CS 525M Mobile & Ubiquitous Computing

EmotionSense:

A Mobile Phones based Adaptive Platform for Experimental Social Psychology Research

Rachuri K., Rentfrow P., Musolesi M., Longsworth C., Mascolo C., Aucinas A.

Mike Shaw

Computer Science Dept.

Worcester Polytechnic Institute (WPI)

OUTLINE

- Motivation
- Related work
- Goal
- Assumptions & limitations
- Methodology
- Benchmarking
- Results
- Future work

MOTIVATION

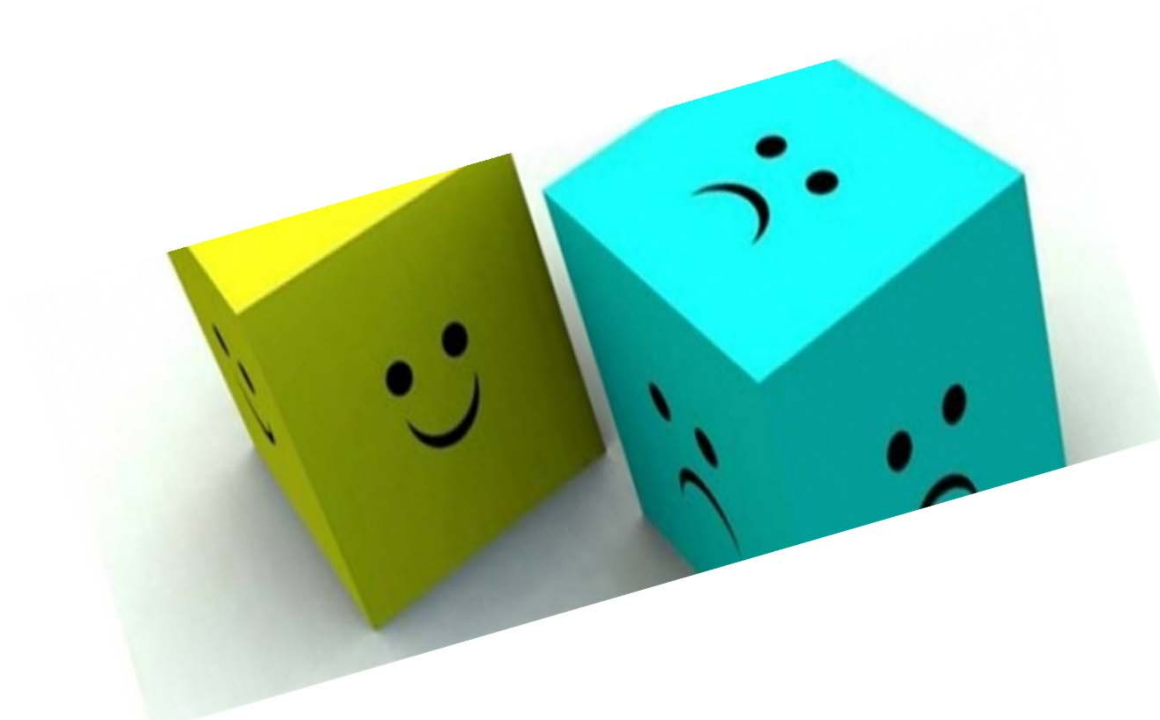
- Study emotions and the relationship to environment
- Provide mental health and social science experts with –
  - Emotional factors with respect to interpersonal relationships
  - Identify locations and emotional responses
  - Evaluate activity vs. emotions
- Smartphones allow field study w/o specialized equipment
  - Past: In-home cameras, attached mics & diaries = *Biased results*
  - Today: Ubiquity of smartphones *desensitizes* users from monitoring activities

RELATED WORK

- Location and activity correlation
  - BeTelGeuse [2] open source framework to gather situational information
  - CenseMe [4] detects activity at a location (e.g., dancing w/friends) and reports activity to social media
- Social science experimentation
  - Environmental activated recorder (EAR) to evaluate sociability contexts [3]
- Self-reporting
  - Use smartphone to report moods throughout the day to suggest therapy options [5]

GOAL

- “The overarching goal of EmotionSense is to exploit mobile sensing technology to study human social behavior.”
  - Evaluate people’s emotions using smartphone sensors and speech-recognition tools to observe behavior patterns in social situations



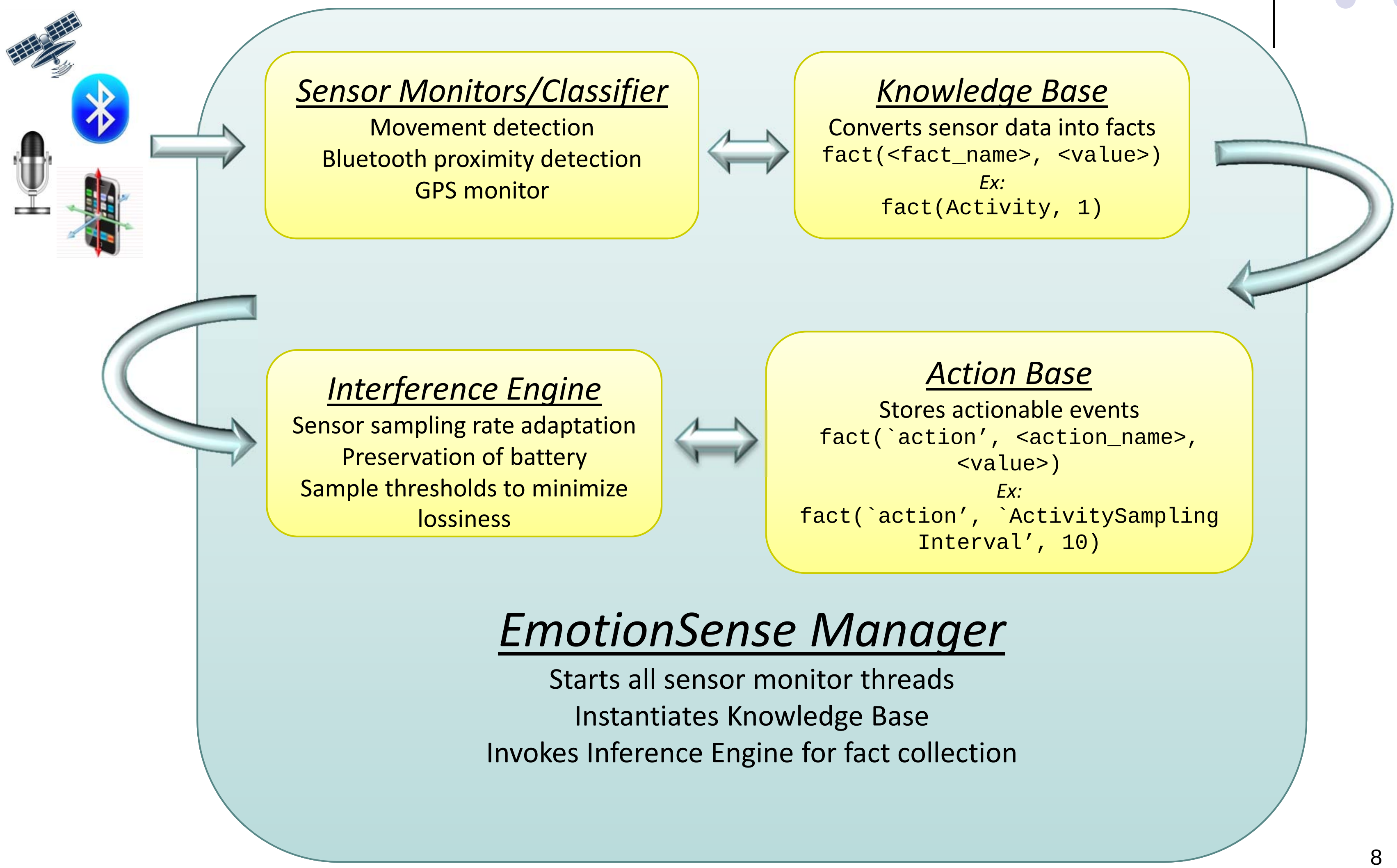
ASSUMPTIONS & LIMITATIONS

- Assumptions
  - Participants will have smartphone with them majority of the time
  - Microphone is unobstructed
  - Participants gather frequently
  - HTK produces correct results (before and after porting to Symbian)
- Limitations
  - How well the participants represent persons who exhibit a wide range of detectable emotions
  - How well the training data represents emotional signatures

METHODOLOGY

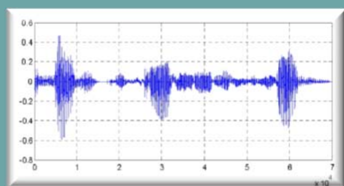
- Information flow
- Speaker recognition
  - Based on Gaussian Mixture Model (GMM) & Maximum A Posteriori (MAP) adaptation
  - Windows/Linux toolkit ported to Symbian OS
- Emotion recognition
  - Also based on Gaussian Mixture Model
  - Narrow emotional types are clustered into a broad classifications
- Adaptation framework
  - Generate rules to govern sensor sampling rates

Information Flow

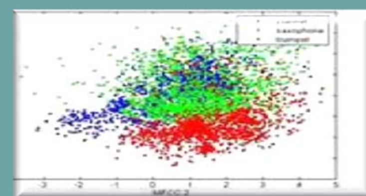


Speaker Recognition

Audio Data Collection & Parameterization



Apply GMM to distinguish study participants from others



MAP is applied to derive user-specific GMMs

Audio sequences are assigned user probabilities at run-time

Emotion Recognition

- Similar method as speaker recognition
  - GMM trained on Emotional Prosody Speech and Transcripts library to classify emotions
  - MAP adaptation is used to generate user specific models
  - Emotional characteristics are assigned to audio sequences
- Emotion clustering
  - Emotion grouping used by social psychologists
  - Narrow emotion classification difficult even for humans

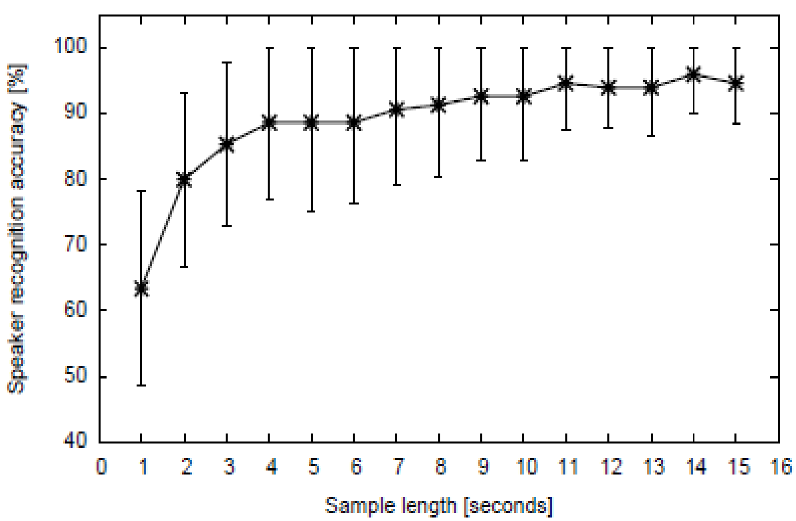
Broad emotion	Narrow emotions
Happy	Elation, Interest, Happy
Sad	Sadness
Fear	Panic
Anger	Disgust, Dominant, Hot anger
Neutral	Neutral normal, Neutral conversation, Neutral distant, Neutral tete, Boredom, Passive

BENCHMARKS

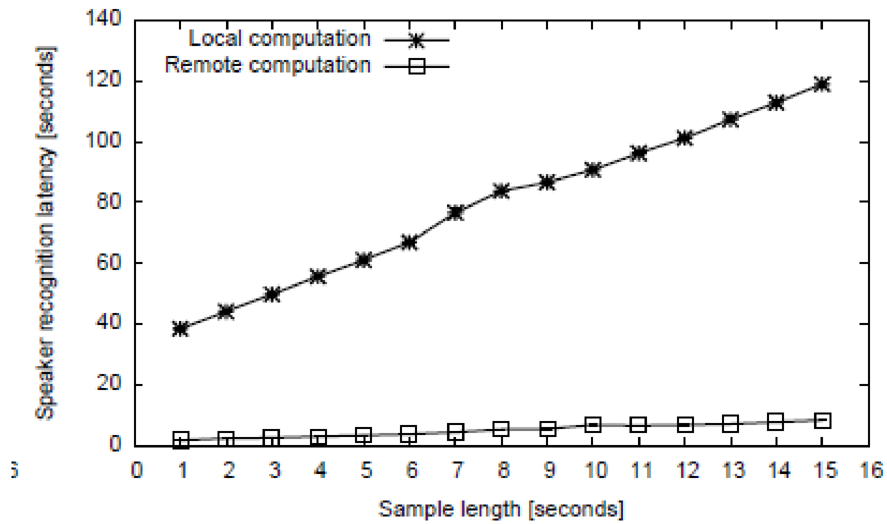
- Micro-benchmarks to evaluated system performance
  - Adaptation rules were collected from 12 users in a 24hr period
  - Tuned framework based on the Nokia's 6210 sensor data captures
- Speaker recognition
  - 10min of training data from 10 users
  - Sample lengths varied from 1 to 15 seconds
  - 90% accuracy with sample lengths greater than 4 seconds
- Emotion recognition
  - Use pre-existing test and training library
  - 350 test samples per-sample length second
  - ~70% accuracy with sample lengths greater than 5 seconds

Benchmarks

Recognition accuracy & latency



Speaker recognition *accuracy* vs. audio sample length



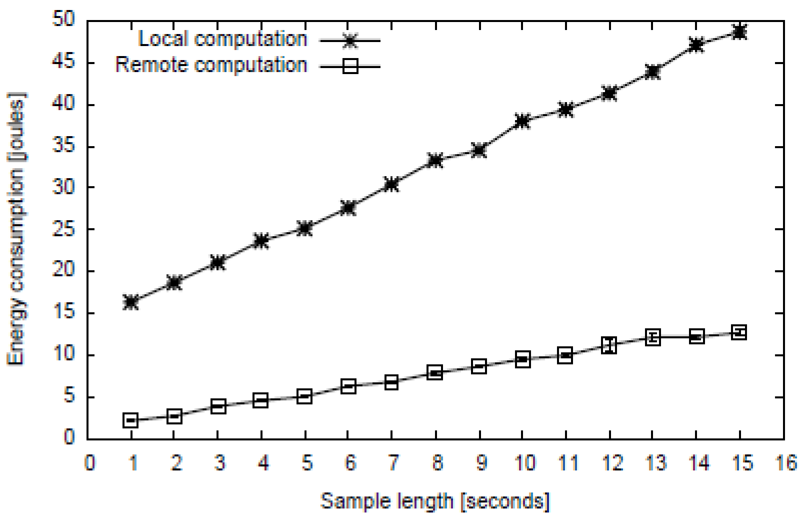
Speaker recognition *latency* vs. audio sample length

Convergence ~90% > 4 seconds

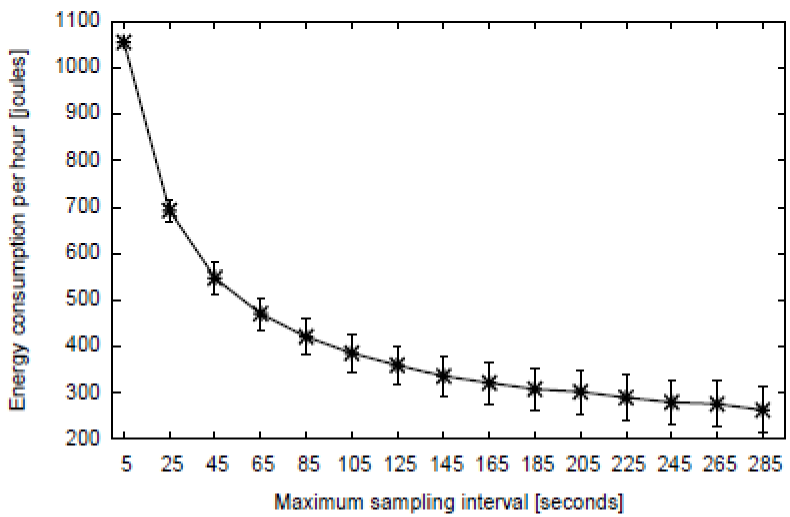
Local benchmark based on 369MHz ARM 11 µP

Benchmarks

Power Consumption



Energy consumption vs. audio sample length



Energy consumption vs. maximum sampling interval

Benchmarks

Confusion Matrix

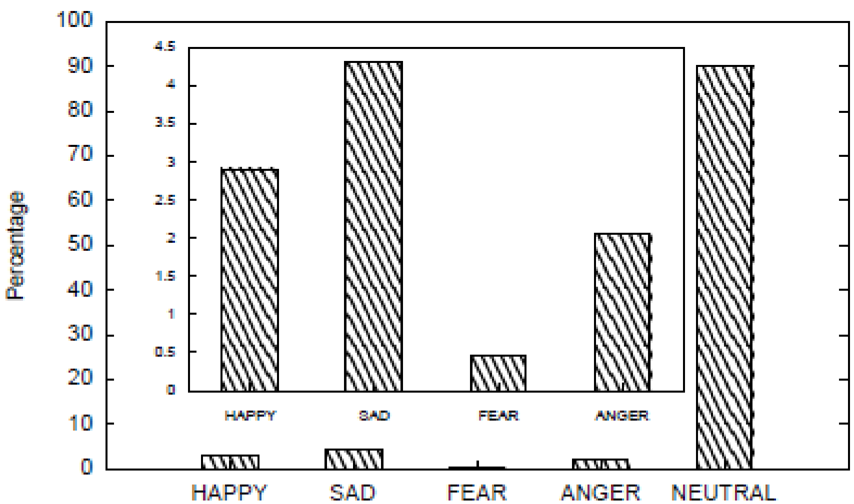
Emotion [%]	Happy	Sad	Fear	Anger	Neutral
Happy	58.67	4	0	8	29.33
Sad	4	60	0	8	28
Fear	8	4	60	8	20
Anger	6.66	2.66	9.34	64	17.33
Neutral	6	5.33	0	4	84.66

RESULTS

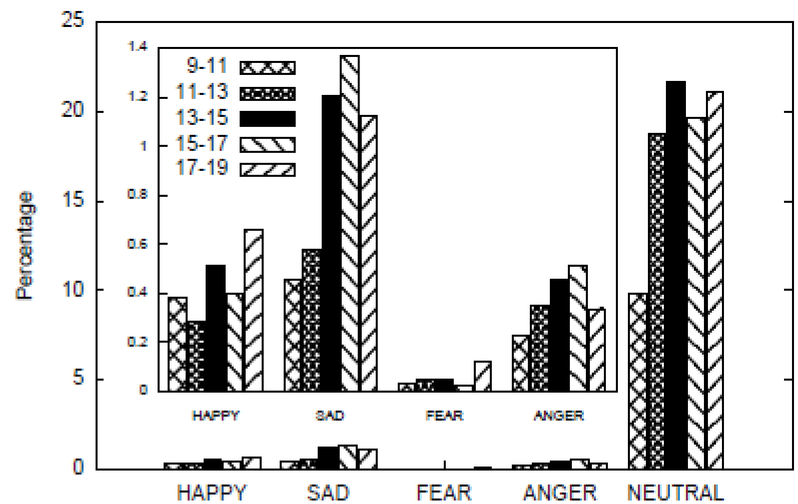
- Trial conducted for 10 days with 18 participants
- Participant daily diaries
  - Activities
  - Who was present
  - Mood
  - Location
- Emotion Distribution
  - Neutral emotions are the most prevalent
  - Fear is the least prevent

Results

Emotion Distribution



Distribution of detected broad emotions



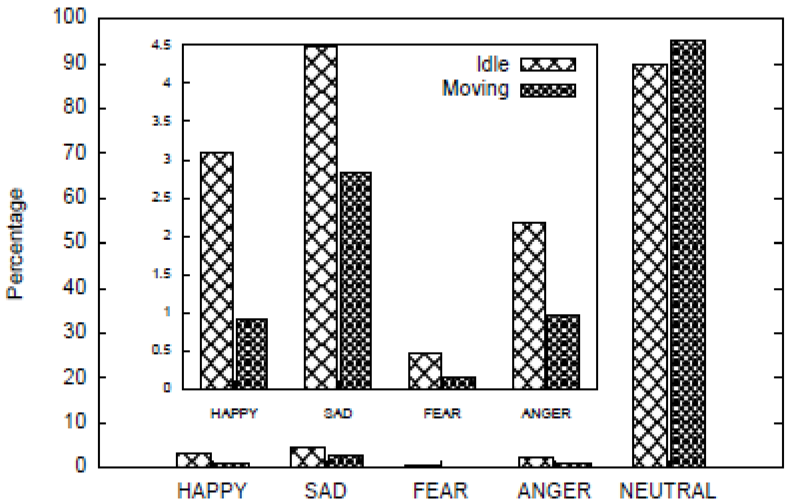
Distribution of detected broad emotions with respect to time of day

Most social activity exhibits neutral emotions

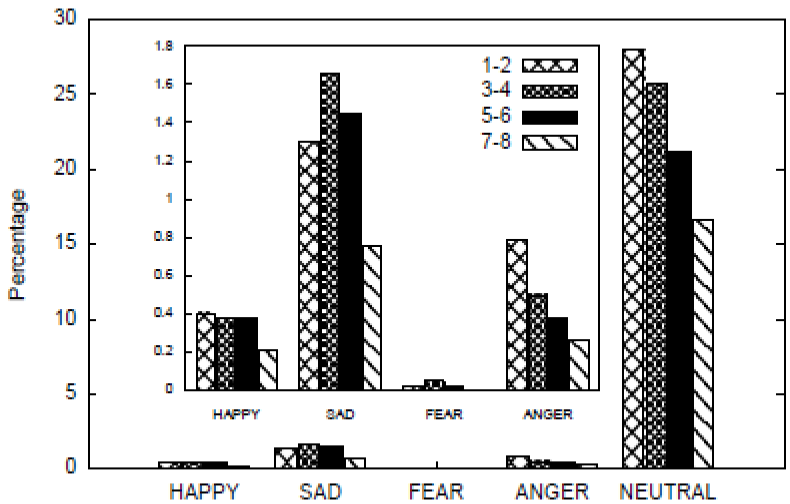
Emotions are more prevalent as the day progresses

Results

Emotion Distribution



Distribution of detected broad emotions within physical state



Distribution of detected broad emotions with respect to number of co-located participants

Non-neutral emotions are more prevalent in the idle state

Why is sadness experience in groups?

CONCLUSIONS

- Demonstrated smartphones are a viable tool for social science research
- Able to identify (to some degree) participant's emotions through speech recognition
- A majority of speech is categorized as neutral
- Emotion categorization algorithm produced underwhelming results

FUTURE WORK

- Galvanic skin response sensor
- Continue optimizing emotional recognition model
- Addition of more realistic noise models
- Real-time feedback, daily monitoring and user interaction options

References

1. J. Froehlich, M. Y. Chen, S. Consolvo, B. Harrison, and J. A. Landay. MyExperience: A System for In situ Tracing and Capturing of User Feedback on Mobile Phones. *In Proceedings of MobiSys '07*, pages 57–70, 2007.
2. J. Kukkonen, E. Lagerspetz, P. Nurmi, and M. Andersson. BeTelGeuse: A Platform for Gathering and Processing Situational Data. *IEEE Pervasive Computing, 8(2):*49–56, 2009.
3. M. R. Mehl, S. D. Gosling, and J. W. Pennebaker. Personality in Its Natural Habitat: Manifestations and Implicit Folk Theories of Personality in Daily Life. *Journal of Personality and Social Psychology, 90(5):*862–877, 2006.
4. E. Miluzzo, N. D. Lane, K. Fodor, R. Peterson, H. Lu, M. Musolesi, S. B. Eisenman, X. Zheng, and A. T. Campbell. Sensing Meets Mobile Social Networks: The Design, Implementation and Evaluation of the CenceMe Application. *In Proceedings of SenSys '08*, pages 337–350, 2008
5. E. M. Morris, Q. Kathawala, K. T. Leen, E. E. Gorenstein, F. Guilak, M. Labhard, and W. Deleeuw. Mobile Therapy: Case Study Evaluations of a Cell Phone Application for Emotional Self-Awareness. *Journal of Medical Internet Research, 12(2):e10*, 2010.
6. A. S. Pentland. Honest Signals: How They Shape Our World. *The MIT Press*, 2008.
7. Allilli M. A Short Tutorial on Gaussian Mixture Models. *Université du Québec en Outaouais*, 2010.