 '공란'으로 표시하며, '1'로 구분된 행 중에서 체류종료시간에 해당하는 행은 제거하도록 구성할 수 있다.
- [0018] 여기에서, 상기 (A)단계와 (B)단계 사이에 원본 통신데이터로부터 시간 변수에서 초 단위를 제거하는 단계가 추가되도록 구성할 수 있다.
- [0019] 여기에서, 상기 (1)단계와 (2)단계 사이에 불려온 콜로그 데이터로부터 시간 변수에서 초 단위를 제거하는 단계가 추가되도록 구성할 수 있다.
- [0020] 한편, 상기의 목적을 달성하기 위한 통신데이터 위치정보에 미분기법을 적용한 효과적인 체류정보 추출장치는, 콜로그가 발생한 원본 통신데이터를 불러오되, 원본 통신데이터로부터 정렬된 단말기ID 정보와 위치정보를 미분하여 같은 위치가 연속되는지 아닌지를 판단하는 데이터 판단부; 같은 위치가 연속해서 나오는 행이 발생되면 같은 위치가 연속으로 발생하는 콜로그의 첫 행(체류 시작)과 마지막 행(체류 종료)을 남기고 제거하는 데이터 제거부; 콜로그의 마지막 행을 열로 변환 생성하여 데이터를 축약하는 데이터 축약부; 축약된 데이터로부터 체류시간을 추출하고, 체류시간을 기준으로 이동 중인지 체류 중인지를 구분하는 체류정보 추출부;를 포함하는 것을 특징으로 한다.
- [0021] 여기에서, 상기 데이터 판단부는, 콜로그가 발생한 원본 통신데이터를 불러와서 원본 통신데이터로부터 시간 변수에서 초 단위를 제거하는 시간변수 처리부; 원본 통신데이터로부터 단말기ID 정보와 시간순으로 정렬하는 데이터 정렬부; 원본 통신데이터로부터 정렬된 위치정보와 단말기 ID 정보를 각각 미분하는 데이터 미분처리부; 미분된 데이터를 이진화 처리하여 이진 코드로 변환하는 이진 변환부; 이진 변환 후, 단말기 측 같은 위치가 연속해서 나오는 행이 있는지를 탐색하는 데이터 탐색부;로 이루어지는 구성일 수 있다.

발명의 효과

- [0022] 본 발명에 따르면, 단말기(암호화된 ID) 정보와 위치(위도, 경도) 정보 및 통신시간 정보 등의 콜로그(Call log) 데이터를 단말기와 시간순으로 정렬한 후 미분기법을 적용하는 방식으로서, 같은 위치에서 중복 발생하는 기록과 위치가 변하는 기록을 구분하여 체류 위치별 시작과 종료 시각으로 데이터를 축약할 수 있으며, 이와 같은 축약된 정보를 토대로 하나의 단말기로부터 위치별 체류정보를 매우 효과적으로 추출할 수 있는 유용한 효과를 제공할 수 있다.
- [0023] 본 발명에 따르면, 정렬된 위치정보를 미분하여 같은 위치가 연속되는지 아닌지를 판단하는 기능과, 같은 위치에서 연속으로 발생하는 콜로그의 첫 행과 마지막 행을 남기고 제거하며 마지막 행을 열로 변환시켜 데이터를 축약하는 기술을 제안하는 것으로서, 기존에 비해 통신데이터를 보다 빠르고 효율적으로 처리 가능하고 실시간 처리가 가능하며, 방대한 통신데이터의 저장을 위한 자원을 줄일 수 있는 유용한 효과를 제공할 수 있다.
- [0024] 본 발명에 따르면, 통신데이터의 축약 또는 처리에 유용하게 활용할 수 있을 뿐만 아니라, 특히 감염병 확진자 등의 발생시 해당자에 대한 이동경로 등 동선을 파악함과 더불어 접촉자를 선별하는데에도 활용할 수 있는 등 응용범위를 넓힐 수 있는 유용한 효과를 제공할 수 있다.

도면의 간단한 설명

- [0025] 도 1은 본 발명의 실시예에 따른 통신데이터 위치정보에 미분기법을 적용한 효과적인 체류정보 추출방법을 설명하기 위해 나타낸 순서도이다.
- 도 2 및 도 3은 본 발명의 실시예에 따른 통신데이터 위치정보에 미분기법을 적용한 효과적인 체류정보 추출방법에 있어 처리과정을 보여주는 예시도이다.

도 4는 본 발명의 실시예에 따른 통신데이터 위치정보에 미분기법을 적용한 효과적인 체류정보 추출방법에 있어 원본데이터로부터 처리 후 축약된 데이터를 보여주는 예시도이다.

도 5는 본 발명의 실시예에 따른 통신데이터 위치정보에 미분기법을 적용한 효과적인 체류정보 추출방법으로 처리한 콜로그 데이터 축약 결과를 나타낸 예제이다.

도 6은 본 발명의 실시예에 따른 통신데이터 위치정보에 미분기법을 적용한 효과적인 체류정보 추출장치를 나타낸 블록 구성도이다.

발명을 실시하기 위한 구체적인 내용

- [0026] 본 발명에 대해 바람직한 실시예를 도면을 참조하여 설명하면 다음과 같으며, 이와 같은 상세한 설명 및 도면을 통해서 본 발명의 목적과 구성 및 그에 따른 특징들을 보다 잘 이해할 수 있게 될 것이다.
- [0027] 도 1 내지 도 6은 본 발명의 실시예에 따른 통신데이터 위치정보에 미분기법을 적용한 효과적인 체류정보 추출 방법 및 장치를 설명하기 위해 나타낸 도면들이다. 이동통신 단말기와 통신 기지국 간의 통신은 단말기 사용시에 이루어지거나 단말기 미사용시에도 일정한 시간 간격으로 이루어지는데 본 발명에서 이동통신 단말기와 통신 기지국 간의 신호 정보인 콜로그(Call log)에 포함된 통신 시간 정보는 단말기 사용시와 단말기 미사용시의 콜로그 모두를 아우르는 것이다.
- [0028] 본 발명의 실시예에 따른 통신데이터 위치정보에 미분기법을 적용한 효과적인 체류정보 추출방법은 도 1에 나타낸 바와 같이, 콜로그가 발생된 원본 통신데이터를 불러온다(S10).
- [0029] 이때, 콜로그(call log)는 단말기(암호화된 ID) 정보와 위치(위도, 경도) 정보 및 통신시간 정보 등의 데이터를 가지며, 이동통신 단말기와 통신기지국 간에 주고 받는 통신데이터이다.
- [0030] 원본 통신데이터로부터 시간 변수에서 초 단위를 제거한다(S20).
- [0031] 즉, 시간 변수는 시 및 분까지만 기록한다. 그러나 상기 S20은 선택적 단계이다.
- [0032] 원본 통신데이터로부터 단말기ID 정보와 시간순으로 정렬한다(S30).
- [0033] 즉, 단말기ID 별로 정렬될 뿐만 아니라 각각의 단말기ID의 위치정보에 대해 시간순으로 정렬한다.
- [0034] 이렇게 정렬된 위치정보와 단말기ID 정보를 각각 미분 처리한다(S40).
- [0035] 본 발명에서 미분은 상태변화(단말기 및 위치변경)를 의미한다.
- [0036] 미분된 위치정보와 단말기ID 정보를 각각 이진 변환함으로써 이진 미분 열을 생성하여 기록한다(S50).
- [0037] 이렇게 생성된 이진 미분 열로부터 시간 순서대로 단말기 측 같은 위치가 연속해서 나오는 행을 탐색한다(S60).
- [0038] 단말기 측 같은 위치가 연속해서 나오는 행이 탐색되면, 같은 위치에서 연속으로 발생하는 콜로그의 첫 행(체류 시작)과 마지막 행(체류 종료)을 남기고 나머지는 제거한다(S70),
- [0039] 그리고, 콜로그의 마지막 행을 열로 변환 생성하여 데이터를 축약한다(S80).
- [0040] 이렇게 축약된 데이터로부터 체류시간을 추출하고, 체류시간을 기준으로 이동 중인지 체류 중인지를 구분한다(S90).
- [0041] 이와 같은 상세한 단계로 이루어지는 본 발명에 따른 통신데이터 위치정보에 미분기법을 적용한 효과적인 체류정보 추출방법에 대해 도 2 내지 도 4를 참조하여 더욱 상세하게 설명하면 다음과 같다.
- [0042] (1) 일정 지역 및 일정 기간 동안에 발생한 콜로그 데이터를 불러온다.
- [0043] (1') 이렇게 불러온 콜로그 데이터로부터 시간 변수에서 초 단위를 제거한다. 그러나 상기 (1')단계는 선택적 단계이다.
- [0044] (2) 불러온 콜로그 데이터로부터 단말기ID를 1차 조건으로 하고 통신시간을 2차 조건으로 하여 정렬한다.
- [0045] 도 2의 왼쪽 표에서 단말기별 단말기ID 정보와 통신시간 정보 및 위치(경도, 위도) 정보를 정렬한 예시를 보여주고 있다.
- [0046] (3) 이렇게 정렬된 콜로그 데이터로부터 단말기ID, 위치정보 1(위도 또는 X), 위치정보 2(경도 또는 Y) 각각에

대해 열을 미분함으로써 시간에 따른 변화를 계산하여 저장한다.

- [0047] 도 2의 오른쪽 표에서 단말기별 단말기ID 정보와 위치(경도, 위도) 정보를 미분한 예시를 보여주고 있다. 도 2에서 통신시간은 단말기별로 동일 날짜에 해당하는 시간이다.
- [0048] 도 2의 오른쪽 표에서 e)는 도 2의 왼쪽 표에 정렬한 a)단말기ID를 미분한 예시이고, f)는 c)위치1를 미분한 예시이며, h)는 d)위치2를 미분한 예시이다.
- [0049] (4) 미분된 단말기ID, 위치정보 1(위도 또는 X), 위치정보 2(경도 또는 Y) 각각에 대해 '0' 또는 '1'의 이진 변환을 수행한다.
- [0050] 즉, 미분 후 '0'이 아닌 값은 '1'로 변환하고 콜로그 데이터 중 맨 위 첫행의 미분값과 이진 코드는 '1'로 채운다.
- [0051] 도 2의 오른쪽 표에서 미분된 단말기ID, 위치정보 1(위도 또는 X), 위치정보 2(경도 또는 Y) 각각에 대해 이진 변환을 수행한 예시를 보여주고 있다.
- [0052] 도 2의 오른쪽 표에서 g)는 f)를 이진화 처리한 예시이고, i)는 h)를 이진화 처리한 예시이다.
- [0053] (5) 이진 미분 단말기ID(e), 이진 미분 위치정보 1(g), 이진 미분 위치정보 2(i) 중에서 어느 하나라도 '1'이면 '1', 모두 '0'이면 '0'으로 하여 행을 추가함으로써 단말기 또는 위치의 변화를 표시한다.
- [0054] 도 2의 오른쪽 표에서 j)열이 단말기 또는 위치의 변화를 이진 코드로 표시한 예시이다.
- [0055] g)는 f)를 이진화 처리한 예시이고, i)는 h)를 이진화 처리한 예시이다.
- [0056] (6) 상기 (5)단계에서 계산된 이진 미분 위치정보 열(j)을 1행씩 위로 시프트시키고 빈칸은 '1'로 채워 다음 열(k)을 생성한다.
- [0057] 도 2의 오른쪽 표에서는 상술한 k)열을 생성한 예시를 보여주고 있다.
- [0058] (7) 상기 (5)단계에서 계산된 이진 미분 위치정보 열(j) 측 행과 다음 열(k) 측 행을 더한 값으로 차기 열(l)을 생성한다.
- [0059] 도 2의 오른쪽 표 및 도 3에서는 상술한 l)열을 생성한 예시를 보여주고 있다.
- [0060] 이때, 상기 (5)단계에서 계산된 이진 미분 위치정보 열(j) 측 행과 다음 열(k) 측 행을 더한 값으로 생성된 차기 열(l)에 있어서는, 같은 단말기 및 같은 위치에서 3건 이상 콜로그 발생하고 이전 행과 다음 행에서 위치변화가 없는 경우에는 '0'으로 기록하고, 단말기가 변경되거나 같은 단말기 및 같은 위치에서 2건 이상 콜로그 발생하고 이전 행 또는 다음 행에서 위치가 변화한 경우의 첫 행(체류 시작)과 마지막 행(체류 종료)은 '1'로 기록하며, 같은 단말기 및 같은 위치에서 발생한 콜로그가 1건인 경우에는 '2'로 기록하여 구분한다.
- [0061] 도 2의 오른쪽 표 및 도 3에 표시된 l)열에서, '0'은 같은 단말기, 같은 위치에서 3건 이상 콜로그 발생하고 이전 행과 다음 행에서 위치변화가 없는 때 '0'인 행을 제거한 상태를 의미한다.
- [0062] l)열에서, '1'은 단말기 변경 또는 같은 단말기, 같은 위치에서 2건 이상 콜로그 발생하고 이전 행 또는 다음 행에서 위치가 변화한 때 처음과 마지막 행으로서, 처음 행은 체류시작시간을 의미하고, 마지막 행은 체류종료시간을 의미한다.
- [0063] l)열에서, '2'는 같은 단말기, 같은 위치에서 1건의 콜로그가 발생하여 체류시간의 계산이 불가능한 상태를 의미한다.
- [0064] (8) 상기 (7)단계에서 생성된 차기 열(l)에서 '1'로 구분된 행들만 추출하되, 같은 단말기 및 같은 위치의 행 중에서 첫 행은 체류시작시간으로 하고, 마지막 행은 체류종료시간으로 하여 같은 단말기 및 같은 위치로 짝(체류시작과 체류종료)을 이루는 콜로그 정보를 한 개의 행으로 축약함으로써 (m)열을 생성한다.
- [0065] 도 3의 표에서는 상술한 (m)열을 생성한 예시를 보여주고 있다.
- [0066] m)열은 체류종료시간을 의미하며, l)열에서 1인 값 중 같은 위치 2번째 행의 b)시간을 나타낸다.
- [0067] 여기에서, 상기 축약 생성된 (m)열에서는, 상기 (7)단계에서 생성된 차기 열(l)에서 '0'으로 구분된 행은 제거하고, '2'로 구분된 행은 체류종료시간을 '공란'으로 표시하며, '1'로 구분된 행 중에서 체류종료시간에 해당하는 행은 제거한다.

- [0068] (9) 상기 (8)단계에서 생성된 (m)열의 체류종료시간과 체류시작시간인 (b)열의 차를 계산하여 체류시간 (n)열을 생성한다.
- [0069] 도 3의 표에서는 상술한 (n)열을 생성한 예시를 보여주고 있다.
- [0070] n)열은 체류시간을 의미하며, m)열의 체류종료시간에서 b)시간 즉 체류시작시간의 차를 나타낸다.
- [0071] 이와 같이 미분기법을 적용한 상술한 처리 순서를 통해 단말기별 발생된 콜로그 데이터를 축약할 수 있으며, 단말기별 체류시간 등의 체류정보를 효과적으로 추출할 수 있는 장점을 제공할 수 있다.
- [0072] 도 4에서는 본 발명에 따른 통신데이터 위치정보에 미분기법을 적용한 효과적인 체류정보 추출방법을 통해 콜로그 원본데이터로부터 처리 후 축약된 데이터 예시를 보여주고 있다.
- [0073] 또한, 도 5는 본 발명의 실시예에 따른 통신데이터 위치정보에 미분기법을 적용한 효과적인 체류정보 추출방법으로 콜로그 원본데이터를 축약 처리한 결과를 나타낸 예제로서, 도 5의 왼쪽에 나타낸 데이터가 콜로그 원본데이터이고, 오른쪽이 축약된 결과를 보여주는 데이터이다.
- [0074] 도 5에서는 예제로 콜로그 원본데이터 691행에 있어 처리 후에 208행으로 축약되었음을 보여주고 있으며, 같은 위치에서의 체류시간을 추출할 수 있음을 보여주고 있다.
- [0075] 또한, 본 발명의 적용예로 56만 개의 단말기에서 발생한 63백만 건의 콜로그 데이터에서 체류정보를 추출하였을 때 약 7분이 소요되었는데, 종래 "groupby" 알고리즘으로 단말기별로 처리하였을 때는 4시간 이상이 소요됨을 보여주고 있어 기존에 비해 통신데이터를 보다 빠르고 효율적으로 처리 가능함을 확인할 수 있었다. 즉, 본 발명은 단말기별로 그룹화할 필요가 없이 모든 단말기를 열로 세워 열 단위로 처리되기 때문에 처리 속도가 매우 빠르다.
- [0076] 한편, 상술한 통신데이터 위치정보에 미분기법을 적용한 효과적인 체류정보 추출방법은 알고리즘에 의해 처리되는 것으로서, 하기에 설명하는 통신데이터 위치정보에 미분기법을 적용한 효과적인 체류정보 추출장치가 사용될 수 있다.
- [0077] 본 발명에 따른 통신데이터 위치정보에 미분기법을 적용한 효과적인 체류정보 추출장치는 도 6에서 보여주는 바와 같이, 데이터 판단부(10)와 데이터 제거부(20), 데이터 축약부(30), 및 체류정보 추출부(40)를 포함한다.
- [0078] 상기 데이터 판단부(10)는 콜로그가 발생된 원본 통신데이터를 불러오되, 원본 통신데이터로부터 정렬된 단말기 ID 정보와 위치정보를 미분하여 같은 위치가 연속되는지 아닌지를 판단한다.
- [0079] 상기 데이터 판단부(10)는 콜로그가 발생된 원본 통신데이터를 불러와서 원본 통신데이터로부터 시간 변수에서 초 단위를 제거하는 시간변수 처리부(11, 상기 S10과 S20 및 (1)과(1')단계 수행)와, 원본 통신데이터로부터 단말기ID 정보와 시간순으로 정렬하는 데이터 정렬부(12, 상기 S30 및 (2)단계 수행)와, 원본 통신데이터로부터 정렬된 위치정보와 단말기 ID 정보를 각각 미분하는 데이터 미분처리부(13, 상기 S40 및 (3)단계 수행)와, 미분된 데이터를 이진화 처리하여 이진 코드로 변환하는 이진 변환부(14, 상기 S50 및 (4)단계 수행), 및 이진 변환 후 단말기 즉 같은 위치가 연속해서 나오는 행이 있는지를 탐색하는 데이터 탐색부(15, 상기 S60 및 (5)내지 (7)단계 수행)로 이루어진다.
- [0080] 상기 데이터 제거부(20, 상기 S70 및 (8)단계 수행)는 같은 위치가 연속해서 나오는 행이 발생되면 같은 위치가 연속으로 발생하는 콜로그의 첫 행(체류 시작)과 마지막 행(체류 종료)을 남기고 제거한다.
- [0081] 상기 데이터 축약부(30, 상기 S80 및 (8)단계 수행)는 콜로그의 마지막 행을 열로 변환 생성하여 데이터를 축약한다.
- [0082] 상기 체류정보 추출부(40, 상기 S90 및 (9)단계 수행)는 축약된 데이터로부터 체류시간을 추출하고, 체류시간을 기준으로 이동 중인지 체류 중인지를 구분한다.
- [0083] 이에 따라, 본 발명에 따른 방법과 장치를 통해서는 단말기(암호화된 ID) 정보와 위치(위도, 경도) 정보 및 통신시간 정보 등의 콜로그(Call log) 데이터를 단말기와 시간순으로 정렬한 후 미분기법을 적용함으로써 같은 위치에서 중복 발생하는 기록과 위치가 변하는 기록을 구분하여 체류 위치별 시작과 종료 시각으로 데이터를 축약할 수 있으며, 이와 같은 축약된 정보를 토대로 하나의 단말기로부터 위치별 체류정보를 매우 효과적이면서도 빠른 처리속도로 추출할 수 있는 장점을 제공할 수 있다.
- [0084] 또한, 본 발명은 통신데이터의 축약 또는 처리에 유용하게 활용할 수 있을 뿐만 아니라, 특히 감염병 확진자 등

의 발생시 해당자에 대한 이동경로 등 동선을 파악함과 더불어 빠른 시간 내에 접촉자를 선별하는데에도 활용할 수 있는 장점을 제공할 수 있다. 즉, 매일 생성되는 빅데이터에서 의미 있는 정보를 태깅하여 적은 용량으로 저장 가능하다는 것은 활용 가치가 다양하다. 예를 들어 한 달 동안 전국의 이동통신데이터를 저장하는 것이 저장 공간의 한계라면 본 발명을 활용한다면 1년치의 데이터가 저장이 가능해지며 태깅 정보를 활용하여 단말기별 체류한 위치들을 추출하여 감염병 확진자로 확인 시 과거의 저장된 데이터를 활용하여 같은 위치에 같은 시간에 있었던 다른 단말기의 목록을 제공할 수 있는 일종의 블랙박스를 구축할 수 있는 것이다.

[0085] 이상의 설명은 본 발명을 예시적으로 설명한 것이고, 명세서에 게시된 실시예는 본 발명의 기술사상을 한정하기 위한 것이 아니라 설명하기 위한 것이므로 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자라면 본 발명의 기술사상을 벗어나지 않는 범위에서 다양한 수정과 변형 또는 단계의 치환 등이 가능할 것이다. 그러므로 본 발명의 보호범위는 청구범위에 기재된 사항에 의해 해석되고, 그와 균등한 범위 내에 있는 기술적 사항도 본 발명의 권리범위에 포함되는 것으로 해석되어야 할 것이다.

부호의 설명

- [0086] 10: 데이터 판단부
- 20: 데이터 제거부
- 30: 데이터 축약부
- 40: 데이터 추출부

도면

도면1



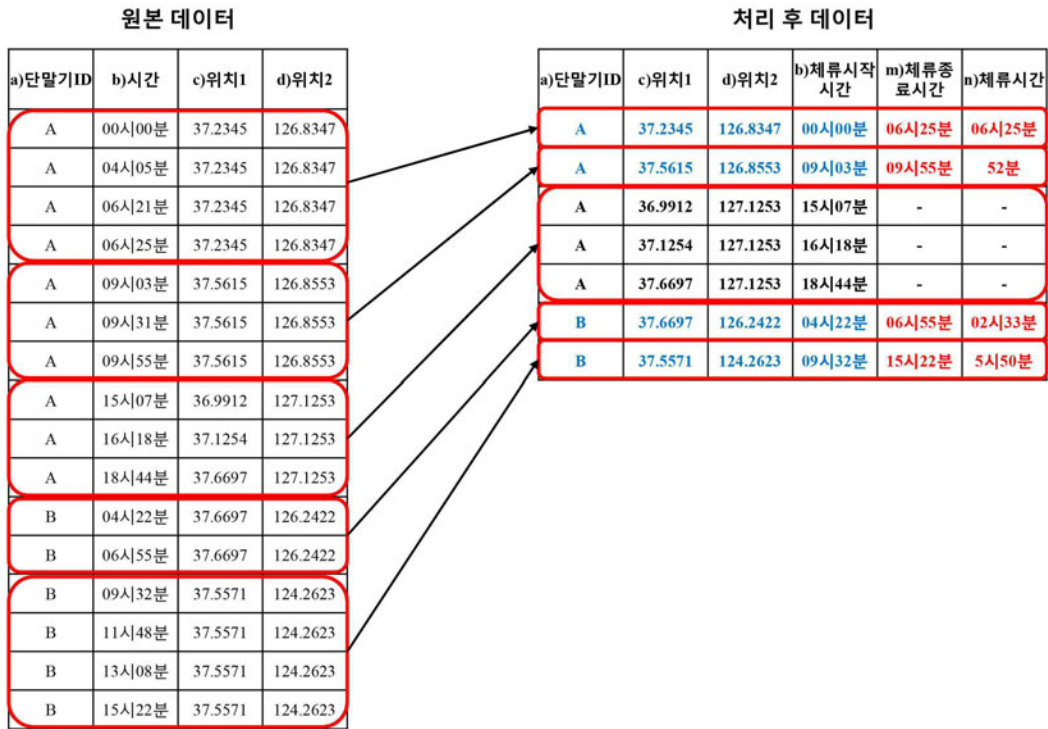
도면2

a)단말기ID	b)시간	c)위치1	d)위치2	e)	f)	g)	h)	i)	j)	k)	l)
A	00시00분	37.2345	126.8347	1	1	1	1	1	1	0	1
A	04시05분	37.2345	126.8347	0	0	0	0	0	0	0	0
A	06시21분	37.2345	126.8347	0	0	0	0	0	0	0	0
A	06시25분	37.2345	126.8347	0	0	0	0	0	0	1	1
A	09시03분	37.5615	126.8553	0	0.327	1	0.0206	1	1	0	1
A	09시31분	37.5615	126.8553	0	0	0	0	0	0	0	0
A	09시55분	37.5615	126.8553	0	0	0	0	0	0	1	1
A	15시07분	36.9912	127.1253	0	-0.5703	1	0.27	1	1	1	2
A	16시18분	37.1254	127.1253	0	0.1342	1	0	0	1	1	2
A	18시44분	37.6697	127.1253	0	0.5443	1	0	0	1	1	2
B	04시22분	37.6697	126.2422	1	0	0	-0.8831	1	1	0	1
B	06시55분	37.6697	126.2422	0	0	0	0	0	0	1	1
B	09시32분	37.5571	124.2623	0	-0.1126	1	-1.9799	1	1	0	1
B	11시48분	37.5571	124.2623	0	0	0	0	0	0	0	0
B	13시08분	37.5571	124.2623	0	0	0	0	0	0	0	0
B	15시22분	37.5571	124.2623	0	0	0	0	0	0	1	1

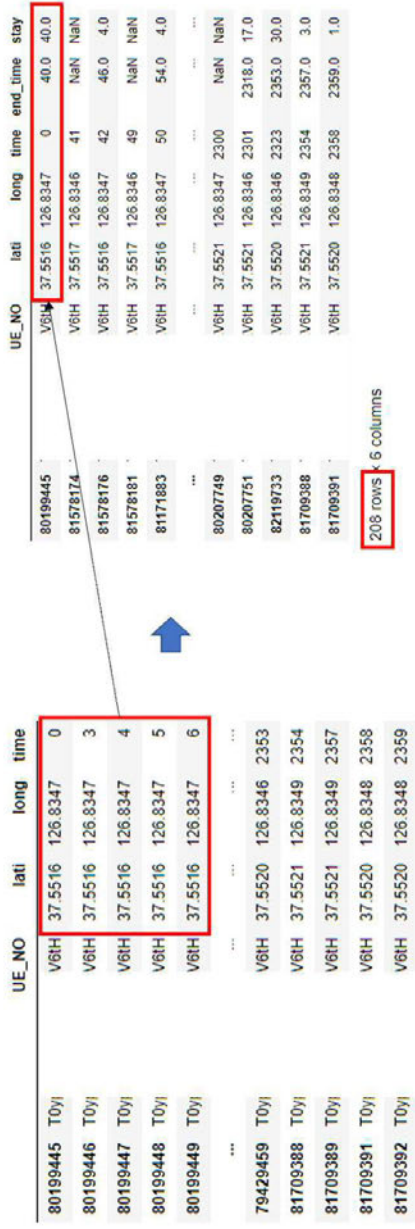
도면3

a)단말기ID	b)시간	c)위치1	d)위치2	l)	m)	n)
A	00시00분	37.2345	126.8347	1		
A	04시05분	37.2345	126.8347	0		
A	06시21분	37.2345	126.8347	0		
A	06시25분	37.2345	126.8347	1	06시25분	06시25분
A	09시03분	37.5615	126.8553	1		
A	09시31분	37.5615	126.8553	0		
A	09시55분	37.5615	126.8553	1	09시55분	52분
A	15시07분	36.9912	127.1253	2	-	-
A	16시18분	37.1254	127.1253	2	-	-
A	18시44분	37.6697	127.1253	2	-	-
B	04시22분	37.6697	126.2422	1		
B	06시55분	37.6697	126.2422	1	06시55분	02시33분
B	09시32분	37.5571	124.2623	1		
B	11시48분	37.5571	124.2623	0		
B	13시08분	37.5571	124.2623	0		
B	15시22분	37.5571	124.2623	1	15시22분	5시50분

도면4



도면5



도면6

