Risk Priority Estimation for Major Corridors by using PCA Application of Traffic Mobility in Urban Areas: A Case Study of Gaddiannaram Municipality Area Hyderabad, A.P., India

Awari Mahesh Babu¹, K.M. Lakshmana Rao² and Arti Khaparde³

¹Associate Professor, Civil Engineering Department, Tirumala Engineering College, Hyderabad, A.P., India  
E-mail: maheshbabuawari@gmail.com

²Professor, Department of Transportation Engineering, JNTU College of Engineering, Hyderabad, A.P., India  
E-mail: kmrlraoin@yahoo.com

³Professor and HOD, Department of Electronics and Communication Engineering, Tirumala Engineering College, Hyderabad, A.P., India  
E-mail: artikhaparde@gmail.com

Abstract

Risk priority for major corridors of traffic mobility was done with the objective of knowing the exact criterions and the links which lead to congestion. The Gaddiannaram municipality of Ranga Reddy district, Andhra Pradesh, INDIA has been delineated as the study area. Road network characterization surveys and traffic characterization surveys are conducted in the study area. Travel demand analysis was done using GIS where desire line diagrams, user preferred paths and the functionality of road system has been identified. Identifying congestion in the links of the network is an important task in integrated planning. The performance and functionality of a link can be attributed to many characteristics including the neighborhood characteristics of the link and there is a research gap in finding the exact criterions which are the real factors of risk generation. So the link has to be evaluated based on these conflicting multi attributes at one point of time. The present paper deals with the finding of risk for the Gaddiannaram municipality area, Hyderabad AP., India using Principal Component Analysis(PCA).

Keywords: Risk generation; Road network characterization; Travel demand analysis; Desire line diagrams; User preferred paths; Principal component analysis.
Introduction

Urbanization is the physical growth of urban areas from rural areas as a result of population immigration to an existing urban area. Urbanization is closely linked to modernization, industrialization, and the sociological process of rationalization. Rapid Urbanization causes haphazard and unplanned growth of urban centers which becomes more complicated with the fact that it must take place within the built up area. This pressure of continuously growing population results in overcrowding and becomes burden to limited civic cycle amenities which forces the middle class as well as builders to move to outlying suburbs, phenomenon called Urban Sprawl or Growth. Urbanization operating in the fringe brings a number of Transportation problems of safety, congestion, accidents, parking, management and enforcement. [1]

The study is attempted, with reference to the geometric, traffic, utility and land use characteristics of the study area, to identify the major corridors with traffic characteristics as independent characteristics of risk generation as basis to identify the links leading to congestion.

Principal component analysis

The central idea of principal component analysis (PCA) is to reduce the dimensionality of a data set consisting of a large number of interrelated variables, while retaining as much as possible of the variation present in the data set. This is achieved by transforming to a new set of variables, the principal components (PCs), which are uncorrelated, and which are ordered so that the first few retain most of the variation present in all of the original variables. [4]

Principal Component Analysis is a well-established technique for dimensionality reduction and multivariate analysis. Examples of its applications include data compression, image processing, visualization, exploratory data analysis, pattern recognition, and time series prediction.

PCA summarizes the variation in correlated multivariate attributes to a set of non-correlated components, each of which is a particular linear combination of the original variables. The extracted non-correlated components are called Principal Components (PC) and are estimated from the eigenvectors of the covariance matrix of the original variables. Therefore, the objective of PCA is to achieve parsimony and reduce dimensionality by extracting the smallest number components that account for most of the variation in the original multivariate data and to summarize the data with little loss of information.

In PCA, the extractions of PC can be made using either original multivariate data set or using the covariance matrix if the original data set is not available. In deriving PC, the correlation matrix may be used, instead of the covariance matrix, when different variables in the data set are measured using different units or if different variables have different variances. Using the correlation matrix is equivalent to standardizing the variables to zero mean and unit standard deviation.

The PCA model can be represented by:

\[(1)\]
where $U$, an $m$-dimensional vector, is a projection of $X$ - the original $d$-dimensional data vector ($m \ll d$). It can be shown that the $m$ projection vectors that maximize the variance of $U$, called the principal axes, are given by the eigenvectors $e_1, e_2, \ldots, e_m$ of the data set’s covariance matrix $S$, corresponding to the $m$ largest non-zero eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_m$.

The data set’s covariance matrix $S$ can be found from:

$$
\Sigma
$$

(2)

Where $\mu$ is the mean vector of $x$. The eigenvectors $e_i$ can be found by solving

The set of equations:

$$
O 1, 2, \ldots, \ldots
$$

(3)

Where $e_i$ are the eigen values of $S$. After calculating the eigenvectors, they are sorted by the magnitude of the corresponding eigen values. Then the $m$ vectors with the largest eigen values are chosen. The PCA projection matrix is then calculated as:

where $E$ has the $m$ eigenvectors as its columns. Here $W$ is a $m \times d$ matrix.

One of the motives behind the selection of PCA for the detection of network traffic anomalies is its ability to operate on the input feature vector’s space directly without the need to transform the data into another output space as in the case with self-learning techniques. For example, in Self-Organizing Maps, the transformation of a high-dimensional input space to a low-dimensional output space takes place through the iterative process of training the map and adjusting the weight vectors. The weight vectors are typically selected randomly which makes the process of selecting the best initial weight vectors a trial-and-error process. In PCA, dimensionality reduction is achieved by calculating the first few principal components representing the highest variance in the components of the input feature vector, without the need to perform any transformations on the input space. The input data is analyzed within its own input space, and the results of the transformations are deterministic and do not rely on initial conditions.

**Generalizations and adaptations of principal component analysis**

The basic technique of PCA has been generalized or adapted in many ways. Material that has presented some definitions of ‘non-linear PCA’ and ‘generalized PCA,’ respectively have connections with corresponding analysis. Non-linear extensions of PCA include the Gifi approach, principal curves, and some types of neural network, while the generalizations of cover many varieties of weights, metrics, transformations and centering. Modifications of PCA are useful when secondary or ‘instrumental’ variables are present, and there are some possible alternatives to PCA for data that are non-normal. These include independent component analysis (ICA). The ideas of three-mode and multi-way PCA are useful. These analyses are appropriate when the
data matrix, as well as having two dimensions corresponding to individuals and variables, respectively, has one or more extra dimensions corresponding, for example, to time. Considering some ideas from neural networks and goodness-of-fit, we can present some other modifications of PCA.

The flowchart showing the step by step procedure of conducting Principal component analysis is presented in Fig. 1

**Figure 1:** Flowchart showing the step by step procedure of Principal component analysis.

**Study area**
The Gaddiannaram municipality of Ranga Reddy district has been delineated as the study area (Fig.2). The study area comprises of 21 wards. Based on the survey, 5 major corridors have been identified. The map shows the major corridors and major intersections in the study area. The map showing the major intersections can be viewed in Fig.3

**Figure 2**
Figure 3

Figure 4: Risk priority graphs for velocity criterion.

Figure 5: Risk priority graphs for headway criterion.
Data collection and Processing
The field surveys that were carried out to collect the data in the study area have been grouped into two broad categories. They are Road network characterization studies and Traffic characterization studies. The data collected from these surveys has been analyzed to evaluate the network characteristics of the study area.

The broad criterions are categorized in four groups – Geometric characteristics, Traffic characteristics, Land use or road side characteristics and Utility characteristics. The attributes considered under each criterion are as follows.

**Geometric characteristics**
Geometric characteristics represent the geometric features of the roadway affecting the Level of service of the link. These are the static characteristics of the road infrastructure.

They are:
1. Roadway width in meters (RW)
2. Carriageway width in meters (CW)
3. Stopping sight distance in metres (SSD)
4. Number of curves on the link (NC)
5. Pavement Condition index determined from the rating of the pavement based on the pavement condition and riding comfort experienced by the user to the scale of 1 to 5, 5 being an excellent pavement and 0 being an impassable pavement. – (PCI)
6. Number of access points on the link (NA)

**Traffic characteristics**
Traffic characteristics are the dynamic characteristics of the road that influence the level of service of the link. They are
1. Headway in seconds (H)
2. V/C Ratio (VCR)
3. Speed in kmph (V)
4. Delay in seconds (D)

**Land use or Road side Characteristics**
These are also the static elements of the link which influence the operational efficiency of the link. They are:
1. Commercial area along the road side of the link in sq.km (CA)
2. Residential area along the road side of the link in sq.km (RA)
3. Semi Residential area along the road side of the link in sq.km (SRA)
4. Industrial area along the road side of the link in sq.km (IA)
5. Intensity Parking, business activities and road side activities in a point scale (PBE indicating parking, business activities and encroachments)

**Utility characteristics**
Utility characteristics are the characteristics of the link indicating the degree of utility of the link with reference to the static analysis and dynamic analysis. They are
1. Overlap size of the link from static analysis (OS)
2. Trip intensity on the link (TI) in trips/day

**Methodology**

**Step-1: Data acquisition**
All the geometric, traffic, land use and utility characteristics are taken into consideration for the analysis.

**Step-2: Standardization of input values**
The objective function is optimized to identify the links in the network which are leading to congestion and the factors which are responsible for congestion in the network evaluation criteria for attending the objective function are the path characteristics that influence more in relation to the congestion or risk occurrence. A lead is the identification of functional characteristics is obtained from (Nesamani K.S. et al 2005) and Stephen r. Alderson & Yorgos Stepharedes 1986 which evaluates the links based on the geometrics, traffic, land use and utility characteristics and showed a strong influence on the overall performance of the network. The main criterions with their objective functions which provide ideal values for congestion are linked below.

<table>
<thead>
<tr>
<th>Geometric</th>
<th>Traffic</th>
<th>Land Use</th>
<th>Utility</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carriage way width (Min)</td>
<td>Headway (Min)</td>
<td>CA (Max)</td>
<td>Trip intensity (Max)</td>
</tr>
<tr>
<td>Roadway width (Min)</td>
<td>V/C ration (Min)</td>
<td>RA (Max)</td>
<td>Overlap Size (Max)</td>
</tr>
<tr>
<td>Shopping right distance (Min)</td>
<td>Speed (Min)</td>
<td>SRA (Max)</td>
<td></td>
</tr>
<tr>
<td>Number of curves (Max)</td>
<td>Delay (Max)</td>
<td>IA (Max)</td>
<td></td>
</tr>
<tr>
<td>PCI (Min)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of access points (Max)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
As per the above table, input values are standardized by subtracting either the minimum value or maximum value from the respective characteristic vector and dividing it with the difference between maximum and minimum value.

**Step-3:** Calculation of the covariance matrix
The covariance of the above matrix is found using the software MATLAB.

**Step-4:** Calculation of the eigenvectors and Eigen values of the covariance Matrix

**Step-5:** Calculation of the Principal components

**Step 6:** Deriving new Dataset
The Final values are obtained by multiplication of the one transposed matrix i.e Principal components matrix and Standardized matrix. The final values obtained are the standardized values.

**Step-7:** Finding the Risk
After finding the final values of speed, delay, and Headway are compared with actual values. Therefore the risk generated links are found out

**Findings and Recommendations**
The study links the complexity of the road systems with the land use and demand profiles. Identifying the risk generated in the major links is the crucial task in the integrated planning. The performance and functionality of a link can be attributed to many characteristics including the neighborhood characteristics of the link. Risk analysis is done through Principal component analysis and to find the links which are leading to congestion. From Principal component analysis, it is found that the risk is observed because of the three criterions, speed, delay and headway.

With speed as major criterion, it is found that the risk is generated in the links and distributed in the following order

<table>
<thead>
<tr>
<th>S.No.</th>
<th>Link id</th>
<th>Name of Link id</th>
</tr>
</thead>
<tbody>
<tr>
<td>1)</td>
<td>M3</td>
<td>Dilsuknagar – Sankeswar Bazaar</td>
</tr>
<tr>
<td>2)</td>
<td>M5</td>
<td>Sankeswar Bazaar - Saroornagar</td>
</tr>
<tr>
<td>3)</td>
<td>M1</td>
<td>Dilsuknagar - Chaithanyapuri</td>
</tr>
</tbody>
</table>

With Headway as main criterion, it is found that risk is generated in order in the following corridors

<table>
<thead>
<tr>
<th>S.No.</th>
<th>Link id</th>
<th>Name of Link id</th>
</tr>
</thead>
<tbody>
<tr>
<td>1)</td>
<td>M2</td>
<td>Chaithanyapuri – Kothape</td>
</tr>
<tr>
<td>2)</td>
<td>M1</td>
<td>Dilsuknagar - Chaithanyapuri</td>
</tr>
<tr>
<td>3)</td>
<td>M3</td>
<td>Dilsuknagar – Sankeswar Bazaar</td>
</tr>
<tr>
<td>4)</td>
<td>M5</td>
<td>Sankeswar Bazaar - Saroornagar</td>
</tr>
</tbody>
</table>
With Delay as main criterion it is found that the risk is generated in the links.

S.No.    Link id    Name of Link id
1) M1    Dilsuknagar - Chaithanyapuri

It is concluded that the risk is found, generated and distributed from one point to multiple points.

The following measures are proposed in order to control the risk generation produced on the major corridors of the study area based on the three criterions-speed, delay and headway.

1. Promotion of land use must be discouraged along the corridors.
2. Speed monitoring measures are to be taken care.
3. Parking, loading and unloading activities are restricted and regulated throughout the corridors
4. Raised footways, barriers and foot over bridges are recommended to discourage the pedestrians moving on the main corridors.
5. Road delineators are recommended along the lane markings.
6. Face-lifting of the junctions are recommended.

Table 1: Link Input attributes of the characteristics in each.

<table>
<thead>
<tr>
<th>S.No</th>
<th>Link name</th>
<th>Lin k id</th>
<th>Geometric characteristics</th>
<th>Traffic characteristics</th>
<th>Land use or Road side Characteristics</th>
<th>Utility characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>RW in m</td>
<td>CW in m</td>
<td>SSD in m</td>
<td>N C</td>
</tr>
<tr>
<td>1.</td>
<td>Dilruk Nagar - Chaithan</td>
<td>M1</td>
<td>21.5</td>
<td>18</td>
<td>50.5</td>
<td>0</td>
</tr>
<tr>
<td>2.</td>
<td>Chaithnagar - Ko</td>
<td></td>
<td>M2</td>
<td>21.5</td>
<td>18</td>
<td>50.5</td>
</tr>
<tr>
<td>3.</td>
<td>Dilruk Nagar - Sankeswar Bazaar</td>
<td>M3</td>
<td>10</td>
<td>7</td>
<td>52.1</td>
<td>2</td>
</tr>
<tr>
<td>4.</td>
<td>Chaithnagar - P&amp;T</td>
<td></td>
<td>M4</td>
<td>9.5</td>
<td>7</td>
<td>48.0</td>
</tr>
<tr>
<td>5.</td>
<td>Sankeswar Bazaar - Saroornagar</td>
<td>M5</td>
<td>10</td>
<td>8</td>
<td>55.4</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 2: New data sheet.

<table>
<thead>
<tr>
<th>RW in m</th>
<th>CW in m</th>
<th>SSD in m</th>
<th>NC</th>
<th>PCI</th>
<th>H in sec</th>
<th>VCR</th>
<th>PB E</th>
<th>V in km/h</th>
<th>D in sec</th>
<th>NA</th>
<th>CA in sq.km</th>
<th>RA in sq.km</th>
<th>SRA in sq.km</th>
<th>IA in sq.km</th>
<th>TI in trips / day</th>
<th>O S</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.83</td>
<td>2498</td>
<td>8.21</td>
<td>391</td>
<td>4.01</td>
<td>357</td>
<td>2.4</td>
<td>2066</td>
<td>3.572</td>
<td>0.0024</td>
<td>31</td>
<td>0.067</td>
<td>7.33E-05</td>
<td>7.22E-05</td>
<td>56.9</td>
<td>3021</td>
<td>6789.94</td>
</tr>
<tr>
<td>12.3</td>
<td>555</td>
<td>9.40</td>
<td>6062</td>
<td>3.26</td>
<td>578</td>
<td>5.72</td>
<td>1057</td>
<td>0.0048</td>
<td>0.0514</td>
<td>19</td>
<td>0.0514</td>
<td>7.78E-05</td>
<td>7.78E-05</td>
<td>49.7</td>
<td>7691</td>
<td>6283.34</td>
</tr>
<tr>
<td>14.7</td>
<td>782</td>
<td>5.80</td>
<td>591</td>
<td>3.68</td>
<td>357</td>
<td>4.68</td>
<td>2070</td>
<td>0.0028</td>
<td>0.0399</td>
<td>5</td>
<td>0.0399</td>
<td>3.57E-05</td>
<td>3.57E-05</td>
<td>44.4</td>
<td>3398</td>
<td>7507.7</td>
</tr>
<tr>
<td>5.20</td>
<td>298</td>
<td>5.61</td>
<td>382</td>
<td>5.11</td>
<td>357</td>
<td>4.68</td>
<td>2188</td>
<td>0.0023</td>
<td>0.0353</td>
<td>33</td>
<td>0.0353</td>
<td>3.60E-05</td>
<td>3.60E-05</td>
<td>63.9</td>
<td>8206</td>
<td>24561.2</td>
</tr>
<tr>
<td>13.3</td>
<td>0786</td>
<td>1.23</td>
<td>9404</td>
<td>4.79</td>
<td>3565</td>
<td>4.68</td>
<td>2675</td>
<td>0.0033</td>
<td>0.0513</td>
<td>33</td>
<td>0.0513</td>
<td>3.60E-05</td>
<td>3.60E-05</td>
<td>18.7</td>
<td>9004</td>
<td>27556.7</td>
</tr>
</tbody>
</table>
If estimated value of velocity of the link is less than original input value, risk will be generated at the node. Hence risk priority links are assigned as follows

**Table 3: Velocity risk analysis**

<table>
<thead>
<tr>
<th>S.No</th>
<th>Link name</th>
<th>Link Id</th>
<th>Speed after analysis</th>
<th>Initial Speed</th>
<th>Difference</th>
<th>Risk Priority</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Dilsuknagar - Chaithanyapuri</td>
<td>M1</td>
<td>32.00684</td>
<td>32.59</td>
<td>-0.58316</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>Chaithanyapuri – Kothapet</td>
<td>M2</td>
<td>33.54173</td>
<td>30.6</td>
<td>2.94173</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>Dilsuknagar – Sankeswar Bazaar</td>
<td>M3</td>
<td>32.36854</td>
<td>33.28</td>
<td>-0.91146</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>Chaithanyapuri – P&amp; T</td>
<td>M4</td>
<td>32.31442</td>
<td>31.56</td>
<td>0.754421</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>Sankeswar Bazaar - Saroornagar</td>
<td>M5</td>
<td>28.00714</td>
<td>34.61</td>
<td>-6.60286</td>
<td>1</td>
</tr>
</tbody>
</table>

If the estimated value headway of the link is more than original input value, risk will be generated at the link. Therefore risk priority links are assigned as given below

**Table 4: Headway risk analysis.**

<table>
<thead>
<tr>
<th>S.No</th>
<th>Link name</th>
<th>Link Id</th>
<th>Headway After analysis</th>
<th>Initial Headway</th>
<th>Difference</th>
<th>Risk Priority</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Dilsuknagar - Chaithanyapuri</td>
<td>M1</td>
<td>5.720395</td>
<td>2.93</td>
<td>2.790395</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>Chaithanyapuri – Kothapet</td>
<td>M2</td>
<td>5.793951</td>
<td>2.56</td>
<td>3.233951</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>Dilsuknagar – Sankeswar Bazaar</td>
<td>M3</td>
<td>4.48226</td>
<td>2.11</td>
<td>2.37226</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>Chaithanyapuri – P&amp; T</td>
<td>M4</td>
<td>4.686635</td>
<td>5.1</td>
<td>-0.41137</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>Sankeswar Bazaar - Saroornagar</td>
<td>M5</td>
<td>6.482675</td>
<td>4.5</td>
<td>1.982675</td>
<td>4</td>
</tr>
</tbody>
</table>

If the estimate delay of the link is less than original input value, risk will be generated at the link. Hence risk priority links are assigned as follows

**Table 5: Delay risk analysis.**

<table>
<thead>
<tr>
<th>S.No</th>
<th>Link name</th>
<th>Link Id</th>
<th>Delay after analysis</th>
<th>Initial Delay</th>
<th>Difference</th>
<th>Risk Priority</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Dilsuknagar - Chaithanyapuri</td>
<td>M1</td>
<td>17.00932</td>
<td>17.85714</td>
<td>-0.84782</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>Chaithanyapuri - Kothapet</td>
<td>M2</td>
<td>11.23669</td>
<td>9.722222</td>
<td>1.514468</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>Dilsuknagar – Sankeswar Bazaar</td>
<td>M3</td>
<td>8.725838</td>
<td>5.813953</td>
<td>2.911885</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>Chaithanyapuri – P&amp; T</td>
<td>M4</td>
<td>9.565714</td>
<td>6.060606</td>
<td>3.505108</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>Sankeswar Bazaar - Saroornagar</td>
<td>M5</td>
<td>19.73673</td>
<td>19.73673</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
References

LIST OF PUBLICATIONS

1. **Title:** Risk priority estimation for major corridors by using PCA applications of traffic mobility in urban areas; A case study of Gaddiannaram municipality area Hyderabad, A.P. India.
   **Journal Name:** International of Applied Engineering Research
   ISSN 0973-4562, Volume 6, Number5 (2011) pp.615-625

2. **Title:** GIS and PCA based Risk generation in traffic mobility in urban areas
   **Journal Name:** Institute of Engineers IE(I) Journal-CV
   Volume9, February 2011.

3. **Title:** GIS and PCA based risk generation in traffic mobility in urban areas
   **Journal Name:** JOURNAL OF DEFENCE TECHNOLOGIES, BUSHOFTU, ETHIOPIA,
   Volume 5, Number1 June 2010.