<table>
<thead>
<tr>
<th>1. Name of the candidate</th>
<th>Ms. Riju Rekha Sen</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. Name of the Department/Center</td>
<td>Computer Science &amp; Engineering</td>
</tr>
<tr>
<td>3. Name of the Supervisor</td>
<td>Prof. Bhaskar Raman</td>
</tr>
<tr>
<td>Co-Supervisor</td>
<td>Prof.</td>
</tr>
<tr>
<td>4. Thesis Title</td>
<td>Different Sensing Modalities for Traffic Monitoring in Developing Regions</td>
</tr>
<tr>
<td>5. Name of the examiner, his/her designation and address</td>
<td>Venkat Padmanabhan Principal Researcher, Microsoft Research India, &quot;Vigyan&quot; #9, Lavelle Road, Bangalore - 560023</td>
</tr>
</tbody>
</table>

6. **Recommendation of the examiner (Mark appropriate box)**
   a) The thesis be accepted
   
   b) The thesis be accepted after clarification of the minor points listed in the report at the time of viva-voce
   
   c) The thesis be accepted after minor modifications/revision as suggested. After modification, the thesis need not be referred to me again
   
   d) The thesis requires major modifications/revision. The Nature of the modifications required are indicated. It is recommended that the revised thesis be referred again by an external examiner
   
   e) The thesis be rejected

Date: 10 Sep 2013

(Signature of Examiner)
7. Detailed report on the thesis: (Comment on the salient features of the work reported in the thesis and the basis of the recommendations - use additional sheets, if necessary.)

This thesis considers the important problem of monitoring road traffic in developing regions such as India. While there has been much prior work on traffic monitoring in developed regions with orderly, lane-based traffic, the developing region setting poses unique challenges which have received much less attention in the research literature. These challenges arise, both from the heterogeneous and chaotic nature of traffic and from the unaffordability of expensive roadside infrastructure. What makes the thesis stand out is the combination of an innovative approach that leverages various, inexpensive sensing modalities and extensive real-world experimentation and deployment to validate the ideas. Each of the ideas is examined in depth, with both the merits as well as the shortcomings pointed out. The work presented in the thesis has led to publications at highly-selective, top conferences, including ACM MobiSys and ACM SenSys, which also speaks to the quality of the work.

In summary, the work has all the elements of a solid PhD thesis — a compelling problem, novel and substantial ideas, and a thorough experimental evaluation. My main observation regarding the writing is that the three parts of the thesis come across as disconnected. A coherent story that motivates and stitches together the three parts is desirable.

8. Questions to be asked to the candidate at the time of the viva-voce examination. (use additional sheets, if necessary.)

1. Page 57: given that Doppler shift does not affect the amplitude of a signal, what is the explanation for the highest peak recorded at one node to correspond to the second highest one at another node?

2. Pages 70-72: even if the difference in the speed distributions for the congested and free-flowing traffic states is statistically significant, the fact remains that the two distributions overlap heavily. Would this not be problematic with regard to classification accuracy?

3. What is the advantage of Kyun Queue over a vision-based system, especially considering the elaborate setup needed for the former? Under what circumstances would you recommend the former?

4. Given the low height at which the Kyun Queue sensors are mounted, these would seem susceptible to local crowding in the vicinity of one or more sensors, unrelated to road traffic conditions. How does this affect the robustness of Kyun Que?

5. How would road-specific calibration be made practical for the purposes of honk-based speed estimation and RSSI-based traffic estimation?

(Signature of Examiner)
use this audio based approach for synchronization because we have to record anyway.

Other synchronization mechanisms like Wi-Fi, Bluetooth or GPS based techniques, which is available on the N-79, will need extra effort with no added advantage. In actual

ments, with automated data collections, no common sound playing mechanism will be

Then we will need to design a different synchronization mechanism, maybe based on

audio like Wi-Fi, Bluetooth or Zigbee.

**Synchronization error**

Error in synchronization is the error in detecting the start of the common sound. We seek
tify this error in the following way. The problem of detecting the start of the common
same as the problem of detecting the start of any of the 10 square waves. We know
ected difference between the start times of two consecutive square waves. We take the

e value of the difference between this expected value and the calculated value as error

![Diagram](image)

**Figure 3.5: Method to calculate synchronization error**

We do the above procedure for 70 pairs of square waves in eight different recordings. The
s are shown in Fig. 3.6. The maximum error is of concern here which is 62 or 63 microsecs.

ity of the errors are zeros.
Honk matching

Honk matching can be done independently by each recorder. After detection, the same honk matched across the two recordings. In our honk-matching step, we also seek to match only honks in the “zone of interest” (Fig. 3.1).

TimeDiff: For two honk windows $h_1$ and $h_2$, at recorders R1 and R2 respectively, originated from the same honk, within the zone of interest, the difference between the of $h_1$ and $h_2$ must be bounded. For instance, in Fig. 3.1, suppose the honking vehicle of $x_1$ and $x_2$ respectively from the two recorders, when it starts honking. And if the within the zone of interest at this time, then $|x_1 - x_2| < 20 m$. So ideally, the start and $h_2$ must differ by not more than $D = \frac{20}{v}$, where $v$ is the speed of sound.

RunRatio: This criterion bounds the ratio of honk durations in the two recorders to be ideally, if $d_1$ and $d_2$ are the honk durations at recorders R1 (honking vehicle receding) and R2 (honking vehicle approaching this recorder) respectively, $d_1 f_1 = d_2 f_2$, number of wavelengths (lambda) seen by both the recorders is the same (also same number of wavelengths generated at source). So, $\frac{d_2}{d_1} = \frac{F_2}{F_1} = \frac{\frac{20}{v} + l_x}{\frac{20}{v} - l_x}$ where $v$ is speed of vehicle. Since $v$ is fixed, this ratio will decrease with increasing $v_s$. 

Figure 3.8: Equidistant peaks
Since \((f_1 - f_2)\) is negative, in the former case, we are under estimating speed and in the latter case, over estimating it. The lower the value of \(N\), higher is the value of \(n\) and more is the error in speed calculation. So here we should use high value of \(N\). The corresponding decrease in time resolution is not an issue now, as we have already detected the honk windows precisely. So we can compute a high point FFT on the honk windows to extract \(f_1\) and \(f_2\) as precisely as possible.

![Figure 3.12: Effect of FFT points on speed](image)

**N to use?** - In choosing high value of \(N\), we need to fulfill two criteria mandated by FFT computation – (a) \(N\) should be a power of 2 and (b) each time window passed the FFT computation algorithm should have \(N\) samples. If we choose \(N = 4096\), we find time window of 256 ms as our sampling frequency is 16 KHz. From Fig.3.7, about 5 of the honks in each sound clip is less than 250 ms in length, so we will have very few honks with 256 ms time window. Hence we choose \(N = 2048\), which needs 128 ms window. If a honk has more than one 128 ms windows, then we do 2048-point FFT on each individual window and average out the amplitudes of each frequency across the multiple windows. According to Section 3.4, the minimum honk duration for us is 112 ms. So for the few honks with duration 112 ms or 120 ms (our honk duration always is a multiple of 8 as detection uses time window of 8 ms), we will use \(N = 1024\).

**Bays_peak** - Spectrum of several pairs of matched honks in *Audacity* show local peaks remain similar across matches (see Fig. 3.13). Based on this observation, we argue that if a certain frequency has highest amplitude in a honk and another frequency has
y is all that is needed to remove this second factor.

**Prototype design**

Discuss the custom hardware prototype that we built for our acoustic sensing system. Our sensor based congestion detection technique, that this prototype has to implement, has the following requirements - (1) **Sensing** - sampling road noise, (2) **Processing** - filtering of the noise, detecting honks, matching honks between R1 and R2 after time synchronization of speeds and (3) **Remote communication** - sending various metric values like honks and speeds to remote server.

**Design choices**

The choices available to us for Recorder1 (R1) and Recorder2 (R2), as shown in Fig. 3.1 above, are:

- Both R1 and R2 having sensing and remote communication capabilities. Sampled raw data will be sent to a central server which will do the processing. The primary disadvantage is that, the amount of data to be communicated to the central server will be huge, as the central server would need only the honk related information. This will unnecessarily increase communication cost and delay.

- Only R1 and R2 having sensing, computation and remote communication capabilities. Sampled metric values will be sent to the server. The two units will also need local computation capabilities to communicate between each other, as honk matching and speed synchronization cannot be done individually. This seems plausible but has scope of further optimization in the next design choice.

- Only sensing and R2 having sensing, computation and remote communication capabilities. The units will have local communication capabilities for R1 to send sampled data to R2 over a small inter-sensor distance. R2 will do processing and remote communication. This choice seems good, provided we can handle the various technical challenges.

The quality of audio signal should not be affected during communication from R1 to R2.
Real-time classification of traffic states

Above, wireless logs of 5 minutes duration show good visual difference in the distribu-
tive wireless characteristics, between free-flowing and congested traffic. Applications like
don maps and bottleneck detection can be handled at a time scale as large as 5 minutes,
traffic light control would intuitively need faster inputs. From our observations
of Chennai and Bengaluru traffic lights, traffic signal cycles typically last for about a minute.
A minute cycle time is divided into slots, in which different contending flows get their
green times. Green time for any flow lasts for about 10-30 secs, though it can go over
for critical flows.

Figure 4.4: Mismatch of sensing and traffic state

A system like ours, aiming to provide traffic state information to traffic lights, would
require a parameter about the frequency at which traffic queue estimates are needed. We
termined the lowest classification time window at which a sufficiently high accuracy
can be obtained. Very low time windows give noisy predictions. This noise comes
from several sources - (a) the inherent stochastic nature of wireless links which causes link quality
to be inherently bad though the tx-rx are in perfect line of sight (b) the instantaneous traffic
between the tx-rx are contrary to the actual traffic state. For e.g. tx-rx may be in line
with several standing vehicles in congestion (Fig. 4.4(a)) or several heavy vehicles
blocking tx-rx pair simultaneously in free-flow obstructing line of sight (Fig. 4.4(b)). Thus
we chose the classification time window, henceforth referred to as \( t \), carefully.
than using additional percentile features based on LQI and PRR. The reason for this
positive observation is that RSSI is much more strongly correlated with line-of-sight than
PRR.

This strong correlation of RSSI with line-of-sight is also evident from an experiment that
we perform on the narrow road. The goal of the experiment is to observe the effect of distance
between transmitter and receiver, on the measured wireless link characteristics. The experiment
is shown in Fig. 4.8. Directly measuring $d'$ is difficult on a busy road with vehicles
constantly moving. So we measure $d''$.

![Figure 4.8: Sensing on narrow road](image)

By keeping the transmitter fixed, we move the receiver from $d''=5m$ to $d''=40m$ in steps of
5m on the receiver log for 5 minutes at each position. We do this both in free-flowing and
congested conditions. The CDF of reception and RSSI, for both traffic states, are shown in
Figs 4.9 and 4.10 respectively. LOS indicates line-of-sight condition in free-flow and NLOS,
no-line-of-sight condition in congestion. We show the plots only for 15m, 20m, 25m and 30m
distances to avoid the figures from getting cluttered.

![Figure 4.9: CDF of PRR vs. $d''$](image)

![Figure 4.10: CDF of RSSI vs. $d''$](image)

As can be seen, both free-flow (LOS) and congested (NLOS) traffic show almost identical