

---

Department of Computer  
Science and Engineering,  
IIT Bombay

# Department Overview

---

- Premier Department of Computer Science and Engineering in India.
  - 39 full-time faculty
  - 87 students to the B.Tech program/year
  - 115 students to the M.Tech program/year
  - Ph.D. strength is about 67 students
  - Students are high rankers in the qualifying exams and admission tests.

# Areas of Research

---

- Graph Theory and Algorithms
- Natural language processing
- Machine Learning
- Database management, Data mining
- Formal verification
- Graphics and Vision

# Areas of Research

---

- Networking and Internet technologies
- Telecommunications
- Operating Systems
- Programming languages
- Computational Biology
- Security
- Forecasting

# Academic Highlights

---

- Between 80-100 research publications a year.
- Several text books: Databases, Architecture, Compilers, Operating Systems, Security.
- Several Reference Books: Web Mining, Compiler Optimization, Networks.
- Some of the most highly cited authors from India are from CSE, IITB.
- Running projects of ~ Rs 100 Crores sponsored by MHRD, DIT, DST, DAE, HP, IBM Research, Infosys, TCS, WIPRO, Yahoo.

# Awards and Honors

---

- ACM and IEEE Fellow
- Fellow of INAE, IASc
- Fellow of American Association for Advancement of Science
- Young Scientist Award from the Indian National Academy of Engineering
- Swarnajayanti Fellow.
- DAE-SRC outstanding research investigator award
- Young Scientist Award from BRNS
- Distinguished Alumnus Award from IIT Madras
- Career Award from AICTE
- Several IBM Faculty Awards
- On editorial boards, program committee chairs of conferences

# Facilities

---

- HP Cluster for research on data intensive problems
- Coming up: A GPU cluster for research on parallel computing, bio-computing.
- Coming up: A Cloud Collaboratory for research on cloud computing
- Graphics lab with high end workstations
- Embedded systems lab with instructional robots

---

# Service Initiatives



# Spreading Education

---

- Contributor to the NPTEL and CDEEP initiatives
  - Development of video and web based courses in Engineering.
- National Mission on Education Through ICT
  - Enhancement of Engg College Teachers
    - Geographically distributed interactive workshops over 45 remote centres
  - Development of learning aids
- Short courses for industry and academia offered in-house and via distance learning modes.
- Text books on computers for school children

# Technology in local use

---

- Rural area networking
- aAqua
  - Multi media community forum for discussions on agricultural issues
- ICT in education of Nomadic tribes
  - Mobile classrooms, Anchored Instructions
- Digital library of agricultural documents
  - With provisions for cross-lingual search
- Mumbai Navigator: travel planner for Mumbai

# Contributions to Open Source Software

---

- GNU Compiler Collection
  - Documentation of GCC Internals
  - Annual workshop on GCC as test-bed for compiler teaching and research

# Contact:

---

- Department of Computer Science and Engineering, IIT Bombay, Powai, Mumbai 400076.
- [www.cse.iitb.ac.in](http://www.cse.iitb.ac.in)
- (91) 22 25767901

---

# Research Highlights

# Algorithms and Complexity

---

## Members

- Sundar Vishwanathan
- Ajit A Diwan
- Abhiram Ranade
- Milind Sohoni
- Bharat Adsul
- Nutan Limaye

# Algorithms and Complexity

---

- Graph Theory and Combinatorics
  - Study of graphs and other discrete structures as mathematical objects.
- Complexity Theory
  - GCT: Attempt to solve the most well-known open problem in CS through techniques from mathematics
  - Language classes in P: Solution of problems under resource bounds.
- Algorithms Design under various constraints
  - Online algorithms, Algorithmic game theory, Parallel Algorithms, Approximation algorithms, algorithms in the streaming model.

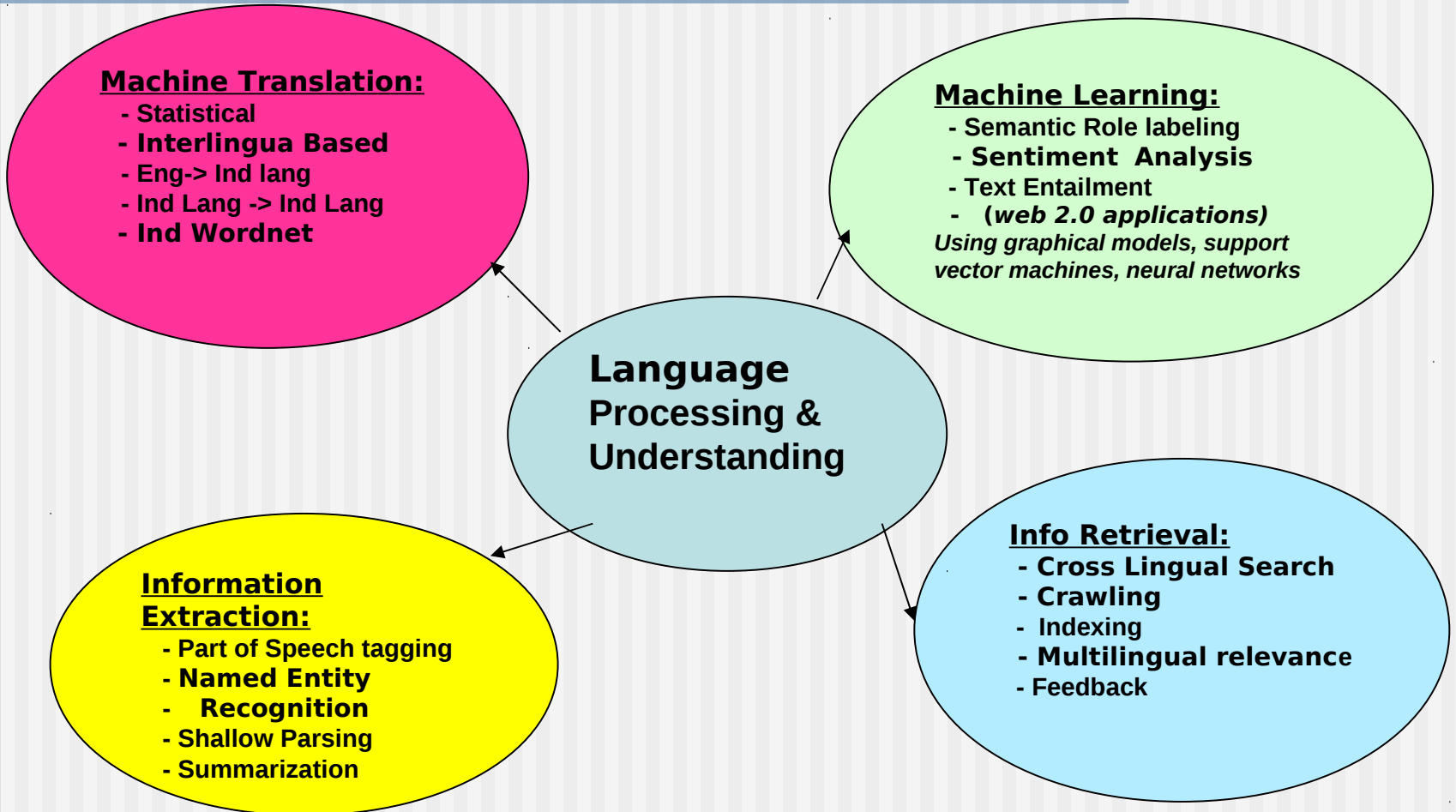
# Natural Language Processing

---

- Members
  - Pushpak Bhattacharyya
  - Saketha Nath
  - Ganesh Ramakrishnan
  - Malhar Kulkarni (HSS)



# Natural Language Processing



# Natural Language Processing

---

- *Currently* 4 faculty members, 22 graduate students, 9 research staff
- 14 sponsored research projects in last 5 years (amounting to 8 crores):
- Major achievements:
  - Hindi wordnet (Patwardhan award, Mathan Award)
  - Technology Transfer to Google
  - Crowdsourcing based Machine Translation in collaboration with Xerox
  - Leadership in large consortium based national projects on Machine Translation and Cross Lingual Search: publicly used portals: Ministry of IT funded
  - New directions in pan-Indian language sense diambiguation

# Data Management Research Group

---

## Faculty

- Soumen Chakrabarti
  - Saketa Nath
  - D. B. Phatak
  - Ganesh Ramakrishnan
  - 8 PhD students
  - Numerous MTech students
- Krithi Ramamritham
  - Sunita Sarawagi
  - N. L. Sarda
  - S. Sudarshan

# A few of our Research Areas

---

- Intelligent Information Retrieval
  - Information Extraction
  - From unstructured to structured data
  - From text to entities
  - From organizational data integrated with Web data
- Data mining
- Database systems
  - Query optimization
  - Database application development tools
- Geographical Information Systems

# Intelligent Search

- Question answering using (some) semantics

- E.g. **How many students are there in IIT Bombay?**

<NUMEX TYPE="ORDINAL">

selector match

WN Lin similarity

**IIT Bombay**, which has **5000 students**, is located in Powai, ...

- Which universities are known for Web search?
  - Step 1: identify “entities” in Web pages
  - Step 2: aggregate information from multiple Web pages, and present ranked answers
- Search on organizational data
  - Which professors are experts on Web search?

# Intelligent Search Technologies

---

- Text indexes on data
  - Standard for Web search
- Graph indexes on linked data (in-memory)
- Indexes on entity types
  - E.g. person, university, ...
- Handling Web-scale data
  - Indexing and search on large clusters of machines

# Querying the Semi-Structured Web

## Table

Ouagadougou	 Burkina Faso
Pago Pago	 American Samoa
Palikir	 Federated States of Micronesia
Panama City	 Panama
Papeete	 French Polynesia
Paramaribo	 Suriname
Paris	 France

## Regular page

### Airports in Germany

#### Berlin-Tegel Airport

Berliner Flughafen-Gesellschaft mbH, Flughafen Tegel, 13405 Berlin, Germany

**Airport Code:** TXL

#### Cologne - Bonn Airport

Postfach 98 01 20, 51129 Cologne, Germany

**Airport Code:** CGN

#### Munich Airport

PO Box 23 17 55, 85326 Munich, Germany

**Airport Code:** MUC

#### Hamburg Airport

Flughafenstrasse 1-3, 22335 Hamburg, Germany

**Airport Code:** HAM

## List

1. Gulf War oil spill, Persian Gulf, January 23 1991
2. Ixtoc oil well, SGulf of Mexico, June 3, 1979
3. Nowruz oil field, Persian Gulf, February, 1983
4. Atlantic Empress and Aegean Captain collision, off Trinidad and Tobago,
5. Castillo de Bellver, off Cape Town, South Africa, August 6, 1983
6. Amoco Cadiz (BP/Amoco, USA) - Brittany, France, March 16 1979
7. Torrey Canyon, South England, March 18 1967
8. Sea Star, Gulf of Oman, December 19, 1972
9. Urquiola, La Coruna, Spain, May 12, 1976
10. Hawaiian Patriot, N Pacific February 26, 1977
11. Othello, Tralhavet Bay, Sweden, March 20, 1970

## Formatted list

- [Braer - Shetland Islands, January 5, 1993](#)
- [Prestige - Galicia, Spain, November 13, 2002](#)
- [Aegean Sea, off N Spain, December 3, 1992](#)
- [Sea Empress - Wales, February 15, 1996](#)
- [World Glory, off South Africa, June 13, 1968](#)
- [Corinthos Delaware River, Marcus Hook, Pennsylvania, January 31, 1975](#)
- [Burmah Agate Galveston Bay, Texas, November 1, 1979](#)
- [Exxon Valdez \(Exxon, USA\) - Prince William Sound, Alaska, March 24, 1989](#)
- [Keo, off MA, November 5, 1969](#)
- [Storage Tank, Sewaren NJ, November 4, 1969](#)
- [Ekofisk oil field, North Sea April 22, 1977](#)
- [Erika - Bay of Biscay, December 12, 1999](#)
- [Tasman Spirit, Karachi, Pakistan, July 28, 2003](#)

# World Wide Tables Search

- Query by content or by description

**Alan Turing**   **Turing Machine**

**Inventor**   **Computer science concept**

**E. F. Codd**   **Relational Databases**

- Answer

<b>Person</b>	<b>Concept/Invention</b>
Alan Turing	Turing Machine
Seymour Cray	Supercomputer
E. F. Codd	Relational Databases
Tim Berners-Lee	WWW
Charles Babbage	Babbage Engine

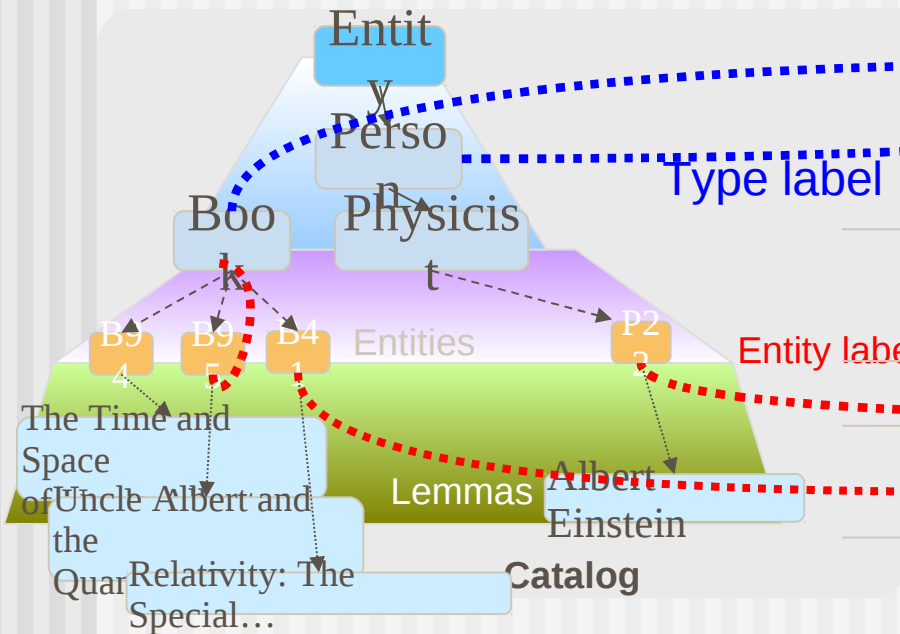


# Entity Recognition and Annotation

Relation label

**Writes**(Book,Person)  
 bornAt(Person,Place)  
 leader(Person,Country)

Type hierarchy



Type label

Entity label

Title	Author
Petros and the Goldback conjecture	A Doxiadis
Albert and the Quantum Quest	R. Stannard
Relativity: The Special and the General Theory	A Einstein

# A Few Other Projects

---

- Support for database applications
  - How to generate data to “kill query mutants”
    - i.e. to differentiate erroneous queries from correct ones
  - Holistic optimization of database applications

# Group Strengths

---

- Research recognition
  - Program committee chairs of top conferences
    - WWW 2010, KDD 2009, VLDB 2011, WSDM 2008
  - Editorial board of leading journals
    - ACM TODS, ACM Trans. Web, IEEE TKDE, IEEE DMKD, Found. And Trends in IR, ...
  - Regular publication in top int'l conferences
    - VLDB, ACM SIGMOD, IEEE ICDE, KDD, WWW Conference, ...
  - Regularly on PCs of these conferences
- Authors of textbooks used internationally
- Consultants for leading companies
  - NSDL, Microsoft, Yahoo, SBI, ...

# Formal Methods

---

- **Members**
  - G. Sivakumar
  - Supratik Chakraborty
  - Krishna S.
  - Paritosh Pandya (TIFR)

# Formal Methods

Most activities undertaken in Centre for Formal Design and Verification of Software (CFDVS)

- Collaborative setup with TIFR, BARC, BRNS

## High-level categorization of activities



# Formal Methods

---

## Accomplishments

- Publications in top conferences and journals, contributions to book chapters
- Tools for practical formal verification
  - DCUPPAAL, DCCHECK, SplnE, DAGGER, QuaLMoDE, ...
- Formal verification projects for govt & private agencies
  - ADA, ISRO, DRDL, BARC, Intel, General Motors, Texas Instruments, Microsoft Research, TCS, ...

# Formal Methods

---

- **The Road Ahead**

BRNS funding of Rs. 3+ crore for 5 years (from Apr 2011)  
Focus on improving precision and scalability of reasoning about systems with very large ( $> 2^{1000}$ ) or infinite state spaces

Theoretical and engineering aspects

Potential applications in formal verification, systems biology, ...

# Graphics and Vision

---

- Members
  - Sharat Chandran
  - Parag Chaudhuri



# Graphics and Vision

---

## ■ Indian Digital Hampi Project:

- Creation of 3D virtual visual and auditory reality from photographs, laser scans, archives, paintings, murals, video recording

## ■ National project on perception engg:

- Perception based multimedia video-conferencing system capturing emotions.

## ■ Healthcare Initiative Consortium:

- Cancer prognosis through feature extraction from pathological data

## ■ Video Mining:

- Automatically dividing a video recording of into a sequence of meaningful events.

# SYNERG Systems and NetwoRks Group

Varsha Apte  
Umesh Bellur  
Kameswari Chebrolu  
Sridhar Iyer  
Purushottam Kulkarni  
Bhaskaran Raman  
Anirudha Sahoo

*Crossing boundaries to connect*



# Cognitive Radio Based Wireless Network

## Project

---

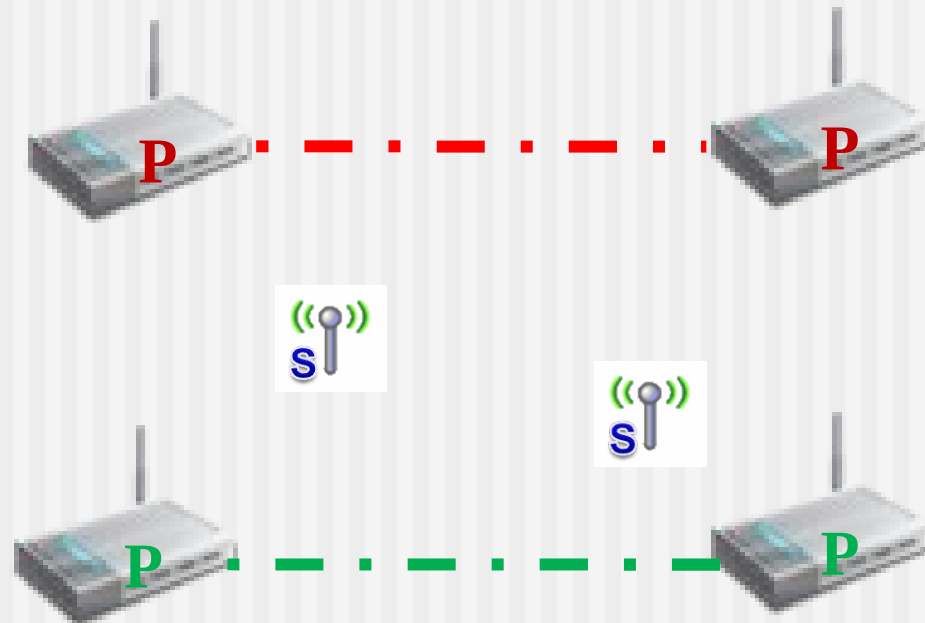
**Aim: Develop theoretical model which can be used by secondary wireless devices to opportunistically access wireless channel when it is not used by primary devices**

**PhD Student: Manuj Sharma**

---

## Motivation:

- 1) Scarcity of wireless spectrum
- 2) under-utilization some of the allocated spectrum



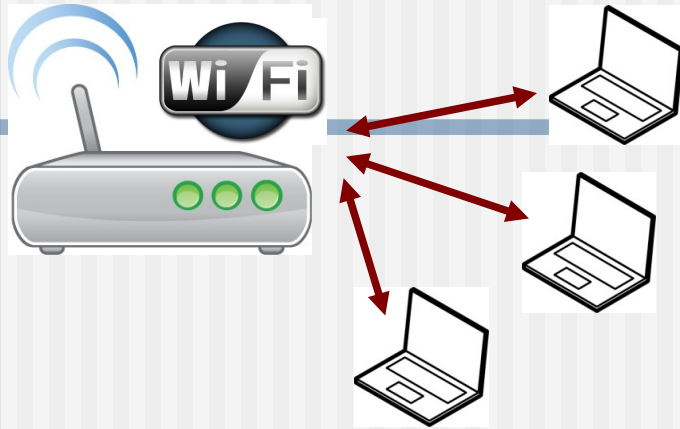
# The FRACTEL Project

**FRACTEL: wiFi-based Rural data ACcess  
and TELephony**

**Aim: To provide low cost broadband  
connectivity in rural regions**

**PhD Student: Vishal Sevani**

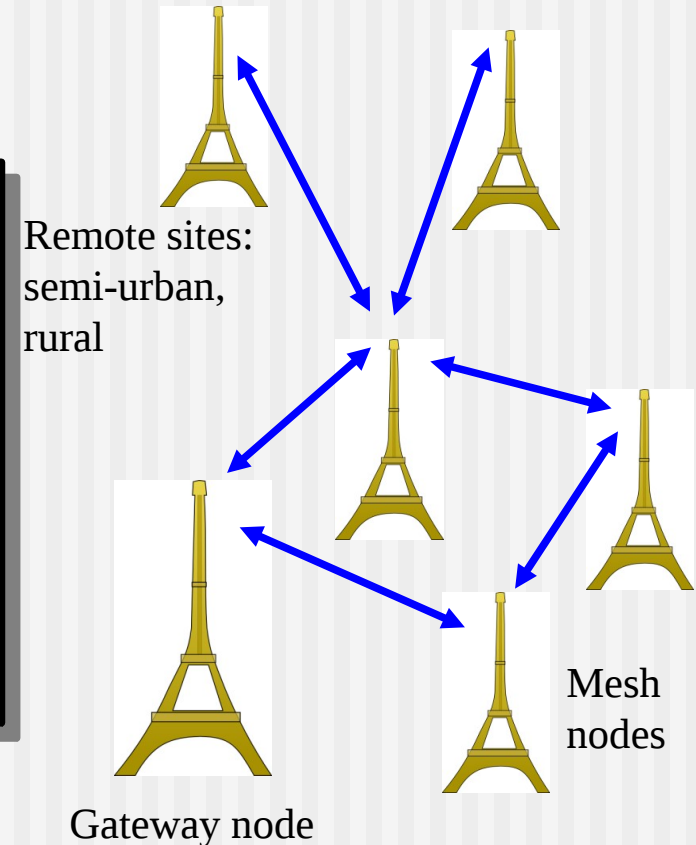
# FRACTEL: Mesh Network



- Inexpensive radio: <\$50
- 2.4 Ghz license-free band
- **But, designed for indoor**

## → Practical & technical contributions:

- *Outdoor, long-distance mesh operation*
- TDMA MAC: **MADWIFI** modifications
- Tested on single links with TTSL
- Next step: work on **ath9k** driver + low-cost hardware platform



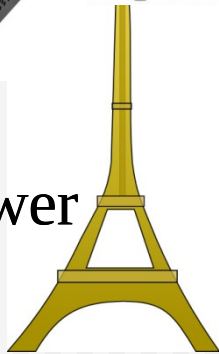
# A Long-Distance FRACTEL Link

RB433AH  
with FRACTEL  
MAC protocol

Parabolic dish  
antenna



Tower



Long-distance  
wireless link: up to 10s of kms



# The CARTS Project

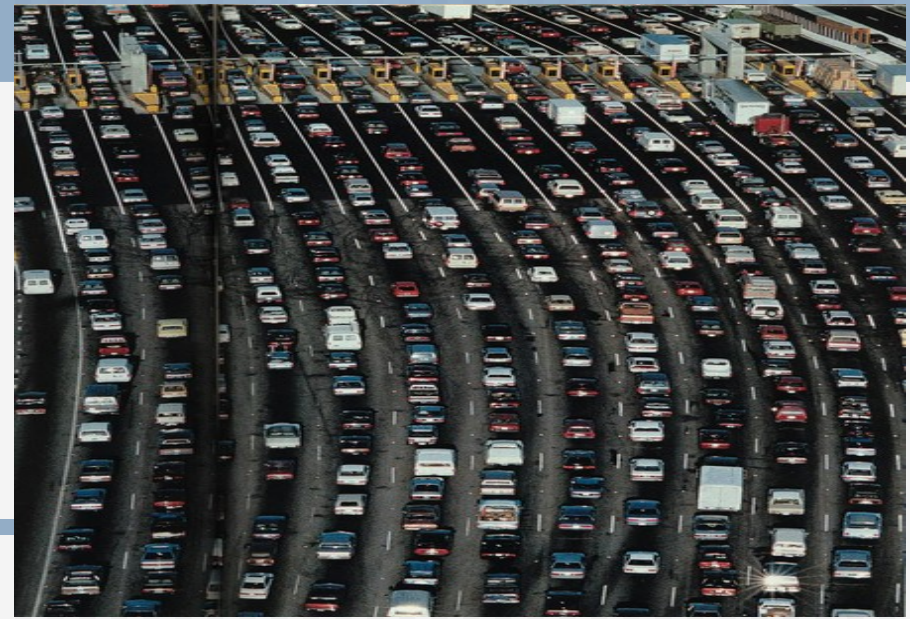
**CARTS: Communication Assisted Road  
Transportation Systems**

**Specific work: WirelessAcrossRoad**

**Aim: Develop intelligent transportation  
techniques for chaotic Indian roads**

**PhD Student: Rijurekha Sen**





**Traffic is a real pain !**



# Existing Techniques: not for Indian Roads

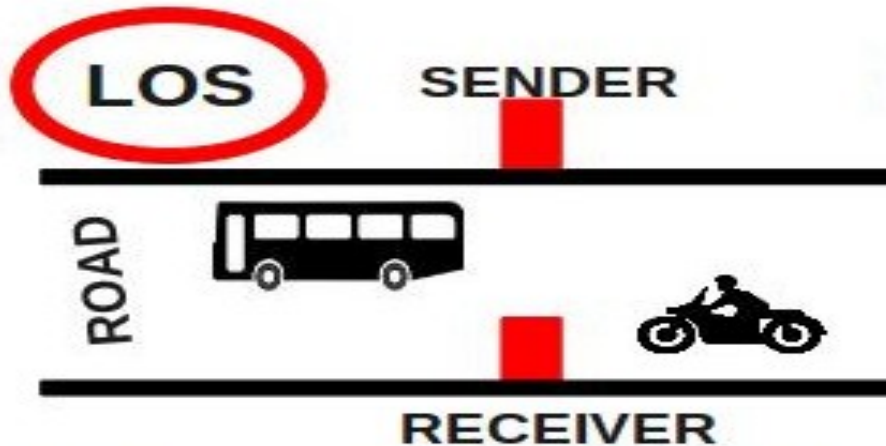
## Existing Solutions

- expensive (e.g., dual loop detectors)
- some of them will not work in non-lane based traffic

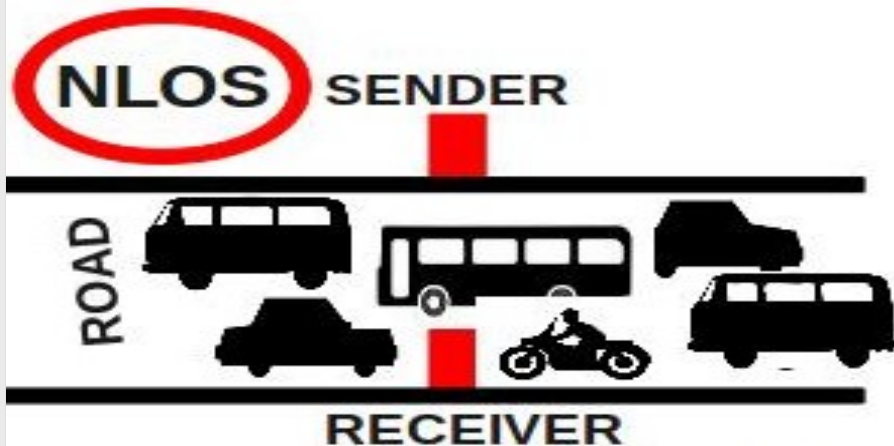
## Our solution

- Cost effective (uses inexpensive wireless devices)
- Works on non-lane based traffic
- No disruption during installation

# Wireless Across Road



(a) Free-flowing traffic



(b) Congested traffic

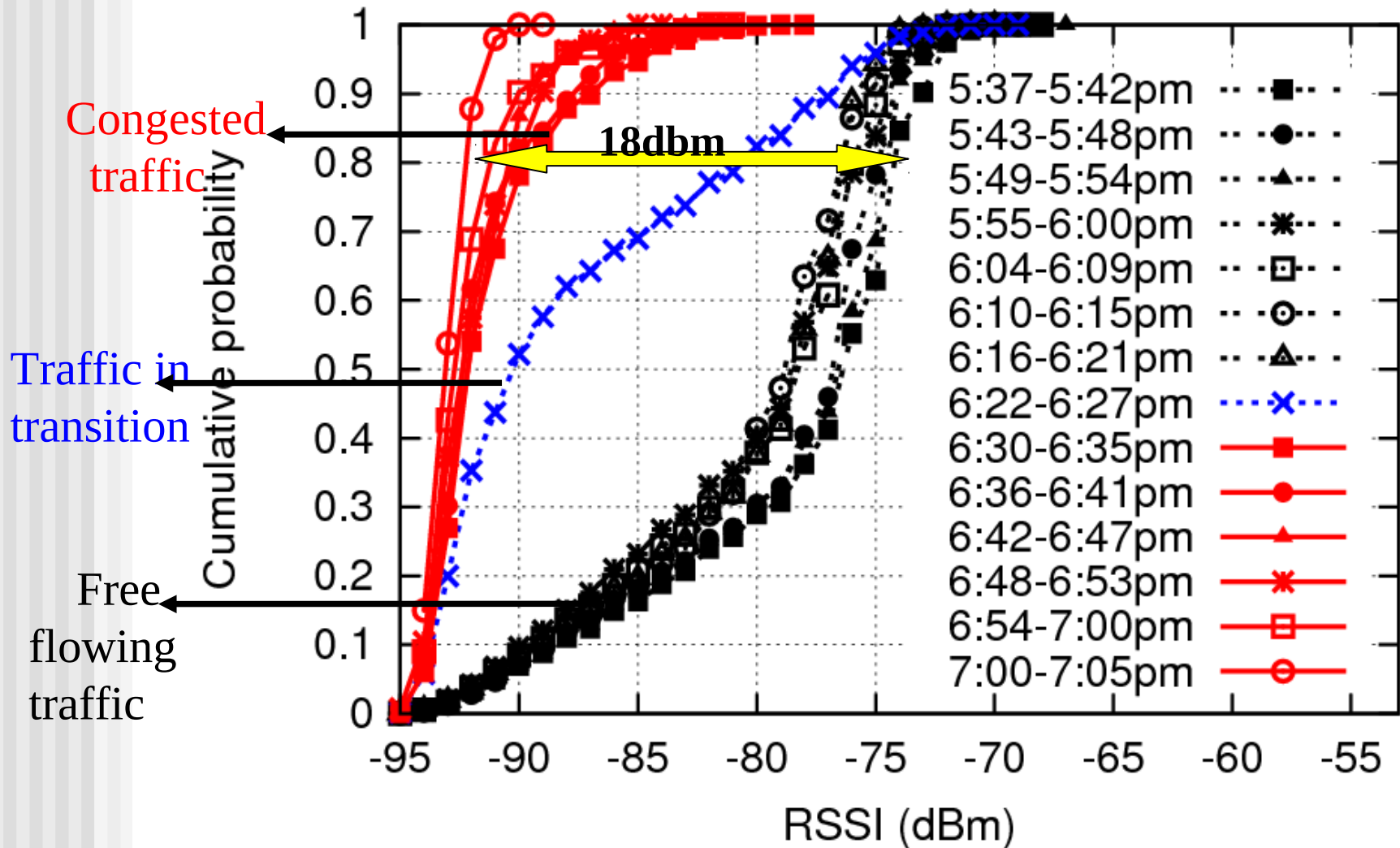
→ *Hypothesis:*  
→ *Wireless link characteristics =  $F(\text{road traffic})$*

→ Can we **classify** road traffic state?

→ Estimated cost:

~INR 5K

# Link Signal Strength & Traffic State



Ongoing work: automated classification, queue length estimation

# The Lo<sup>3</sup> Project



**Lo<sup>3</sup>: 802.15.4-based  
Low-Cost,  
Low-Power,  
Local Voice and  
Messaging**

**Aim: To provide low cost, low power voice  
connectivity in villages**

**PhD Student: Vijay Gabale**

# Rural India: GSM/CDMA Inadequate



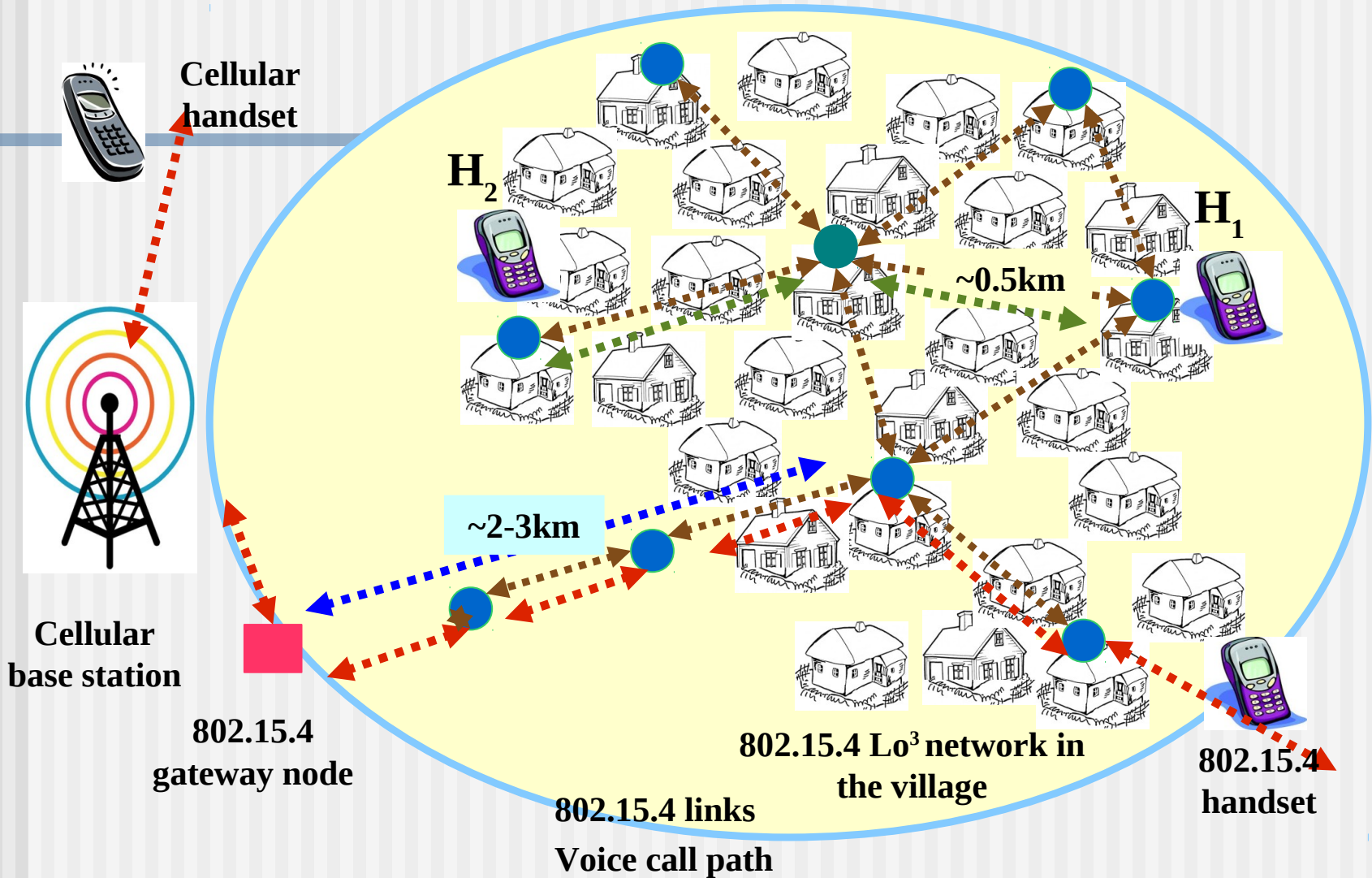
Rural teledensity in India = 22%  
(DoT report Jul. 2010)

Electrification of India  
40%+ households lack proper  
electricity connection (< 4 hrs/day)

Several “no/weak coverage”  
areas in rural-India.

**GSM/CDMA base-stations: high cost \$10-50K, high power consumption (0.5-1KW), huge power backup, backhaul connectivity is costly**

# Lo<sup>3</sup>: Local Voice, Voice Through Gateway



**Low Cost, Low Power WIRELESS INTERCOM/PABX**

# Technical Issues Addressed

- TDMA based MAC protocol design (DEV'10)
- Scheduling for voice calls (COMSNETS'11)
- Pipelined data transfer (SenSys'10)

## Pilot Test: Ahupe, near Pune





# Programming Languages

---

- Members
  - D. M. Dhamdhere
  - Supratim Biswas
  - Amitabha Sanyal
  - Uday Khedker
  - Rushikesh Joshi

# Programming Languages

---

- Program Analysis
  - Optimization of programs
  - Automatic Parallelization of programs
  - Data race detection for parallel programs
  - Validating correctness of compilation
  - Dynamic updation of programs in execution
- GCC Resource Centre
  - Test bed for compiler research

# Operating Systems

---

- Cloud computing
- Real Time Operating Systems (RTOS)
  - For safety critical systems.
- Fair scheduling in Operating Systems

# Security

---

- Members
  - Bernard Menezes
  - G. Sivakumar

# Security

---

- Optimization of Elliptic Curve Cryptographic (ECC) operations
- Web Application Security (New generation attacks and defenses)
- Intrusion Detection
- Malware obfuscation techniques
- Secure Group Communications
- Security Protocols
- Network and System security