Open Traffic Data Exchange and Collaborative Platform

Jatuporn Chinrungrueng* and La-or Kovavisaruch
National Electronics and Computer Technology Center
Pathumthani, THAILAND
*jatuporn.chinrungrueng@nectec.or.th

Lalita Narupiyakul and Konlakorn Wongpatikaseree
Department of Computer Engineering
Faculty of Engineering, Mahidol University
Nakhon Pathom, THAILAND

***Abstract*—Traffic information is a necessity information for people within a smart city context. Whether we travel by personal cars or public transportations, up-to-date information on traffic situations helps make our travel much more convenient. Thailand is no different from the rest of the world. This is evident that many cities including Bangkok, Chiang Mai, Khon Kaen and Phuket built mobile applications for disseminating their existing transportation data to public. Also, several universities implement mobile applications to give information on campus transportation. However, these implementations are independent and their data cannot be shared or exchanged between them.**

This paper shows an implementation of an open traffic data exchange that is ongoing between Nectec and Mahidol University. Both Nectec and Mahidol University have their proprietary traffic systems to inform of their transportation. We agree on some sets of simple standards so that collaboration and sharing of information is possible and efficient. It is our intention to make this platform available to all community including R&D and business enterprises.

Keywords—traffic data exchanget; open traffic data platform;

I. INTRODUCTION

Public transportations in major cities are vital for rapid growth. One important aspect that makes traveling by train, bus, or boat convenient is their location information especially when they are not very punctual due to unpredictable traffic situations. In Thailand, cities like Bangkok, Chiang Mai, Khon Kaen, and Phuket have GPS sensors installed on public transportations so location information can be readily obtained through servers with mobile applications. Department of Land and Transportation (DLT) of Thailand has issued a safety measure for taxis, transport vans, and buses to install GPS so they can collect relevant information. Also several private companies with their own fleets install GPS to get their own information to track their fleets and analyze data for efficient management.

It is evident that these organizations or companies realize that GPS data are essential and are valuable to their operation. However, these available systems do not readily collaborate or share between them. The designs, protocols, and data schema are mostly proprietary even though they share the same purposes for collecting the data, which are for public. There is no consensus for different parties to agree upon traffic data standards.

However, we can see many attempts to create open platform for traffic data all around the world. OpenTraffic [1] is a global data platform that collects and processes anonymous positions of vehicles and smartphones to obtain real-time and historical traffic statistics. In their project, open-source softwares are freely distributed to obtain data among collaborative partners. SharedStreets [2] is a non-profit organization attempt to provide a data standard and a platform to serve for public-private collaboration that share information such that it is not dependent on any proprietary based maps. Their work focuses on a format of data that cuts loose the community of proprietary based maps. These two examples and many more still pose their own sets of standards that other must follow and somehow gain a slow popularity.

It is known that many data sources scatter without collaboration. It is not easy for data owners to change or adjust their systems to follow some standards that they are not even sure which to choose and why. Therefore, in order to make huge amount of available sources of traffic data genuinely useful and beneficial for public and to promote Thailand 4.0 economic growth, it is essential that a common standard for open traffic data exchange and collaborative platform to be realized. Nectec and Mahidol University are two of many organizations that involve in research work in this transportation field and have come together and made an initial attempt to form a working group and promote the idea of the open platform. In this paper, we will discuss about our work and give details of our proposed platform. In Section II, Nectec and Mahidol Traffic Information Systems are described. We also include GTFS in this section as it is widely adopted in the US as an example of open data platforms. Our proposed platform is discussed in Section III. Then we conclude of our platform in Section IV.

II. NECTEC AND MAHIDOL TRAFFIC INFORMATION SYSTEM

A. Nectec Traffic Information System

Thailand Science Park under National Science and Technology Development Agency provides transportation as benefits to its employees and science park communities. More than 20 bus lines are provided for transportation between home and office both in the morning and evening. In addition, an electric tram is serviced within Science Park weekdays for commuting between 9 a.m. to 5 p.m. More trams are running during yearly exhibition.

Nectec has developed a monitoring system to provide information of the position of these transportation between home and office and within the Science Park. Either mobile application installed on the driver's mobile phone or an embedded device with GPS sensor can be used to track the vehicles. Data are submitted over cellular network and then relayed to passengers. MQTT messaging is employed such that efficient transmission of information is ensured.

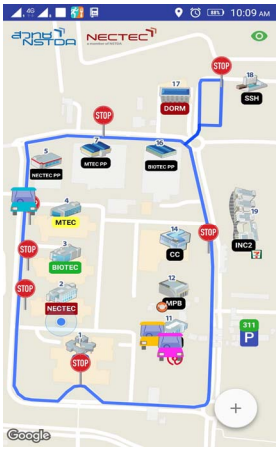

Figure 1. NSTDA Tram App

This system can be extended to support several areas of service. For example, the system also monitors serviced bus within and between campuses of Silapakorn University. The challenge we face is that how we exchange data or share our data not based on our data structure or format. In other words, what format of data should we use?

B. Mahidol Traffic Information System

Mahidol University at Salaya Campus covers the area approximately 1,984,000 square meters including 10 faculties and more than 20 administrative and service units. More than 3,000 people including students, faculties and staffs commute by using electric trams, called MU Tram, daily. Currently, MU Tram operates on 4 routes and there are 16 electric trams in services between 6:00 AM and 8:00 PM.

To improve the quality of transportation services and facilitate tram passengers at Salaya campus, the department of Computer Engineering, Mahidol University, has developed a MU Tram tracking system funding by Innovation Hub-Smart City. This tram tracking system is divided into 3 main parts as listed below.

(i) "MU Tram App", illustrated in Figure 2, is a mobile application for end-user to search the direction, check tram schedule, receive a notification and so on.

(ii) "MU Tram Driver" is a mobile application for electric tram's driver to send the GPS data of tram's current location.

(iii) "MU Tram Management System" is a web application for administrator to collect the system logs and GPS data using Firebase as a real-time database. It can be used to update the information on the mobile application.

Figure 2. MU Tram App

This tram tracking system can provide the information of the tram's current positions. The system also has the data of the tram schedule at each tram station, the GPS location of all buildings and stations at Salaya campus. All information is in form of JSON format and information services is provided via APIs.

As mentioned in section I, this transportation data from MU Tram is designed for a particular usage without referring to any data standards which cannot be shared or exchanged to other systems. This is a reason why we propose the open traffic data exchange and collaborative platform in this paper. The detail about this platform will be described in section III.

C. General Transit Feed Specification

General Transit Feed Specification (GTFS) formally known as Google Transit Feed Specification is widely adopted in US as an open data platform for transportation [3]. GTFS defines a common format for public transportation schedules and associated geographic information in the zip file. GTFS consists of a 6 mandatory text files and another 7 optional text files in csv format. These files provide a static information about the transit as shown in Table 1.

There is another standard called extended GTFS which provide the up-to-date information about position of vehicle. GTFS first adopted by TriMet Portland Transportation in 2005. Later, many public transportations in US major cities have been adopted GTFS to provide the transit information. During early adoption, some suggest that GTFS data should be charge upon requesting. However, they later discover that sharing data give more benefit than charging.

Table 1. The feed files of GTFS

file	required	Defines
agency.txt	Required	One or more transit agencies that provide the data in this feed.
stops.txt	Required	Individual locations where vehicles pick up or drop off passengers.
route.txt	Required	Transit routes. A route is a group of trips that are displayed to riders as a single service.
trips.txt	Required	Trips for each route. A trip is a sequence of two or more stops that occurs at specific time.
stop_times.txt	Required	Times that a vehicle arrives at and departs from individual stops for each trip.
calendar.txt	Required	Dates for service IDs using a weekly schedule. Specify when service starts and ends, as well as days of the week where service is available.
calendar_dates.txt	Required	Exceptions for the service IDs defined in the calendar.txt file. If calendar.txt includes ALL dates of service, this file may be specified instead of calendar.txt.
fare_attributes.txt	Optional	Fare information for a transit organization's routes.
fare_rules.txt	Optional	Rules for applying fare information for a transit organization's routes.
shapes.txt	Optional	Rules for drawing lines on a map to represent a transit organization's routes.
frequencies.txt	Optional	Headway (time between trips) for routes with variable frequency of service.
transfers.txt	Optional	Rules for making connections at transfer points between routes.
feed_info.txt	Optional	Additional information about the feed itself, including publisher, version, and expiration information.

III. Open Platform For Traffic Data

It is obvious from Section II that both Nectec and Mahidol's systems cannot exchange data that they collected. For our idea, this situation happens throughout the country that huge amount of data are scattered and are not beneficial as it should. However, it is not possible either to reinvent everything and come up with a standard that everyone agrees. To compromise the situation, it is possible to agree upon a set of simple standards to permit data exchange between two or more organizations.

A. Platform Architecture

Platform architecture illustrated in Fig. 3 shows the following major components.

i) The Internet – the global computer network that connects entities via the TCP/IP protocol.

ii) Public platform-complied hosts – any platform that open to public via proprietary APIs for traffic data collection. These platforms can provide services to its subscribers. Data are collected with embedded devices or mobile devices via proprietary or open APIs.

iii) Private platform-complied hosts – any platform that is not open for public access. These platforms are set up for private usage such as company fleet's management. Usually, the usage of this information is limited within the organization. Data are collected with embedded devices or mobile devices mostly via proprietary APIs.

iv) Google platform – google platform aka General Transit Feed Specification is included here to represent well known and widely access proprietary platforms that are open as a service. Users of these platforms can sign up and can access for traffic information as needed

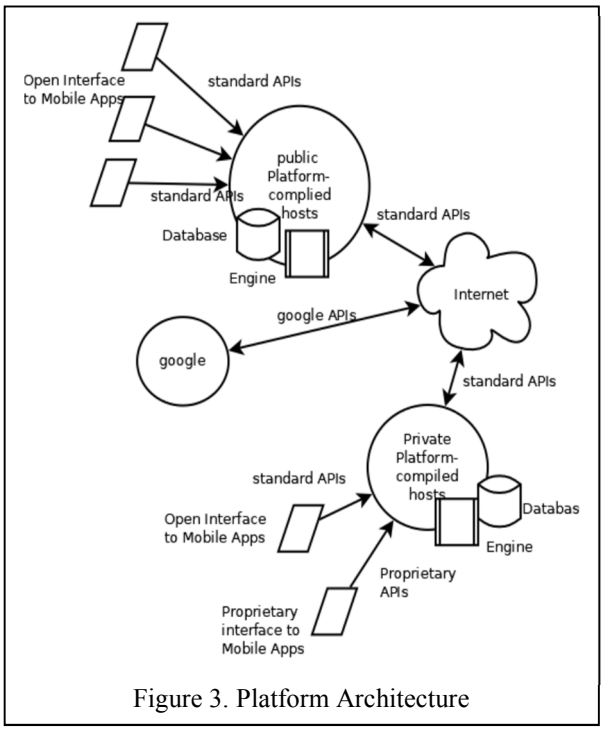

Figure 3. Platform Architecture

For these components to communicate via the Internet, we therefore propose to have a set of open standard APIs between them.

B. Data Schema

In this section, we design a minimal set of data schema agreed between Nectec and Mahidol University. This set of data schema is based on what users commonly request to access. The data schema is in form of JSON which is a kind of the data structure representation.

At this early stage of discussion between Nectec and Mahidol University, we agreed on setting up a set of static data schema that focuses on describing public transportation data such as bus and train. Thus, four kinds of the traffic data schema are proposed as listed below.

- รหัสสถานี, ชื่อ (ไทย หรือ อังกฤษ หรือ อื่นๆ) , ตำแหน่ง (lat,lon)
  - {id: xxxxx}
  - {name: {th:'มหิดล'},{en:'mahidol'},{jp:'xxx'}…}
  - {location: {lat,lon}}
- รหัสเส้นทาง, ชื่อ (ไทย หรือ อังกฤษ หรือ อื่นๆ)
  - {id:xxxxx}
  - {name:{ {th:'อนุสาวรีย์-ศาลายา'},{en:'Victory monument-Salaya'},{jp:'xxx'}…}
  - {station: {station.id1,station.id2,….}}
  - {polyline: [{lat1:aaa,lon1:bbb},{lat2:aaa,lon2:bbb},…]}
  - {operation_time:{start:6.00, stop:17.00}}
- รหัสรถ, ทะเบียน, สี, ประเภท
  - {id:xxxx}
  - {license_plate:xxxx}
- รหัสคนขับ, ชื่อ
  - {id:xxx}
  - {name:xxx, lname:yyy}
- รหัส poi (landmark), ชื่อ (ไทย หรือ อังกฤษ หรือ อื่นๆ)
  - {id:xxx}
  - {name:{ {th:'อนุสาวรีย์'},{en:'Victory monument'},{jp:'xxx'}…}
  - {location: {lat,lon}}

C. *Stake holders*

Who will benefit from this platform? The only answer is everybody. Figure 4 shows all the stake holders as they play different roles in this platform.

i) Platform hosts are individual groups that usually have enough resources and expertise to setup their platform either privately for their own business or operation or for serving other businesses. Platform hosts are categorized into public hosts and private hosts. Public hosts are defined as hosts that provide service for public usage with a possible return of income. Some public hosts may be supported by Government. Private hosts however are hosts that setup platform to serve their own purpose, for example, to do fleet management. Usually, they do not want to publish their data in the open.

ii) Providers represent stake holders who have data but do not have system to collect and share them such as bus operator. They usually have their main business other than providing technological expertise. They essentially will require help from technical people.

iii) Developers represent those people who have special expertise but lack of useful data. What they can offer is their expertise to provide services to Platform hosts or Providers.

iv) Regulators represent government or independent organizations that in some way regulate or monitor activities of other stake holders in the platform.

v) Users are mainly people who consume analyzed data or information for their own travelling or transport purposes.

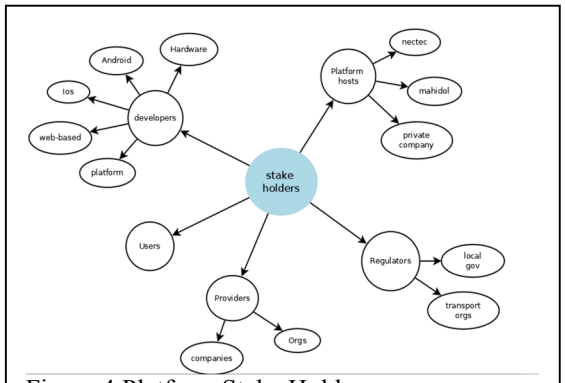

Figure 4 Platform Stake Holders

IV. CONCLUSIONS

We provide details of the starting set of data schema agreed between Nectec and Mahidol University. This set of data schema is based on what users commonly want to have access and in form of JSON format of data structure representation.

ACKNOWLEDGMENT

We sincerely express our thanks to Dr.Waresara Weerawat for her initiating this collaboration between Nectec and Mahidol.

REFERENCES

[1] G. Tamminga, P. Knoppers and J.W.C van Lint, "Open Traffic: a toolbox for traffic research," Procedia Computer Science 32 (2014), pp. 788-795.

[2] The National Association of City Transportation Officials (NACTO), "National Association of City Transportation Officials and the open transport partnership launch transportation data standard and platform, buiding foundation for public-private partnerships in the digital age [Press Release]," February 2018. Available: https://sharedstreets.io/public/SharedStreets_Release.pdf.

[3] https://developers.google.com/transit/gtfs/ retreived on July 2018