**A Bayesian Network Approach to Improve Mass Transit Operations in the Developing World**

**Authors**: Offei Adarkwa (Grad Student, University of Delaware)
Nii Attoh-Okine (Professor, University of Delaware)

## Abstract

With a rapidly growing population, the city of Accra, Ghana needs transportation reforms that will ensure sustainable development. The Metro Mass Transit Company (MMT) was established by the government in 2003 to serve as a mass transportation mode that will address the negative externalities of having too many automobiles on the roads. However, the service is under-performing. The underlying factor for this is that a greater proportion of the population is not satisfied with the services being rendered. This study uses a Bayesian belief network approach to model the patronage of the service which is linked to the level of satisfaction of the population. The level of satisfaction is categorized into three namely, satisfaction with station, satisfaction with bus and satisfaction with other service characteristics. The model reveals critical policy areas such as bus safety and station orderliness that must be addressed by the government in order to boost patronage of the service and in the process ensure good performance.

**Keywords**: Developing Countries, Transport Planning, Transport Management

**List of Mathematical Terms:**

G= Bayesian Network

V (G) = ($V_1$, $V_2$, …$V_n$), $n \geq 1$ represents the finite set of variables or nodes

A (G) = finite set of links or arcs that show relationships between variables

## Introduction

Current transportation trends in Ghana indicate a heavy reliance on automobile transport. This includes private automobiles as well as the minibuses known as tro-tros. With a rapidly growing economy (Ghana, $14^{th}$ Fastest Growing Economy, GBN), projections suggest a growing middle class which will result in demand for more improved transportation services. The ideal situation will be for government to improve mass public transportation services in order to lessen the stress on the environment resulting from the use of a large number of private automobiles. However, this has not been the case as attempts at providing a better form of mass urban transportation in the country has failed due to poor management and low patronage from the general public. In order to establish and maintain successful mass public transportation companies in Ghana, there must be adequate research into the preferences of the general public and all other stakeholders in order to guide policy makers towards creating a system that is sustainable, economically viable and highly patronized by the public. This paper focuses on modeling the preferences of the Ghanaian public and service components using a Bayesian approach in order to improve the performance of the Metro Mass Transit Company (MMT) as well as the services they provide. The model will serve as a decision-making guide for policy makers.

## Background

The first public transport bus company in Ghana was the Omnibus Service Authority (OSA) which started operations in 1927 (The Establishment of MMT). Due to various reasons, the assets of the company had to be put into divestiture in 1995. Metro Mass Transit was established by the government in October 2003 as a public-private partnership with the government holding a 45% share. The MMT has not been able to meet performance targets albeit 85% of Ghanaians rely on public transport. A vast majority of the public use the local minibuses known as “trotro”. As a result, there is a low patronage for the MMT by the general public. In order to ensure that the MMT thrives, there is the need for management to make well-informed decisions regarding its operation and design. This will require the use of expert decision support systems.

Decision support systems refer to computer-based solutions that are used as a guide by management in decision making. A typical decision support system comprises of a database management component, modeling component and a user-friendly interface (Shim J.P. et al, 2002). In this study, the software that will be used is Netica from Norsys Software Corporation. Its modeling component is based on Bayes’ Rule for inference. The software is used to optimize decisions that will be made by urban transportation policy makers.

## Bayesian Networks

Bayesian Networks refer to probabilistic graphical models which are directed and acyclic. Bayesian networks are an expressive language that reveals uncertain knowledge about causal and associational relationships between variables in a system (Attoh-Okine et al, 2006). The nodes in the graph represent variables and the links represent relationships in a given system. Bayesian networks can be used for inference and as a result a decision support tool. Bayesian networks can be classified as belief networks or decision networks. The difference between the two is that decision networks have decision and utility nodes in addition to nature nodes in a typical Bayesian network. A Bayesian network is a graph G, such that

$$G= [V(G), A(G)] \qquad [1]$$

where $V(G) =(V_1, V_2,\dots V_n)$, $n\geq 1$ represents the finite set of variables or nodes

A(G) represents the finite set of links or arcs that show relationships between variables

Figure 1 below illustrates a simple Bayesian network.

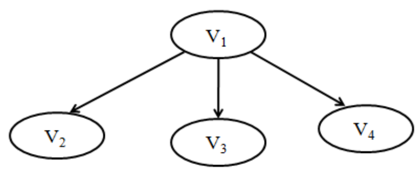


Figure 1. Simple Bayesian network

In figure 1, the node $V_1$ is the parent while $V_2$, $V_3$ and $V_4$ are known as the children. The relationships between the nodes are shown by the direction of the arrows. The parent is also known as the influencer node. Arrows originate from the parent to the child nodes also known as influenced nodes. Bayesian networks can be used for both predictive and diagnostic reasoning. Analysis made from cause to effect is more predictive while analysis from effect to cause is diagnostic (Bedeley et al, 2013).

Interestingly, Bayesian networks have been used as a decision support tool in public-private partnerships (Xie et al, 2013), improving rapid rail transit service quality (Ugurlu et al, 2011), among others. Bayesian networks are chosen as the tool for analysis since they address both the complexity and uncertainty involved in major decision making.

Bayesian networks give two levels of information namely quantitative and qualitative information. At the qualitative level, the structure of the network reveals the relationships and interdependencies of all variables in the model. At the quantitative level, the relationships between the variables are represented by conditional probabilities for each variable. There are two approaches to constructing Bayesian networks; data-based and knowledge based approach (Nadkarni S. et al, 2004). The knowledge-based approach is preferred since it uses a causal knowledge of the system from an expert to create the network. This is very similar to how humans think about problems. As such, construction of a Bayesian network usually involves first developing causal maps for the system before converting to a Bayesian network. A causal map is therefore a cognitive map representing experts' domain knowledge. Causal maps reveal cause-effect relationships underlying experts' opinions.

Ugurlu et al (Ugurlu et al, 2011) applied Bayesian Belief networks as a policy decision support tool to improve multimodal rail transit services. In the study, network coverage, ticket price, travel time, reliability, service schedule and station accessibility were identified as the indicators of travelers' willingness to make trips using transit services. Variables that affect transportation systems were also listed from extensive research. Expert opinions were used to develop causal maps for the variables after which these were converted to Bayesian networks used for the analysis. The prior probabilities and conditional probabilities used were largely based on data obtained customer satisfaction surveys. The Bayesian Networks were used to identify policy areas that have the potential to improve the transit services rendered to the public.

**Urban Transit**

There are so many factors that affect the ridership of urban transit. In a study by Harold Kohn (Kohn, 2000), he noted that transit ridership decline in Canada between 1990 to 1996 may have been due to increased suburbanization, perceived lower costs of operation of automobiles, increased fares, decreasing subsidies and an aging population. Kohn also identified factors that affect demand and supply as family size, economic changes, ridership loyalty, convenience and community size. Studies into factors that affect mode choice in urban areas using probit and structural equation models (Tyrinopoulos et al, 2013) showed that crowding, bad accessibility to transit stations and inadequate information about transit services discouraged people from using transit services.

Also, individuals between the ages of 35-44 preferred passenger cars while a larger proportion of the female population preferred public transit. This showed that mode choice was also influenced greatly by demographic and socio-economic factors. In the same study, the models indicated that the main factor affecting preference for passenger cars was availability of parking.

In a review and analysis of ridership literature, researchers identified population density, topography, freeway network extent, parking availability and cost, transit fares, safety and cleanliness as some of the factors that affected transit ridership (Taylor B. D., et al, 2003). Research based on surveys conducted in Ghana in May 2012 indicated that passengers' satisfaction with public transport was greatly affected by bus traffic safety record (Aidoo et al, 2013). In the study, the researchers identified safety, control of crime at stations and fare policies as service components that users considered the most. Binary Logistic Regression Models were used to determine the effect of passengers' satisfaction level of individual service components on the overall service quality rating.

**Methodology**

The first step in this study involved using information gathered from the literature review to identify the factors affecting the patronage of the MMT. See Table 1. In a population with 85% of the people relying on public transport, the major factor that will affect patronage of the MMT will be the level of satisfaction that the general population derives from using the MMT's services. The level of satisfaction is also dependent on various factors which can be broadly categorized into two namely; service characteristics and population preferences. Figure 2 shows the conceptual framework based on which further analysis was conducted.

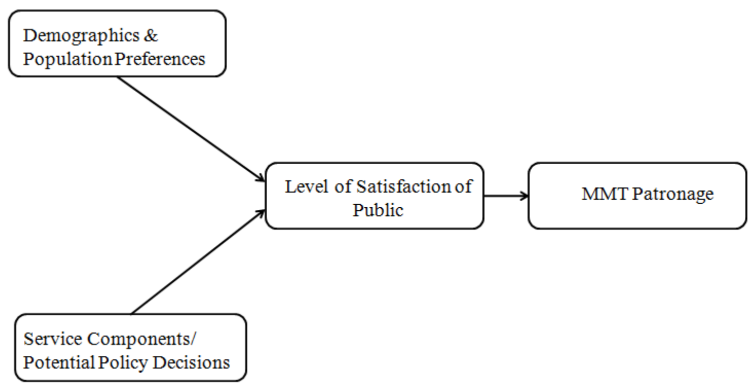


Figure 2. Conceptual Framework for Bayesian Network

Table 1. Factors affecting patronage of Urban Public Transport

| REFERENCE | FACTORS AFFECTING DEMAND FOR PUBLIC URBAN TRANSPORT |
|---|---|
| Taylor B.D., 2003 | population density, topography, freeway network extent, parking availability and cost, transit fares, safety and cleanliness |
| Aidoo E. N. et al, 2013 | Bus Traffic safety record, fare structure, waiting time for bus before departure, announcement and information on services, schedule adherence, cleanliness of station, cleanliness of bus, availability of shelters, bus comfort, availability of shelters, crime rate at station, frequency of bus breakdown |
| Kohn H.M., 2000 | Family size, economic changes, convenience, community size |
| Ugurlu et al, 2011 | Consistency to declared travel plans, travel time, waiting time, access to the station, safety and security of station and vehicles, attitudes of personnel, gender, level of income, car availability |

The framework in figure 2 illustrates that MMT patronage will reach desired levels depending on the public's level of satisfaction with the service. The public refers to the general urban population in Accra, Ghana's capital city. Service components refer to elements of the MMT service that can be changed by management in order to improve performance and patronage. The service components that were selected for this study were fare levels, network coverage, service schedule, station accessibility, station security, safety, cleanliness, bus safety, bus comfort, among others. The service components are also known as potential policy decisions since they can be altered by management to effect change in the patronage of the service. Demographics and population preferences refer to location-specific characteristics of the population. For this study, income levels and perception on the cost of using private automobiles were considered as demographic data.

The second step is to create a causal map for the patronage of the MMT system in Ghana. The causal map will reveal the cause-effect relationships between the various service components, demographics and the satisfaction level of the general public. Figure 3 shows the causal map for the study. The map shows all the variables in the study and how the parent nodes affect the child nodes. The positive sign indicates that an increase in the child node is caused by an increase in the parent node whilst the negative sign indicates otherwise. For example, an increase in bus comfort, station security and station accessibility will result in an increase in the level of satisfaction of the general population. On the other hand, it may be generally difficult to satisfy populations with higher income levels.

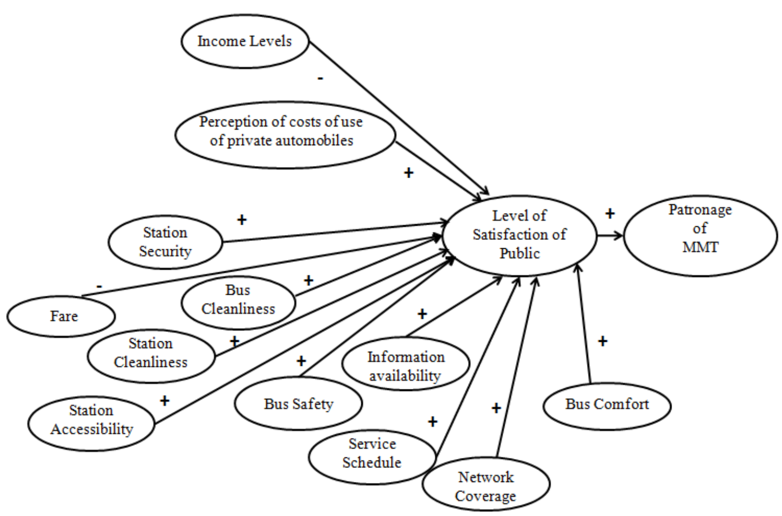


Figure 3. Causal Map for Patronage of Metro Mass Transit (MMT) System

The third step involves the modification of the causal map into a form that is compatible with Bayesian network implementation. Issues taken into consideration include eliminating circular relations, making distinctions between direct and indirect relations, conditional independencies and reasoning underlying links between variables. The network is the developed using Netica version 5.12 from Norsys Software Corporation. See figure 4. Station accessibility and station cleanliness were combined into one variable known as station orderliness. Bus comfort and cleanliness were also combined into one variable.

The fourth step is probability encoding which involves providing the probabilities for the states of the various nodes as well as filling all the necessary conditional probability tables (CPTs). Probability encoding refers to the extraction and quantification of individual judgment about uncertain quantities (Spetzler et al, 1975). It requires deriving numerical parameters for the node states in the network based on which propagation algorithms are used to make inferences. This involves identifying the state space of the variables and derivation of the conditional probabilities. The state space refers to all the possible states of a variable and these must be mutually exclusive and exhaustive of the variable (Nadkarni S., 2004). Ideally, the data must be obtained from survey results. In the absence of field data, assumptions and expert opinions about the local population were used. Belief networks are more suited for these types of problems in which a limited amount of knowledge is available. Figure 4 shows the compiled belief network using Netica.

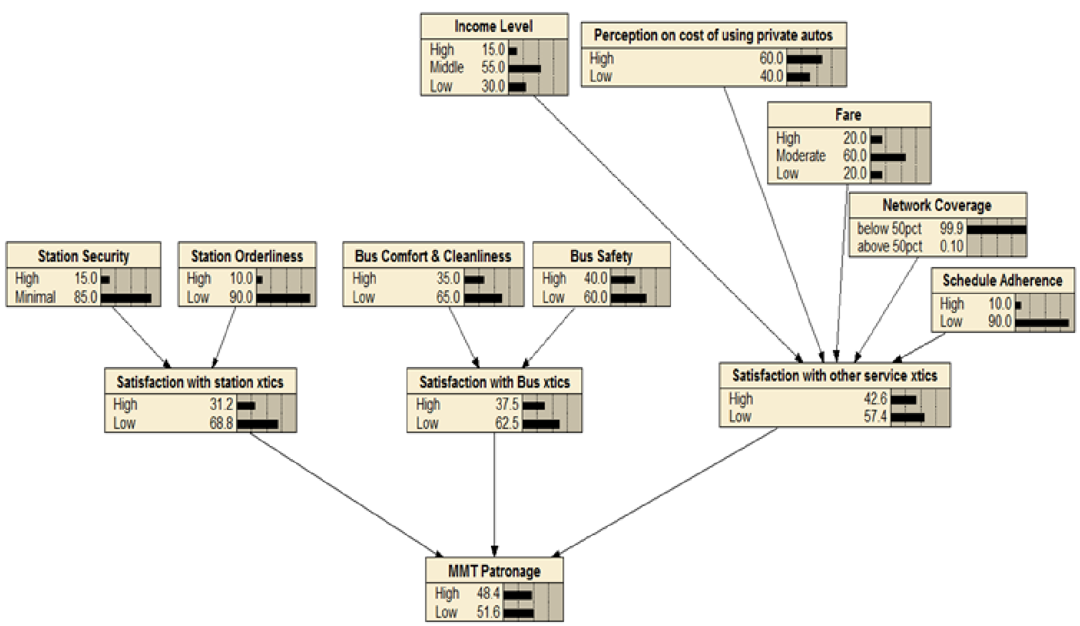


Figure 4. Compiled Bayesian Network for MMT Patronage developed in Netica.

Finally, analyses on the network are done based on posterior probabilities obtained after findings are entered. The findings used to analyze posterior probabilities, in this study, will be the policy decisions that will be made by the government to effect a change in the perception of the population on various service characteristics.

**Results**

The belief network created is based on the fact that MMT patronage in Accra, Ghana will increase when the public is satisfied with the station characteristics, bus characteristics as well as other service characteristics. After compiling, the probability of low patronage of the service is 51.6%. The probability of high patronage increases from 48.4% to 99% when high satisfactions with bus, station and service characteristics are entered as findings. See figure 5. This is to mimic a scenario when the public is pleased with all aspects of the services being offered. However, this is a hypothetical case since the government cannot afford to address all these aspects simultaneously given its financial constraints.

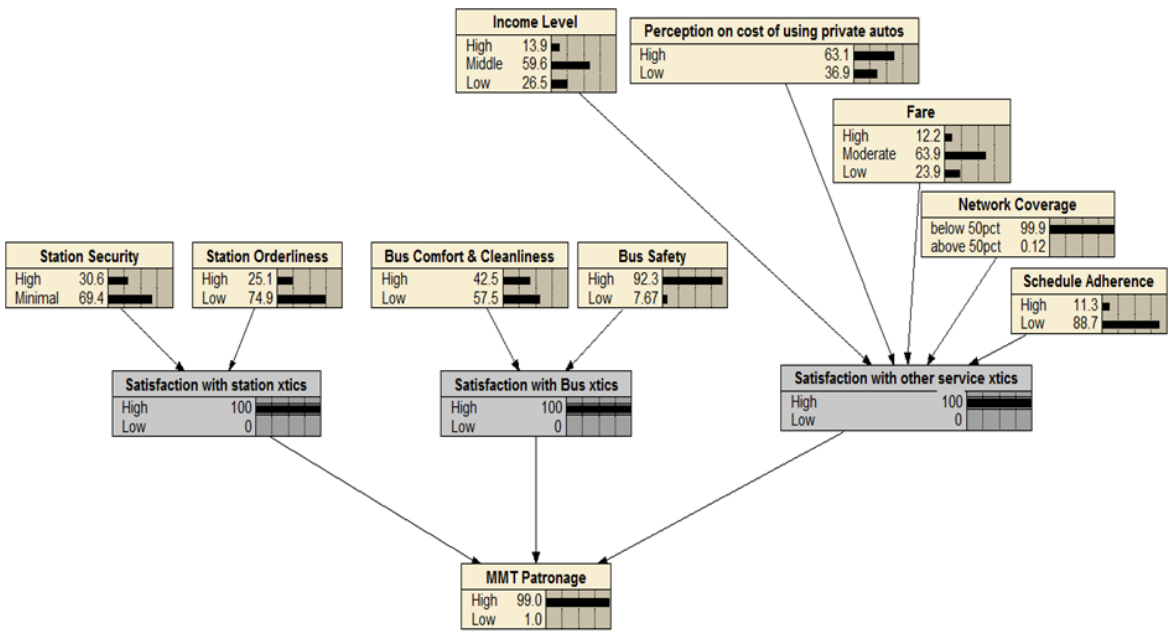


Figure 5. Posterior probabilities with high levels of satisfaction findings

A more practical approach will be to address individual satisfaction nodes which results in different posterior probabilities for patronage. Table 2 shows the increase in the probability of high patronage with a high level of satisfaction for bus characteristics, station characteristics and service characteristics.

Table 2. Change in chance of high patronage of MMT service given a high satisfaction node

| SATISFACTION NODE | % CHANGE IN PROBABILITY OF HIGH PATRONAGE |
|---|---|
| Station Characteristics | +10.6% |
| Bus Characteristics | +11.6% |
| Service Characteristics | +33.8% |

Based on the assumptions made, it was found out that the satisfaction with service characteristics which included fare structure, network coverage and schedule adherence had the most significant probability change in high patronage. This will imply that the governing agency needs to address these issues because they are likely to have a greater effect on the level of patronage of the service. The satisfaction with service characteristics was also affected by income levels and perception on the cost of use of private automobiles. Perception on the cost of private automobiles can be affected by government policies such as higher parking fees. A perception of high cost results in an increase of about 1.3% in the chance for high patronage. Table 3 also shows the new probabilities of high patronage given that policy decisions affect station security station orderliness, bus safety, bus cleanliness, fares and network coverage one at a time. Government solutions to the issue of patronage can be in the form of a combination of two or more of the policy decisions.

Table 3. Probability of high patronage given a policy decision to improving a specific service characteristic

| SERVICE CHARACTERISTIC | Probability of High Patronage (%) | Change in Probability of High Patronage (%) |
|---|---|---|
| Station Security | 53.4 | 5.0 |
| Station Orderliness | 55.7 | 7.3 |
| Bus Comfort & Cleanliness | 49.9 | 1.5 |
| Bus Safety | 57.5 | 9.1 |
| Fare (Low) | 53.3 | 4.9 |
| Network Coverage | 53.1 | 4.7 |

From the analyses, the probability of high patronage is highly sensitive to bus safety record. As such management must make decisions to ensure that will ensure good safety records for its fleet. Also, patronage is affected greatly by orderliness at the stations. This refers to cleanliness and availability of information to customers. Policies must therefore be directed at improving and maintaining station facilities.

## Concluding Remarks

MMT has the potential to solve urban transportation problems in most of the major cities in Ghana. However, this has not been possible due to a few management mishaps. In order to improve the service provided and place it ahead of its competition, a Bayesian network approach has been used to capture the uncertainties that are associated with MMT decision-making. The use of the Netica software eliminates the level of complexity that will also be encountered when dealing with a large number of variables and dependencies. The use of the belief networks will allow mass transportation officials to prioritize service characteristics that need to be addressed to ensure high patronage of the system.

This study made use of assumptions in the absence of real data. To improve the accuracy of results derived from the network, surveys must be conducted in order to obtain the perception of the public on the various factors. These surveys must be repeated periodically to ensure the network represents the current situation.

## References


(1) Aidoo E.N. et al (2013). Passenger's Satisfaction with Public Bus Transport Services in Ghana: A Case Study of Kumasi-Accra Route. *Theoretical and Empirical Researchers in Urban Management*. Volume 8 Issue 2

(2) Attoh-Okine N.O., Bowers S (2006). A Bayesian Belief Network Model of Bridge Deterioration. *Proceedings of Institution of Civil Engineers*, Bridge Engineering 159, Issue BE2, Pages 69-76

(3) Bedeley R.T., Okine N. O., Lee E (2013). Modelling pedestrian crossing behavior using Bayesian networks. *Proceedings of Institution of Civil Engineers*. Transport 166, Issue TR5

(4) Ghana is world's 14th fastest growing economy in 2013-EIU. *Ghana Business News*. See http://www.ghanabusinessnews.com/2013/10/31/ghana-is-worlds-14th-fastest-growing-economy-in-2013-eiu/ (accessed 12/12/13)

(5) Kohn, H. M (2000). Factors Affecting Urban Transit Ridership. See http://www.statcan.gc.ca/pub/53f0003x/4225128-eng.pdf. (accessed: 12/20/13)

(6) Nadkarni S., Shenoy P. (2004). A Causal Mapping Approach to Constructing Bayesian Networks. *Decision Support Systems* Vol 38, Issue 2, pp 259-281

(7) Shim J. P. et al (2002). Past, present and future of decision support technology. *Elsevier. Decision Support Systems* 931

(8) Spetzler C.S., Stael Von Holstein C. (1975). Exceptional Paper- Probability Encoding in Decision Analysis. *Management Science* 22(3):340-358.

(9) Taylor B.D., Fink C.N.Y, The Factors Influencing Transit Ridership: A Review and Analysis of the Ridership Literature, UCLA Dept. of Urban Planning, 2003

(10) The Establishment of MMT. *Metro Mass Transit Ltd*. See http://metromass.com/establish_mmt.htm (accessed 12/12/13)

(11) Tyrinopoulos Y., Antoniou C. (2013) Factors Affecting Modal Choice in Urban Mobility, *Eur. Transp. Res.* Rev. 5:27-39, DOI 10.1007/s12544-012-0088-3

(12) Ugurlu S., Aktas E., Topcu I (2011). A Decision Support System to Improve Service Quality in Multimodal Rapid Rail Systems: A Bayesian Perspective. *Proceedings of the 41st Internatonal Conference on Computers & Industrial Engineering*. P872-877

(13) Xie, J. and Thomas Ng, S. (2013). Multiobjective Bayesian Network Model for Public-Private Partnership Decision Support. *Journal of Construction Engineering and Management*. 2013.139: 1069–1081