Accident Analysis and Prevention 50 (2013) 635–644

Contents lists available at SciVerse ScienceDirect

Accident Analysis and Prevention

journal homepage: www.elsevier.com/locate/aap

# Safety models incorporating graph theory based transit indicators

Liliana Quintero[a,1], Tarek Sayed[b,*], Mohamed M. Wahba[b,2]

[a] Department of Civil Engineering, University of British Columbia, 2002-6250 Applied Science Lane, Vancouver, BC, Canada V6T 1Z4
[b] Department of Civil Engineering, The University of British Columbia, Vancouver, BC, Canada V6T 1Z4

ARTICLE INFO

*Article history:*
Received 26 February 2012
Received in revised form 30 April 2012
Accepted 13 June 2012

*Keywords:*
Collision prediction models
Safety planning models
Safety performance functions
Network techniques
Transit indicators
Connectivity
Coverage
Directness
Complexity of transit networks

ABSTRACT

There is a considerable need for tools to enable the evaluation of the safety of transit networks at the planning stage. One interesting approach for the planning of public transportation systems is the study of networks. Network techniques involve the analysis of systems by viewing them as a graph composed of a set of vertices (nodes) and edges (links). Once the transport system is visualized as a graph, various network properties can be evaluated based on the relationships between the network elements. Several indicators can be calculated including connectivity, coverage, directness and complexity, among others. The main objective of this study is to investigate the relationship between network-based transit indicators and safety. The study develops macro-level collision prediction models that explicitly incorporate transit physical and operational elements and transit network indicators as explanatory variables. Several macro-level (zonal) collision prediction models were developed using a generalized linear regression technique, assuming a negative binomial error structure. The models were grouped into four main themes: transit infrastructure, transit network topology, transit route design, and transit performance and operations. The safety models showed that collisions were significantly associated with transit network properties such as: connectivity, coverage, overlapping degree and the Local Index of Transit Availability. As well, the models showed a significant relationship between collisions and some transit physical and operational attributes such as the number of routes, frequency of routes, bus density, length of bus and 3+ priority lanes.



## 1. Introduction

The need for tools to enable the evaluation of the safety of transit networks at the planning stage is growing. Recent research efforts include the development of techniques and decision tools which facilitate a proactive safety approach for the design of public transport networks and services. One example of these decision tools is macro-level collision prediction models which were shown to facilitate the evaluation of safety proactively at the planning stage (deLeur and Sayed, 2003; Hadayeghi et al., 2003; Lovegrove and Sayed, 2006a,b; Lovegrove et al., 2010). However, most of these collision prediction models have been developed for auto collisions only. Although, some vehicle-to-vehicle collisions could be attributed to the presence of transit vehicles (i.e. stop-and-go behavior of transit vehicles, transit vehicles blocking the view of the road for auto drivers), vehicle collision prediction models traditionally do not include transit related explanatory variables (Cheung et al., 2008) and collision prediction models accounting for transit elements are scarce in the literature. Only a small number of studies were found related to the development of transit safety models, such as Jovanis et al. (1989) and Cheung et al. (2008). Although these studies represent important efforts in associating some transit characteristics to collision occurrence, additional research in transit safety is needed to enhance the validity of existing models and add additional relevant variables.

The study of networks is an area of research under which transit related variables can be developed. Network techniques involve the analysis of systems by viewing them as a graph composed by a set of vertices (or nodes) and edges (or links). Once the transport system is visualized as a graph, various network properties can be computed and evaluated based on the relationships between the network elements (i.e. relationships between vertices and edges). In the case of transportation systems, network properties of interest include connectivity, coverage, and complexity, among others. This study describes the development of collision prediction models that explicitly incorporate both physical transit elements and transit network properties as explanatory variables. The transit network properties and the collision prediction models are developed for the Greater Vancouver Regional District (GVRD) public transportation system and its 479 traffic analysis zones located in urban areas.

* Corresponding author. Tel.: +1 604 822 4379.
*E-mail addresses:* lilianaq@civil.ubc.ca (L. Quintero), tsayed@civil.ubc.ca (T. Sayed), mwahba@civil.ubc.ca (M.M. Wahba).
[1] Tel.: +1 778 847 3860.
[2] Tel.: +1 604 822 2949.

0001-4575/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.aap.2012.06.012

## 2. Previous work

### 2.1. Graph theory based transit network indicators

Graph theory has a wide range of applications in transportation as it was developed from an urban transportation problem (Derrible and Kennedy, 2010). The basic concepts of graph theory originated in the 18th century with the solution of the "The Seven Bridges of Konigsberg" problem performed by the famous mathematician Leonard Euler. In 1962 the French mathematician Claude Berge proposed the first connectivity indicator, called the cyclomatic number "$\mu$", which represents the number of circuits or alternative paths in a graph. Garrison and Marble (1962) tried to apply graph theory principles to transportation networks and developed three indices for connectivity, labeled alpha $\alpha$, gamma $\gamma$ and beta $\beta$. Kansky (1963) worked on transportation network indicators and tried to relate them with economic development.

Gattuso and Mirello (2005) proposed new indicators termed "node range of influence" $R_i$ and network covering. They calculated several indicators for 13 metro networks (located in cities in Europe and New York). Another recent and significant contribution was by Derrible and Kennedy (2010). Their contributions included: a methodology to redraw metro networks into graphs, the creation of two indicators (directness $\tau$ and structural connectivity $\rho$) and the characterization of 33 metro networks around the world. Finally, Quintero-Cano et al. (2011) proposed a methodology to redraw bus networks into graphs and new connectivity indicators, which incorporates the influence of bus operational characteristics (i.e. frequency of routes) to estimate connectivity.

### 2.2. Collision prediction models and transit

There have been recent efforts to include safety at the planning stage. These efforts have been focused mostly on developing techniques and decision tools to facilitate a proactive safety approach (deLeur and Sayed, 2003). An important example of these techniques are macro-level collision prediction models (CPMs). However, most CPMs have been developed for vehicle collisions and do not include explanatory variables related to transit characteristics (Cheung et al., 2008). A few exceptions include the study of Jovanis et al. (1989) which used data from the metropolitan Chicago area to develop models that relate transit collisions frequency on a route to annual revenue miles, weekday average ridership, average weekday morning headway, annual revenue hours, speed of the route and bus driver attributes. The models, however, did not include variables related to geometric design or road characteristics. Another study, Cheung et al. (2008) used a data set from Toronto, Canada to develop collision prediction models using the general linear regression (GLM) technique and assuming a negative binomial error structure. The results indicated that transit collisions were positively correlated with vehicle kilometers traveled, bus kilometers traveled, arterial road kilometers, bus stop density and percentage of near sided stops. Low transit collision occurrence was associated to zones with high average posted speed and far-sided stops. As well, the results indicated that increased collision frequency (vehicle and transit) was related to increased traffic volume (AADT), transit frequency, arterial road segment length, percentage of near sided stops and presence of on street parking.

## 3. Methodology

### 3.1. Data

The data used for the models development was obtained from the Greater Vancouver Regional District (GVRD) transit network. The aggregation units were based on the 479 traffic analysis zones (TAZ) used by the GVRD in their Emme/2 transportation planning model. This data aggregation level was selected in order to develop zonal-level CPMs including explanatory variables such as transit physical elements and transit network indicators. The list of the variables used and their summary statistics are presented in Table 1. The data was extracted and compiled from three main sources:

(1) TransLink, the Greater Vancouver Regional District (GVRD) transportation authority provided geocoded files of land use, road network, and zone – census tract boundaries, as well as geocoded files for transit data (bus routes and stops) and rail data (lines and stations). Additionally, TransLink provided two data sets.
   (a) The first data set included Emme2 transportation planning model outputs, which consists of travel demand (i.e. vehicle and transit kilometers traveled) for the AM morning peak scenario for base year 1996. Similar to other planning models, the GVRD's Emme/2 model had been constructed to model only morning rush hours. Therefore, PM exposure data was not available. It is assumed that total exposure will be directly related to morning peak exposure. While this assumption can produce some errors, this was felt reasonable for the purposes of this research for two reasons: first, it was expected to wash out across zones and the region without introducing a systematic bias; and, second, it was consistent with methodology followed by many other studies (Hadayeghi et al., 2003; Lovegrove and Sayed, 2006a,b, 2007), and therefore the results would be at least comparable with other macro-level CPMs.
   (b) The second spreadsheet included: frequencies, type of vehicles and service time span information for each bus route.
(2) Census Canada (1996): provided socioeconomic data (population, employment, etc.) and mode split data for each zone from the census made in 1996.
(3) The Insurance Corporation of British Columbia (ICBC), a public automobile insurance company, provided geocoded files of collision claims in the GVRD for the years 1996–1998.

### 3.2. Characterization of a bus network using transit network indicators

The first step to characterize any transportation network is to represent it as a graph. The objective is to collect the basic measurements of the graph representation (e.g. number of vertices, number of edges, length of edges, etc.). The methodology proposed by Derrible and Kennedy (2010) to redraw metro networks into graphs was modified and applied to the case of bus networks. The methodology was also modified to allow for a zonal level analysis. Briefly, the proposed methodology considered as vertices of the graph only two types of bus stops: (1) transfer bus stops (stops where it is possible to change routes) and (2) terminal bus stops (the last stop of a bus route, where there is no possibility to transfer to another route). Two types of edges are considered: directed edges which are represented as arrows indicating the direction for the movement or flow; and undirected edges which are represented as lines with flow or movement always occurring in both directions. Edges were counted as: (1) directed edges: one directed edge equals 0.5 of an edge and (2) undirected edges: one undirected edge counts as one edge. In order to account for the specific characteristics of the bus system, directed graphs were used to represent the predominantly one-way bus routes (Quintero-Cano et al., 2011).

In order to represent the existence of route transfers by having riders walk from one bus stop to another, stops connected by such "walking links" were considered as part of the graph. Walking