HORN-OK-PLEASE

An *Acoustic* Sensor Based Road *Congestion Detection* technique in *Developing Regions*

Rijurekha Sen, Bhaskaran Raman, Prashima Sharma
Department of Computer Science,
IIT Bombay
Why is solving this problem important?

“Congestion in Developing Regions”
Armenia
Bangladesh: Dhaka
Brazil : Sao Paolo
China
Egypt : Cairo
Pakistan : Karachi
Mexico
Nepal : Kathmandu
Sri Lanka : Candy
India: Delhi
India : Mumbai
India : Chennai
India : Bangalore
India : Hyderabad
Why is solving this problem interesting?

“Issues with Existing Solutions”
Intelligent Transport Systems (ITS)

- Infrastructure growth slow due to lack of funds & space, bureaucracy
- ITS uses technology to alleviate problems

State of the art in ITS

- **Fixed sensor based**
  - Sensors placed on or under road
  - *Eg. - Dual loop detector, Image sensor, Magnetic sensor*

- **Mobile sensor based**
  - Sensors placed in probe vehicles
  - *Eg. - GPS receiver, smartphone's accelerometer & microphone*

Challenges in developing regions

- ❌ High installation and maintainance costs
- ❌ Assumption of lane based system
- ❌ Assumption of low variability in vehicle speed
- ❌ *Low proliferation* of GPS and smartphones
- ❌ *Lack of incentive* in participatory sensing
- ❌ *Power drainage* issue of phones
- ❌ *Privacy* issues

No one technique can be used to solve an issue of this magnitude. Techniques have to be used in conjunction with one another, based on *applicability* and *ease of deployment*
Our Approach: Using *Acoustic* Sensors

Chaotic Traffic is Noisy!!

Can we exploit this??
System Architecture: Doppler Shift of Honks

**Envisioned Architecture**

- Raw or processed data from deployments on other roads
- Dynamic traffic info to mobiles (pushed/pulled)
- Honking zone of interest
- Recorder 1: 5 m
- Recorder 2: 20 m
- Local radio (optional): 802.15.4 or Bluetooth or 802.11

**Underlying Theory**

**Receding honk**

\[ f_1 = \frac{v}{(v + v_s)} \times f_0 \]

**Approaching honk**

\[ f_2 = \frac{v}{(v - v_s)} \times f_0 \]

**Vehicle speed**

\[ v_s = \frac{(f_2 - f_1)}{(f_2 + f_1)} \times v \]
Work Done (Jul-Dec, 2009)

Empirical Data Collection
- Are there enough honks on road?
- What is the honk frequency range?
- What is the average honk length?

Algorithm Design & Evaluation
- How to detect honks in the presence of road noise?
- How to match honks across the two recorders?
- How to extract $f_1$ & $f_2$ from a pair of matched honks?
- How accurate are our speed estimates?

Testing Applicability on Real Roads
- Do speed estimates from city roads represent traffic state?
- Can some metrics distinguish congestion from freeflow?
- Can we detect traffic state for individual directions on a bidirectional road?
- Will metric values for congested and freeflow be statistically different?
- Can we classify new traffic data into based on historical data? What is the classification accuracy?
- Can we detect the onset of congestion?

Experimental approach used all through
Hardwares and Softwares used

- Voice recorder of Nokia N79
- 16 KHz sampling frequency
- Mono channel
- 16 bit encoding
- Wav format
- Audio based synchronization

Empirical Data

3 hours of data = 18 clips of 10 mins each
manually detected 257 honks

- Honk frequency range – 2-4 Khz
- Average number of honks per clip - 30
- Honk length - CDF
Phone Synchronization

Method

- Audio based synchronization
- Square wave pattern generated in matlab
- Evaluation done across 70 pairs

Evaluation

Expected difference = 700 ms
Calculated difference = t1 ms
Error = |700 – t1| ms

Expected difference = 1 s
Calculated difference = t2 s
Error = |1 – t2| s

Part of recording to be clipped from recording1
Recording started early
Synchronization pattern in recording1

Recording started late
Synchronization pattern in recording2
Synchronization pattern in recording
Honk Detection

**Preprocessing**
1) Bandpassing
2) Windowing time
3) FFT computation

**Algorithms**
1) PeakVsAvgAllFreq
2) PeakVsAvgHonkFreq
3) PeakAbsAmp

**Postprocessing**
1) Length bounding
2) Honk merging

**Evaluation:**
- 257 honks from empirical data used
- An 8 ms window marked as **honk & not detected** -> fn++
- An 8 ms window marked as **non-honk & detected** -> fp++
- $\text{fn}/(\text{total number of honk windows}) \times 100 = \text{fn(\%)}$
- $\text{fp}/(\text{total number of non-honk windows}) \times 100 = \text{fp(\%)}$

<table>
<thead>
<tr>
<th>Stage</th>
<th>PeakVsAvgAllFreq</th>
<th>PeakVsAvgHonkFreq</th>
<th>PeakAbsAmp</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>22.3</td>
<td>0.2</td>
<td>9.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>43</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>18.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.3</td>
</tr>
<tr>
<td>length bounding</td>
<td>5.6</td>
<td>0.7</td>
<td>2.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>74.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1.04</td>
</tr>
<tr>
<td>honk merging</td>
<td>5.7</td>
<td>0.4</td>
<td>2.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>73.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>10.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1.01</td>
</tr>
</tbody>
</table>

**fp(\%) falls**
- 22.3%
- 5.6%
- 5.7%

**fn(\%) falls**
- 0.2%
- 0.7%
- 0.4%

**T = 10**
**T = 2**
Experimental Setup

Honk Matching

Criteria:
1) starttime_difference
2) duration_ratio

\[
\frac{d_1 f_1}{d_2 f_2} = \frac{d_2}{d_1} = \frac{f_1}{f_2} = \frac{v - v_s}{v + v_s}
\]

Start time difference should be 28-30 ms at 10 m and 57-59 ms at 20 m with sound speed 340-350 m/sec

Duration ratio should be 1 for stationary bike, which is not so. So only starttime_difference is used for matching.

To be done: Analyze cause
Frequency Extraction

2048 point FFT used (1024 point FFT for honks < 128 ms length)

Local maxima same after Doppler shift

Exchange of top two local maxima
How accurate are the speed estimates?
Setup

Speeds in Kmph

<table>
<thead>
<tr>
<th>Type</th>
<th>V12</th>
<th>V13</th>
<th>V23</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>STA</td>
<td>0.89</td>
<td>7.57</td>
<td>3.77</td>
<td>4.41</td>
<td>2.89</td>
</tr>
<tr>
<td>SLOW</td>
<td>11.2</td>
<td>14.79</td>
<td>7.53</td>
<td>11.17</td>
<td>3.63</td>
</tr>
<tr>
<td>SLOW</td>
<td>13</td>
<td>18.38</td>
<td>7.5</td>
<td>12.96</td>
<td>5.44</td>
</tr>
<tr>
<td>SLOW</td>
<td>14.88</td>
<td>18.38</td>
<td>11.3</td>
<td>14.85</td>
<td>3.54</td>
</tr>
<tr>
<td>SLOW</td>
<td>14.93</td>
<td>18.43</td>
<td>11.33</td>
<td>14.9</td>
<td>3.55</td>
</tr>
<tr>
<td>SLOW</td>
<td>16.82</td>
<td>18.43</td>
<td>15.16</td>
<td>16.8</td>
<td>1.64</td>
</tr>
<tr>
<td>MED</td>
<td>20.55</td>
<td>19.03</td>
<td>22.03</td>
<td>20.53</td>
<td>1.5</td>
</tr>
<tr>
<td>MED</td>
<td>20.56</td>
<td>22.05</td>
<td>19.01</td>
<td>20.54</td>
<td>1.52</td>
</tr>
<tr>
<td>MED</td>
<td>20.56</td>
<td>19.01</td>
<td>18.93</td>
<td>19.5</td>
<td>0.92</td>
</tr>
<tr>
<td>MED</td>
<td>22.39</td>
<td>22.81</td>
<td>21.99</td>
<td>22.4</td>
<td>0.41</td>
</tr>
<tr>
<td>MED</td>
<td>22.39</td>
<td>22.81</td>
<td>21.99</td>
<td>22.4</td>
<td>0.41</td>
</tr>
<tr>
<td>MED</td>
<td>23</td>
<td>15.62</td>
<td>25.19</td>
<td>21.27</td>
<td>5.01</td>
</tr>
<tr>
<td>MED</td>
<td>23</td>
<td>25.08</td>
<td>26.27</td>
<td>24.78</td>
<td>1.65</td>
</tr>
<tr>
<td>MED</td>
<td>24.22</td>
<td>22.81</td>
<td>25.58</td>
<td>24.2</td>
<td>1.38</td>
</tr>
<tr>
<td>MED</td>
<td>25.5</td>
<td>26.04</td>
<td>24.98</td>
<td>25.51</td>
<td>0.53</td>
</tr>
<tr>
<td>MED</td>
<td>25.5</td>
<td>20.83</td>
<td>29.98</td>
<td>25.44</td>
<td>4.57</td>
</tr>
<tr>
<td>HIGH</td>
<td>27.99</td>
<td>20.83</td>
<td>20.23</td>
<td>23.02</td>
<td>4.32</td>
</tr>
<tr>
<td>HIGH</td>
<td>27.99</td>
<td>20.83</td>
<td>34.83</td>
<td>27.88</td>
<td>7</td>
</tr>
<tr>
<td>HIGH</td>
<td>30.73</td>
<td>29.98</td>
<td>31.52</td>
<td>30.74</td>
<td>0.77</td>
</tr>
<tr>
<td>HIGH</td>
<td>33.36</td>
<td>34.97</td>
<td>31.66</td>
<td>33.33</td>
<td>1.66</td>
</tr>
<tr>
<td>HIGH</td>
<td>33.36</td>
<td>31.66</td>
<td>34.97</td>
<td>33.33</td>
<td>1.66</td>
</tr>
<tr>
<td>HIGH</td>
<td>33.36</td>
<td>29.98</td>
<td>36.93</td>
<td>33.42</td>
<td>3.48</td>
</tr>
</tbody>
</table>

Error Measures

- 30 honks
- 25 matched across all three
- stddev(|v12|,|v13|,|v23|)
- max(|v12-v13|,|v12-v23|,|v13-v23|)
- avg(|v12-v13|,|v12-v23|,|v13-v23|)
- |v12|/(|v13|+|v23|)*100
road_speed_vary

- 36 honks
- 4 lost due to annotation errors
- 26 matched across all three

<table>
<thead>
<tr>
<th>type</th>
<th>v12</th>
<th>v13</th>
<th>v23</th>
<th>mean</th>
<th>s.d</th>
</tr>
</thead>
<tbody>
<tr>
<td>sta</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>slow</td>
<td>9.50</td>
<td>8.75</td>
<td>11.46</td>
<td>9.9</td>
<td>1.4</td>
</tr>
<tr>
<td>slow</td>
<td>13.95</td>
<td>10.58</td>
<td>17.96</td>
<td>14.16</td>
<td>3.69</td>
</tr>
<tr>
<td>slow</td>
<td>13.95</td>
<td>10.58</td>
<td>17.24</td>
<td>13.92</td>
<td>3.33</td>
</tr>
<tr>
<td>slow</td>
<td>14.11</td>
<td>17.43</td>
<td>15.15</td>
<td>15.56</td>
<td>1.7</td>
</tr>
<tr>
<td>slow</td>
<td>14.11</td>
<td>17.44</td>
<td>10.71</td>
<td>14.08</td>
<td>3.37</td>
</tr>
<tr>
<td>slow</td>
<td>14.15</td>
<td>13.99</td>
<td>14.32</td>
<td>14.15</td>
<td>0.16</td>
</tr>
<tr>
<td>slow</td>
<td>14.15</td>
<td>13.99</td>
<td>14.32</td>
<td>14.15</td>
<td>0.16</td>
</tr>
<tr>
<td>slow</td>
<td>14.15</td>
<td>13.99</td>
<td>14.32</td>
<td>14.15</td>
<td>0.16</td>
</tr>
<tr>
<td>slow</td>
<td>15.30</td>
<td>12.39</td>
<td>17.29</td>
<td>14.99</td>
<td>2.46</td>
</tr>
<tr>
<td>med</td>
<td>17.49</td>
<td>17.24</td>
<td>17.74</td>
<td>17.49</td>
<td>0.25</td>
</tr>
<tr>
<td>med</td>
<td>19.10</td>
<td>17.63</td>
<td>17.95</td>
<td>18.23</td>
<td>0.77</td>
</tr>
<tr>
<td>med</td>
<td>21.04</td>
<td>18.93</td>
<td>17.84</td>
<td>19.27</td>
<td>1.63</td>
</tr>
<tr>
<td>med</td>
<td>22.00</td>
<td>24.27</td>
<td>21.60</td>
<td>22.62</td>
<td>1.44</td>
</tr>
<tr>
<td>med</td>
<td>22.00</td>
<td>24.20</td>
<td>21.54</td>
<td>22.58</td>
<td>1.43</td>
</tr>
<tr>
<td>med</td>
<td>22.96</td>
<td>20.75</td>
<td>21.80</td>
<td>21.84</td>
<td>1.11</td>
</tr>
<tr>
<td>med</td>
<td>24.60</td>
<td>24.14</td>
<td>21.47</td>
<td>23.4</td>
<td>1.69</td>
</tr>
<tr>
<td>high</td>
<td>26.42</td>
<td>24.14</td>
<td>28.80</td>
<td>26.45</td>
<td>2.33</td>
</tr>
<tr>
<td>high</td>
<td>26.42</td>
<td>27.51</td>
<td>28.80</td>
<td>27.57</td>
<td>1.19</td>
</tr>
<tr>
<td>high</td>
<td>29.00</td>
<td>34.19</td>
<td>28.80</td>
<td>30.66</td>
<td>3.06</td>
</tr>
</tbody>
</table>

Errors Measures

- Frequency (Hz)
- Time (secs)

Annotations
Wrong matches at 1 and 2

R2 does not detect honks at 7

Vehicle honks are **directional**

To be done:
- Improve matching algorithm
- Ascertain accurate speed ground truth
Can we apply these speed estimates on real city roads to detect congestion?
Road Experiments

- 18 hours of road data collection
- 2 different roads

- Different times of the day
- Different weather conditions

4.30 pm : Freeflowing

7.30 pm : Highly Congested

Adi Shankaracharya Marg (outside IITB, notorious for congestion)
Empirical Speed CDFs

Percentile speed < 10 Kmph: clearly distinguishes congested from freeflow.

High speed in congestion??

Zero speeds in freeflow??

70th percentile speed: clearly distinguishes congested from freeflow.

Freeflowing Traffic
Congested Traffic
(1) Speed estimates give traffic direction (2) Show freeflow and congestion in opposite directions on a normal day (3) On a rainy day, both directions show congestion, as rain causes vehicles to be slower.
### Congested vs Freeflowing : Metrics

#### 70th Percentile Speed (Kmph)

<table>
<thead>
<tr>
<th>Metric</th>
<th>Hira Congested mean (s.d) [24 samples]</th>
<th>Hira Free-flow mean (s.d) [54 samples]</th>
<th>Adi Congested mean (s.d) [27 samples]</th>
<th>Adi Free-flow mean (s.d) [27 samples]</th>
</tr>
</thead>
<tbody>
<tr>
<td>70th perc. speed (kmph)</td>
<td>12.2 (4.0)</td>
<td>18.2 (6.2)</td>
<td>7.7 (6.1)</td>
<td>21.1 (6.1)</td>
</tr>
<tr>
<td>Perc. speed &lt; 10 Kmph</td>
<td>65.6 (11.6)</td>
<td>51.1 (16.3)</td>
<td>79.5 (16.1)</td>
<td>37.6 (20.2)</td>
</tr>
</tbody>
</table>

**Speed based metrics**

#### Number of honks

<table>
<thead>
<tr>
<th>Metric</th>
<th>Hira Congested mean (s.d) [24 samples]</th>
<th>Hira Free-flow mean (s.d) [30 samples]</th>
<th>Adi Congested mean (s.d) [27 samples]</th>
<th>Adi Free-flow mean (s.d) [27 samples]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Num. Honks</td>
<td>113 (30.4)</td>
<td>55.5 (21.1)</td>
<td>149.4 (27.8)</td>
<td>57.6 (21.2)</td>
</tr>
<tr>
<td>Honk duration (sec)</td>
<td>45.1 (12.4)</td>
<td>21.8 (9)</td>
<td>71.5 (21.4)</td>
<td>21.7 (9.2)</td>
</tr>
</tbody>
</table>

**Non speed based metrics**

Noise level (db) can be used in Hira but not in Adi.
Statistical divergence of congested vs freeflowing data, based on all four metrics, is verified at 99.9% confidence using the Mann-Whitney U and two sample Kolmogorov-Smirnov tests.

<table>
<thead>
<tr>
<th>Metric</th>
<th>Mann-Whitney U test</th>
<th>Kolmogorov-Smirnov test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>Hira</em></td>
<td><em>Adi</em></td>
</tr>
<tr>
<td>70th perc. Speed</td>
<td>2.00E-006</td>
<td>7.48E-007</td>
</tr>
<tr>
<td>Perc. Speed &lt; 10 Kmph</td>
<td>1.05E-005</td>
<td>2.28E-004</td>
</tr>
<tr>
<td>Num. Honks</td>
<td>5.33E-015</td>
<td>2.13E-014</td>
</tr>
<tr>
<td>Honk duration</td>
<td>3.86E-014</td>
<td>3.89E-014</td>
</tr>
</tbody>
</table>

Statistical divergence of congested vs freeflowing data, based on all four metrics, is verified at 99.9% confidence using the Mann-Whitney U and two sample Kolmogorov-Smirnov tests.

Threshold based congestion detection

<table>
<thead>
<tr>
<th>Metric</th>
<th><em>Hira</em></th>
<th><em>Adi</em></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>*Fp (%)</td>
<td>*Fn (%)</td>
</tr>
<tr>
<td>70th perc. Speed</td>
<td>24.1</td>
<td>8.3</td>
</tr>
<tr>
<td>Perc. Speed &lt; 10Kmph</td>
<td>20.9</td>
<td>25.3</td>
</tr>
<tr>
<td>Num. Honks</td>
<td>10.7</td>
<td>17.4</td>
</tr>
<tr>
<td>Honk duration</td>
<td>7.1</td>
<td>19.6</td>
</tr>
</tbody>
</table>

Maximum false positive is 27.2% and maximum false negative is 25.3%
Continuous recording of road sound from 6 pm – 8 pm on 4th December, 2009, in Adi Shankaracharya Marg, showed transition from freeflow to congestion, based on all four metrics.

- Directionality of honks, road fork
- Non-speed based metrics show earlier rise than speed based
- Metrics to be used in conjunction (clip 10)
Ongoing Work
Sensor Platform Design

- **C5505 EZDSP** will sample and process road noise.
- Results will temporarily be stored in flash memory.
- Stored results will be sent to server over GPRS.
Enhancements

1. Reducing hardware
   - Mic + FM receiver with C5505 stick (bottom) instead of first phone.
   - Mic + FM transmitter (left) at 30 m instead of second phone.

2. Beyond acoustic sensor
   - Using RSSI and LQI variation, packet error rates between Zigbee tx-rx pair (bottom) across road as metrics to measure congestion.
   - Using magnetic sensor module SBT80 from Honeywell (right).

3. Beyond honks
   - Using characteristic road sounds other than honks.
   - Auto engine (left) and heavy vehicle sound (right).
Thank You

Questions??