Urban Planning and Building Smart Cities Based on IoT Using Big Data Analytics

M. Mazhar Rathore, Awais Ahmad, Anand Paul, Seungmin Rao

Presented By – Sahil Puri
Outline

- Background of IoT, Smart Cities and Urban Planning
- Proposed 4-tier architecture for IoT based Smart City
- Implementation model
- Data set information
- Analysis conclusions and inferences
- System Implementation
In IoT, internet technologies provide a way of integrating and sharing a common communication medium between devices.

- IoT empowers an object to see, listen and communicate at the same time.

- In 2008, CISCO reports that the number of things connected to the internet has surpassed the number of people living on the earth, whereas in 2020, it will reach the limit of 50 billion, resulting in the enrichment of the digital world.

- More (Devices on internet) = More (Data generated)
  - Yay, Its Big Data!
Background – Internet Of Things

Applications of Internet of Things:

- Wearables
- Smart Grids
- Smart Farming
- Smart Homes
- Smart City
Background – Smart City

• Cities equipped with electronic equipment such as:
  • Street Cameras, Censors for transportation systems, and now individual mobile devices.

• Inhale the recent developments in ICT to achieve:
  • Smart homes
  • Vehicular traffic
  • Environment pollution
  • Surveillance systems
Background – Urban Planning Development

• New field in which devices are integrated by geographic location and analyzed.

• Analysis of disjointed output datasets of sensors is done (not in real time) to design a new system for various services in the city.

• Impact areas of IoT in urban development:
  • Health and safety of citizens
  • Transportation systems in terms of availability, mobility and pollution
  • Public car parking
  • Air and noise pollution
  • Surveillance systems in cities
Underlying Questions

• Can the analysis results be incorporated into decision planning? Will the results always map to the desirable output?

• Is it okay to automate real time decision making systems? What happens in case of failures or hacks?

• How to ensure the quality of information?

• How to minimize the cost of data collection that is being generated by some devices?
## Proposed IoT based Smart City

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Houses</td>
<td>Temperature sensors.</td>
<td>Temperature</td>
<td>Electric Companies can manage their distribution</td>
<td>Load can be distributed efficiently</td>
</tr>
<tr>
<td>Cities</td>
<td>Smoke sensors at many places.</td>
<td>Smoke</td>
<td>Pollution level can be tracked</td>
<td>Warn citizen if pollution level above threshold.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Instances of fire can be tracked</td>
<td>Emergency procedures can be taken.</td>
</tr>
<tr>
<td>Parking spots</td>
<td>Track vehicles coming in and out of parking zones.</td>
<td>Parking space availability</td>
<td>Citizens can plan parking location beforehand.</td>
<td>Reduce fuel wastage and time.</td>
</tr>
<tr>
<td>Weather &amp; Water information</td>
<td>Installing sensors in reservoirs and open places.</td>
<td>Rain, Humidity, Pressure, Wind speed, Water level at reservoirs, rivers, lakes</td>
<td>To predict city water requirement. Weather predictions.</td>
<td>Water load and storage can be handled efficiently.</td>
</tr>
<tr>
<td>Vehicular traffic information</td>
<td>Road surveillance, Installing sensors on vehicles windshield.</td>
<td>Time required to reach destination, speed etc.</td>
<td>Predict road blockages, traffic circumstances.</td>
<td>Citizens can plan and take alternate routes.</td>
</tr>
<tr>
<td>Gas Pollution levels</td>
<td>Gas detecting sensors at open places.</td>
<td>Monitor harmful gas levels such as carbon monoxide etc.</td>
<td>Gas pollution can be tracked.</td>
<td>Citizens can be warned to stay at home.</td>
</tr>
</tbody>
</table>
Proposed Urban Planning and Development

- Electric usage for upcoming years can be predicted.
  - Dams, power plants etc. can be built.
- Smart parking and vehicular traffic.
  - New parking lots and new roads / flyovers can be constructed.
- Weather and Water.
  - Planning for agricultural usage, floods and safe drinking water.
- Security surveillance
  - Identify most crime prone areas and take remedial actions.
  - Policing deployments can be managed.
Unified Model – Smart City and Urban Development
4-Tier Analytical Architecture for IoT based Smart City
Data Collection Level

Data is in heterogeneous format with varying formats, different point of origin and periodicity.

Data Aggregation Level

Contains a lot of metadata associated with sensors.

Data Filtration Level

Filtration technique is applied which removes unnecessary metadata and discards duplicate data.
Communication technologies are used between sensors and relay nodes to send data from aggregation level to backend for processing.
4-Tier Analytical Architecture – Tier 3

Analysis and data processing

Real Time processing of data

Storm / SPARK / VoltDB can be used.

Hadoop & MapReduce for storage and analysis
Summary of analysis is prepared and report is generated.

Report results is announced and disbursed to different departments such as flood detection etc.
Implementation model of 4-Tier Analytical Architecture
Data Set Analysis

Dataset details.

<table>
<thead>
<tr>
<th>S#</th>
<th>Datasets</th>
<th>Size</th>
<th>No. of parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Floods</td>
<td>16MB</td>
<td>30</td>
</tr>
<tr>
<td>2</td>
<td>Water usage</td>
<td>5MB</td>
<td>11</td>
</tr>
<tr>
<td>3</td>
<td>Madrid Highway vehicular traffic</td>
<td>450MB</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>Vehicular mobility traces</td>
<td>4.03GB</td>
<td>5</td>
</tr>
<tr>
<td>5</td>
<td>Parking lots</td>
<td>294KB</td>
<td>7</td>
</tr>
<tr>
<td>6</td>
<td>Pollution</td>
<td>32GB + 570MB</td>
<td>8</td>
</tr>
<tr>
<td>7</td>
<td>Social network (twitter)</td>
<td>8 + 8MB</td>
<td>7</td>
</tr>
<tr>
<td>8</td>
<td>Aarhus city traffic</td>
<td>33GB</td>
<td>9</td>
</tr>
<tr>
<td>9</td>
<td>Weather</td>
<td>3MB</td>
<td>7</td>
</tr>
</tbody>
</table>

- **Flood**
  - Collected from Officials and TV news channels of flooded countries, archived by Dartmouth Flood Observatory, University of Colorado.
  - Dataset contains Date of flood, area of flood, severity, deaths caused etc.

- **Water usage**
  - Details of water meter readings of 61263 households of Surrey, Canada.
  - Contains complete address and water usage.
Analysis of Data Set (Cont.)

• Vehicular traffic of Madrid highway.
  • Contains location of each vehicle between the two ends of the highway and speed of the vehicle.

• Vehicle Mobility
  • Dataset generated by ITS, German Aerospace containing the mobility of 700 cars in Cologne, Germany (approx. 400km$^2$ for 24 hours).

• Parking Lots
  • Information of 8 parking lots of Aarhus, Denmark from May, 2014 to Nov 2014 from 55 points / sensors.

• Pollution.
  • Gases and smoke concentration were monitored at various locations of Aarhus, Denmark using 449 sensors.

• Social Media Dataset
  • Twitter data contains information (City, location, time) from 13674 tweets from Sept 2013 to Dec 2013.

• Weather Datasets
  • Data containing Rain, humidity, Wind, Pressure etc. from Feb 2014 to June 2014 and August, September 2014.
Data Analysis – Vehicular Traffic

• In Fig. 4, when number of cars is high, the speed decreases.
  • At sample 106 and 121, when number of cars is maximum -> Speed is minimum.

• In Fig. 5, number of cars is 25 to 35 considered to be high intensity traffic.

• Thus, the above information can help in planning for road management as per the car density.
In Fig. 6, two types of traffic classes are used.  
- Class 1: 1 – 15 cars.  
- Class 2: 25 – 35 cars  
- At certain samples, time required for cars to reach destination is quite high.  
- These signal road blockages at intersections. For e.g. road near university, colleges, schools etc.  

This can help identify blocking intersections which might need wider roads / flyovers.  

Also, places where traffic is usually less, roads with less number of lanes can be constructed.  

Fig. 7 suggests peak traffic time (8:25 to 11:25) so that alternative routes can be taken by citizens.
Data Analysis – Vehicular Traffic (Cont.)

- Analysis of Madrid Highway shows maximum number of cars passing through various locations.

- Suggests that location 500 is a major intersection through which majority of cars pass.

- Analysis can also help in determining road conditions. If during non peak hours if traffic time is consistently higher, suggests bad road condition or blockage.

- Speed violations can be tracked using these sensors.
Data Analysis – Parking Data

- Free spaces at various parking spots at various time samples.
- Bruuns has the highest free spaces at some samples, it has the maximum utility too.
- Citizens are informed of free parking spaces at various locations.
- Greater profit and equilibrium between vendors.
- Government can use this data to plan for new parking lots.
Data Analysis – Smart Home

- Water usage of all houses in Surrey, Canada was analyzed (in Cubic meters).

- Can be used to determine the number houses using more than average quantity of water.

- Used Skewness measure to find out that more than 50% household use more than average amount of water and the usage is positively skewed.

- Can help Government determine water billing rates based on normal usage.

Fig. 12. Total water usage for Surrey.
• Water usage of various areas of Surrey, Canada was analyzed.

• Areas 101B and 102B consume very less water in comparison to areas 102A and 103A consume very high water, hence the flow can be balanced.

• Consumption in industrial areas is quite high as compared to residential areas.

• Effective drainage system can be developed.

• Similar initiatives can be taken up for electricity, gas etc.
Data Analysis – Floods

- Each flood is associated with factor ‘M’ where $M = \log(\text{duration} \times \text{area affected} \times \text{death})$

- **M** represents the magnitude of flood.
  - If $M \geq 4$, Flood is of higher intensity.
  - If $M \geq 6$, Flood is of dangerous intensity.

- Each category of flood is analyzed and
  - 35% of floods have occurred due to rain followed by 1.5% of floods due to snow.

- Hence we can define threshold for rain, if crossed, Citizens can be alerted for emergency procedures.

- High intensity drainage pipes can be installed and sensors can be installed on mountains to check for snow melting.
Data Analysis – Environmental Pollution

- Ozone value is greater than normal at time sample 70 - 90.
- Particulate matter higher at 185-215 and above 245.
- Nitrogen dioxide at the start and end of the sampling interval.
- Carbon monoxide at 90-115.

- Children should not be allowed to be outside during these times.
- Adults should avoid deep breathing activities.
- People with respiratory diseases should be careful.

- Authorities can also use this data and warn the citizens if the gas amount is quite high and dangerous.
- Government can use this data for urban planning to analyze the changes in pollution in different seasons and months.
- Yearly analysis can help government move out industries from residential areas.
System Implementation – Smart City

- Single node Hadoop at Ubuntu with 3.2GHz * 4 processors and 4GB memory was used.

- Circles outside green boundary are inputs to the system.

- Circles inside green boundary are various modules implemented in the system.

- Boxes outside green boundary are the entities to which the results are supplied.
System Evaluation

- Processing time and dataset size are proportional to each other.
- Throughput is also proportional to the dataset size.
Conclusion

• Smart cities and urban development can have a major impact on national development.

• The proposed system benefits citizens and authorities to take quick and intelligent decisions.

• The proposed system takes in a lot of factors and generates result for developing a smart city and for urban population.

• Proposed system is tested for efficiency considering time and throughput and is found efficient even with larger data set.
References

• S.S. Han, Global city making in Singapore: a real estate perspective, Prog. Plan. 64 (2) (2005) 69–175.
Discussion - Future Work

• Devise a platform for consumption of the analysis as generated by the system.

• Check the efficiency of the system for real time usage and analysis of data.

• Check system performance when actual large amount of data is used. Would need scaling up of database and system.

• Plan for security and privacy concerns of user data for analysis.
Discussion - Related Papers / Articles

• CISCO, *The Internet of Things*,
  - Blog Post depicting the outreach of Internet Of Things.

• A. Ahmad, A. Paul, M. Mazhar Rathore, H. Chang, *Smart Cyber Society: Integration of capillary devices with usability based on Cyber – Physical system*
  - Research paper gives introduction to Web of things and application for Smart Homes and touches base on smart community.

• Rob Kitchin, *The real – time big city? Big data and smart urbanism*.
  - Introduces a political angle in smart city.
  - Points out challenges in the implementation of smart cities.
Discussion - Paper Review

• Strengths
  • Good background information about the topic.
  • Includes end to end details on architecture to implementation.
  • Goals and motivation were clearly explained.

• Weaknesses
  • Security and privacy information not considered.
  • Graphs not clear and concise.
  • Explanations are digressive at some points.
  • Dataset extraction method not explained.
  • Social Media analysis not shown
Discussion - Questions

• Can the analysis results be incorporated into decision planning? Will the results always map to the desirable output?

• What do you think are the limitations of the architecture that the authors propose?

• How to tackle uncertainty induced due to the real time and offline dynamics and ensure the quality of information?

• How to minimize the cost of data collection that is being generated by some devices?

• Is it okay to automate real time decision making systems? What happens in case of failures or hacks?