Crowdmap and Ushahidi: to obtain and visualize traffic congestion information in Mexico City

Position paper
Karla Ivon López Guillén, Uriel Flores Mendoza, Larissa Welti Santos
Intelmex it
Uruguay 58 Col. Centro
México, D.F.
(52) 55-55216179
klopez@telmex.com, dreck_nrk@hotmail.com, lwelti@telmex.com

ABSTRACT
Our research group is creating a tool that uses information technology and geographic information systems to map spatial information in order to assist people with decision-making related to the traffic congestion problem in Mexico City.

We are creating a novel application that focuses on improving vehicular mobility through the Crowdmap platform and we are asking for participation from Mexico City’s society. This application will not provide users with a dynamic shortest path, but will help them with information to take an alternative route to their destination.

We will present the most important challenges and results of our in-progress research, the solution approach to visualize traffic flow in real-time based on Crowdmap and Ushahidi, and the most important next steps to be taken in our research.

We are looking forward to the adoption of this application to evaluate what kind of impact it has in Mexico City’s society.

General Terms

Keywords
Mexico City, traffic congestion, visualization, localization, real-time, spatial information.

1. INTRODUCTION
Traffic congestion is one of the biggest problems in different cities all over the world. Nowadays, with the help of information technologies and geographic information systems, we can have access to real time information about traffic congestion in some cities in USA and in Stockholm, Dublin, Singapore and Brisbane, amongst others. Navteq and IBM have done important progress in creating systems for this purpose.

People who live in Mexico City need easy access to this kind of information as well. Although there are some Twitter accounts that tweet about traffic congestion in specific avenues, web pages showing accidents or places where demonstrations are taking place, and radio stations communicating congestion traffic events, these do not provide information about the entire areas affected by the particular events.

We are using information technologies, geographic information systems and the participation of people to create an application to map traffic congestion in real-time.

Normally, everybody has 2 or even 3 routes to reach a destination. Our goal is to create a solution over a period of time no longer than a year that is able to provide key information to make a quick decision about the route to take to reach a destination in Mexico City.

We aim to build an easy to use and real-time visualization application of the different pieces of the traffic network in the city. The main objective is to make an on-line application accessible from any device, whether mobile, laptop, desktop, etc., that graphically describes with colors the average speed of traffic flow on the main avenues. People will benefit from using this application through reducing their time-commute when knowing about the less congested avenues which may result in reducing stress levels, increasing the time spent with their family or friends and therefore improve their quality of life.

We have explored several ideas. Throughout our research and developing periods we have done changes according to raising issues and partnerships that we have encountered.

2. FIRST STEPS
2.1 Survey
We ran a survey in the main frame of Mexico City’s downtown area to confirm whether there was any interest for these kinds of applications and their usefulness for people of various ages.

The objective of this survey was to determine how the population would prefer to be informed about recommendations and alternatives to avoid traffic congestion.

We took a representative sample of the population (400 surveys) with a 97.5% confidence interval. In the survey, people had to choose one of the following options as their preferred one:

- Real-time displays located at different streets
- Displays on mobile devices
- GPS displays.
Streets and avenues included in the study were Allende, Argentina, 5 de Mayo, Venustiano Carranza, Izazaga, República del Salvador, Mesones, Uruguay, 20 de Noviembre and Isabel la Católica (all are principal routes in the downtown area).

Some of the main results were: on average, 198.896 cars get into the main frame of Mexico City’s downtown every day and 49% of people prefer having access to a display located on the main streets and avenues, followed by 35% of them who prefer having that information through their mobile devices.

We inferred that since most of the people surveyed were older than 48 years, they preferred the real-time displays located at different streets option since some of them do not feel comfortable using mobile devices and/or a GPS. Most of the people also said that they would not spend money on a GPS device.

This application may assist these 198.896 drivers in choosing the streets and avenues that are less congested and therefore, reach their destinations in less time, speeding the traffic flow.

After this survey, we had 3 main ideas on how to get traffic congestion data given that Mexico City lacks of open databases with this kind of information:

- GPS data from coworkers traveling throughout the day.
- GPS and GSM data from customers of a mobile company.
- Coordinates from mobile devices from citizens.

These three options, although feasible for implementation were going to demand many weeks for data collection, building and management. Therefore, impacting our prototyping and evaluation times, which are restricted to a one-year development period within our laboratory. For that reason, in order to build a geographic database in a short period of time, we looked for an alternative solution to be discussed later in this paper.

3. FIRST MONITORING

We hired a service to track a group of coworkers (18 subjects) during working-hours. Tracking was done through their mobile devices every 15 minutes during 3 months, from March through May 2011. Data obtained was their mobile phone number, day, time, coordinates x and y, state, neighborhood and street.

We chose people who normally travel around the city during their working hours in order to get as much data as possible.

3.1 First results

From our study with these 18 subjects, we collected about 28,519 localizations from around 1203 different streets which we divided into five different categories: downtown, north, south, east and west areas, providing interesting results about distances and displacements depending the time of the day and of the week day as well.

Data collection was not flawless and even with a percentage of error of 35.16% we decided to analyze the data to obtain both information and knowledge from it. Table 1 details the number of successful localizations obtained from the tracking service we hired.

Table 1. Total number of successful localizations by month.

<table>
<thead>
<tr>
<th>Month</th>
<th>Number of successful localization data points</th>
</tr>
</thead>
<tbody>
<tr>
<td>March</td>
<td>8328</td>
</tr>
<tr>
<td>April</td>
<td>5541</td>
</tr>
<tr>
<td>May</td>
<td>4623</td>
</tr>
</tbody>
</table>

In this table we can see that successful localization data dropped month after month. The reduction of the number of data points obtained during the month of April may be partially due to holidays, where subjects may have turned off their mobiles. During the month of May, the supplier had problems with his server. This information reinforces our argument that GPS data handling is not easy.

Collected data was also analyzed based on the schedule of successful localization data and errors (see Table 2.) We identified 3 main clusters where data collection was successful: from 7 to 9 am, from 12 to 3 pm and from 6 to 8 pm.

Table 2. Successful localization data points and errors by clusters.

<table>
<thead>
<tr>
<th>Day</th>
<th>Time frame of successful localization data points</th>
<th>Time frame of errors in localization data points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monday</td>
<td>From 7 to 9 am</td>
<td>From 12 to 3 pm</td>
</tr>
<tr>
<td>Tuesday</td>
<td>From 12 to 3 pm</td>
<td>From 6 to 8 pm</td>
</tr>
<tr>
<td>Wednesday</td>
<td>From 12 to 3 pm</td>
<td>From 12 to 3 pm</td>
</tr>
<tr>
<td>Thursday</td>
<td>From 12 to 3 pm</td>
<td>From 6 to 8 pm</td>
</tr>
<tr>
<td>Friday</td>
<td>From 7 to 9 am</td>
<td>From 6 to 8 pm</td>
</tr>
</tbody>
</table>

We shared this information with the supplier so that they can help us understand whether the system was working properly or not. This information may help in finding the time of the day when network activity is bigger.

Regarding displacements, we obtained that Mondays are the days with less commutes around the entire city than the rest of the week. Tuesdays are the days with more commutes in the north than the rest of areas of the city, Thursdays are the days with more commutes in the west than the other days of the week and Fridays are the most congested days in downtown. Fridays are also the days with more commutes around the entire city than the rest of the days of the week.

We identified a particular event in the southern part of the city. There is an avenue called Aztecas, located in front of a neighborhood called Nueva Díaz Ordaz, and none of our research subjects who have to travel from north to south or from downtown to the south pass by this neighborhood (see Figure 1). We asked people in the company if they knew about this neighborhood and some of them answered that it is a conflict area. This kind of finding, points out that a system like the one we are building, may be a useful tool for potentially identifying conflict areas within the city and therefore, through congestion traffic information, quickly find those neighborhoods in need for interventions.

![Figure 1. Neighborhood Nueva Díaz Ordaz where nobody passes through.](image-url)
4. PROTOTYPES

4.1 First prototype

The very first prototype shows a schematic map of the main avenues of Mexico City in different colors to describe the average speed of traffic flow (see Figure 2) and some zooms on specific locations on the map to show with more detail more avenues and streets on that perimeter. We are simulating speed data on the map.

Unlike other forms of traffic information, it covers the main avenues of the city based on a visual reading through colors for a quick real-time understanding, making it an information system enough for people of all ages to understand it.

The map of the main avenues was designed in a schematic form and its development was in PHP with the use of the Xampp application for Windows. The map is resized according to the device being used for visualization. It has additional images that locate the exact position of the avenues in the map, for a better comprehension (see Figure 2). It also has 5 links that show the perimeter with further detail. The web page is refreshed every 30 seconds to simulate real-time data actualization.

4.2 Second prototype: Crowdmap and Ushahidi

Ushahidi is a system to collect and visualize data on a map and Crowdmap has spatial data management capability needed to develop an easy and quick platform for real-time visualization. These systems are open source and anyone can use them to build an application according to their needs. In our case, we built an application to visualize traffic congestion in real-time in Mexico City (see Figure 3). We chose using these systems because we were looking forward to obtaining information from a large number of people without increasing the monitoring cost. Information from 18 subjects provided us with a glimpse of traffic congestion, yet, it was not enough data for deeper analysis.

The second prototype is filled with information provided by people in cars, taking the buses or subway and anyone who is seeing traffic congestion in the city and wants to prevent others from getting stuck in the same situation.

As any other online application, its success depends entirely on people’s participation. We launched a full version on June 9, 2011, 3 months after we have collected around 1000 events from social networking applications and around 3000 twitter messages. It automatically obtains every message containing a hashtag of #traficoDF, #trafico889, #reportevial, among others, and the user can comment on every traffic event via Facebook.

The application we have built has 6 submenus:

1. Inicio (Start): It is the main page that includes the map of Mexico City and the events reported in real-time. It has a box describing the 3 congestion categories defined for this application according to speed: Speed > 60 km/h, Speed between 60 and 30 km/h and Speed < 30 km/h. There are also: one box with the instructions for creating a new event and one box that includes statistics of created events per day.

   The application is able to get information from Twitter with hashtags related to #TraficoDF and they appear in the bottom of this section.

2. Eventos (Events): It is a list of all the reports obtained through the web page or Twitter and there are 2 categories of events: verified and unverified. The administrator of the application is the person responsible for categorizing the events depending on the source and the information described.

3. Nuevo Evento (New Event): This is the section where people can create a new event and provide information. On the map, they must click on the exact location of the event to be posted, write the name of the street followed by the words: Mexico DF (Mexico City), and then, they must choose one category of speed and write down a brief title and a clear and correct description of the event.

4. Recibir alertas (Getting alerts): If people want to be updated about the information loaded on the application, in this section they can fill a form to get alerts when a new event is published.

5. Contacto (Contact): This is a form for sending a message to the web page administrator.

![Figure 2. Mexico City’s schematic map with the most important avenues.](image)

![Figure 3. Ushahidi system for visualizing traffic congestion in Mexico City.](image)
5. CURRENT PROGRESS
With only 7 months of research, surveying, monitoring and prototyping, we tested different tools for data collection and visualization, and we learned about application preferences for monitoring traffic congestion. We have obtained data on specific routes, number of cars entering downtown on daily basis, rush hours on certain streets, amongst other interesting facts, and have developed a system using the platform of Crowdmap to collect data from people and systems like Facebook and Twitter.

6. FUTURE WORK
6.1 Partnerships
During our research and development, we have concluded that the most appropriate next step is to make partnerships in order to improve the application.

We have identified specific partners we would be interested in collaborating with and the reason why such collaboration is important:

- Navteq: To get historical traffic congestion data in Mexico City;
- Texas Transportation Institute: To get advice and knowledge on effective techniques for tracking traffic congestion;
- Txteagle: To get data from their mobile users’ network which is well identified and already trained in Mexico City;
- IBM Mexico: To obtain and use specialized systems for tracking traffic congestion;
- Casa Vecina: To get help with the promotion and improvement of our application;
- Radio Stations: To ask for collaboration to enter new events with their traffic congestion information.

Through partnerships we can gain data, improve the application and important feedback so we can work towards developing an intuitive application in order to obtain as much participation as possible.

6.2 Challenges
To manage, store and process large amount of spatial data, a good number of developers and a big enough database is needed, amongst other resources.

In terms of programming, the most urgent challenges are:
- transforming clustered data into lines that describe the speed;
- visualizing real-time data only instead of historical data;
- finding a solution to automatically enter a new event without the supervision of the system administrator;
- developing an algorithm for discriminating truthful information provided by the users from the untruthful information;
- developing mobile apps for IOS and Android operating systems.

The most important challenge is getting enough participants feeding the system to give users the correct information about traffic congestion in the most important avenues of Mexico City.

7. CONCLUSION
Making ideas happen is the most important guideline in our laboratory. We work on projects that may have social impact; therefore, rapid prototyping and constant evaluation are key elements in our research.

For our project, collecting real-time data directly from users through Crowdmap and Ushahidi allowed us to approach the problem in a fast way. These systems turned out to be an effective solution, yet their usability remains challenging because they are not simple and intuitive for users.

Since our on-line application depends on the number of people feeding it with data, our objective at this moment is to improve some of the design principles of our Ushahidi platform so that people can intuitively feed it with truthful information and check it every time they need traffic congestion information. After that, we plan to develop apps for different mobile operating systems to reach more people and therefore obtain more data.

The next step is to reach out to the largest possible number of people using the application through partnerships already identified, to then be able to measure its impact.

We will continue networking with people or companies interested in this kind of applications, so that we can foster and develop more ideas in order to improve people’s quality of life through geographic information systems.

8. REFERENCES
[8] PosteaTraficoDF http://twitter.com/PosteaTraficoDF