

# Research and Implementation of the Tolling Audit Analysis Platform in Yunnan Smart Highway Project

Xiuzhen Sun

Yunnan Communications Investment & Construction Group CO., LTD, Kunming 250061, China.

**Keywords:** smart highway, big data, operation analysis, toll audit, decision support.

**Abstract.** Based on the status quo and actual requirement of Yunnan highway operation and management, this paper focuses on the solution of the comprehensive governance of the highway toll evasion and operation decision support service by means of the development of the toll audit analysis Platform in first phase of Yunnan Intelligence-expressway Project.

## 1. Introduction

Highway operation management covers all aspects of levy, monitoring, operation and maintenance, emergency command and dispatch, etc., which has always been the focus of attention. In the past ten years, Yunnan highway operation management with road network charging and network monitoring and other information projects has made considerable progress, daily operation management has entered into the preliminary orbit, and formed a three-level management mechanism including provincial operation and management center, administrative office and road section center, and achieved good results.

With the expansion of Yunnan province highway mileage, traffic flow, logistics quantity rapid increase year by year and the ETC implementation of national network, geographic scope, business management involves the daily management work of highway operation difficulty, complexity is increased rapidly, and operation management from the day-to-day supervision and inspection checks, passive response to active perception, early warning and forecast, monitoring the whole process of change direction, and cooperative ability exist for a long time in the daily management of highway departments, system integration capability is not strong, it is low for data integration in the multi-application domain, the data utilization rate is not high, which is also highlighted problem, which mainly reflected in the following aspects:

(1) It is not enough for the tolling audit, lack of the behavior pattern recognition method on tax evasion charges vehicle, audit efficiency is not high, difficult to obtain evidence, not a comprehensive audit vehicle tax evasion charges behavior.

(2) Although tolling and monitoring service system provided their self-management application, application scope is narrow and single function, video, vehicle trajectory, the charge system and et al. cannot achieve effective sharing and linkage. In the practical application of toll audit, it is impossible to form a linkage check between the charge record and the charge video and the charge image.

(3) The established toll system, monitoring system and maintenance system and so on of highway have provided massive and abundant operation data resources for managers. However, there is a lack of relevant valuable information recognition methods for behavior pattern recognition, analysis and decision support information.

Facing the above-mentioned problems, in order to adapt to the development trend of highway, during the phase one and two construction of smart highway, we have targeted the planning and construction of the highway data center, some business and operation data from the established information systems were unified into data quality management system and formed a standardized data view. Based on this, integrated information management platform consisting of the core of the analysis system of toll audit was built to solve operational data integration, information sharing, resources service, operation analysis and decision support problems, with video monitoring system improvement of the whole road network integration, to achieve multi-sectoral coordination, multi-system linkage, multi-application domain information resource sharing, big data support,

highway toll, audit operation platform fusion analysis of emergency command security, operation and maintenance of multiple business areas, to solve the vehicle tax evasion charges have long plagued the operation management work inspection management and relevant analysis and decision support of hot and difficult problems.

This paper focuses on the key technologies involved in highway toll audit and operation analysis based on "smart highway", and the design and implementation of Yunnan highway toll audit operation analysis platform based on the above technology.

## 2. Technical Researches

### 2.1 Investigations of Vehicle Tax Evasion Charges.

Analyzed from the management experience and tax evasion charges inspection work in the toll audit departments in recent years, we summarize the following typical vehicle fee evasion means:

#### Exchange Card between Two Vehicles

When the two cars enter the highway mainline, they change their cards on the way (or change license plate and cards), there will be at least one vehicle pay less fees using exchange card, such as the exchange cards between two trucks, or the exchange cards between truck and car.

#### Single Vehicle Reverse Card

There are two (or two more) pass cards to the entrance issued for a special vehicle, during the highway running, and the use of a station adjacent to the station entrance exit card, and so forth using inverted card types, common fixed line private transport vehicles

#### Escaping Charges through Fake Cards

It is common to criminal gangs who sell counterfeit cards to heavy goods vehicles by obtaining the information of highway toll pass media after obtaining the highway toll media key through illegal means.

#### Escaping Charges through Free Vehicle

It is a common mean to fake free vehicle to escape charges, such as faking Green pass vehicle.

#### Escaping Charges through Downgraded Vehicle Type

Some vehicles intended to conceal or fake downgraded vehicle type to escape charges, such as operating vehicle with ETC recharge card applied in other provinces.

### 2.2 The Key Algorithms of Vehicle Escaping Charges.

According to the actual situation, combined with the charging data, this paper put forward a kind of new data mining algorithm fitting for the vehicle running on highway with toll management and abnormal behavior characteristics through the key toll elements and the expert experience into data mining algorithm, it was proved the algorithm is quite efficiency to catch and locate escaping changing vehicle by the establishment of various types of escaping charges behavior recognition model.

1) The definition of the new data mining algorithm: Take advantage of the accumulated knowledge formed by combating tax evasion management departments at all levels in the long-term road network to acquire characteristic value and related parameter from credible training data based on expert experience, its core idea is: selecting 3-month transaction data as the training sample data and using classification algorithm as a data mining tool, the training results will overlap into characteristic value with escaping charges and abnormal behavior.

2) The algorithm flow is shown in the following diagram.

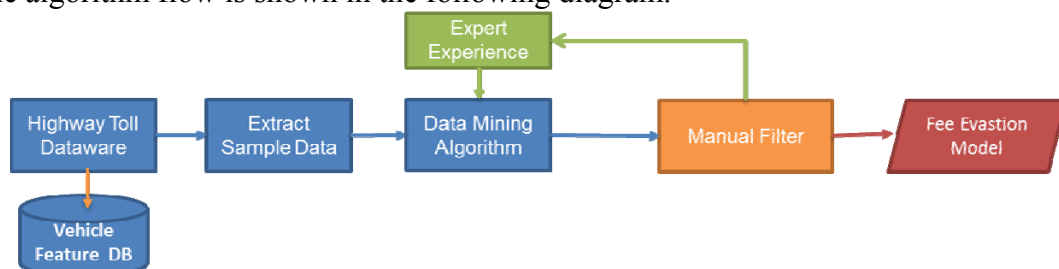


Fig 1. The algorithm flow of data mining Flow

### 3) Algorithm steps

#### Step 1: Selection of characteristic parameters

According to the charge mode on the existing road network, we choose the following objects as the characteristic parameters to research, which is mainly related to charge as follow.

Entrance toll station  $Se$ , exit toll station  $Sx$ , entrance vehicle information  $Ve$ , exit vehicle information  $Vx$ , entrance time  $Te$ , exit time  $Tx$ , exit free information  $Vkx$ , entrance license plate  $Pe$ , exit license plate  $Px$ , entrance card information  $Ce$ , exit card information  $Cx$ , vehicle load  $Wx$ , driving interval mileage  $D$  ( $Se \rightarrow Sx$ ), average travel time of driving interval mileage in the latest twenty-day period  $Tavg$  (equal to  $\sum$  travel time of each vehicle in the interval mileage/number of vehicles).

#### Step 2: Characteristic value for vehicle with abnormal behavior

Through the selected characteristic parameters, the independent characteristics of the related conditions are derived.

Inconsistent access license plate:  $F1 = \text{Diff}(Pe, Px)$ ; if  $Pe = Px$ ,  $\text{Diff}(Pe, Px)$  equals to 0, otherwise, the value is 1.

Different access card information:  $F2 = \text{Diff}(Ce, Cx)$ ; if  $Ce = Cx$ ,  $\text{Diff}(Ce, Cx)$  equals to 0, otherwise, the value is 1.

Different access to the vehicle information:  $F3 = \text{Diff}(Ve, Vx)$ ; if  $Ve = Vx$ ,  $\text{Diff}(Ve, Vx)$  equals to 0, otherwise, the value is 1.

Serious time out for vehicle:  $F4 = \text{Comp}(\lambda * D / (Tx - Te), (D / Tavg))$ ;  $\lambda \geq 3$ , if  $\lambda * D / (Tx - Te) > (D / Tavg)$ ,  $F4$  equals to 0, otherwise, the values is 1.

Vehicle collection and free change:  $F5 = \text{Diff}(V_{kx}, prev(V_{kx}))$ ; where  $prev(V_{kx})$  is the type of the last free collection, if  $V_{kx} = prev(V_{kx})$ ,  $F5$  equals to 0, otherwise equals to 1.

Vehicle load changes:  $F6 = \text{Diff}(w)$ ; where  $prev(Wx)$  is the loading of the last vehicle, if  $abs|Wx - prev(Wx)| / Wx \leq 5\%$ ,  $F6$  equals to 1, otherwise the value is 0.

Access to high-speed time frequency:  $F7 = \text{Fr}(Te - prev(Tx))$ ; where  $prev(Tx)$  is the time of the last leaving highway, if  $Te - prev(Tx) \geq 1$  hour,  $F7$  equals to 0, otherwise the value is 1.

Vehicles continuous entrance:  $F8 = S((Next(Te) - Te), (Tx - Te))$ ; where  $Next(Te)$  is the next time entrance time, if  $(Next(Te) - Te) \geq (Tx - Te)$ ,  $F8$  equals to 0, otherwise the value is 1.

#### Step 3: Sample Training

According to the above analysis, it introduces the relevant characteristics on audit into the Bayesian formula is as follow:

$$P(C | O) = P(C | F1, \dots, Fn) = P(C) \prod_{i=1}^n P(Fi | C)$$

Where  $C$  is the escaping charge vehicle abnormal event classification (such as dumped cards, exchange cards, etc.).  $Fi$  Is the characteristics value of previous summarization, the left side of the equation  $P(C | O) = P(C | F1, \dots, Fn)$  is the probability of occurrence of the characteristics

value  $Fi$  in the band sample. The right side of equation  $P(C) \prod_{i=1}^n P(Fi | C)$  is the combination probability of characteristics value  $Fi$  occurring in the classification  $C$  based on sample training.

After calculation, through artificial filter to establish a certain type of escaping charges and the combination of the characteristics binding relationship, and thus output related to escape pattern recognition model.

## 2.3 Platform Integration Technology.

### 1) Data Integration Technology

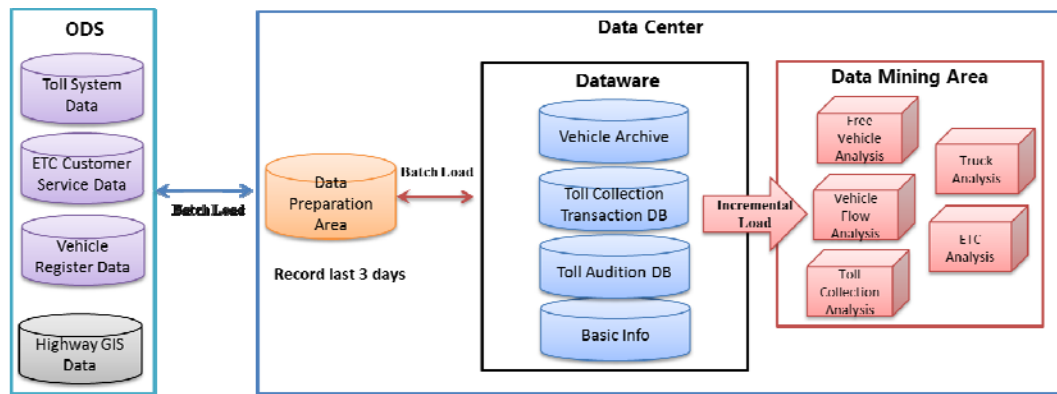


Fig 2. ETL Process

According to the requirements of operational analysis and operational auditing system for mass data analysis and mining, an operational data center independent of each production system was established. Business-related operational data was integrated from production systems according to operational analysis and audit application requirements using a dedicated data integration tool. Through unifying and consolidation at the data center level, it forms uniform standard data architecture for the whole highway road network.

In order to improve the integration efficiency of operational data, ensure the quality of integrated data and reduce the pressure on production business systems, the overall data integration strategy and flow of operational data center are implemented in three phases according to business data extraction, data import and data processing conversion.

**Data Extraction:** The ODS data preparation area is set in the data center. The data format of the buffer is exactly the same as the data format of the business system. The incremental data is periodically extracted from the business data to the data buffer. In principle, the data buffer only records 3 to 7 days.

**Data Import:** The daily data of various types of business data preparation area, according to the type of business into the operational data center business data warehouse. The quality of data is governed during import, corrected and filtered for incorrect data.

**Data processing and transformation:** Each day according to various types of operational themes of the data analysis model of the operational data warehouse business database of incremental data conversion processing, the processed data is saved to various types of data analysis library, to meet the operational analysis Demand cubes market.

## 2) Video Integration Technology

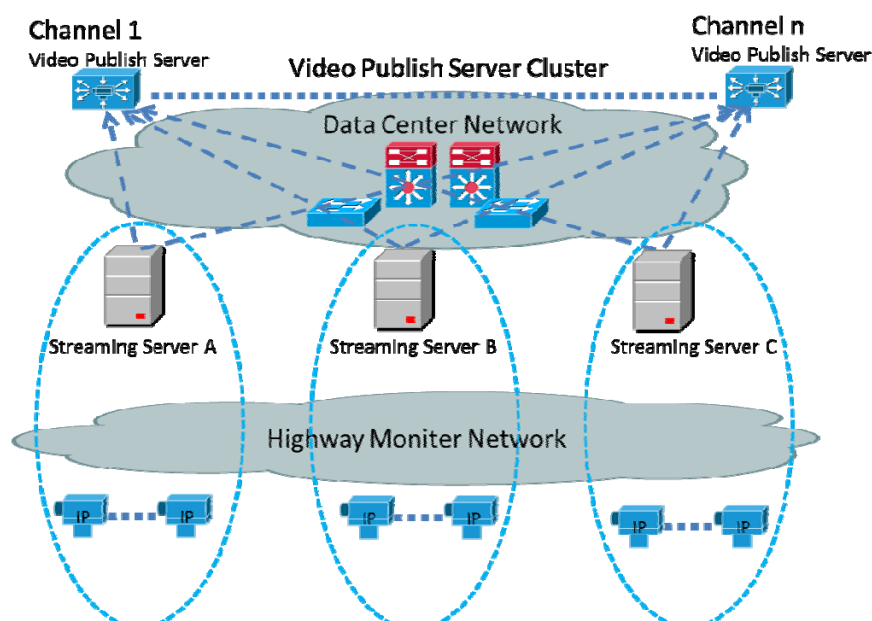


Fig 3. The architecture of video integration system

First of all, according to the status quo of multi-party construction of video resources on the entire network, heterogeneous protocols and difficulties in integration, the Operation Center has formulated the "Requirement for Digital Video Networking in Yunnan Highway" to standardize the video coding, interface, distribution and access, unified standards of storage management, and directly upgrade and monitor all video access devices and platforms on the road network according to the standard, and fundamentally solve the problem that video resource protocol interfaces are not uniform, thereby creating favorable conditions for subsequent video integration.

In the operation center video integration platform implementation, we realized the platform according to the multi-level cascade, large-scale networking technology ideas. The logical structure of the video integration platform is divided into a video access layer and a distribution layer. The access layer is classified according to the standard format of the road network according to the existing resource structure, and the sub-nodes are connected to the video streams of video platforms and terminals at all levels in the road network, the distribution layer is deployed in the network operation center in a unified way and adopts the multicast technology of the cluster to satisfy the multi-user concurrent access.

Based on the already integrated video, it creates the conditions for the establishment of audiovisual video recording inspection.

### **3. System Implementation**

#### **3.1 Operating & Audit System Implementation.**

Charges operating inspection system around the vehicle evasion behavior management and charging work assessed two core functions of construction. Based on the three-level management system in Yunnan Province, the system establishes a set of systems related to charging inspection and personnel evaluation at the provincial trading levy office, the management office and the road section sub-centers, and information-related operations such as operational audit management and personnel performance evaluation are implemented Management, to further enrich the means of inspection, reform work methods, enhance the efficiency of inspection, to ensure the orderly conduct of operating charges, the specific features include the following four aspects to be considered

(1) Through the processing and integration of charging, settlement and settlement data, it is planned and constructed in the operation data center that audit information resources such as toll collection fee operation behavior database, traffic character database, card circulation database and vehicle tracking database are established on the basis of this Analysis of the toll collection data mining, regular (daily) increment generated escape charges suspicious vehicle depot.

(2) Through the integration of the entire network charges, video surveillance resources, the construction of fee data inspection, video inspection front-end application through the integration of the entire network charges, video surveillance resources, construction fee data inspection, video inspection front-end application system, the management office, points Central audit department for the data center regularly generate suspicious vehicles combined with charging data, video recording of suspicious vehicles cross-sectoral, cross-section of the inspection lock.

(3) Through the toll collection operations fees for the conduct of statistics, the establishment of performance evaluation on the charging business assessment system.

(4) Through the operation of audit system for toll collection and operation, a three-level operational inspection management process of operation center, management office and road sections sub-centers as well as a fair, open and impartial fee-for-service business star rating system.

#### **3.2 Operating Analysis System Implementation.**

At the provincial level, a fee-based operation analysis system is established. Based on business data integration of business systems such as charging and monitoring, it initially integrates data resources such as high-speed traffic, monitoring, toll collection, video and operation and maintenance, and integrates the production of related information systems Data into a unified data quality management system, the initial formation of high-speed operation of Yunnan data center. Based on the operational themes of high-speed toll revenue, high-speed toll booth traffic, high-speed truck



performance and high-speed relief vehicles, the Company utilized the multi-dimensional analysis of big data to establish a data analysis model to form a multi-dimensional data market for operation managers. Multi-angle, multi-dimensional, customized flexible analysis capabilities, while facing the charge of operational management decision-makers, the use of various visualization of the visualization of the way to achieve business KPI dashboard display, intuitive reflection of the charging operation of the thematic analysis of statistical data to facilitate Policy makers quickly grasp the status quo of toll collection operation of the entire road network and provide leadership support for the formulation and revision of operating policies and the promotion of new businesses. Targeted operational analysis system specific goals are as follows

(1) Build a unified provincial operation data center. Realize the effective integration of business data resources such as "provincial center-sub-center-toll station" three-level traffic network charging, monitoring, operation and maintenance, unstructured pictures and video resources as well as data resources shared by external unit data centers. Provide data support for operation audit, decision analysis, public service, data exchange and sharing. It laid the foundation for realizing the wisdom high-speed big data platform and the overall framework of Yunnan smart high-speed data based on the unified information resource specification system

(2) Build a data exchange integration platform. Establish a complete provincial data integration strategy and data governance system.

(3) Construction related fee analysis system. Based on the high-speed charging business data, the daily statistical report data analysis, and customized report analysis, cross-statistical analysis of various data such as operation and maintenance data and audit data are provided to provide operation and management personnel with mobile terminal access. Leadership decision-making to provide data support.

#### **4. Conclusion**

Through the design and implementation of the high-speed toll collection and operation analysis platform of Yunnan Wisdom, this paper discusses the application of support technologies such as data integration, data analysis, data mining, video integration and data exchange and sharing to comprehensively integrate the high-speed operation of various types of business data to achieve operation-oriented operation Income, audit fees, audit performance appraisal and other related topics of mass data mining, highway managers for the analysis and decision making the necessary support. In actual use, we have achieved good results in innovating operational inspection methods, strengthening operation and auditing work, scientifically formulating operation policies and improving comprehensive information services.

#### **References**

- [1]. Han J, Kimber M, Pei J. Data Mining: Concepts and Techniques [M]. Burlington: Morgan Kaufmann Publishers, 2011.
- [2]. Highway Toll System Audit Data Mining Technology Research [D]. Hangman University, 2014.
- [3]. ZHAO Yan, WU Shelling, LIN Zheng, et al. Study on prediction model of highway card's escape behavior [J]. Chinese Science paper, 2015 (19): 2245-2251.
- [4]. Wang Xuan. Highway Toll and Audit System Based on Data Mining [J]. China ITS Journal, 2015 (12): 97-98.
- [5]. QIAN Chao, XU Honked, DAI Liang, et al. ETC Data Mining Based on Hybrid Markov Model [J]. Journal of Transportation Systems Engineering and Information Technology, 2012, 12 (4): 35-42.
- [6]. Gao Xuedong, Zhang Shoji. Application of Data Mining and Analysis System for Network Charging in Beijing-Shenyang Highway [J]. China ITS Journal, 2016 (12): 97-97.

- [7]. Guo Waxing. Research of Classification Algorithms in Data Mining [D]. Hangzhou: Zhejiang University, 2008.
- [8]. Li Chan, Wu Lie yang. Application of Data Mining in Highway Escape Analysis [J]. China ITS Journal, 2017 (S1): 124-126.
- [9]. Richard A Johnson, Dean W. Lichen, Lu Xuan translation. Applied Multivariate Statistical Analysis [M]. Tsinghua University Press, 2001.
- [10]. Li Xu, Miao Duo Qian, Feng Qianlong. Rough Clustering Algorithm Based on Data Field [J] Computer Science, 2009, 36 (2): 203-206.
- [11]. Agrawal R, Imielinski T. Mining Association Rules between Sets of Items in Large Databases [J]. Acme Sigmoid Record, 1999, 22 (2): 207-216.