



Análisis de los métodos de Big Data, desafíos y Aplicaciones en sistemas de transporte inteligentes

MAHENDRA.G¹, ROOPASHREE.H.R², YOGESH.A.C³

¹ Investigador académico, Departamento de CSE, GSSSIETW, Mysuru, VTU, Belagavi, gg.mahendra@gmail.com

² Profesor asociado, Departamento de CSE, GSSSIETW, Mysuru, VTU, Belagavi, roopashreehr@gsss.edu.in

³ Profesor Adjunto, Departamento de Ingeniería Civil y Ambiental, Facultad de Ingeniería del Gobierno, Kushalnagar, yogesh13@gmail.com

ABSTRACTO

El Big Data es un conjunto de datos que supera la capacidad de los métodos tradicionales de minería de datos. Se trata de un conjunto de datos voluminosos, heterogéneos y rápidos generados por numerosos sensores instalados cerca de infraestructuras viales.

Dispositivos asociados a los vehículos. Necesitamos diseñar modelos de Big Data altamente escalables, con técnicas para tipos de datos complejos, robustez al ruido, insensibilidad al orden de entrada, capacidad para manejar datos de alta dimensión, interpretabilidad y usabilidad. Existen numerosos desafíos en la enorme cantidad de datos que almacenar, procesar y operar.

Apache Spark, Hadoop y Flink son las plataformas de análisis de Big Data.

El procesamiento masivo de información en los sistemas de transporte enfrenta dificultades en términos de privacidad, seguridad, almacenamiento y calidad de los datos. Se realizan diversos análisis de datos de Big Transportation, como análisis de video y audio, análisis de imágenes y análisis históricos. El mundo se monitorea y gestiona de manera eficiente y eficaz mediante la Web de Dispositivos (WoD).

Palabras clave: Big Data, Hadoop, Spark, Flink, Stockpile, WoD.

1. INTRODUCCIÓN

En el mundo, los datos se generan e integran a partir de diversas fuentes, como redes sociales, datos geoespaciales y textos de comunicación lingüística. Los datos pueden ser homogéneos o heterogéneos. Algunas herramientas de bases de datos antiguas solo son compatibles con datos homogéneos. Vivimos en la era de los datos, donde terabytes o petabytes de datos generados por diversas empresas, ciencia, ingeniería, sociedad y medicina se vierten en las redes informáticas. El análisis de datos global se ha desarrollado desde diversos dominios, como el análisis visual, el procesamiento masivo de datos y el análisis perceptivo. La minería de datos también se define como la búsqueda de información a partir de datos brutos. Las técnicas e ideas básicas de la minería de datos consisten en obtener patrones o datos interesantes de una gran cantidad de datos en diversas aplicaciones. El procesamiento de datos puede ser...

Se aplica a cualquier dato, siempre que esté destinado a una aplicación específica. El rápido crecimiento de los datos masivos ha otorgado al procesamiento de datos un papel mucho más importante, además de mayores desafíos. Los datos masivos deben incluir tanto datos no estructurados como semiestructurados, además de los predominantemente estructurados. Los desafíos relacionados con el tipo de datos y otros, como la cantidad y la velocidad de su producción, se han acumulado enormemente en los últimos años. El volumen, la variedad y la velocidad de los datos miden las tres características principales de los datos masivos. Los datos globales se han acumulado de cincuenta a doscientos petabytes entre 2008 y 2015[1]. El análisis de datos desempeña un papel especialmente importante en el sistema de vehículos inteligentes. Se desarrollan numerosas aplicaciones de software, como los sistemas de alerta ante desastres, la gestión automatizada de los controles de velocidad, la programación activa de semáforos, la transmisión y la colaboración entre vehículos, y la previsión y redirección del tráfico vial.

Muchos de los métodos de información incluyen métodos de big data, así como procesamiento de información, automatización, IA, fusión de conocimientos y, por consiguiente, redes sociales, etc. [2]. Los ITS buscan brindar un mejor servicio a conductores y viajeros en los sistemas de transporte [3-5]. El sistema de transporte analizará big data en los siguientes aspectos:

El análisis de big data se gestiona con grandes cantidades de información diversa y sofisticada generada en el sistema de transporte. El análisis de big data ha resuelto tres problemas: almacenamiento, análisis y gestión de datos. Existen plataformas de datos de gran envergadura como Apache Spark y Hadoop capaces de procesar grandes cantidades de datos [6].

- i) El análisis de grandes volúmenes de información mejorará la eficacia del funcionamiento del sistema de transporte.
- ii) El análisis de grandes volúmenes de información mejorará el nivel de seguridad del sistema de transporte.

1.1 Funciones del análisis de datos

Hay cuatro aspectos del análisis de datos. Son:

- Análisis explicativo (descriptivo)
- Análisis proyectivo (predictivo)

- Análisis investigativo (diagnóstico)
- Análisis prescriptivo

Estos cuatro aspectos son útiles para representar la secuencia de pasos del proceso de análisis de datos. El análisis descriptivo de datos describe las propiedades de los datos en un conjunto de datos de aplicación muy específico para describir las características de un individuo específico. El análisis descriptivo analiza datos históricos para generar datos tanto cuantificables como cualitativos. El análisis de datos de diagnóstico determina la frecuencia con la que se ha descubierto algo dentro del análisis de datos descriptivos o si está relacionado con otro. El análisis de diagnóstico utiliza diversas técnicas, incluyendo técnicas de almacenamiento y procesamiento de datos.

El análisis de diagnóstico utiliza las operaciones de desglose de OLAP (es decir, desglose de datos). El análisis de diagnóstico es laborioso y exploratorio. El análisis predictivo se emplea para generar predicciones basadas en datos actuales mediante numerosos modelos matemáticos y estadísticos aplicados.

El análisis predictivo predice lo que podría ocurrir en el futuro basándose en eventos pasados. Los paradigmas de proyección más populares son los árboles de decisión y las redes neuronales.

Los paradigmas de proyección se crean con información de entrenamiento y son probabilísticos. El análisis prescriptivo proporciona recomendaciones inteligentes sobre los resultados más comunes. El análisis de datos prescriptivo y predictivo genera confusión debido a su gran superposición.

El análisis prescriptivo aborda las cuestiones que surgen en el análisis diagnóstico. Se basa en análisis informativos, investigativos y de proyección.

El análisis prescriptivo se aplica mediante análisis de características perceptivas.

1.2 Análisis de datos de transporte

Los datos asociados con el transporte, recopilados de diversas fuentes, contienen información valiosa. Se pueden presentar diversos análisis sobre el transporte de información, que incluyen datos en tiempo real, históricos, predicciones, videos e imágenes, y análisis visual [7].

1.2.1 Análisis de datos históricos de transporte

Se podrían investigar los datos relacionados con el transporte histórico para encontrar tendencias y ejemplos que ayuden a crear opciones a largo plazo y a impulsar la expansión urbana. Por ejemplo, ahora es posible examinar los flujos de tráfico, el comportamiento del conductor y el uso de las infraestructuras viales antes de desarrollar una infraestructura de transporte moderna [8]. Además, los datos del transporte pueden ser analizados y utilizados por los fabricantes para regular sus políticas, como el transporte gratuito o la señalización vial [9].

La expansión urbana puede incluso beneficiarse de la información recopilada de la población general [10]. Este esfuerzo exploró la guía de transporte público de colaboración colectiva utilizando datos de geolocalización, donde el usuario mejoró las guías de transporte público a través del mapa web. Estos datos permiten identificar vías de consulta importantes, horas punta e información relevante para el desarrollo del transporte y las rutas. Esto se utiliza para estudiar reglas algorítmicas de aglomeración según las preferencias del usuario y aplicarlas.

Estos algoritmos guían a la constelación para mejorar las conexiones de transporte óptimas a la comunidad.

1.2.2 Investigación de datos reales de transporte

La investigación de datos de transporte reales representa el proceso de capacidad, cuyo análisis y difusión son breves, ya que la información pertinente existe en el marco. Ofrecerá los datos a los segundos creadores de evaluaciones justo antes de permitirles elaborar una mejor opción comercial con mayor rapidez. La necesidad de actualizaciones manuales impulsa la necesidad de analizar el tiempo, cuya expectativa puede variar desde segundos hasta minutos. Por ejemplo, la técnica actual de gestión adaptativa necesita ejecutar inmediatamente el análisis y obtener resultados para determinar si el orden temporal del indicador debe modificarse para adaptarse a los diferentes datos del estado del tráfico. La difusión de datos a los usuarios en un intervalo de tiempo muy significativo, o incluso más largo, genera información adicional correcta y evita la infobesidad que pueda confundir a los usuarios o generar una reacción inesperada. La diferencia entre el suministro de datos y la necesidad de actualizar al usuario verifica la latencia adecuada del análisis, lo que requiere la elección de los métodos de investigación utilizados. La investigación dentro de la instancia de análisis de datos puede ser vital en el espacio de big data dentro del análisis del dominio del transporte, se realizará para medir el flujo de tráfico para observar el accidente o la condición de la aventura del robo en la carretera y

Ayuda a los vehículos de emergencia a encontrar la mejor ruta [11-13]. El procedimiento de selección de ruta, que analiza la ruta más corta según las necesidades de cada experto, también se encuentra en las obras [14]. La percepción de la situación del flujo de tráfico en tiempo real se basará en datos recopilados mediante la explotación de dispositivos de señalización vial, el Sistema de Posicionamiento Global (SGP), dispositivos inteligentes y vehículos conectados.

1.2.3 Análisis predictivo

El análisis predictivo, basado en grandes cantidades de datos sin procesar, y los marcos de big data facilitan el desarrollo de modelos proféticos. Un modelo de predicción utiliza el procesamiento de datos mediante aprendizaje automático (ML) para obtener información sobre los datos y generar predicciones de la posible actividad de tráfico. Por ejemplo, predicciones de obstáculos para el tráfico [15].

Este sistema de previsión utiliza una plataforma de minería de datos conocida como KNIME para analizar la información climática y del flujo de tráfico recopilada por dispositivos de señalización vial en Santander, España. Mediante modelos de previsión, el sistema predice el tráfico en los próximos quince minutos. El análisis predictivo permite a la organización adecuada prepararse para los eventos de tráfico previstos. Por ejemplo, si se prevé una afluencia de vehículos debido a un evento importante, la autoridad de transporte ofrecerá vehículos específicos para atender el evento e instará a la población a solicitar transporte.

Cuando se prevé que un mal clima pueda provocar algunos accidentes, se recomienda ser más cauteloso al conducir.

La policía podría desplegar una fuerza adicional en cualquier lugar donde sea probable que se produzcan delitos. Las pruebas de visión autorizan la aplicación de las normas pertinentes.

posición para necesitar actividades dinámicas y defensivas para ayudar y mejorar el bienestar de la red.

1.2.4 Análisis de datos de transporte visual

El análisis visual depende de la visualización individual para reconocer modelos de datos, que difieren significativamente de la apariencia de la información de planificación de renovación en un formulario. Una imagen visual bien diseñada permitirá a los clientes obtener información sobre la información utilizando características perceptivas de bajo esfuerzo. Los grupos de información que permanecen solo para que se vuelvan evidentes en el gráfico se muestran con mayor claridad. Por ejemplo, la función asociada, conocida como analizador de accidentes de tráfico, analiza información sobre accidentes de tráfico en la ciudad de Nueva York y visualiza la información en una guía [16]. La imagen muestra una visión completa del área del accidente y las variedades del objeto correspondiente (por ejemplo, ciclistas, camiones, furgonetas y conductores de automóviles), junto con diferentes características como la hora, la ubicación y el motivo de la lesión o muerte. Los clientes verificarán el razonamiento detrás de la tasa de impacto o las variedades del objeto correspondiente en el gráfico y el área del accidente en el mapa. En el análisis visual, es común generar varias vistas coordinadas para permitir a los clientes comprender la complejidad del conocimiento. Por ejemplo, se predice la reflexión sobre las redes sociales y el crédito de información de transporte móvil en Tokio, presentando tres vistas coordinadas: Mapa de Calor, Cinta de Animación y Burbuja de Tweet [17]. Cualquier lectura se centra en términos de información seleccionados. La lectura del mapa se centra en las dimensiones temporales de los datos, es decir, si está activo o no con una calificación mayor que la cantidad restante en el análisis diario. La Cinta de Animación utiliza simulación y cinta codificada por colores sobre la vía para visualizar el número de viajeros y la variación del público en ciertas rutas.

1.2.5 Análisis de datos de transporte de imágenes y videos

El análisis de video del transporte con imágenes es cada vez más común en el ámbito del transporte y el análisis de big data. Tareas básicas de análisis de video, como el código de matrícula y la clase de vehículos, impulsan fotografías de transporte [18], se utilizan para el control social del tráfico, ya que aumenta el riesgo de fraudes en la emisión de documentos de viaje. El propósito de este plan es detectar y analizar fallas cercanas y comprender sus causas. El objetivo de la aplicación, con la previsión 0, es reducir a cero las muertes y los daños significativos en el tráfico.

La información del video también facilita la investigación de seguridad del tráfico de datos [19]. El proceso consiste en analizar cerca de 473 periodos de datos de video de veinte rotondas en Canadá, concretamente Quebec. La investigación utiliza diversas reglas, como la extrapolación de la portabilidad de la probabilidad de colisión y la colaboración de vehículos. Esta investigación ayudará a identificar zonas con alta incidencia de accidentes y permitirá a la oficina de transporte formular las medidas necesarias para reducir el número de impactos en dichas zonas.

2. ESTRATEGIAS DE BIG DATA EN SU

La IA es la hipótesis de demostración e investigación preferida en entornos de gran volumen de información, lo que facilita la obtención de ejemplos y paradigmas a partir de una enorme cantidad de datos. Los paradigmas de IA se clasifican habitualmente en procedimientos de aprendizaje supervisado, no supervisado y de refuerzo. La información de configuración etiquetada se utiliza en un programa algorítmico de aprendizaje regulado. Los paradigmas se utilizan para introducir documentos y, en consecuencia, obtener un resultado real para realizar el trabajo o guiar uno de ellos. Junto con el paradigma académico y, por lo tanto, el registro de datos, se anticipan regularmente los resultados ocultos.

La regresión lineal se utiliza para explicar la asociación entre factores de variación única y factores autónomos únicos o adicionales. La regresión lineal es el aprendizaje supervisado más utilizado. Es sumamente sencilla, potente, fácil de interpretar y de programar. A pesar de su simplicidad, la regresión lineal es especialmente útil en numerosas proyecciones de ITS, como la predicción del flujo de tráfico, la estimación de la velocidad del tráfico y el análisis de rutas de transporte [20].

Un árbol de decisiones puede ser el resultado de un dispositivo que funciona como una cadena de mando, como un diagrama, para mostrar las decisiones y sus resultados alcanzables [21]. Debido a su tamaño, densidad y simplicidad, los árboles de decisiones

Los árboles se utilizan ampliamente en diversas proyecciones de ITS, como la detección de accidentes de tráfico, el análisis de la gravedad de los accidentes [22] y la selección de modos de transporte. Las RNA (Redes Neuronales Artificiales) pueden ser un ejemplo innovador de aprendizaje administrado flexible y robusto tanto para la caracterización como para la recaída. Con suficientes capas ocultas de centros de procedimientos y datos de preparación, las RNA comprenderán cualquier relación indirecta entre la información y los datos objetivo. Como dispositivo de visualización de datos, también se ha adoptado en ITS, como la predicción del flujo de tráfico [23] y los viajes. Pronóstico de tramo, descubrimiento automático de colisiones y pronóstico de regiones de detención restantes.

SVM (Máquina de Vectores de Soporte) es una estrategia de aprendizaje supervisado más conocida que utiliza información marcada para la recaída y la organización. Entre todas las herramientas de modelos de investigación de datos grandes en ITS, las SVM han obtenido excelentes resultados en el área de estudio. Se han utilizado eficazmente en la predicción de la longitud del movimiento, la predicción del tiempo de tránsito de una aparición y la identificación de accidentes. El aprendizaje no supervisado, también conocido como aprendizaje de enfoque de aglomeración, se basa en el aprendizaje de grupos naturales a partir de información plana no etiquetada. K-means es la herramienta de aprendizaje no supervisado más popular y se ha utilizado ampliamente en el desarrollo del transporte público básico y la predicción de la duración del viaje [24].

El aprendizaje de refuerzo espera reducir el valor del final del día a través de la investigación y ganar competencia con la metodología menos compleja comunicándose con el

Información exploratoria. El aprendizaje de fortificación es muy pertinente para la hipótesis de tablero y mejora, y se ha evaluado como muy sólido en el tablero ligero en ITS [25-28]. El análisis de series temporales podría ser un conjunto de métodos matemáticos aplicados para paradigmizar y justificar series de puntos de datos relativos a instancias. La formulación estadística utiliza un paradigma para obtener pronósticos de extrapolaciones para eventos futuros basados en eventos pasados identificados. Los datos estadísticos presentan una colección secuencial habitual. Esto difiere de las aplicaciones típicas de minería de datos/aprendizaje automático, donde cada dato es la realización independiente asociada del concepto a adquirir y, por lo tanto, el orden de los puntos de datos dentro de un conjunto de datos es irrelevante.

El análisis de conglomerados es un conjunto de métodos matemáticos aplicados para agrupar datos, dividiendo un conglomerado en grupos más pequeños de datos comparables, cuyas características de relación no son evidentes de antemano. Esto suele ser un tipo de aprendizaje no supervisado, ya que no se aplican los datos de entrenamiento [29]. La clasificación es un conjunto de métodos para identificar las clases a las que pertenecen los nuevos puntos de datos, basándose en un conjunto de entrenamiento que contiene datos previamente clasificados. Estas técnicas se consideran típicamente como reconocimientos de aprendizaje supervisado de la realidad de un grupo de datos de entrenamiento [30]. El análisis de puntos de cambio podría ser una herramienta potente y versátil que caracteriza bien los cambios, determinando si se ha producido una modificación. Es capaz de Variaciones delicadas en el trabajo de investigación; caracteriza las variaciones del trabajo de investigación proporcionando intervalos y niveles de confianza. Es adecuado además para la detección de un esquema. El análisis de puntos de cambio es compactable para datos históricos de procesos, especialmente al abordar big data [31].

3. ENTORNO DE BIG DATA PARA SU

El análisis de grandes volúmenes de datos en ITS se está desarrollando gracias al innovador ecosistema de big data. La plataforma de datos masivos utiliza la capacidad de computación paralela y un sistema de archivos distribuidos para cambiar la estrategia de datos rápidos. Es ideal para generar un sentido de desarrollo de sistemas de gran alcance y, además, para el soporte de grandes volúmenes de datos. Apache Hadoop es la plataforma de software libre preferida para la distribución.

Técnicas y la capacidad de gestionar una gran cantidad de índices de información. Hadoop podría ser una plataforma integral de estrategia de datos a gran escala, donde se difundirán diversos tipos de técnicas de datos o actividades de sistemas de información. La técnica adecuada permite a Hadoop desglosar la información en sistemas de información de sistemas (ITS), como datos de tarjetas inteligentes, diversos dispositivos electrónicos, redes sociales, datos de GPS, etc.

Apache Spark es la base más reciente del código abierto. Punto de acceso para una gran cantidad de procesos de índices de información que se adaptan perfectamente a las tareas de IA. Spark adopta una tecnología Hadoop similar a la del almacenamiento distribuido y permite que los programas de usuario se carguen en un conjunto de memoria de grupo, y se consulta periódicamente. Spark es ideal para...

Enfoques de IA. La investigación de datos masivos se acerca; en la última sección, presentaremos unidades de IA basadas principalmente en IA, que se implementarán en las etapas Spark y Hadoop. La etapa de datos masivos, con los enfoques de análisis de datos que se implementarán posteriormente, puede desempeñar un papel crucial en la investigación de datos masivos en ITS.

4. APLICACIONES DE BIG DATA EN SU

4.1 Análisis de la Justificación de Accidentes de Tránsito en Carreteras revela que, a nivel mundial, casi 120.000 personas mueren y 500.000 resultan heridas cada año por accidentes de tráfico [32]. En [33], se adoptan la inducción bayesiana y los bosques aleatorios en un modelo de pronóstico de accidentes de período para reducir los riesgos de colisión y los peligros de accidente. En [34], se utilizan árboles de clasificación y regresión (CART), la recaída logística y las conexiones de la recaída adaptativa variable para implementar el procedimiento sistemático sobre la información mecanizada de las lesiones por accidentes de tráfico.

4.2 Pronóstico del flujo de tráfico en carreteras. La información ideal y precisa sobre el flujo de tráfico es fundamental para el transporte. Investigar a fondo la información en ITS resulta beneficioso para predecir los atascos en horas punta [35-36].

4.3 Administraciones de Transporte Abierto Coordinando La investigación de la enorme cantidad de información sobre transporte abierto facilitará la comprensión de los patrones de viaje de los usuarios del transporte en las coordenadas del mismo. Los ejemplos de viaje de los usuarios servirán para orientar a los administradores de transporte en la planificación de sus servicios. Además de diversas fuentes de información sobre estimación de tráfico, en [37] se describe un flujo de ruta basado esencialmente en un paradigma de desarrollo no lineal para evaluar la demanda activa de origen-destino que no...

Quieren información de conexión explícita y potente. Abusan de registros portátiles triangulados de numerosos clientes desconocidos. En [38] se describe un enfoque para prever visitas diarias normales de origen a destino.

4.4 Diseño de rutas de viaje personales Teniendo en cuenta la información de teléfonos razonables y la información del GPS del automóvil, ciertas aplicaciones de transporte brindan a los viajeros información de tráfico genuina [39], mientras que otras ofrecen los cursos de conducción más apropiados con el menor tiempo posible. [40]. El análisis de grandes volúmenes de información en ciertas aplicaciones de transporte crea enormes ventajas económicas al reducir los tiempos de viaje, el flujo de vehículos, la contaminación y las emisiones de sustancias que agotan la capa de ozono.

4.5 Junta y Controlador de Transporte Ferroviario Con la ayuda de tecnología informática avanzada, se están remodelando los sistemas de transporte ferroviario. Estos son los principales receptores de grandes volúmenes de datos. Por ello, los sistemas de transporte ferroviario suelen contar con sistemas seguros que procesan con precisión grandes volúmenes de datos, como la posición y la velocidad.

Información sobre la hora de llegada y salida de un tren a una estación y los datos de origen-destino (OD) de los viajeros. Los operadores de la junta de transporte ferroviario mejoran su funcionamiento mediante el análisis de big data.

4.6 Modelos predictivos de pronóstico del tiempo de viaje

Los tiempos de viaje tienen una influencia crucial en el transporte y la coordinación. Desde la perspectiva de los exploradores, la información sobre el intervalo de viaje ayuda a reducir los retrasos y a optimizar la carga de trabajo gracias a una mayor variedad de rutas. En la coordinación, una aproximación correcta del intervalo de viaje puede contribuir a reducir los costos de transporte y a aumentar la calidad del servicio de transporte al mover los artículos en el período ideal. Para los responsables de tráfico, el tiempo de viaje es fundamental para la productividad de la actividad del sistema de transporte [41].

4.7 Estados de tráfico del clustering

Los problemas más generales de agrupamiento y clasificación en el análisis de tráfico son el agrupamiento del estado del tráfico [42] y el agrupamiento del comportamiento de los participantes para la formación de agrupamientos [43].

El agrupamiento de datos de tiempo de viaje buscó encontrar patrones de tráfico homogeneizados con el modelo de predicción común utilizado. Un agrupamiento, que podría ser una técnica prometedora de estadística de procedimientos no paramétricos, permite exponer agrupamientos formados arbitrariamente.

4.8 Problema de enrutamiento del tráfico

Se investigó un problema de movilidad vial con la decisión localizada de los gestores de vehículos en el sistema de flujo de tráfico urbano. El método de diseño, a cargo de los gestores de vehículos, se divide en dos fases: diseño clave para la decisión del mejor curso y desarrollo del plan de acción para la circulación de la vía actual. Se utilizaron un diseño de sistema multiagente y técnicas basadas en información de procesos clave para...

Se evaluaron dos rutas en una visualización muy aleatoria [44][45] y se desarrolló la búsqueda de la ruta más cercana [46], las cuales se asignaron en la mesa de diseño estratégico. Los modelos se autorizaron a partir de información genuina y se integraron en una función de territorio de tráfico; en este caso, se evaluó la eficacia de los métodos. Un método de desarrollo distribuido para la vía de movimiento del tráfico fue la idea propuesta en [47].

4.9 Modelado de requisitos de transporte

Para predecir el comportamiento y representar las necesidades de viaje, es necesario evaluar tres niveles: estratégico, que examina las vías de mejora de la infraestructura a largo plazo; plan de acción, además de la planificación y gestión; y, por último, operativo, así como políticas de gestión de la congestión a corto plazo. Los modelos de necesidades de transporte son métodos estadísticos que predicen las necesidades de calidad a largo plazo, basándose en proyecciones futuras y condiciones actuales. Requieren capacidades analíticas específicas, como la predicción de las necesidades de calidad, la elección de la ruta, la elección de la propia ruta y la representación del flujo de tráfico dentro de la red vial. Estos sistemas se limitan a evaluar con precisión los cambios en los parámetros (por ejemplo, velocidad, retrasos y líneas que se están formando) debido a...

ejecución de formularios de obtención de información, vulnerabilidad ante sospechas y estructura del modelo.

4.10 Diseño / Dirección

Decidir las rutas ideales en los sistemas viales desde una fuente propuesta hasta un área objetivo determinada es un problema frecuente en la vida diaria. También existen herramientas y

Aplicaciones como organizadores estratégicos y sistemas de pruebas de tráfico que requieren abordar innumerables consultas de rutas. Los métodos de organización de rutas han avanzado rápidamente en términos de productividad y precisión en los últimos años. Uno de los dos cálculos de dirección propuestos considera la separación (Dijkstra, Bellman-Ford, utilizando medidas predefinidas) entre un nodo y otro en un plan. Los grandes procedimientos amplían estos cálculos con datos de refuerzo y una metodología inagotable para mejorar sus resultados. El big data metaheurístico en estrategias de transporte vial y exploración de versatilidad también se ha utilizado eficazmente en el cálculo de la ruta más corta. Sin embargo, estos sistemas producen tiempos de consulta cortos cuando se gestionan sistemas viales inteligentes, lo que dificulta su uso en aplicaciones continuas o intuitivas.

Por otra parte, la aplicación de heurísticas rigurosas no suele producir resultados precisos. Los sistemas de calles presentan propiedades básicas; por ejemplo, son escasos, estratificados, prácticamente planos y presentan estructuras progresivas, lo que facilita diversos métodos de aceleración. Estos procedimientos dependen de una fase de preprocesamiento donde se obtienen datos auxiliares sobre el sistema, se comentan y posteriormente se utilizan para acelerar los resultados.

Consultas. Los sistemas de aceleración se pueden organizar en procedimientos de búsqueda objetiva coordinada que involucran diversos cálculos, como los geométricos.

5. DESAFÍOS ABIERTOS

Las dificultades fundamentales abiertas al utilizar el análisis de grandes datos en ITS son las siguientes.

5.1 Surtido de información

Debido al continuo desarrollo de vehículos y peatones, los datos recopilados en el transporte pueden ser incorrectos, deficientes o problemáticos en ciertas zonas o en ocasiones específicas. Por ejemplo, no todos los vehículos cuentan con los procedimientos necesarios para proporcionar datos de área continuos, y la información del tráfico vial procedente de los dispositivos viales puede estar ausente. Algunas posibles estrategias para gestionar la competencia son introducir avances innovadores en la recopilación de datos y mejorar su capacidad. Gracias al avance del IoT, cada año se desarrollan nuevos sistemas de sensores, lo que puede contribuir a mejorar la recopilación y la calidad de los datos. Asimismo, la opción de automatizar la recopilación de datos para limitar la entrada manual de datos es clave para mejorar la calidad de los mismos.

5.2 Seguridad de la información

En tiempos de enorme información, el problema más preocupante y complejo es la seguridad [48]. La seguridad singular puede ser

Se filtra durante la transmisión, el almacenamiento y el uso de información [49]. La información acumulada de las estructuras de transporte debe ser no singular, por ejemplo, información sobre el flujo de tráfico en las zonas de vehículos. Sin embargo, se han generado problemas de seguridad, ya que la recopilación de información singular por parte de todos y de las zonas privadas se crea con el tiempo. Por ejemplo, la superficie de personas y vehículos se recopila fácilmente.

Ante la posibilidad de que esta información no se proteja cuidadosamente de quienes la obtienen, podría perjudicar a su propietario. Por lo tanto, la protección de la seguridad es fundamental para el uso masivo de información en los sistemas de información de la información (ITS). Para evitar la exposición no autorizada de datos confidenciales individuales, los gobiernos deben establecer regulaciones integrales de seguridad de la información que integren la información que se puede difundir, el grado de transmisión y el uso de las normas básicas de difusión, transparencia y diversas zonas [50]. Las oficinas de transporte deben controlar cuidadosamente la definición de información individual, fortalecer la organización de la insistencia en la seguridad de la información y utilizar mediciones adicionales para mejorar el nivel de seguridad de la información.

5.3 Almacenamiento de información

Actualmente, la capacidad de información ha aumentado de Terabytes a Petabytes, por lo que el avance en la capacidad de almacenamiento de información está muy lejos de la mejora de la información. En particular, en ITS, se generará una recopilación de información de los diferentes sensores de forma constante. Las bases de datos y los instrumentos de almacenamiento de información convencionales no han podido adaptarse a la innegable cantidad y complejidad de datos masivos [51]. Por lo tanto, es fundamental planificar la información más sensata.

La ingeniería de almacenamiento se ha convertido en un gran desafío.

Los principales proveedores de almacenamiento distribuido abierto, como Google y Microsoft, siguen mejorando sus servicios de forma coordinada. Las enormes capacidades de datos, el almacenamiento multidistribuido y el almacenamiento híbrido se están consolidando como áreas clave para el almacenamiento masivo de datos. Las capacidades de análisis estadístico ofrecen condiciones favorables para diversos tipos de tareas de análisis de datos. Además, integrar la información con la capacidad también es una buena solución. Las empresas buscan herramientas de gestión inteligentes que puedan proporcionar análisis integrados dentro de la capacidad. Esto les permite gestionar la monitorización de recursos y utilizar la base de datos.

5.4 Procesamiento de datos

La practicidad es vital para las aplicaciones de big data en ITS. Esta aplicación incorpora el preprocesamiento de datos de tráfico, el reconocimiento del estado del tráfico, el control continuo del tráfico, la dirección dinámica del curso y la planificación continua del transporte. Los datos de tráfico, que contienen diversas configuraciones de diferentes fuentes, deben contrastarse y, a continuación, los datos auténticos deben procesarse en un plazo breve [52]. El sistema de procesamiento de datos debe ser capaz de procesar progresivamente.

Datos enredados y progresivamente extendidos, paso a paso

Las instrucciones para garantizar la practicidad del procedimiento, y debido a la gran cantidad de información disponible, constituyen una prueba importante. Varios sistemas de información masivos comunes, que gestionan recursos de información actuales, como Spark y Samza,

Flink, Kafka Streams y Apache Storm han aparecido recientemente. Además, se han desarrollado sistemas dedicados al manejo de big data para ITS, como una plataforma para el control del tráfico continuo y la evaluación de la velocidad normal y los tramos congestionados de una autopista. Estas estructuras de entrenamiento ofrecen excelentes soluciones para el manejo continuo de datos.

5.5 Apertura de datos

Para que los clientes de la administración del transporte y los diseñadores de aplicaciones puedan descubrir y reutilizar datos con éxito, estos deben registrarse y estar disponibles gratuitamente con la máxima calidad. La excelencia de la información se refiere a...

Consistencia en el cumplimiento de la precisión y calidad inquebrantable [53][54].

Sin una excelente calidad de información, la información masiva desviará la dinámica e incluso producirá resultados desastrosos. En cualquier caso, la información emergente con una excelencia extraordinaria puede requerir tiempo y dinero. Existe un intercambio rápido de información clave que requiere poco o ningún esfuerzo y permite acceder a datos de primera calidad a un costo considerable, lo que convierte la recopilación de datos de alta calidad en un desafío aún mayor. Las soluciones viables incluyen la captura automática de datos o, potencialmente, el uso de razonamiento artificial para analizar los datos.

Además, las divisiones de transporte deben tener información sobre el procedimiento de embarque establecido para garantizar datos impecables y exactos.

6. CONCLUSIÓN

Los sistemas de vehículos públicos son determinantes básicos de la satisfacción personal. Varias naciones enfrentan dificultades para enriquecer la idea de sus sistemas de vehículos abiertos. Desarrollar estrategias eficientes es fundamental para la estabilidad de los sistemas de vehículos. El informe del Consejo Mundial de Transporte [55] demostró que el desarrollo y la accesibilidad de una gran cantidad de información en el sector del transporte podrían impulsar estrategias innovadoras, importantes conjuntos de datos y mejoras operativas del tráfico. En concreto, el desarrollo de diferentes tipos de tecnologías, en particular la creciente digitalización de los sistemas de infraestructura de transporte, podría aumentar la calidad de la información de los expertos en evaluación y la rentabilidad del aumento. Además, la accesibilidad de diferentes tipos de datos en el sector del transporte permite a los líderes recopilar y triangular diferentes tipos de evidencia de forma progresiva o casi continua para tomar mejores decisiones. Es previsible que la cantidad de información relacionada con el transporte aumente debido al creciente movimiento urbano y a la globalización, así como al avance general de las personas y los productos, lo que conlleva un aumento en la proporción de tráfico en las grandes ciudades. Los vehículos generan más información a partir de PDA y sus acompañantes.

Transpondedores. Las técnicas tradicionales para asuntos sociales y el análisis de esta información para generar encuentros significativos serán un desafío. Nuevos tipos de adquisición, asociación y planificación de información son necesarios para que los representantes de las regiones urbanas y los coordinadores municipales incrementen la comprensión del apoyo y contribuyan a la dinámica. La información y el análisis a gran escala ofrecen diversas oportunidades para iluminar el progreso en cuanto a estructuras de transporte viables. La gran cantidad de información de transporte recopilada por una estructura de información importante permite diversos tipos de análisis de información, incluyendo videos con visión de futuro persistentes grabados y análisis visuales e fotográficos. El resultado de este análisis de información puede ayudar a la planificación urbana a brindar asistencia constante en las rutas y mejorar las precauciones de tráfico. Además, los vehículos asociados amplían la increíble posibilidad de impulsar el análisis de diversidad y el uso de las estadísticas de transporte.

Los problemas identificados con la utilización de grandes volúmenes de información incluyen la gestión de diversas fuentes de información y el límite de transmisión de correspondencia obligatoria, la gestión de la veracidad intrínseca de la información de transporte mediante diferentes estructuras de límites y la garantía de que los seguros y la seguridad de la información se regulan adecuadamente en diferentes sistemas.

REFERENCIAS

- [1]. Vasant Dhar, "Ciencia de Datos y Predicción", Comunicaciones de la ACM, vol. 56, n.º 12, págs. 64-73, diciembre de 2013.
- [2]. Gema Bello-Orgaz, Jason J. Jung, David Camacho, "Big Data Social: Logros Recientes y Nuevos Desafíos", Information. Fusion, vol. 28, págs. 45-59, 2016.
- [3]. SH An, BH Lee y DR Shin, "Un estudio de los sistemas de transporte inteligente", Conferencia internacional sobre inteligencia computacional, sistemas de comunicación y redes, IEEE Computer Society, págs. 332-337, 2011.
- [4]. NE El Faouzi, H. Leung, A. Kurian, "Fusión de datos en sistemas de transporte inteligentes: avances y desafíos - Una encuesta", Information Fusion, vol. 12, págs. 4-10, 2011.
- [5]. Junping Zhang, Fei-Yue Wang, Kunfeng Wang, Wei-Hua Lin, Xin Xu y Cheng Chen, "Sistemas de transporte inteligentes basados en datos: una encuesta", IEEE Transactions on Intelligent Transportation Systems, vol. 12, No. 4, págs. 1624-1639, diciembre de 2011.
- [6]. Matei Zaharia, Mosharaf Chowdhury, Tathagata Das, Ankur Dave, Justin Ma, Murphy McCauley, Michael J. Franklin, Scott Shenker e Ion Stoica "Análisis rápido, e interactivo de datos de Hadoop con Spark" Sistemas en red; login, vol. 37, No. 4, pp. 45-51, agosto 2012.
- [7]. Alex Neilson, Indratmo, Ben Daniel, Stevanus Tjandra. "Revisión sistemática de la literatura sobre Big Data en el Dominio del transporte: conceptos y aplicaciones", Big Data Research, vol. 17, págs. 35-44, 2019.
- [8]. Cristian Chilipirea, Andreea -Cristina Petre, Loredana-Marsilia Groza, Ciprian Dobre, Florin Pop, "Una arquitectura integrada para estudios futuros en procesamiento de datos para ciudades inteligentes", Microprocesadores Microsistemas, vol. 52, págs. 335-342, 2017.
- [9]. Davide Tosi, Stefano Marzorati, "Big Data de redes celulares: escenarios reales de movilidad para futuras ciudades inteligentes", en: Actas de la Segunda Conferencia Internacional del IEEE sobre Servicios y Aplicaciones de Computación de Big Data, págs. 131-141, 2016.
- [10]. Alexey Golubev, Ilya Chechetkin, Konstantin S. Solnushkin, Natalia Sadovnikova, Danila Parygin, Maxim Shcherbakov, "Strategway: Soluciones web para construir rutas de transporte público utilizando análisis de grandes geodatos", en: Actas de la 17.ª Conferencia internacional sobre integración de información y aplicaciones y servicios basados en la web, 2015.
- [11]. Paula Ta-Shma, Adnan Akbar, Guy Gerson-Golan, Guy Hadash, Francois Carrez, Klaus Moessner, "Una arquitectura de ingestión y análisis para IoT aplicada a casos de uso de ciudades inteligentes", IEEE Internet of Things Journal, vol. 5, n.º 2, págs. 765-774, 2018.
- [12]. Sam Aleyadeh, Sharief MA Oteafy, Hossam S. Hassanein, "Monitoreo escalable del transporte utilizando el sistema de monitoreo de carreteras con teléfonos inteligentes (SRoM)", en: Actas del 5.º Simposio de la ACM sobre desarrollo y análisis de redes y aplicaciones vehiculares inteligentes, págs. 43-50, 2015.
- [13]. Qi Shi, Mohamed Abdel-Aty, "Aplicaciones de big data en la operación del tráfico en tiempo real y el monitoreo y mejora de la seguridad en autopistas urbanas", Transportation Research Part C, vol. 58, págs. 380-394, 2015.
- [14]. Alessandro Attanasi, Edmondo Silvestri, Pietro Meschini, Guido Gentile, "Aplicaciones del mundo real que utilizan técnicas de computación paralela en la asignación dinámica de tráfico y búsqueda de ruta más corta", en: Actas de la 18.ª Conferencia Internacional IEEE sobre Sistemas de Transporte Inteligente, págs. 1-6, 2015.
- [15]. Jerome Treboux, Antonio J. Jara, Luc Dufour, Dominique Genoud, "Un modelo predictivo basado en datos para la previsión de atascos de tráfico en un banco de pruebas a escala de ciudad inteligente de Santander", en: Actas de los talleres de la Conferencia de comunicaciones y redes inalámbricas del IEEE (WCNC), págs. 64-68, 2015.
- [16]. Eyad Abdullah, AAhmed Emam, "Analizador de accidentes de tráfico utilizando Big Data", en: Actas de la Conferencia Internacional sobre Ciencias Computacionales y Inteligencia Computacional, CSCI, pp.392-397, 2015.
- [17]. Masahiko Itoh, Daisako Yokoyama, Masashi Toyoda, Yoshimitsu Tomita, Satoshi Kawamura, Masaru Kitsuregawa, "Fusión visual de big data de megaciudades: una aplicación al análisis de datos de tráfico y tuits de pasajeros del metro", en: Actas de la Conferencia Internacional IEEE sobre Big Data, págs. 431-440, 2014.

- [18]. Wei Yuan, Pan Deng, Tarik Taleb, Jiafu Wan, Chaofan Bi, "Un modelo de identificación de taxis sin licencia basado en análisis de big data", IEEE Transactions on Intelligent Systems de Transporte, vol.17, No. 6, pp. 1703–1713, junio de 2016.
- [19]. P. St-Aubin, N. Saunier, L. Miranda-Moreno, "Análisis de seguridad vial proactivo automatizado a gran escala utilizando datos de video", Transportation Research Part C, vol. 58, págs. 363–379, 2015, <https://dx.doi.org/10.1016/j.trc.2015.04.007>.
- [20]. N. Zenina y A. Borisov, "Análisis de regresión para la evaluación de la generación de viajes de transporte", Information Technology and Management Science, vol. 16, págs. 89–94, 2013, doi: 10.2478/itms-2013-0014.
- [21]. J. R. Quinlan, "Inducción de árboles de decisión", Machine Learning., vol. 1, n.º 1, págs. 81–106, 1986.
- [22]. J. Abellan, G. Lopez y J. De Ofia, "Análisis de la gravedad de accidentes de tráfico utilizando reglas de decisión mediante árboles de decisión", Expert Systems with Applications, vol. 40, n.º 15, pp. 6047–6054, 2013.
- [23]. El Vlahogianni, MG Karlaftis y JC Golias, "Redes neuronales optimizadas y metaoptimizadas para la predicción del flujo de tráfico a corto plazo: un enfoque genético", Transportation Research Part C: Emerging Technologies, vol. 13, n.º 3, págs. 211–234, 2005.
- [24]. RPD Nath, H.-J. Lee, NK Chowdhury y J.-W. Chang, "Agrupamiento de K-medias modificado para la predicción del tiempo de viaje basado en datos históricos de tráfico", en Actas de la Conferencia internacional sobre sistemas de información e ingeniería inteligentes basados en el conocimiento, págs. 511–521, 2010.
- [25]. B. Abdulhai, R. Pringle y GJ Karakoulas, "Aprendizaje de refuerzo para un control de señales de tráfico verdaderamente adaptativo", Journal of Transportation Engineering, vol. 129, n.º 3, págs. 278–285, 2003.
- [26]. I. Arel, C. Liu, T. Urbanik y AG Kohls, "Sistema multiagente basado en aprendizaje de refuerzo para el control de señales de tráfico de red", IET Intelligent Transport Systems, vol. 4, n.º 2, págs. 128–135, 2010.
- [27]. Ana LC Bazzan, "Oportunidades para sistemas multiagente y aprendizaje por refuerzo multiagente en el control de tráfico", Agentes autónomos y sistemas multiagente, vol. 18, págs. 342–375, junio de 2009.
- [28]. L. Li, Y. Lv y F.-Y. Wang, "Sincronización de señales de tráfico mediante aprendizaje de refuerzo profundo", IEEE/CAA J. Automatic Sinica, vol. 3, n.º 3, págs. 247–254, julio de 2016.
- [29]. J Manyika, M Chui, B Brown, J Bughin, R Dobbs, C Roxburgs, AH Byers, "Big Data: La próxima frontera para la innovación, la competencia y la productividad", McKinsey Global Institute, mayo de 2011, URL: www.mckinsey.com/.
- [30]. D. Michie, DJ Spiegelhalter y CC Taylor, "Aprendizaje automático: clasificación neuronal y estadística", 1994.
- [31]. Jie Chen, Arjun K. Gupta, "Análisis de puntos de cambio estadísticos paramétricos: con aplicaciones a la genética, la medicina y las finanzas", segunda edición, Birkhäuser. Junio de 2011.
- [32]. X. Meng et al., "MLlib: Aprendizaje automático en Apache Spark", Journal of Machine Learning Research, vol. 17, págs. 1–7, 2016.
- [33]. "Informe sobre la situación mundial de la seguridad vial", Organización Mundial de la Salud Organización, págs. 1–20, 2018.
- [34]. Q. Shi, M. Abdel-Aty, "Aplicaciones de Big Data en la operación del tráfico en tiempo real y el monitoreo y mejora de la seguridad en autopistas urbanas", Transport Research Part C, vol. 58, págs. 380–394, 2015.
- [35]. G. Xiong, F. Zhu, H. Fan, X. Dong, W. Kang y T. Teng, "Nuevos sistemas de transporte inteligentes basados en macrodatos recopilados desde el espacio, el aire y la tierra", Big Data and Smart Service Systems, págs. 115–137, 2017.
- [36]. R. Li, C. Jiang, F. Zhu y X. Chen, "Pronóstico de datos de flujo de tráfico basado en la teoría de conjuntos difusos de tipo intervalo 2", IEEE/CAA Journal of Automatic Sinica, vol. 3, n.º 2, págs. 141–148, abril de 2016.
- [37]. S. Jeon, B. Hong, "Pronóstico de velocidad de tráfico basado en simulación de Monte Carlo utilizando macrodatos históricos", Future Generation Computer Systems, vol. 65, págs. 182–195, 2016.
- [38]. C.-C. Lu, X. Zhou y K. Zhang, "Estimación dinámica del flujo de demanda origen-destino en condiciones de tráfico congestionado", Transportation Research, Parte C, vol. 34, págs. 16–37, 2013.
- [39]. L. Alexander, S. Jiang, M. Murga y MC González, "Viajes de origen a destino según propósito y hora del día inferidos a partir de datos de teléfonos móviles", Transportation Research Part C, vol. 58, págs. 240–250, 2015.
- [40]. [En línea]. Disponible: <http://inrix.com/mobile-apps/>
- [41]. [En línea]. Disponible: <https://www.waze.com/>.
- [42]. Hong-En Lin, R. Zito y MAP Taylor, "Una revisión de la predicción del tiempo de viaje en transporte y logística". Actas de la Sociedad de Asia Oriental para Estudios de Transporte, vol. 5, págs. 1433–1448, 2005.
- [43]. W. Weijermars y E. van Berkum. "Análisis de patrones de flujo de autopistas mediante análisis de conglomerados". Actas de la 8.ª Conferencia Internacional del IEEE sobre ITS. Viena, págs. 831–836, junio de 2005.
- [44]. J.-G. Lee, J. Han y K.-Y. Whang, "Agrupamiento de trayectorias: un marco de partición y agrupación", Actas de la Conferencia Internacional ACM SIGMOD sobre Gestión de Datos. Pekín, junio de 2007.
- [45]. M. Fiosins, J. Fiosina, JP Müller y J. Görmer, "Toma de decisiones integrada basada en agentes para vehículos autónomos en el tráfico urbano", Avances en computación inteligente y blanda, 2011.
- [46]. M. Fiosins, J. Fiosina, JP Müller y J. Görmer, "Reconciliación de la toma de decisiones estratégicas y tácticas en la simulación orientada a agentes de vehículos en el tráfico urbano", Actas de la 4.ª Conferencia Internacional ICST sobre herramientas y técnicas de simulación (SimuTools'2011), págs. 144–151.
- [47]. J. Fiosina y M. Fiosins, "Selección del itinerario más corto en una red de movilidad distribuida basada en la nube", Computación distribuida y computación artificial

- Inteligencia, Avances en Sistemas Inteligentes y Computación (AISC), vol. 217, págs. 103–110, 2013.
- [48]. M. Fiosins, "Enrutamiento descentralizado estocástico de flujos de vehículos no divisibles mediante optimización de restricciones", Computación distribuida e inteligencia artificial. Avances en sistemas inteligentes y computación (AISC), vol. 217, págs. 37-44, 2013.
- [49]. M. Smith, C. Szongott, B. Henne y G. von Voigt, "Big Data Privacy Issues in Public Social Media", en Actas de la Conferencia Internacional IEEE sobre Tecnología de Ecosistemas Digitales, págs. 1–6, junio de 2012.
- [50]. Q. Wang, C. Wang, K. Ren, W. Lou y J. Li, "Habilitación de la auditariedad pública y la dinámica de datos para la seguridad del almacenamiento en la computación en la nube", IEEE Transaction on Parallel and Distributed Systems, vol. 22, n.º 5, págs. 847–859, mayo de 2011.
- [51]. O. Tene y J. Polonetsky, "Big Data para todos: Privacidad y control del usuario en la era de la analítica", Northwestern Journal of Technology and Intellectual Property, vol. 11, n.º 5, págs. 240-273, abril de 2013, artículo n.º 1.
- [52]. M. Hilbert y P. Lopez, "La capacidad tecnológica del mundo para almacenar, comunicar y computar información", Science, vol. 332, núm. 60, págs. 60–65, abril de 2011.
- [53]. MD Assuncao, RN Calheiros, S. Bianchi, MAS Netto y R. Buyya, "Big Data Computing y nubes: tendencias y direcciones futuras", J. Parallel Distrib. Comput., vols. 79–80, págs. 3–15, 2015.
- [54]. J. Liu, J. Li, W. Li y J. Wu, "Replanteando el Big Data: una revisión sobre la calidad y el uso de los datos", ISPRS Journal of Photogrammetry Remote Sensing, vol. 115, págs. 134-142, 2016.
- [55]. "Big Data y transporte: comprensión y evaluación de opciones", Foro Internacional de Transporte, OCDE/ITF 2015, www.Internationaltransportforum.org.