Universidad del Bío-Bío. Sistema de Bibliotecas - Chile

UNIVERSIDAD DEL BÍO-BÍO

FACULTAD DE INGENIERÍA

DEPARTAMENTO DE INGENIERÍA CIVIL Y AMBIENTAL

Profesor Patrocinante: Patricio Álvarez Mendoza PhD.

"METODOLOGIA PARA LA CARACTERIZACIÓN DE LA CONGESTIÓN NO RECURRENTE EN BASE A REGISTROS DE TRAYECTORIAS DE TRANSPORTE PÚBLICO"

Proyecto de Título presentado en conformidad a los requisitos para obtener el Título de Ingeniero Civil

ÁLVARO ANDRÉS OLIVARES AVENDAÑO

Concepción, Julio 2014.

Universidad del Bío-Bío. Sistema de Bibliotecas - Chile

A mi familia por el apoyo incondicional

y por enseñarme a ser lo que soy.

Universidad del Bío-Bío. Sistema de Bibliotecas - Chile

TABLA DE CONTENIDOS

Universidad del Bío-Bío. Sistema de Bibliotecas - Chile

Universidad del Bío-Bío. Sistema de Bibliotecas - Chile

INDICE DE FIGURAS.

Universidad del Bío-Bío. Sistema de Bibliotecas - Chile

INDICE DE TABLAS.

Universidad del Bío-Bío. Sistema de Bibliotecas - Chile

METODOLOGÍA PARA LA CARACTERIZACIÓN DE LA CONGESTIÓN NO RECURRENTE EN BASE A REGISTROS DE TRAYECTORIAS DE TRANSPORTE PÚBLICO.

Álvaro Olivares Avendaño
Departamento de Ingeniería Civil y Ambiental, Universidad del Bío.Bío.
oli_vares23hotmail.com

Patricio Álvarez M.MSc.PhD.
Departamento de Ingeniería Civil y Ambiental, Universidad del Bío-Bío.
palvarez@ubiobio.cl

Resumen

La congestión de tránsito tiene costos de diversa magnitud dependiendo del origen de la congestión observada, según lo cual se pueden distinguir dos tipos: congestión recurrente y congestión no recurrente. Esta última se refiere a un tipo de congestión que ocurre de manera irregular e independiente del aumento de la demanda asociada a las horas punta, y asociada a incidentes de tránsito.

A diferencia de la congestión recurrente, las características de la congestión no recurrente, esto es, frecuencia, duración y severidad, no han sido abordadas en el medio local, y por ende no existen antecedentes suficientes que permitan estimar el costo de estas externalidades en la operación del sistema de transporte. Dado lo anterior, es posible razonar que existirán ciertos umbrales a partir de los cuales es posible proponer planes de manejo de incidentes tales que el costo de su implementación sea inferior a los ahorros obtenidos producto de la operación de dichos planes, y por esta vía justificar planes de gestión de tránsito que puedan mitigar los efectos de la congestión no recurrente.

En este sentido, se plantea una metodología en base a minería de datos para la detección y caracterización de incidentes de tránsito, mediante la utilización de datos GPS contenidos en archivos del transporte público del Gran Concepción, de manera de poder generar los

Universidad del Bío-Bío. Sistema de Bibliotecas - Chile

antecedentes que permitan la cuantificación del problema de la congestión no recurrente y en consecuencia faciliten la evaluación económica de medidas de gestión de tránsito orientadas a mitigar los efectos en la operación de la red que tienen que ver con eventos inesperados o de rara ocurrencia.

Palabras Claves: Congestión no recurrente, incidentes de tránsito, clasificación de trayectorias.

7467 Palabras Texto + 8*Figuras/Tablas*250 + 1*Figuras/Tablas*500 = 9967 Palabras Totales

Universidad del Bío-Bío. Sistema de Bibliotecas - Chile

METHODOLOGY FOR THE CHARACTERIZATION OF NON RECURRING CONGESTION RECORDS BASED ON PUBLIC TRANSPORT PATHS.

Álvaro Olivares Avendaño
Department of Civil and Environmental Engineering, University of Bio-Bio.
oli_vares23hotmail.com

Patricio Álvarez M.MSc.PhD.
Department of Civil and Environmental Engineering, University of Bio-Bio.
palvarez@ubiobio.cl

Abstract.

Traffic congestion has costs of various sizes depending on the origin of the observed congestion, whereby there are two types: recurring congestion and non-recurring congestion. The latter refers to a kind of congestion that occurs irregularly, and in general, independent of the increase in demand associated with peak hours, associated with traffic incidents.

Unlike recurrent congestion, the characteristics of the non-recurring congestion, ie, frequency, duration and severity have not been addressed in the local environment, and thus there is not enough history to estimate the cost of these externalities in operation of the transportation system. Given this, it is possible to reason that there will be certain thresholds above which it is possible to propose management plans for such incidents that the cost of implementation is less than the savings obtained product of the operation of such plans, and in this way justify traffic management plans that the effects of non-recurring congestion can be mitigated.

Here, a methodology is proposed based on data mining for the detection and characterization of traffic incidents, using GPS data files contained in Gran Concepción public transport, so as to generate the background to enable the quantify the problem of non-recurring congestion and thus

Universidad del Bío-Bío. Sistema de Bibliotecas - Chile

facilitate economic evaluation of traffic management measures to mitigate the effects on network operation that deal with unexpected events or rare occurrence.

Keywords: Non-recurring congestion, traffic incidents, classifying trajectories

Universidad del Bío-Bío. Sistema de Bibliotecas - Chile

1. INTRODUCCIÓN

La congestión de tránsito ha ido en aumento en gran parte del mundo, y todo indica que seguirá agravándose, constituyendo un peligro que se cierne sobre la calidad de vida en la ciudad. Su principal manifestación es la progresiva reducción de las velocidades de circulación, que se traduce en incrementos de tiempos de viaje, tiempos de viaje menos confiables, aumento del consumo de combustibles, y otros costos de operación y polución atmosférica. Se distinguen básicamente dos tipos de congestión: congestión recurrente y congestión no recurrente.

La congestión no recurrente se refiere a un tipo de congestión que ocurre de manera irregular, generalmente asociada a eventos como accidentes de tránsito, reparaciones en las vías, malas prácticas de conducción, etc. En general estos eventos reducen la capacidad del sistema de transporte y suceden independientemente del aumento de la demanda asociada a las horas punta (congestión recurrente).

A diferencia de la congestión recurrente, las características de la congestión no recurrente, esto es, frecuencia, duración y severidad, no han sido abordadas en el medio local, y por ende no existen antecedentes suficientes que permitan estimar el costo de estas externalidades en la operación del sistema de transporte. Dado lo anterior, es posible razonar que existirán ciertos umbrales a partir de los cuales es posible proponer planes de manejo de incidentes tales que el costo de su implementación sea inferior a los ahorros obtenidos producto de la operación de dichos planes, y por esta vía justificar planes de gestión de tránsito que puedan mitigar los efectos de la congestión no recurrente.

En este sentido, en este proyecto se plantea una metodología para la detección y caracterización de incidentes de tránsito, mediante la utilización de datos GPS contenidos en archivos del transporte público del Gran Concepción, de manera de poder generar los antecedentes que permitan la cuantificación del problema de la congestión no recurrente y en consecuencia faciliten la evaluación económica de medidas de gestión de tránsito orientadas a mitigar los efectos en la operación de la red que tienen que ver con eventos inesperados o de rara ocurrencia.

Para ello, utilizando técnicas de minería de datos, se busca identificar patrones y anomalías en datos censados por los dispositivos GPS, a través de 2 funciones que se aplican a los datos

Universidad del Bío-Bío. Sistema de Bibliotecas - Chile

contenidos en los archivos históricos de operación. La primera función consiste en la extracción automatizada de la información según un segmento o corredor de la red vial que se desee estudiar. Estos datos constituyen el insumo de la segunda función, en la cual se categorizan (clasifican) las diferentes trayectorias que efectúan los vehículos de transporte público a través de un algoritmo denominado K-MEANS, con el fin de identificar la variabilidad de las trayectorias en función del tiempo y en particular aislar aquellas trayectorias y estudiar las posibles causas que la originan.

1.1 Objetivo General.

Proponer una metodología para la identificación y caracterización de incidentes de tránsito, en base a datos GPS del transporte público del Gran Concepción.

1.2 Objetivos Específicos.

- Describir y analizar la base de datos ELESIS que contienen registros de trayectorias de vehículos del transporte público del Gran Concepción.
- Proponer una metodología para la depuración y extracción de información desde la base de datos de ELESIS.
- Proponer una metodología para la clasificación de trayectorias extraídas desde la base de datos, de manera de identificar aquellas que potencialmente estén asociadas a eventos de congestión no recurrente.
- Aplicar ambas metodologías a una sección de la red de transporte público del Gran Concepción, de manera de identificar y caracterizar incidentes de tránsito.

Universidad del Bío-Bío. Sistema de Bibliotecas - Chile

2. MARCO TEÓRICO.

Este capítulo se orienta principalmente a la revisión de la literatura de modo de identificar los elementos que enmarcan el resto de la investigación, así como también algunos trabajos realizados por otros autores en relación al análisis y caracterización de la congestión no recurrente de tránsito.

2.1 La Congestión de Tránsito.

En los últimos años el aumento de la demanda de transporte y del volumen de tránsito han causado aumentos significativos en la congestión, y todo indica que este problema seguirá agravándose. Su principal manifestación es la progresiva reducción de las velocidades de circulación, que se traduce en incrementos de tiempos de viaje, tiempos de viaje menos confiables, aumento del consumo de combustible, polución atmosférica y aumento de otros costos de operación. Además, la lentitud de desplazamiento aumenta la frustración de los conductores y fomenta el comportamiento agresivo de ellos.

Dependiendo del origen que pueda tener la congestión de tránsito, esta se puede clasificar en dos tipos: congestión recurrente y congestión no recurrente.

La congestión recurrente, se refiere principalmente al hecho que la demanda por usar la infraestructura de transporte excede la capacidad de la misma y por ende a menudo es considerada un problema de dimensionamiento que es lógicamente combatido aumentando la capacidad del sistema. En general este tipo de congestión tiende a concentrarse en períodos cortos de tiempo, típicamente conocidos como “horas punta”. Por lo tanto este tipo de congestión se refiere a un fenómeno de carácter repetitivo y predecible, y por ende los costos asociados han sido bien estudiados, existiendo una amplia gama de estudios cuantificando los costos sobre el sistema de transporte. Comúnmente, la congestión recurrente es analizada por medio de procesos de planificación con metodologías probadas y bien conocidas en el ámbito nacional.

Por otro lado, la congestión no recurrente se refiere a un tipo de congestión que ocurre de manera irregular, generalmente asociada a eventos que reducen la capacidad del sistema de transporte y que suceden de forma independiente del aumento de la demanda en las horas punta. La

Universidad del Bío-Bío. Sistema de Bibliotecas - Chile

congestión no recurrente es el resultado de accidentes de tránsito, vehículos en "panne", malas prácticas de conducción (estacionar en doble fila, etc.), presencia de basura o elementos extraños en el sistema de transporte, actividades de mantención del sistema de transporte (bacheos, mantención de semáforos, etc.), faenas de construcción, actividad policial, clima adverso, y en general cualquier otra actividad no rutinaria en el sistema de transporte.

Las causas mencionadas anteriormente se denominan en general incidentes de tránsito y se pueden agrupar en tres categorías principales: incidentes de tránsito (50%), faenas constructivas (15%-25%) y clima adverso (10%).

Estudios previos han indicado que los incidentes son una de las principales causas de pérdida de tiempo y aumentos de costos en las redes de transporte, por ejemplo, en los Estados Unidos, se determinó que en el 2003, más del 60% de la congestión en las autopistas urbanas fue causada por incidentes, y ese indicador se estima que fue de más del 70% en 2005 (Schrank et.al., 2003).

A diferencia de la congestión recurrente, las características de la congestión no recurrente, esto es severidad (cantidad de pistas afectada), frecuencia (periodicidad de ocurrencia de incidentes) y duración (tiempo en que se ve afectada la capacidad de la vía), no han sido abordadas en el medio local y por ende no existen antecedentes suficientes que permitan estimar el costo de estas externalidades en la operación del sistema de transporte.

Dado lo anterior, es posible razonar que existirán ciertos umbrales a partir de los cuales es posible proponer planes de manejo de incidentes tales que el costo de implementación sea inferior a los ahorros obtenidos producto de la operación de dichos planes, y por esta vía mitigar los efectos de la congestión no recurrente.

2.2 Incidentes de Tránsito y su Detección.

Los incidentes de tránsito pueden ser definidos como cualquier evento que interrumpe el normal funcionamiento de la infraestructura de transporte, degradando la seguridad y reduciendo la capacidad. Estos eventos incluyen: vehículos averiados, accidentes de tránsito, actividades de mantenimiento de vías, condiciones climáticas adversas, protestas, escombros en la carretera, etc. La congestión de transito relacionada con el incidente (incluyendo los impactos secundarios) tiene efectos perjudiciales en la seguridad pública, la economía local y el medio ambiente. Por lo

Universidad del Bío-Bío. Sistema de Bibliotecas - Chile

tanto la cuantificación y caracterización de estos eventos son esenciales a la hora de la implementación de planes de manejo de estos incidentes y en consecuencia la disminución de sus efectos. El manejo de incidentes conlleva importantes beneficios, tales como la reducción de los retrasos en los tiempos de viaje de los vehículos debido a incidentes, a través de la reducción de la frecuencia de incidentes y la mejora de la respuesta y el tiempo de despacho de unidades de asistencia al incidente.

Los incidentes de tráfico reducen la capacidad disponible de una carretera o degradan su rendimiento, expresado en velocidades de operación más bajas y en una mayor congestión. También pueden aumentar la probabilidad de incidentes secundarios y una degradación del rendimiento en calles que ni siquiera están directamente influenciados por el incidente, a través de circunstancias como el conocido fenómeno rubbernecking (Cambridge Sys Inc. et al., 1998).

2.3 Estudios Previos.

En los últimos años se han implementado diversos sistemas de control de tránsito, que nacieron con el objeto de supervisar, controlar, administrar y mejorar la gestión de tránsito de un sector urbano o vial. Estos sistemas se basan en datos obtenidos mediante diversas técnicas, ya sea mediante el uso de vehículos sonda, equipados con GPS (“probe cars”), o mediante equipos de conteo automático basados en espiras.

La experiencia internacional en relación al uso de la información recolectada con este tipo de dispositivos indica que los detalles adicionales que proporcionan los datos GPS tanto en términos espaciales como temporales permiten un mejor entendimiento y representación de la operación y de las condicionantes que determinan el comportamiento de la red de transporte. En particular, los datos GPS que cubren periodos extensos de tiempo entregan la oportunidad de clasificar los días y eventos del año en diferentes patrones de comportamiento y así por ejemplo considerar que en ciertos corredores pudiese resultar relevante distinguir entre diferentes temporadas del año o reconocer de forma especial aquellos días en donde se realicen eventos especiales (Álvarez et. al., 2010). El uso de los datos GPS permite entre otras aplicaciones determinar los efectos en la operación de la red producto de variaciones de la demanda horaria, estacional, o por eventos especiales, accidentes de tránsito, trabajos en la vía, y condiciones climáticas entre otras (Cambridge Systematics, 2010; U.S. Department of Transportation, 2004; Courage et. al., 2008).

Universidad del Bío-Bío. Sistema de Bibliotecas - Chile

De igual manera, existen estudios en que mediante diversas técnicas de "datamining" o minería de datos, se utilizó información captada por dispositivos GPS para lograr evaluar los niveles de congestión de tráfico, logrando agrupar segmentos de carretera en varios niveles según la velocidad promedio de desplazamiento de los vehículos. (Ahmet can Diker et. al., 2012; Zhang Yong-Chuan, 2011).

Skabardonis et al. (2003) desarrollaron en California una metodología preliminar para cuantificar la congestión recurrente y no recurrente usando datos del Sistema de medición del desempeño de autopistas de California (PeMS) e informes de incidentes de la patrulla de carreteras de California (CHP). Usando estos datos, el estudio fue capaz de estimar ambos tipos de congestión, recurrente y no recurrente, en segmentos de carretera seleccionados, y caracterizar la congestión no recurrente como resultado de incidentes u otras causas. El estudio también encontró que la congestión no recurrente como parte de la congestión total en cualquier segmento estuvo relacionado con las características del segmento y el grado de congestión recurrente.

2.4 Minería de Datos.

Como se puede prever, la cantidad de datos almacenados mediante los dispositivos GPS suelen ser enormes, dependiendo de los días y la cantidad de vehículos de los que se tenga información. Para poder trabajar con estas cantidades de datos, e inferir información relevante a partir de ellos, es que se utiliza la minería de datos o "datamining", que puede ser definida como un conjunto de técnicas y tecnologías que permiten explorar grandes bases de datos, de manera automática o semi automática, con el objetivo de encontrar patrones repetitivos, tendencias o reglas que expliquen el comportamiento de los datos en un determinado contexto. Dentro de las técnicas de minería de datos más representativas se encuentran las técnicas de clasificación, técnicas de análisis de asociaciones y técnicas de detección de anomalías.

En particular, en este proyecto se consideró la utilización de un algoritmo de clasificación denominado k-means, cuyos fundamentos y funcionamiento relativos a la problemática de este proyecto se explican en la siguiente sección.

Universidad del Bío-Bío. Sistema de Bibliotecas - Chile

2.5 Algoritmo de Clasificación K-means.

Tal como se señaló en la introducción, uno de los objetivos de este proyecto tiene que ver con caracterizar la operación del transporte público por medio de la identificación de patrones y anomalías en la operación del sistema. Para lograr este objetivo se desarrolló una función para poder categorizar (clasificar) las diferentes trayectorias que efectúan vehículos de transporte público a lo largo del día. Esta clasificación se basa en una comparación de la similaridad que presentan las diferentes trayectorias registradas por diferentes vehículos en diferentes horas del día y en diferentes días del período de evaluación (Alpaydin, 2004). En particular, para este problema se consideró la utilización del algoritmo de clasificación k-means. Este algoritmo es un procedimiento iterativo de particionamiento, en que se divide un conjunto de n elementos, en k clusters o clases, donde los elementos pertenecientes a una misma clase se asemejan entre ellos, y a su vez se diferencian de los elementos de las clases restantes. El algoritmo comienza seleccionando k elementos aleatoriamente, los cuales representarán el centroide o media de cada clase. Luego, cada elemento restante es asignado a la clase que más se le asemeje, basándose en una medida de la distancia entre el elemento y el centroide del cluster. A continuación se recalcula el centroide de cada clase y se vuelve a asignar cada elemento a la clase de mayor similitud. El algoritmo itera hasta que los centroides no se modifiquen, y por ende, los elementos no cambien de clase. En el anexo C se ilustra este procedimiento.

Para este proyecto en particular, los elementos a clasificar corresponden a las trayectorias realizadas por los buses, descritas mediante curvas de posición versus tiempo. La distancia entre los puntos que pertenecen a una trayectoria y su centroide de clase es medida por medio de la distancia euclidiana definida mediantes la Ecuación 1 y la Ecuación 2.

$$dist((v_j, c_k) = \sum_i (v_j(t_i) - c_k(t_i))^2, \forall j \in k \qquad \text{Ec.(1)}$$

$$c_k(t_i) = \frac{1}{n_k} \sum_j v_j(t_i)^2, \forall j \in k \qquad \text{Ec.(2)}$$

Universidad del Bío-Bío. Sistema de Bibliotecas - Chile

Donde

$v_j(t_i)$ = posición j en el intervalo de tiempo i

$v_k(t_i)$ = centroide de la clase k en el intervalo de tiempo i, y

n_k = total de trayectorias en la clase k

La implementación de esta función permite especificar el número de clases que se espera como resultado del proceso, posibilitándose encontrar una clase en que se aprecien trayectorias realizadas de manera irregular o anómala con respecto al resto de las clases.

La información referente a las trayectorias a clasificar, será extraída desde archivos del transporte público del Gran Concepción, cuya base de datos se pasa a describir en el capítulo siguiente.

Universidad del Bío-Bío. Sistema de Bibliotecas - Chile

3. BASE DE DATOS ELESIS.

Los operadores de transporte público en las rutas licitadas del Gran Concepción están utilizando dispositivos de posicionamiento global (GPS) para controlar de manera más eficiente la operación, frecuencia y recaudación de su flota.

Una de las empresas que administra estos datos es ELESIS. Los datos contenidos en ELESIS son capturados por medio de dispositivos GPS que graban las coordenadas geográficas del vehículo, en un dispositivo de almacenamiento local. Este dispositivo registra de forma continua los datos proporcionados por el GPS. Posteriormente los datos se traspasan inalámbricamente una vez al día desde el dispositivo local a un servidor que almacena los datos recogidos por todos los vehículos que pertenecen a la misma línea de transporte público. La base de datos es accesible por medio de conexión WEB previa verificación de las credenciales del usuario.

En condiciones normales, la frecuencia típica de captura de datos se realiza a una tasa de 6 datos por minuto, esto es: posición global (coordenadas geográficas) y un estimado de la velocidad en Kilómetros por hora (Km/hr). Adicionalmente, en aquellos períodos donde el vehículo está detenido, el registro de datos se detiene y se reinicia una vez que el vehículo está nuevamente está en movimiento. En estos períodos donde el vehículo está detenido, el sistema registra la duración de la detención en segundos.

Los datos se hacen disponibles para el proyecto utilizando archivos .xlsx (MS EXCEL), en donde el nombre del archivo permite identificar la fecha en la cual se recogieron los datos y la máquina a la cual pertenecen, es decir, cada archivo Excel corresponde a todo el recorrido efectuado por una máquina en un día.

La tabla 1 muestra datos extraídos de un archivo Excel típico (Mayo-04-2013Maq10.xlsx) en el cual se pueden visualizar los campos disponibles para este estudio, entre ellos un índice correlativo, fecha, registro temporal, velocidad en Km/hr, tiempo en que el bus está detenido, falla en la recepción de datos “S-GPS”, latitud y longitud.

Universidad del Bío-Bío. Sistema de Bibliotecas - Chile

Tabla 1. Datos de Archivo Excel Típico ELESIS.

Índice correlativo	Fecha	Registro temporal	Velocidad (km/h)	Tiempo detenido	Falla GPS	Latitud	Longitud
1	04-01-2013	14:14:53	20			36500388	73031766
2	04-01-2013	14:15:03	26			36500018	73032026
3	04-01-2013	14:15:13	0	0:00:47		36499867	73032119
4	04-01-2013	14:16:01	9			36499789	73032167
5	04-01-2013	14:16:10			S-GPS		
6	04-01-2013	14:16:16	14			36499273	73032523
7	04-01-2013	14:16:25			S-GPS		
8	04-01-2013	14:16:29	0	0:00:05		36499222	73032468
9	04-01-2013	14:16:35	6			36499237	73032490
10	04-01-2013	14:16:45	30			36498921	73032707
11	04-01-2013	14:16:55	0	0:00:02		36498655	73032875
…	…	…	…	…	…	…	…

Una cuestión importante tiene que ver con la forma en cómo se registra la información correspondiente a diferentes tramos que un vehículo realiza en un mismo día de operación. En este sentido, se observa que los archivos de ELESIS corresponden a un registro continuo de información sin registro explícito del inicio o fin de una vuelta o tramo en específico. Esto es particularmente inconveniente considerando que un período o tramo de análisis adecuado del comportamiento del sistema, correspondería a una vuelta o un corredor de importancia dentro del Gran Concepción. Esta consideración implica que un procedimiento necesario es el desarrollo de una función capaz de separar de forma automatizada los registros de los diferentes tramos o vueltas de una máquina un mismo día.

Otra característica de ELESIS es la falta de depuración de los datos contenidos en los archivos, es decir, archivos con datos tal como son registrados por el dispositivo GPS y por ende sujetos a diversas fuentes de error sistemático y aleatorio. Entre los errores más comunes observados están:

- Períodos de distinta duración sin registro de datos, como se observa en los puntos 5 y 7 de la Tabla 1.
- Períodos en que la hora entre un punto y otro no avanza
- Intervalos de datos repetidos

Universidad del Bío-Bío. Sistema de Bibliotecas - Chile

En el caso de ELESIS, estos períodos sin registro de datos serán imputados, aplicando una regla para rellenar los vacíos de datos. Además, se remueven los registros donde exista algún vacío de información.

En la base de datos, cada archivo Excel va acompañado por un archivo kml, en los que se puede visualizar en mapa el recorrido de cada máquina utilizando Google Earth. Como se desprende de esto, y considerando la cantidad de vehículos para los cuales se cuenta con información, los datos de ELESIS constituyen un registro interesante desde el punto de vista de la densidad espacial y temporal para caracterizar el funcionamiento del sistema de transporte público.

En el Anexo A, se presentan tablas e imágenes referentes al modo en que la Base de Datos ELESIS se hace disponible para este proyecto. Además se muestran tablas que ejemplifican los tipos de error comentados en los párrafos anteriores.

Universidad del Bío-Bío. Sistema de Bibliotecas - Chile

4. METODOLOGÍA.

Uno de los objetivos del proyecto, dice relación con la caracterización de la operación del sistema de transporte público por medio de la identificación de patrones y anomalías a partir de los datos censados por los dispositivos GPS. Para lograr este objetivo se desarrollaron 2 funciones que se aplican a los datos contenidos en los archivos históricos de operación. La primera función consiste en la extracción automatizada de la información según el segmento o corredor de la red vial que se desee estudiar. Estos datos constituyen el insumo de la segunda función, en la cual se categorizan (clasifican) las diferentes trayectorias que efectúan los vehículos de transporte público, con el fin de identificar la variabilidad de las trayectorias en función del tiempo y en particular aislar aquellas trayectorias anómalas para un estudio más profundo de las posibles causas que la originan. En las secciones sub siguientes se detallan las funciones antes descritas y se muestra un caso de estudio que ilustra el potencial de la metodología.

4.1 Función Pre-Procesamiento de Datos y Extracción de Trayectorias.

En este proyecto se desarrollaron herramientas de pre procesamiento, imputación y extracción de datos que fueron posteriormente implementadas en MATLAB. Estas herramientas se aplican a los datos contenidos en el repositorio ELESIS y producen la información base para alimentar las funciones de clustering o clasificación. En este sentido, el pre procesamiento de datos consiste en leer los datos temporales y de posición desde archivos EXCEL que contienen los datos crudos de ELESIS, y luego representar la distancia acumulada en función del tiempo. Lo anterior teniendo presente un área de estudio consistente en un corredor o tramo de la red vial del Gran Concepción la cual queda definida por un punto de inicio y un punto final en la trayectoria del vehículo.

4.1.1 Determinación de Límites del Área de Estudio.

Uno de los propósitos del pre-procesamiento de datos es obtener la distancia y tiempo acumulados producto del recorrido hecho por cada bus, desde el momento en que este ingresa hasta que sale de un corredor o tramo específico. Para esto, en primera instancia se fijan los límites geográficos (en UTM) del corredor en estudio, de manera de descartar los puntos de la trayectoria del vehículo que se encuentran fuera de dicha sección. Para esto fueron

Universidad del Bío-Bío. Sistema de Bibliotecas - Chile

implementados dos tipos de límites. El primer límite corresponde a las ecuaciones de las rectas paralelas a la sección del recorrido de interés (en color rojo). El segundo límite corresponde simplemente a las latitudes o longitudes que definen el inicio y final del tramo de interés (en color azul). La definición de estos límites permite verificar si el punto reportado por el GPS está dentro o fuera del área de estudio y en consecuencia decidir si dicho punto debe ser considerado como parte de la trayectoria acumulada en función del tiempo. En este sentido, y a modo de ejemplificar este proceso, en la figura 1 se muestran los límites trazados para trayectorias realizadas en la Calle San Martín, donde se marca el inicio y fin de las trayectorias a analizar, además de descartar puntos pertecientes a tramos paralelos a la sección en estudio y en igual sentido, como pudiese ser Avenida Los Carrera.

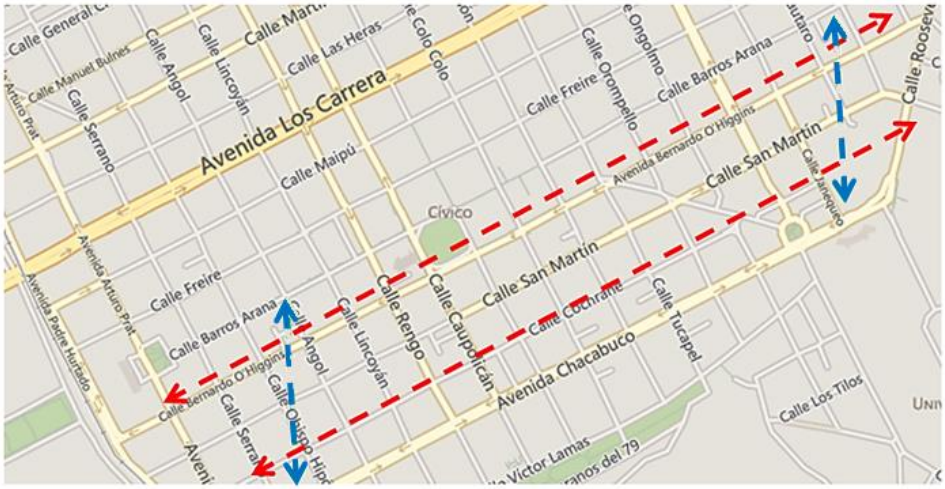


Figura 1. Visualización de Límites Trazados.

4.1.2 Imputación de Datos.

Un problema frecuente en la base de datos ELESIS está dado por la ausencia de datos o repetición de estos en algunos períodos de tiempo. Como se mencionó anteriormente el criterio para manejar este problema fue remover dichos registros, generando archivos compactos sin vacíos de información.

Universidad del Bío-Bío. Sistema de Bibliotecas - Chile

4.1.3 Conversión de Coordenadas.

Los archivos con los datos de ELESIS en su formato original no permiten el cálculo de distancias en un sentido cartesiano. Es por esto que como parte de la función de pre procesamiento se incluyó una rutina que permite la conversión del sistema coordenadas geográficas proporcionado por ELESIS, a coordenadas UTM (Universal Transverse Mercator). Esta conversión permite transformar grados sexagesimales a metros. Este proceso constó de 2 etapas, en la primera se modificó el formato de la posición (formato texto, por ejemplo 36495082, como se observa en la Tabla 1) a formato grados sexagesimales (36,82447°), para luego realizar la conversión a coordenadas UTM (666972,9450 m). Lo anterior queda bien ilustrado por el siguiente ejercicio:

36495082	>	36°49,5082'	>	36,82447°	>	666972,95 m
Sin unidad de medida		Grados, minutos decimales		Grados decimales		Metros

4.1.4 Cálculo de Distancia y Tiempo Acumulados.

Una vez que se conocen las coordenadas X,Y se obtiene el tiempo de viaje entre 2 puntos consecutivos simplemente restando el horario (time stamp) entre dichos puntos. En el caso de trayectorias que crucen los límites del área de estudio, dicho tiempo se obtiene interpolando proporcionalmente a la posición del límite. Análogamente se obtienen las distancias, utilizando para esto la fórmula clásica para la distancia euclidiana, esta vez con la latitud y longitud de 2 puntos consecutivos, e interpolando en caso que las trayectorias crucen los límites del área de estudio. Finalmente, las relaciones distancias acumuladas versus tiempo de viaje se obtienen sumando exclusivamente los tiempos y distancias incluidas dentro del corredor en estudio.

La Tabla 2 muestra el resultado de dicho proceso aplicado a una trayectoria realizada en el Puente Llacolén, donde se observa una distancia recorrida de 2034.5 metros en 98.5 segundos. Además, se observa en los puntos 865 y 874 de la columna "Δtiempo", que el intervalo de tiempo es menor al resto, producto de la interpolación aplicada según lo explicado con anterioridad.

Por otro lado, en los puntos 864 y 876, la distancia y tiempo acumulados son 0, debido a que dichos puntos se realizaron fuera de los límites del tramo en estudio.

Universidad del Bío-Bío. Sistema de Bibliotecas - Chile

Todos los datos presentados en la Tabla 2, son exportados al mismo archivo Excel del que se extraen los datos de posición y tiempo.

Tabla 2 Visualización de Distancias (m) y Tiempos Acumulados (seg.).

Punto	latitud (m)	longitud(m)	Δtiempo (s)	tiempo acumulado (s)	Δdistancia (m)	distancia acumulada (m)
…	…	…	…	…	…	…
864	672284,4	5922252,5	10	0	157,901	0
865	672141,7	5922185,4	6,82	6,820	107,482	107,482
866	671912,8	5922080,7	15	21,820	251,702	359,184
867	671737,5	5921999,2	10	31,820	193,372	552,556
868	671549,7	5921912,5	10	41,820	206,812	759,368
869	671352,1	5921823,1	10	51,820	216,934	976,302
870	671146,3	5921732,9	10	61,820	224,727	1201,029
871	670944,3	5921640,2	10	71,820	222,258	1423,287
872	670733,3	5921543,0	10	81,820	232,299	1655,585
873	670521,2	5921445,5	10	91,820	233,394	1888,979
874	670338,3	5921326,4	6,67	**98,486**	145,519	**2034,498**
876	670257,0	5921194,6	10	0,000	154,810	0
…	…	…	…	…	…	…

En algunos corredores existe doble sentido de tránsito, por lo cual se hace necesario además identificar aquellos recorridos que se hicieron en el sentido que se requiera analizar. Para ello, se extraen los datos de distancia acumulada y tiempo acumulado correspondientes a las trayectorias realizadas por el tramo en estudio, y se verifica en cada una de ellas si el sentido del desplazamiento es el que se desea. Esto se logra comprobando si se cumple un incremento constante de latitud o longitud en el sentido necesario.

El resultado de este proceso se logra con la creación de una matriz cuyas columnas contienen los datos de distancia acumulada correspondientes a todas las trayectoria realizadas dentro del corredor en estudio, y una segunda matriz cuyas columnas contienen el tiempo acumulado de cada trayectoria, asociadas a las columnas de la primera matriz. Ambas matrices son exportadas al Excel desde donde son extraídos los datos, y se presentan en la tabla B2 del Anexo B.

Universidad del Bío-Bío. Sistema de Bibliotecas - Chile

4.1.5 Agregación y Normalización de Datos.

Los datos contenidos en ELESIS son registrados en intervalos del orden de segundos, sin embargo a partir de dicho intervalo es posible construir una descripción de la distancia acumulada versus tiempo para cualquier intervalo de tiempo. Lo anterior debido a que el potencial o la efectividad de la función de clasificación dependerá del nivel de agregación de los datos. De esta forma y pese a disponer de intervalos pequeños de tiempo, el análisis se realizará con períodos que eventualmente y dependiendo de la extensión del área de estudio pudiesen ser de mayor magnitud.

Por otro lado los datos ELESIS se registran entre períodos irregulares de tiempo lo cual dificulta el ingreso de la información a la función de clasificación. Dicho problema se resuelve interpolando a partir de registros irregulares de posición, las posiciones del bus para periodos regulares de tiempo (cada 10 segundos por ejemplo).

El resultado del pre procesamiento de los datos se exporta al Excel del que se extraen los datos de posición y tiempo. Estos datos resultantes se muestran en la Tabla 3, que contiene información para individualizar el bus que originó los datos, la fecha en donde se leyeron los datos, la hora a la cual se cruza el límite donde comienza el diagrama distancia acumulada versus tiempo y finalmente el registro detallado de la posición en metros para sucesivos intervalos regulares de tiempo hasta cruzar el límite de salida del área de estudio.

Tabla 3 . Matriz de Datos Agregados.

Abril-01-2013 Maq9.xlsx						
04-01-2013	**04-01-2013**	**04-01-2013**	**04-01-2013**	**04-01-2013**	**04-01-2013**	**04-01-2013**
8:22:36	**10:26:31**	**12:28:20**	**14:45:33**	**17:03:11**	**19:20:58**	**21:38:34**
0	0	0	0	0	0	0
27,76	57,53	119,30	26,82	48,30	74,83	43,80
55,52	114,16	238,60	52,87	125,36	149,65	87,60
127,21	169,10	377,87	114,49	193,25	176,37	131,40
170,45	228,01	528,57	212,28	203,20	231,40	175,20
177,82	303,50	626,63	329,33	209,97	340,76	219,00
229,65	336,09	653,03	459,15	264,48	473,20	262,79
305,55	361,23	653,25	560,72	341,55	557,59	306,59
329,96	411,16	702,45	573,40	347,84	584,41	352,96
…	…	…	…	…	…	…

Universidad del Bío-Bío. Sistema de Bibliotecas - Chile

Cuando finaliza el pre-procesamiento, MATLAB muestra la representación gráfica de los datos de la matriz de datos agregados, como se indica en la figura 2, donde es posible apreciar las diferencias entre las trayectorias efectuadas en distintos momentos del día. En particular es posible observar las diferencias en los tiempos totales de viaje para la misma sección, cambios locales de velocidad y tendencias globales de velocidad de operación, expresadas por la pendiente de las curvas.

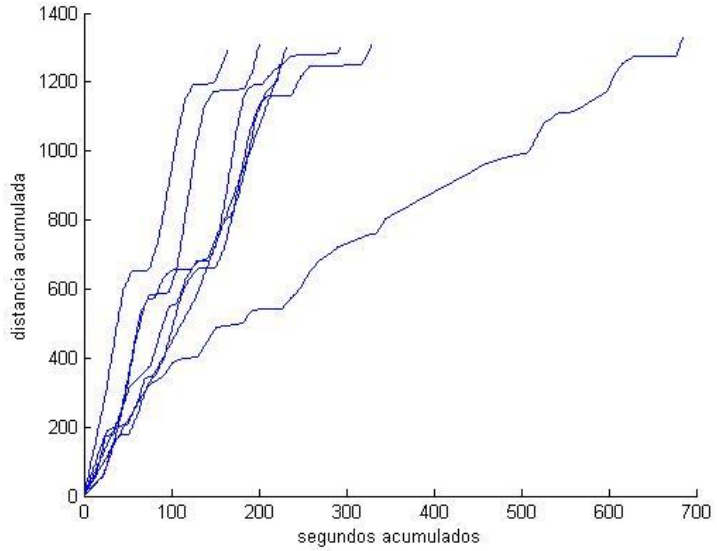


Figura 2. Trayectorias Efectuadas por un Bus, en un Día en un Corredor Determinado.

4.1.6 Síntesis de Base de Datos.

La automatización de los procesos anteriores permite procesar todos los archivos EXCEL disponibles para el análisis, para finalmente consolidar en una planilla única todas las trayectorias de todos los buses que en algún momento transitaron el área de estudio. Para esto, se unen todas las matrices de datos agregados (como la que se muestra en la Tabla 3), correspondientes a cada archivo EXCEL procesado.

Esta nueva matriz resumen, constituye el insumo básico para la función de clasificación, donde cada columna contiene la información necesaria correspondiente a cada trayectoria realizada

Universidad del Bío-Bío. Sistema de Bibliotecas - Chile

dentro de un tramo que se desee estudiar. Un ejemplo de esta matriz se presenta en la Tabla B2 del Anexo B.

4.2 Función de Identificación y Clasificación de Patrones en Trayectorias.

Como ya se mencionó, el objetivo final de esta metodología tiene que ver con caracterizar la operación del transporte público por medio de la identificación de patrones y anomalías en el sistema. Para lograr este objetivo se desarrolló una función para poder categorizar (clasificar) las diferentes trayectorias que efectúan vehículos de transporte público a lo largo del día dentro de un corredor en particular. El objetivo de esta clasificación busca separar las trayectorias que se efectúan en condiciones normales de operación, es decir condiciones que se observan sistemáticamente todos los días en los distintos momentos del día, de aquellas que por una u otra razón se realizan en condiciones diferenciables de las anteriores. La lógica de la separación propuesta tiene que ver con la ocurrencia de incidentes o eventos no típicos asociables a eventos de congestión no recurrente. Aún más se propone que por medio del análisis de la relación distancia acumulada tiempo acumulado de diversos vehículos afectados por el mismo evento se puede inferir información del mismo, por ejemplo: la duración e intensidad de dicho evento. Usando la misma lógica, y por medio del análisis de muchos eventos, es posible determinar la frecuencia de estos eventos no recurrentes.

La función de clasificación se basa en la comparación de un índice de similaridad que puede ser calculado a partir del diagrama posición versus tiempo para las diferentes trayectorias registradas por diferentes vehículos en diferentes horas del día y en diferentes días de un período de análisis dado. En particular, para este problema se consideró la utilización del algoritmo de clasificación k-means (Alpaydin, 2004). Este algoritmo utiliza un procedimiento iterativo de particionamiento, en donde se minimiza la suma de las distancias de los puntos que pertenecen a una misma trayectoria respecto del centroide de su clase; al mismo tiempo que las distancias entre centroides se maximiza.

4.2.1 *Especificación de Número de Clases.*

La implementación de esta función permite al analista especificar el número de clases o clústers (k) que espera como resultado del proceso, es decir en cuántas clases se desea dividir las trayectorias incluidas en la matriz de datos. Por supuesto, a mayor número de clases, estas más

Universidad del Bío-Bío. Sistema de Bibliotecas - Chile

homogéneas son, sin embargo un número elevado de clases no es de interés dado que el objetivo es ayudar a identificar un número limitado de patrones en el sistema estudiado.

Posteriormente, el algoritmo k-means, procede a realizar la clasificación de datos según el número de clústers definido por el usuario entregando en pantalla una descripción gráfica de los patrones encontrados. De la misma forma se exportan a EXCEL la clase de cada trayectoria analizada. Esto último permite posteriormente identificar la fecha y hora en que realizaron las trayectorias que pertenecen a las clases de interés.

El proceso de la identificación de clases o clústers cuyas trayectorias se diferencian notoriamente del resto, es claramente un proceso iterativo que depende de la habilidad del analista para manejar la rutina de clasificación. En consecuencia, si dentro de los clústers resultantes no es posible distinguir alguno que agrupe trayectorias claramente diferenciables de la tendencia, se hace necesario repetir el proceso de clasificación desde el comienzo, aumentando o disminuyendo el número de clases o clústers según los defina el analista.

4.1.2 Análisis de Posibles Incidentes Detectados.

Una vez identificado un clúster cuyas trayectorias presenten características de posibles incidentes de tránsito, se procede a individualizar y analizar los días y períodos de tiempo asociados a las trayectorias en dicho clúster. Para ello basta revisar la marca de clase que la rutina de clasificación registra en el archivo EXCEL con los datos consolidados originales. En dicha sección de la planilla, a cada trayectoria se asigna un número que representa la clase o clúster a la que pertenece.

Con esto es posible analizar en detalle las trayectorias que son candidatas a incidente, y revertir el análisis a partir de la fecha y hora de duración para determinar otras trayectorias que eventualmente pudiesen contener información del incidente y que pueden haber quedado relevadas a un clúster de mayor duración o menor duración. Esto se debe principalmente a que posiblemente estos buses se vieron afectados durante menor cantidad de tiempo por el incidente de ese período, pero de igual manera presentan retrasos con respecto a las condiciones normales.

Además, si el tramo en estudio tuviese doble sentido de tránsito, entonces de manera complementaria es posible hacer un análisis del comportamiento del sentido contrario en el

Universidad del Bío-Bío. Sistema de Bibliotecas - Chile

mismo horario del posibles incidente, de manera de verificar si hay efectos producto en la operación debido a la reducción de velocidad de los usuarios para poder observar lo que sucede en el sentido contrario de circulación.

4.1.3 Estimación de Factores Característicos de la Congestión no Recurrente.

La severidad corresponde al porcentaje de la capacidad afectada o a la cantidad de pistas que pudiesen bloquearse producto de la ocurrencia de un incidente de tránsito. La frecuencia corresponde a la periodicidad con que se produce determinado tipo de evento. Por último, la duración corresponde al tiempo que dura un incidente, desde que este se produce hasta que la capacidad del corredor vuelve a la normalidad.

Con el fin de estimar la duración del incidente, se propone utilizar las ecuaciones (3) y (4) las cuales provienen del análisis clásico de teoría de colas y que entregan la duración media de tiempo en cola y la duración máxima de un vehículo en cola (May, 1990). En el Anexo D se presentan los fundamentos de ambas ecuaciones.

$$\bar{d}_R = \frac{30 t_R (\lambda - \mu_R)}{\lambda} \quad Ec.(3) \qquad , \qquad d_M = \frac{60 t_R (\lambda - \mu_R)}{\lambda} \quad Ec.(4)$$

Donde

t_R = Duración del incidente (horas)

λ = Demanda (Veq/h)

μ_R = Capacidad reducida debido a la presencia de un incidente (Veq/h)

$\bar{d}_R$ = Duración promedio de cada vehículo en cola (minutos)

d_M = Duración máxima de un vehículo en cola (minutos)

Universidad del Bío-Bío. Sistema de Bibliotecas - Chile

A modo de ilustrar gráficamente los factores que caracterizan a la congestión no recurrente, se presenta la Figura 3, donde se grafican 3 trayectorias realizadas durante un incidente de tránsito, y a modo de comparación, una trayectoria realizada en condiciones normales (trayectoria 4).

En primer lugar, la frecuencia de incidentes estará dada por la proporción que haya entre la cantidad de trayectorias con características de incidente, por sobre el total de trayectorias que se generen.

Por otro lado, la severidad de estos incidentes puede ser analizada mediante la velocidad de desplazamiento que tengan los buses, representada por la pendiente de las curvas. Si la velocidad tiende 0 durante la ocurrencia de un incidente de tránsito, se podría concluir que se está ante un bloqueo total de pistas, mientras que si la velocidad solo se viera reducida, se estaría ante un bloqueo parcial de pistas.

Finalmente, para analizar la duración de los incidentes de tránsito, es posible extraer desde estos gráficos la duración máxima (d_M) y duración promedio ($\bar{d}_R$) de los vehículos en cola producto del incidente, siendo ambos datos necesarios para el uso de las ecuaciones 3 y 4. Por una parte, d_M correspondería a d_1, ya que es la mayor duración entre las 3 trayectorias, mientras que $\bar{d}_R$ correspondería al promedio entre d_1, d_2 y d_3.

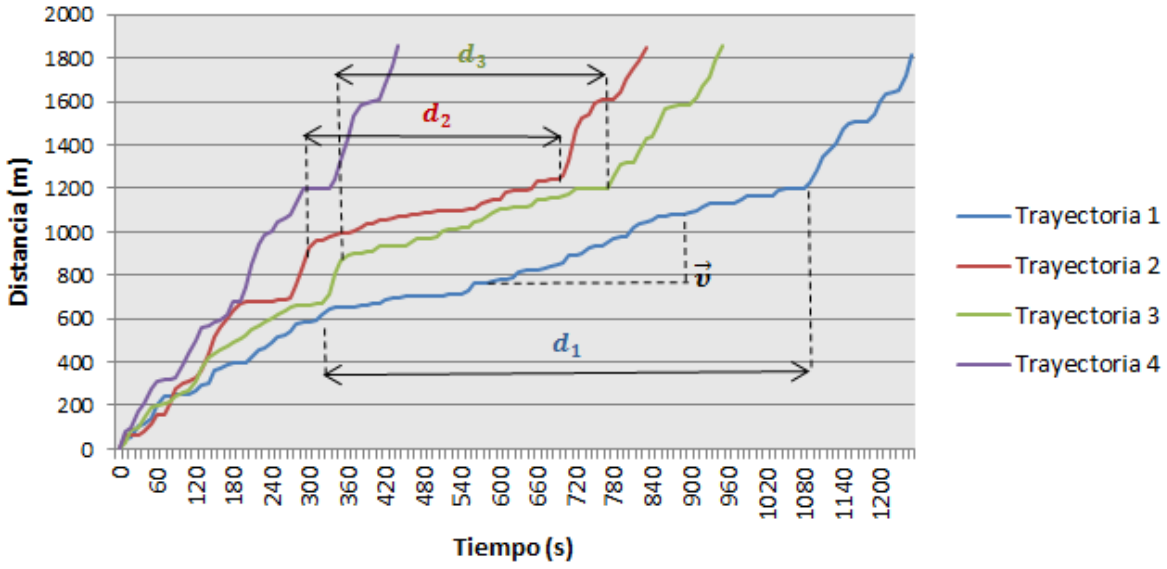


Figura 3. Trayectorias con Características de Congestión No Recurrente.

Universidad del Bío-Bío. Sistema de Bibliotecas - Chile

4.3 Alcances de la Metodología.

Los algoritmos desarrollados en este proyecto, resultan efectivos para la identificación y caracterización de incidentes, sin embargo se ven limitados frente a ciertas situaciones, las que se mencionan a continuación.

- Solo es posible analizar tramos relativamente rectos, en los que la dirección de desplazamiento sea constante, como pudiese ser Avenida Los Carrera, puente Llacolén, camino a Chiguayante, Avenida Paicaví, Avenida Collao, entre otras.

- Las distancias calculadas en este proyecto se basan en un sistema coordenado plano, en el que no es considerada la altura. Por lo tanto las distancias calculadas en tramos largos con pendiente elevada serán erróneas en cierta medida. Sin embargo, aquello no debiese afectar la detección de posibles incidentes de tránsito, ya que las distancias recorridas por todos los buses debiesen verse afectadas en igual medida.

- En calles céntricas, o tramos con variadas posibilidades de desvío frente a un incidente, la confiabilidad de los valores obtenidos podría verse afectada, pues las trayectorias que no se realizan en su totalidad dentro del tramo en estudio son descartadas de manera automática.

Universidad del Bío-Bío. Sistema de Bibliotecas - Chile

5. CASO DE ESTUDIO.

De manera de ejemplificar la metodología propuesta, se analizó una muestra aleatoria de la base de datos para el puente Llacolén, en sentido desde San Pedro de la Paz hacia Concepción, cuyo procedimiento y resultados se presentan a continuación.

5.1 Base de Datos.

Para este proyecto fueron puestos a disposición una muestra aleatoria de datos correspondientes a recorridos efectuados por la línea de buses San Remo S.A. Esta muestra consta de 245 archivos Excel acompañados por sus respectivos archivos kml (Google Earth). Cabe señalar que cada archivo Excel contiene el recorrido efectuado por una máquina dentro de un día entero de operación.

Estas 245 trayectorias diarias de datos, efectuadas por distintas máquinas, están distribuidas aleatoriamente en 44 días entre los meses de enero a mayo del 2013, además del mes de junio del 2012, como se indica en la tabla 4. Téngase en cuenta que la cantidad de archivos Excel por día no es uniforme, por lo que así como hay días en que se cuenta con 10 recorridos hay otros en los que solo se tiene 1.

Tabla 4 Distribución Diaria de Datos.

Junio 2012.						
				1	2	3
4	5	6	7	8	9	10
11	12	13	14	15	16	17
18	19	20	21	22	23	24
25	26	27	28	29	30	

Marzo 2013.						
				1	2	3
4	5	6	7	8	9	10
11	12	13	14	15	16	17
18	19	20	21	22	23	24
25	26	27	28	29	30	31

Enero 2013.						
	1	2	3	4	5	6
7	8	9	10	11	12	13
14	15	16	17	18	19	20
21	22	23	24	25	26	27
28	29	30	31			

Abril 2013.						
1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28
29	30					

Febrero 2013.						
				1	2	3
4	5	6	7	8	9	10
11	12	13	14	15	16	17
18	19	20	21	22	23	24
25	26	27	28			

Mayo 2013.						
		1	2	3	4	5
6	7	8	9	10	11	12
13	14	15	16	17	18	19
20	21	22	23	24	25	26
27	28	29	30	31		

Universidad del Bío-Bío. Sistema de Bibliotecas - Chile

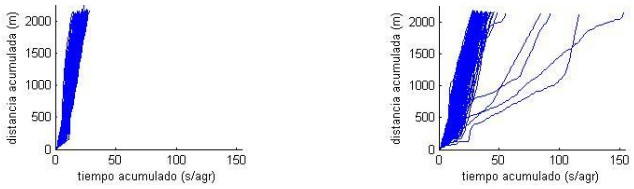
5.2 Clasificación de Trayectorias y Detección de Incidentes.

Como se ha indicado, la metodología comienza con el pre-procesamiento de la base de datos disponible. En este caso, como el corredor es relativamente corto, y los tiempos de viaje también lo son, se ha optado por realizar una agregación de datos correspondiente a 5 segundos, la mitad del intervalo promedio de captación de datos. Producto del pre-procesamiento de datos, se obtuvieron 1204 trayectorias de buses en el puente Llacolén en el sentido antes señalado.

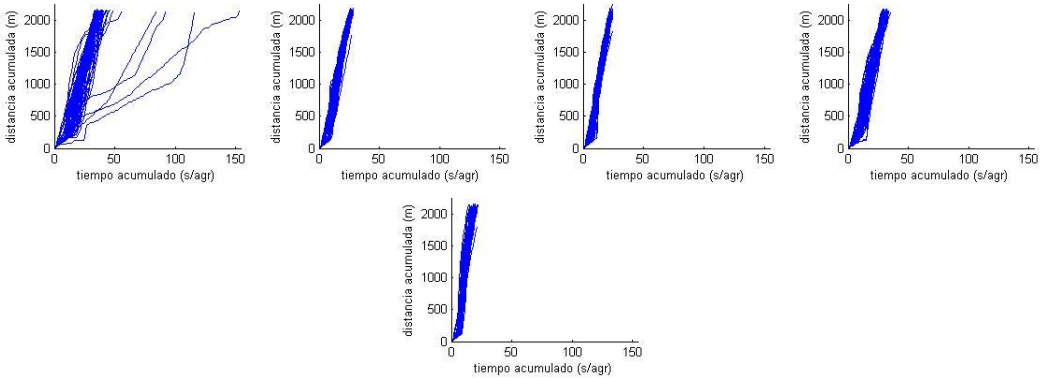
Luego, el analista puede especificar el número de clases que espera como resultado del proceso de clustering. La Figura 4 muestra los resultados de aplicar esta función a los recorridos realizados por la línea de buses San Remo por el puente Llacolén, desde San Pedro de la Paz hacia Concepción, usando diferente número de clases para la categorización. El set de datos inicial contiene registros de volumen en días laborales, fin de semana, días con incidentes de tránsito, diferentes estados del clima, eventos especiales y errores en el registro de datos. Por supuesto, a mayor número de clases, estas más homogéneas son, sin embargo un número elevado de clases no es de interés dado que el objetivo es identificar un número limitado de patrones del sistema estudiado.

Como se ve en la figura 4-a, la especificación de dos patrones no es suficiente, dado que si bien se aprecia un clúster con trayectorias de mayor duración, esta incluye eventos de congestión debido a las horas punta (congestión recurrente) e incidentes de tránsito, ya sean accidentes, protestas, mantención de la pista, etc. (congestión no recurrente). Lo mismo sucede en la primera clase de la figura 4-b. Finalmente la Figura 4-c muestra los resultados del análisis cuando se especifican 10 clases en donde como resultado del criterio elegido es posible observar 4 posibles incidentes de tránsito.

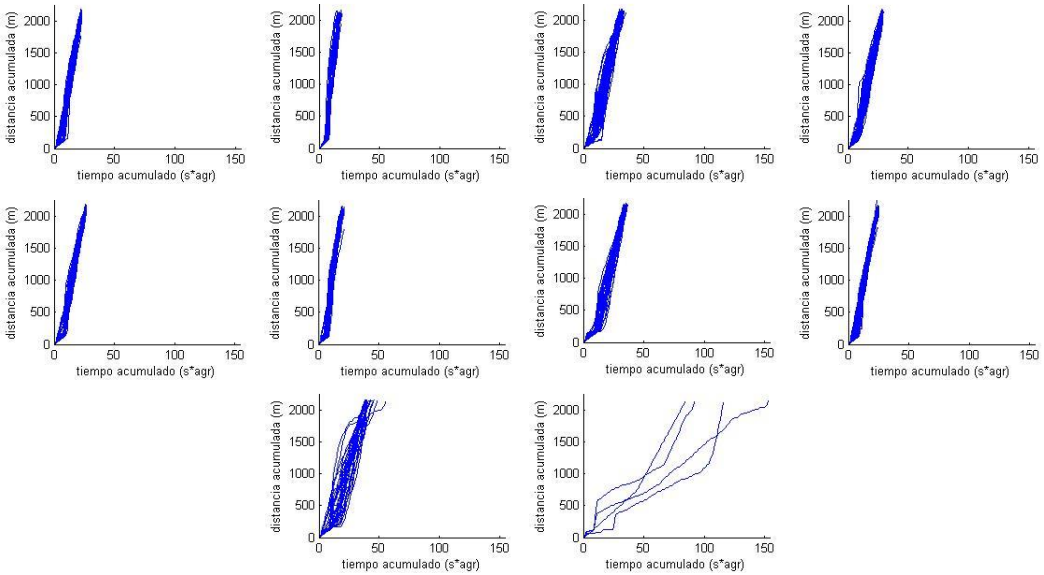
Universidad del Bío-Bío. Sistema de Bibliotecas - Chile



(a)Dos Clases



(b)Cinco Clases



(c)Diez Clases

Figura 4. Resultado de la Clasificación de Datos, Considerando 2,5 y 10 Clases.

Universidad del Bío-Bío. Sistema de Bibliotecas - Chile

Como se puede analizar de la figura anterior, las aplicaciones desarrolladas básicamente dividen los viajes según la velocidad en que estos fueron realizados, y por ende, los tiempos promedio de viajes de cada clúster o clase son similares. Es por esto, que los clúster que posean viajes de notoria mayor duración con respecto a los demás, como se aprecia en la clase 10 de la figura 4-c, eventualmente pudiesen corresponder a incidentes de tránsito. Para la corroboración de estos probables incidentes, se crea una planilla Excel en que a cada trayectoria se asocia el número de clúster al que este pertenece, determinándose que 3 de las trayectorias se produjeron el día 22 de junio del año 2012, entre las 13:00 y las 14:00 horas, todas ellas graficadas en la figura 5. Por lo cual fue descartada la cuarta trayectoria de la clase en análisis, efectuada el día 06 de Abril del 2013 a las 21:54 horas, ya que solo se tenía información de un solo bus para ese día.

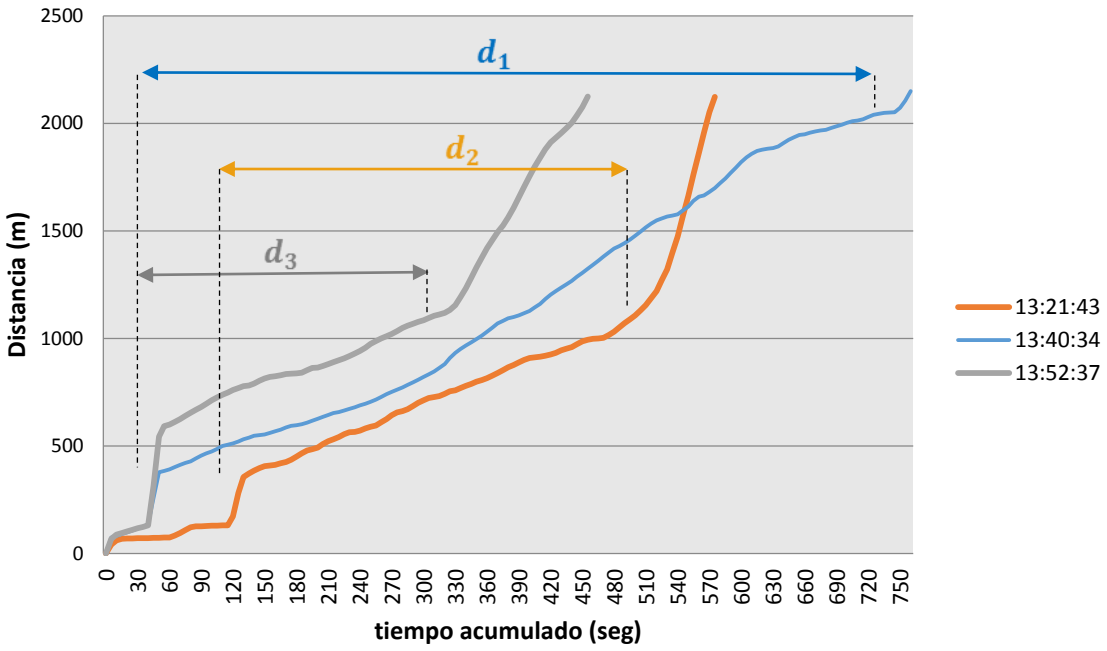


Figura 5. Trayectorias de Incidente de Tránsito en Puente Llacolén. 22 de Junio, 2012.

5.3 Caracterización de Incidentes.

Luego de tener identificados los incidentes, se procede a estimar los factores que caracterizan la congestión, que son la frecuencia, severidad y duración.

Universidad del Bío-Bío. Sistema de Bibliotecas - Chile

5.3.1 Frecuencia de Incidentes.

Como se comentó anteriormente, en la base de datos disponible se detectaron 1204 trayectorias efectuadas por buses de la línea San Remo S.A. por sobre el Puente Llacolén, de las cuales 4 correspondieron a incidentes de tránsito. Por lo tanto, un 0,33% de las trayectorias efectuadas por este puente, desde San Pedro de la Paz hacia Concepción, se ven influenciadas por el fenómeno de congestión no recurrente. Sin embargo este dato es posible de ser enriquecido significativamente si se utiliza una base de datos de tamaño superior.

5.3.2 Severidad de Incidentes.

El Puente Llacolén, posee dos pistas de circulación vehicular en cada sentido, por lo que hay 2 alternativas de bloqueo de pista, bloqueo total, o bloqueo de una pista. Al analizar las trayectorias, es posible detectar que estas jamás velocidad la velocidad es 0, por lo que la severidad del incidente detectado corresponde al bloqueo de una sola pista.

5.3.3 Duración de Incidentes.

Como se propone en este proyecto, la duración (t_R) será calculada mediante las ecuaciones 3 y 4

$$\bar{d}_R = \frac{30 t_R (\lambda - \mu_R)}{\lambda} \quad Ec.(3) \ , \qquad \bar{d}_M = \frac{60 t_R (\lambda - \mu_R)}{\lambda} \quad Ec.(4)$$

Donde $\bar{d}_R$ corresponde al promedio de tiempo en minutos en que los buses se vieron afectados por el incidente. Este valor se calculó desde el gráfico de la figura 5, como promedio entre las duraciones d_1 d_2 y d_3, resultando ser 7,5 minutos.

Por otro lado la capacidad del puente y la demanda en ese horario son 3100 y 2245 [veq/hr] (Bizama, 2012). Sin embargo es requerida la capacidad del puente producto de la severidad del incidente, la que se reduce un 35% de la capacidad en condiciones normales (HCM, 2000). Por lo tanto μ_R corresponde a 1085[veq/hr].

Finalmente, resolviendo para t_R en la Ecuación 3, con los datos presentados en los párrafos anteriores, se obtiene que la duración del incidente o la duración en que la capacidad de la vía se vio afectada por este fue de 29 minutos.

Universidad del Bío-Bío. Sistema de Bibliotecas - Chile

Utilizando la Ecuación 4 para $\bar{d}_M$, (máxima duración) asumiendo que la trayectoria de máxima duración es un buen índice que representa esta condición se considera que = 11.75 min se obtiene un valor de =22.7 min.

Universidad del Bío-Bío. Sistema de Bibliotecas - Chile

6. CONCLUSIONES.

- Producto del análisis y procesamiento de la base de datos ELESIS, se hace notoria una falta de depuración de los datos, encontrándose variados errores tales como intervalos de tiempo de diversa duración sin registro de datos, períodos sin avance de tiempo e intervalos de tiempo en que se repiten los datos.

- Una de las funciones desarrolladas en este proyecto, logra la depuración de la base de datos, generando archivos compactos sin errores ni vacíos de información. Además logra la extracción de información referente a trayectorias realizadas por algún tramo vial que se desee estudiar. De este modo la información de la base de datos queda en condiciones para la aplicación de los procesos de clasificación de trayectorias.

- El uso del algoritmo k-means es efectivo para separar los registros de trayectorias, facilitando la identificación de los potenciales candidatos a incidentes de tránsito, a través de la aislación de aquellas trayectorias realizadas de manera anómala con respecto al resto.

- La aplicación de la metodología desarrollada es efectiva para la caracterización de los parámetros de la congestión no recurrente, esto es, frecuencia, duración y severidad.

- La metodología propuesta y su aplicación, dan cuenta del potencial de información contenida en los registros históricos de la operación del sistema de transporte público. El potencial de información a extraer dice relación con información que hoy en día no está disponible, como lo es la frecuencia, duración y severidad de incidentes de tránsito, y en consecuencia su explotación es de interés para la implementación de planes de manejo de incidentes de tránsito en determinados corredores de transporte público del Gran Concepción

Universidad del Bío-Bío. Sistema de Bibliotecas - Chile

- Para la obtención de valores más representativos en cuanto a la duración, frecuencia y severidad de incidentes de tránsito, se sugiere considerar una base de datos de mayor tamaño, con registros continuos en el tiempo.

Universidad del Bío-Bío. Sistema de Bibliotecas - Chile

7. BIBLIOGRAFÍA

- Lomax T., D. Schrank, S. Turner, and R. Margiotta, (2003) Report for Selecting Travel Reliability Measures.

- Cambridge Sys Inc., H. Cohen and Science Applications International Corporation (1998) Sketch Methods for Estimating Incident Related Impacts. Report No. DTFH61- 95-00060: 21. Office of Environment and Planning, Federal Highway Administration, Washington, D. C.

- Alvarez, P., Hadi, M., Zhan, C. (2010) , Using Intelligent Transportation Systems Data Archives for Traffic Simulation Applications, Journal of the Transportation Research Board (in press), Washington, DC.

- Cambridge Systematics, Texas A&M University, Dowling Associates, Street Smarts, H. Levinson and H. Rakha. (2010) Analytical Procedures for Determining the Impacts of Reliability Mitigation Strategies.

- Federal Highway Transportation (2004), Archived Data Management Systems - A Cross-Cutting Study. Publication FHWA-JPO-05-044. FHWA, U.S. Department of Transportation.

- Courage, K.G. and S. Lee.(2008) Development of a Central Data Warehouse for Statewide ITS and Transportation Data in Florida: Phase II Proof of Concept. Florida Department of Transportation.

- Ahmet can Diker, Elvin Nasibov (2012) Estimation of Traffic Congestion Level via FN-DBSCAN Algorithm by uaing GPA Data.

Universidad del Bío-Bío. Sistema de Bibliotecas - Chile

- Zhang Yong-chuanm, Zuo Xiao-qing, Zhang li-ting, Chen Zhen-ting. (2011). Traffic Congestion Detection Based On GPS Floating-Car Data. Procedi Engineering 15 (2011) 5541-5546.

- Skabardonis Alexander, Varaiya Pravin, Petty Karl. (2002) Measuring Recurrent and Non-Recurrent Traffic Congestion. Washington, D.C.

- Thammasak Thianniwet, Satidchoke Phosaard, Wasan Pattara-atikom. (2009) Classification of Road Traffic Congestion Levels from GPS Data using a Decision Tree Algorithm and Sliding Windows.

- Alpaydin, E. (2004) Introduction to Machine Learning. The MIT Press, Massachusetts.

- May, Adolf D. (1990) Traffic Flow Fundamentals. Prentice-Hall, Inc. Englewood Cliffs, New Jersey.

Universidad del Bío-Bío. Sistema de Bibliotecas - Chile

ANEXO A.

Como se comentó en el capítulo 3, la Base de Datos ELESIS es presentada en 2 formatos complementarios entre sí. El primero consiste en archivos ".xlsx" (MS EXCEL) y el segundo en archivos ".kml". Cada par de estos archivos contienen información del recorrido efectuado por cada máquina en un día.

En este anexo, se presentan imágenes de la visualización del formato ".kml", y de la visualización del formato "xlsx", además de errores en la captura de datos en este último.

Archivos ".xlsx".

La tabla A1 muestra datos extraídos de un archivo Excel típico, denominado en este caso "Abril-01-2013 Maq9.xlsx", en el cual se pueden visualizar los campos disponibles para este estudio, entre ellos un índice correlativo (A), fecha (B), registro temporal (C), velocidad en Km/hr (D), tiempo en que el bus está detenido (E), falla en la recepción de datos "S-GPS" (F), latitud (G) y longitud (H).

Como se puede observar, tanto latitud y longitud son presentados sin formato mediante una sucesión de números, donde los primeros 2 dígitos corresponden a los grados en formato sexagesimal, los siguientes 2 dígitos a minutos, y el resto a minutos decimales.

Tabla A1. Visualización de datos Excel.

A	B	C	D	E	F	G	H
01	04-01-2013	7:21:19	26			36495915	73074580
02	04-01-2013	7:21:30	11			36496065	73074147
03	04-01-2013	7:21:40	0	0:00:05		36496061	73074128
04	04-01-2013	7:21:46	5			36496068	73074128
05	04-01-2013	7:21:56	20			36496170	73073919
06	04-01-2013	7:22:07			S-GPS		
07	04-01-2013	7:22:16	11			36496448	73073254
08	04-01-2013	7:22:23			S-GPS		
09	04-01-2013	7:22:51	0	0:00:26		36496822	73072312
…	…	…	…	…	…	…	…
2897	04-01-2013	22:06:22	31			36492872	73028655

Universidad del Bío-Bío. Sistema de Bibliotecas - Chile

Como fue mencionado en el tercer capítulo, otra característica de ELESIS es la falta de depuración de los datos contenidos en los archivos, sujetos a diversas fuentes de error sistemático y aleatorio. Entre los errores más comunes observados, se pueden encontrar períodos cortos sin registro de datos, como se observa en los puntos 6 y 8 de la Tabla A1, como también períodos largos sin registros, en que la planilla no indica que hubo error en la captura de estos datos, como se puede observar entre los puntos 501 y 502 de la Tabla A2, donde pasaron cerca de 23 minutos sin registro alguno. Este error se hace visible cuando son graficadas las trayectorias, ya que la distancia y tiempo calculado entre ambos puntos serán desmedidamente altos en comparación a los pares de puntos sin este tipo de error.

Tabla A2. Visualización de errores de datos en Excel.

A	B	C	D	E	F	G	H
…	…	…	…	…	…	…	…
500	04-01-2013	9:08:24	48			36495351	73075985
501	04-01-2013	9:08:34	49			36495021	73076825
502	04-01-2013	9:31:44	14			36495283	73076147
503	04-01-2013	9:32:04	18			36495489	73075604
…	…	…	…	…	…	…	…
600	04-01-2013	9:32:34	25			36495914	73074487
601	04-01-2013	9:32:34	7			36496060	73074144
602	04-01-2013	9:32:54	17			36496127	73073985
…	…	…	…	…	…	…	…
670	04-01-2013	10:23:13	22			36495451	73008195
671	04-01-2013	10:23:23	0	0:00:19		36495392	73008342
672	04-01-2013	10:23:43	5			36495421	73008347
673	04-01-2013	10:23:13	22			36495451	73008195
674	04-01-2013	10:23:23	0	0:00:19		36495392	73008342
675	04-01-2013	10:23:43	5			36495421	73008347
…	…	…	…	…	…	…	…

Universidad del Bío-Bío. Sistema de Bibliotecas - Chile

Otro tipo de error, consiste en la repetición de datos en intervalos de tiempo de diversa duración, como se puede apreciar en la Tabla A2, donde el intervalo de puntos 670-672, contiene los mismos datos del intervalo 673-675. La significancia de este error, reside en que para realizar la agregación de datos, se necesita una sucesión de valores que aumenten constantemente, sin embargo la diferencia temporal entre los puntos en que el tiempo "retrocede" arroja un valor negativo, rompiendo la constancia en el aumento de dichos datos.

Archivos ".kml".

En la base de datos, cada archivo Excel va acompañado por un archivo ".kml", en los que se puede visualizar el recorrido de cada máquina utilizando Google Earth. La Figura A1 muestra puntos que el dispositivo GPS ha recogido para una máquina en su recorrido por la calle Bernardo O'Higgins, a la altura de los Tribunales de Justicia.

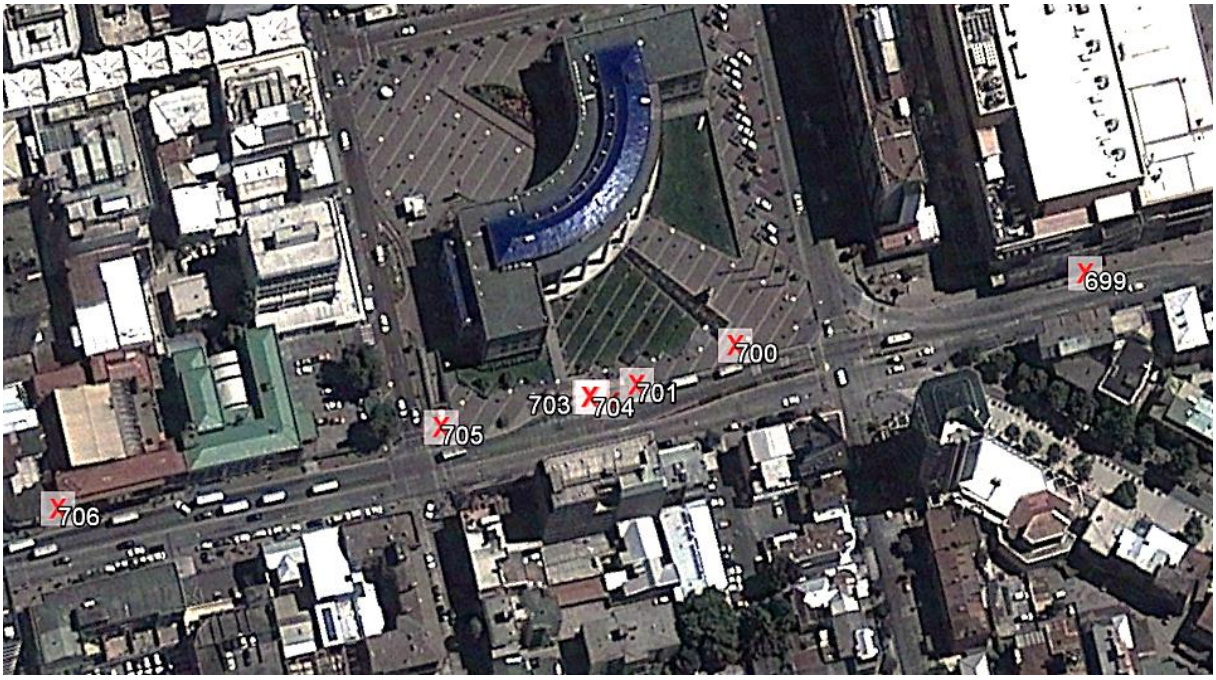


Figura A1. Visualización en de Trayectoria en Google Earth.

Universidad del Bío-Bío. Sistema de Bibliotecas - Chile

A partir de cada punto que se observa en la figura anterior, es posible extraer la información que se muestra en la figura A4, correspondiente a latitud y longitud en Coordenadas Geográficas, la hora y fecha en que los datos fueron capturados, la velocidad en Km/h, y el tiempo en que el bus estuvo detenido, si es que así lo hizo.

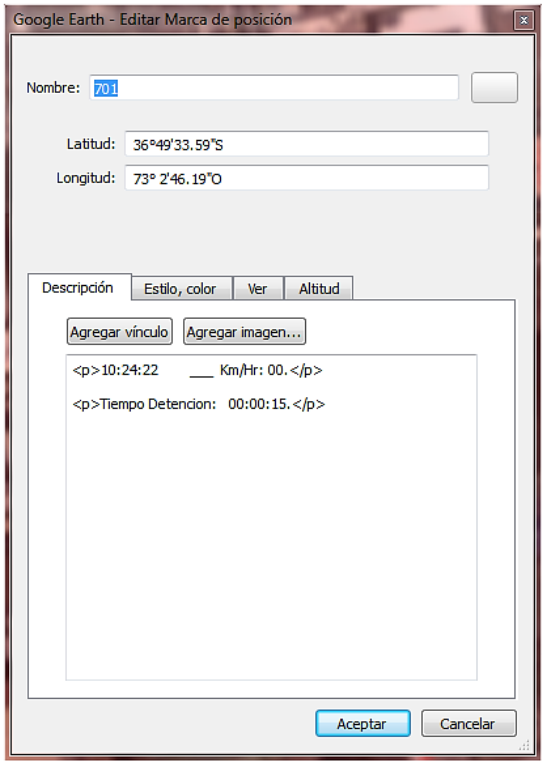


Figura A2. Información Entregada por Archivo .kml

Universidad del Bío-Bío. Sistema de Bibliotecas - Chile

ANEXO B.

En este proyecto se desarrollaron herramientas de pre procesamiento, imputación y extracción de datos que fueron posteriormente implementadas en MATLAB. Estas herramientas se aplican a los datos contenidos en el repositorio ELESIS y producen la información base para alimentar las funciones de clustering o clasificación. En este sentido, el pre procesamiento de datos consiste en leer los datos temporales y de posición desde archivos EXCEL que contienen los datos crudos de ELESIS, y luego representar la distancia acumulada en función del tiempo. Lo anterior teniendo presente un área de estudio consistente en un corredor o tramo de la red vial del Gran Concepción la cual queda definida por un punto de inicio y un punto final en la trayectoria del vehículo.

Todos los valores calculados en este proceso son exportados al Excel desde el que son extraídos los datos necesarios para el propósito de la función de pre-procesamiento de datos. En este anexo se presentan las tablas que son exportadas a cada Excel procesado, según la etapa de la metodología en la que se produzcan.

Datos exportados a Excel.

Como se comentó en la sección 3 del cuerpo principal del presente informe, cada archivo Excel de la Base de Datos ELESIS, contiene una hoja de información, desde la cual a través de programación en MATLAB se extrae la información necesaria para lograr los objetivos de este proyecto. De la misma manera, a cada planilla Excel son exportadas nuevas hojas, cada una con los resultados que van siendo obtenidos.

La Tabla B1, resume los valores calculados con la finalidad de obtener el tiempo y distancia acumulado por un bus desde que este ingresa hasta que sale de un corredor que se quiera estudiar. En ella se puede observar una trayectoria de 2159.6 metros en un tiempo de 104.37 segundos, realizada sobre el Puente Llacolén.

Universidad del Bío-Bío. Sistema de Bibliotecas - Chile

Tabla B1. Cálculo de Distancia y Tiempo Acumulado.

Punto	latitud (grados)	longitud (grados)	latitud (m)	longitud(m)	Δtiempo (s)	tiempo acumulado (s)	Δdistancia (m)	distancia acumulada (m)
…	…	…	…	…	…	…	…	…
41	-36,84	-73,09	670348,33	5921195,56	10,00	0,00	91,6	0,0
42	-36,84	-73,09	670274,86	5921209,42	10,00	0,00	74,8	0,0
43	-36,84	-73,09	670335,13	5921315,88	3,72	**3,72**	45,5	**45,5**
44	-36,84	-73,09	670403,19	5921366,13	37,00	**40,72**	84,6	**130,1**
45	-36,84	-73,08	671094,31	5921699,68	10,00	**50,72**	767,4	**897,5**
46	-36,83	-73,08	671300,67	5921794,13	10,00	**60,72**	226,9	**1124,5**
47	-36,83	-73,08	671511,09	5921890,53	10,00	**70,72**	231,4	**1355,9**
48	-36,83	-73,07	671725,12	5921989,26	10,00	**80,72**	235,7	**1591,6**
49	-36,83	-73,07	671943,66	5922089,55	10,00	**90,72**	240,5	**1832,1**
50	-36,83	-73,07	672162,95	5922190,01	10,00	**100,72**	241,2	**2073,3**
51	-36,83	-73,07	672378,75	5922287,21	3,65	**104,37**	86,4	**2159,6**
52	-36,83	-73,06	672597,42	5922385,82	10,00	0,00	239,9	0,0
53	-36,83	-73,06	672886,81	5922525,17	14,00	0,00	321,2	0,0
…	…	…	…	…	…	…	…	…

Posteriormente se extraen los valores de distancia y tiempo acumulado realizados sobre el tramo que se desee analizar, y son individualizados en 2 matrices, como se muestra en la tabla B2.

La primera matriz contiene la distancia acumulada de cada trayectoria realizada por un bus dentro de un corredor en específico, en un día determinado, mientras que la segunda matriz contiene los tiempos acumulados asociados a la primera matriz. Cada columna de las matrices corresponde a una trayectoria sobre el tramo en estudio, Además es posible observar la fecha y la hora en que fue iniciada cada trayectoria.

Universidad del Bío-Bío. Sistema de Bibliotecas - Chile

Tabla B2. Distancias y Tiempos Acumulados.

Matriz distancia (m)				Matriz tiempo (s)		
22-06-2012	22-06-2012	22-06-2012		22-06-2012	22-06-2012	22-06-2012
15:28:19	17:45:45	20:01:31		15:28:19	17:45:45	20:01:31
0	0	0		0	0	0
45,53	58,13	85,70		3,72	3,67	5,28
130,13	129,49	131,79		40,72	39,67	39,28
897,53	781,75	860,63		50,72	49,67	50,28
1124,48	965,84	1095,06		60,72	59,67	63,28
1355,92	1154,06	1284,04		70,72	69,67	73,28
1591,63	1344,37	1481,56		80,72	79,67	83,28
…	…	…		…	…	…

Luego, mediante agregación de datos de las matrices anteriores, se obtiene una matriz única correspondiente a distancias acumuladas en metros, a intervalos de tiempo constante. Esta matriz se observa en la Tabla B3.

Tabla B3. Matriz de datos agregados.

Abril-01-2013 Maq6.xlsx			
04-01-2013	04-01-2013	04-01-2013	04-01-2013
6:14:51	8:35:07	10:42:40	12:51:39
0	0	0	0
48,45	60,78	81,22	20,01
59,88	70,69	92,11	40,03
71,31	80,60	98,89	60,04
82,75	90,51	105,66	80,06
94,18	100,42	112,44	100,07
105,61	110,33	119,22	120,09
117,05	120,24	126,00	180,11
128,48	151,33	179,76	516,61
458,47	477,46	511,05	822,18
842,17	787,91	842,34	913,99
994,63	879,96	945,82	1006,22
1108,10	972,15	1035,99	1101,43
…	…	…	…

Universidad del Bío-Bío. Sistema de Bibliotecas - Chile

Finalmente, se crea una gran matriz que contiene la totalidad de trayectorias obtenidas desde la Base de Datos ELESIS realizadas por un corredor determinado, como se observa en la tabla B4. A diferencia de las tablas anteriores, esta última es exportada a una planilla Excel individual.

Tabla B4. Matriz resumen de datos.

Abril-13-2013 Maq34.xlsx			**Abril-17-2013 Maq20.xlsx**			**Abril-17-2013 ….xlsx**	
13-04-2013	**13-04-2013**	**13-04-2013**	**17-04-2013**	**17-04-2013**	**17-04-2013**	**17-04-2013**	**17-04-2013**
7:41:32	**10:15:10**	**12:39:55**	**7:08:24**	**9:07:29**	**11:09:54**	**7:28:47**	**9:26:43**
0	0	0	0	0	0	0	0
52,11	58,12	71,09	47,31	90,30	27,42	71,37	75,36
63,33	67,68	123,54	60,31	183,78	43,19	92,07	83,41
74,56	77,25	125,53	73,30	277,27	58,97	98,33	91,47
85,79	86,81	127,52	86,29	370,76	74,74	104,59	99,52
97,01	96,38	129,51	99,29	464,25	90,51	110,85	107,57
108,24	105,94	131,50	112,28	557,73	106,29	117,11	115,62
119,46	115,51	133,49	125,28	651,22	122,06	123,37	123,68
130,69	125,07	156,83	141,84	744,71	137,84	129,63	157,19
413,78	192,69	464,97	541,44	847,21	507,38	471,66	441,87
724,84	322,90	757,61	938,35	952,82	877,15	821,20	708,53
832,17	438,22	843,49	1047,41	1058,43	984,29	915,64	797,46
918,54	537,48	929,40	1156,43	1164,05	1091,26	1004,38	886,22
1004,54	638,58	1015,56	1261,72	1269,66	1202,48	1090,89	973,38
1090,51	741,69	1101,79	1367,00	1375,28	1313,70	1177,34	1060,71
…	…	…	…	…	…	…	…

Universidad del Bío-Bío. Sistema de Bibliotecas - Chile

ANEXO C

Algoritmo de Clasificación K-MEANS.

Como se mencionó en el cuerpo principal del informe, k-means es un algoritmo de clasificación, utilizado en la minería de datos para dividir grandes cantidades de datos en un número limitado k de grupos, en que las características de los elementos de un grupo o clase son similares entre sí, y a la vez diferentes de las de los otros grupos o clases.

El proceso de agrupamiento k-means es relativamente simple. Inicialmente se determina el número de grupos k y se asume el centroide o centro de esos grupos. Para determinar los centroides hay 2 alternativas prácticas: la primera es tomar de forma aleatoria k objetos como centroides iniciales y la segunda es tomar los primeros k objetos en secuencia.

Luego el algoritmo ejecuta los siguientes 3 pasos hasta que alcance el criterio de convergencia, es decir que los objetos no se muevan de grupos.

1. Se determina el o los centroides iniciales de acuerdo al número k de clases esperadas
2. Se determina la distancia de cada objeto con relación a los centroides,
3. Se agrupan los objetos con base en la distancia mínima.

A modo de ejemplo, se presenta la Figura C1 un conjunto de elementos plomos que se dividirán en 3 grupos o clases.

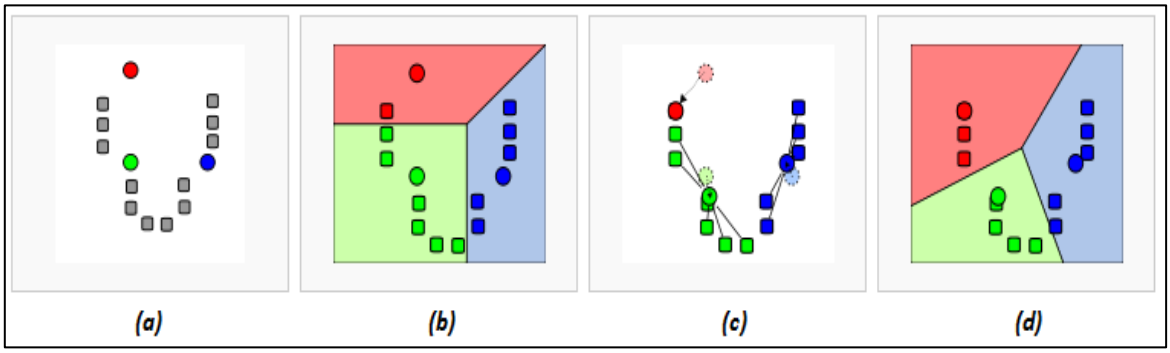


Figura C1. Proceso de ClasificaciónK-means

Universidad del Bío-Bío. Sistema de Bibliotecas - Chile

En la (a), se observan K centroides iniciales (en este caso k=3). Son generados aleatoriamente dentro del conjunto de datos (mostrados en color). Luego en la figura (b), se observan k grupos generados, mediante la asociación de cada elemento con el centroide más cercano. Posteriormente, como se indica en la figura (c), el centroide de cada uno de los grupos generados es recalculado, generando nuevos grupos como se observa en la figura (d).

Los pasos (c) y (d) se repiten hasta lograr la convergencia, en que los centroides no cambiarán de posición, ni tampoco lo harán los elementos de un grupo a otro.

Universidad del Bío-Bío. Sistema de Bibliotecas - Chile

ANEXO D

Duración de Incidentes.

Para calcular un estimado de la duración que puede tener un incidente, como se comentó en el cuerpo principal del informe, se han propuesto 2 fórmulas que provienen del análisis clásico de teoría de colas, que entregan la duración media de tiempo en cola y la duración máxima de un vehículo en cola (May, 1990).

$$\bar{d}_R = \frac{30 t_R(\lambda - \mu_R)}{\lambda} \quad , \quad \bar{d}_M = \frac{60 t_R(\lambda - \mu_R)}{\lambda}$$

Donde

t_R = Duración del incidente (horas)

λ = Demanda (Veq/h)

μ_R = Capacidad reducida debido a la presencia de un incidente (Veq/h)

$\bar{d}_R$ = Duración promedio de cada vehículo en cola (minutos)

d_M = Duración máxima de un vehículo en cola (minutos)

Ambas fórmulas son extraídas desde la Figura D.1a, donde la tasa de llegada o demanda (λ) se especifica en vehículos por hora y es constante para el período de estudio. La tasa de servicio normal (sin incidentes) o capacidad se indica en el diagrama como μ, y puesto que supera la tasa de llegada, no debiesen existir colas normalmente. Sin embargo, si ocurre un incidente se reduce la tasa de servicio a μ_R , que está por debajo de la tasa de llegada, y esta tasa de servicio inferior se mantiene durante t_R horas. Al igual que en la causa de la mayoría de las situaciones de carretera, se supone la disciplina de cola "primero en entrar, primero en salir" (FIFO).

Universidad del Bío-Bío. Sistema de Bibliotecas - Chile

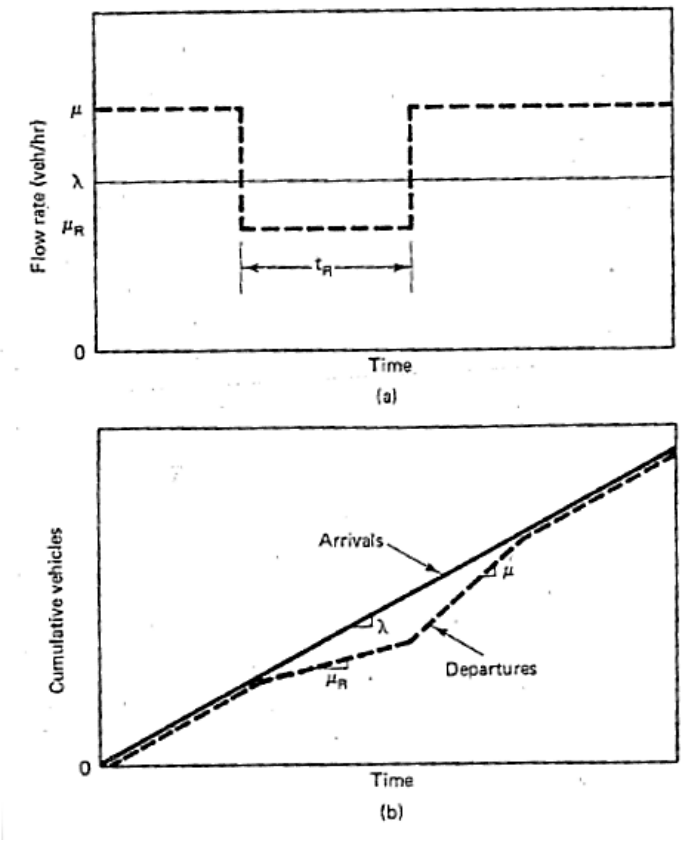


Figura D1. Diagrama de Cola en Situación de Incidentes.

En la figura D.1b, se construye un diagrama de vehículos acumulados versus tiempo. Las llegadas de vehículos se muestran como una línea recta que pasa por el origen con una pendiente hacia arriba y a la derecha equivalente a la tasa de llegada o demanda (λ). Para el primer período de tiempo la línea de servicio sigue la línea de llegada ("Arrivals") hasta que se produce el incidente. En ese punto en el tiempo la tasa de servicio se convierte en equivalente a μ_R y mantiene una pendiente más plana hasta que culmina el incidente. Entonces la tasa de servicio aumenta a μ y la línea de servicio tiene una pendiente más pronunciada. Esto continúa hasta que la línea de llegada y servicio se interceptan, momento en el que la línea de servicio, una vez más se superpone a la línea de llegada.

Universidad del Bío-Bío. Sistema de Bibliotecas - Chile

Como se puede observar, se forma un triángulo con la línea de llegada acumulada, formando la parte superior del triángulo y la línea de servicio acumulado formando los otros dos lados del triángulo. Mediante el análisis de este triángulo, es que se pueden deducir ambas fórmulas indicadas en el principio del presente anexo, que indican la duración media y duración máxima que tienen los vehículos en cola producto de un incidente de tránsito.