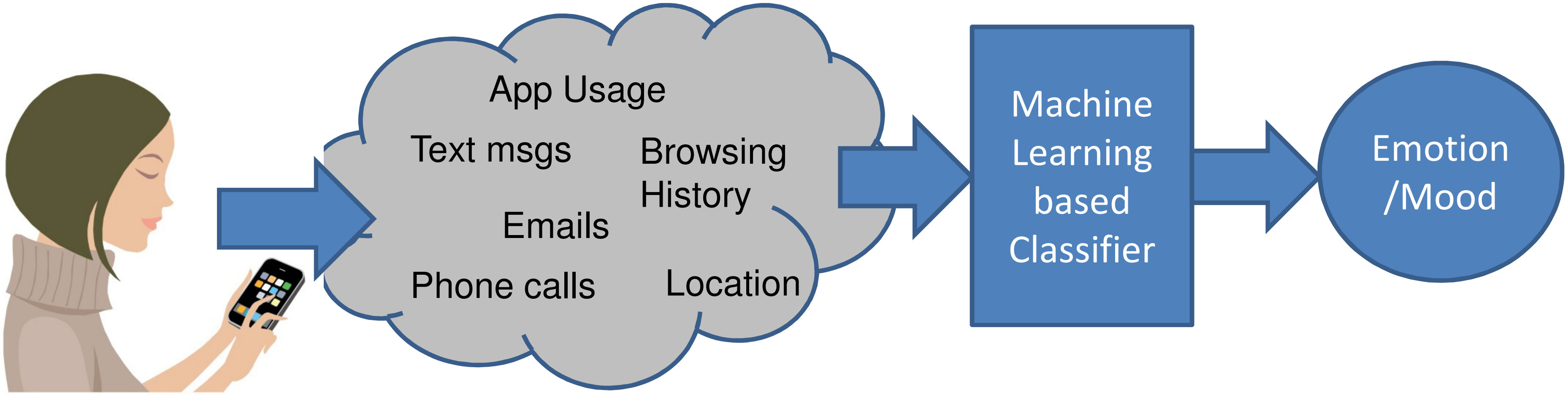
Impact of Experience Sampling Methods on Tap Pattern based Emotion Recognition

Surjya Ghosh, Vatsalya Chauhan, Niloy Ganguly, Bivas Mitra
Indian Institute of Technology, Kharagpur

Pradipta De
SUNY Korea

Smartphone-based Emotion Detection



- MoodScope: detects multiple mood states
- Lee et al. (CCNC 2012): Uses different sensors to collect context, and a modified Twitter app to gather touch behavior
- MouStress: detects stress behavior from mouse usage patterns

Assumption: It is possible to collect the ground truth (or emotion labels) reliably

Collecting Emotion Labels

- Experience Sampling Methods
  - Periodically ask the user to record the emotion
  - Detect a context (or event) to trigger a questionnaire to record emotion

- What if the requests are too frequent or too intrusive
  - User may respond falsely
  - User may not respond at all
  - User will drop off from the study

What is the impact of poor quality ground truth data ?

TapSense App

- An app that tracks the typing pattern of a user
  - Records the inter-tap distance (ITD)
  - Correlate the ITD to the emotion labels from user

- Focus is on
  - Interpretability of the result ➔ relationship between ESM and accuracy
    - How different is the result with respect to differnet ESMs ?
  - Not on flexibility (or raising the accuracy bar)
    - Multiple features may improve accuracy, but makes it harder to isolate the impact of an ESM approach

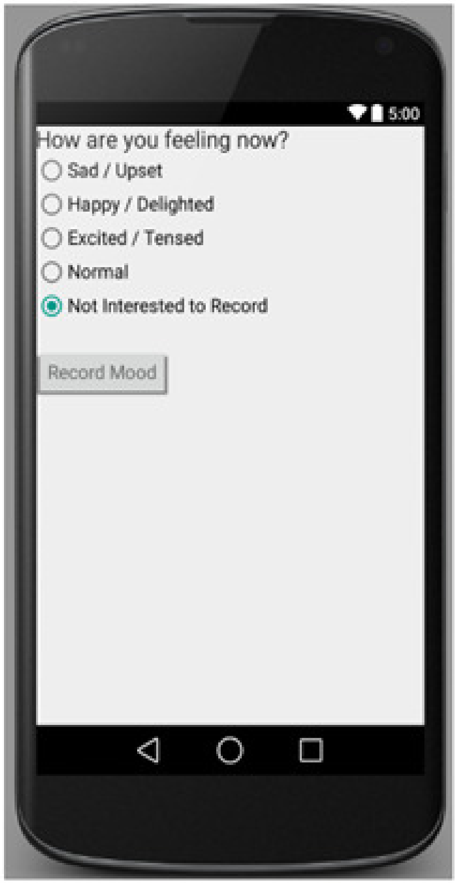
Outline

- TapSense
  - Experience Sampling Methods
  - Architecture
- DataSet
- Evaluation
- Conclusion and Future Work

Experience Sampling Method

- How to collect user response ?
  - We use questionnaire (to avoid ambiguity)

- When to collect user response ?
  - Time-based (TB)
    - *At predefined intervals*
  - Event-based (EB)
    - *Whenever a specific event occurs*
    - *User switches to a different app*
  - Signal-based (SB)
    - *Based on some signal*
    - *Inactive period in typing*



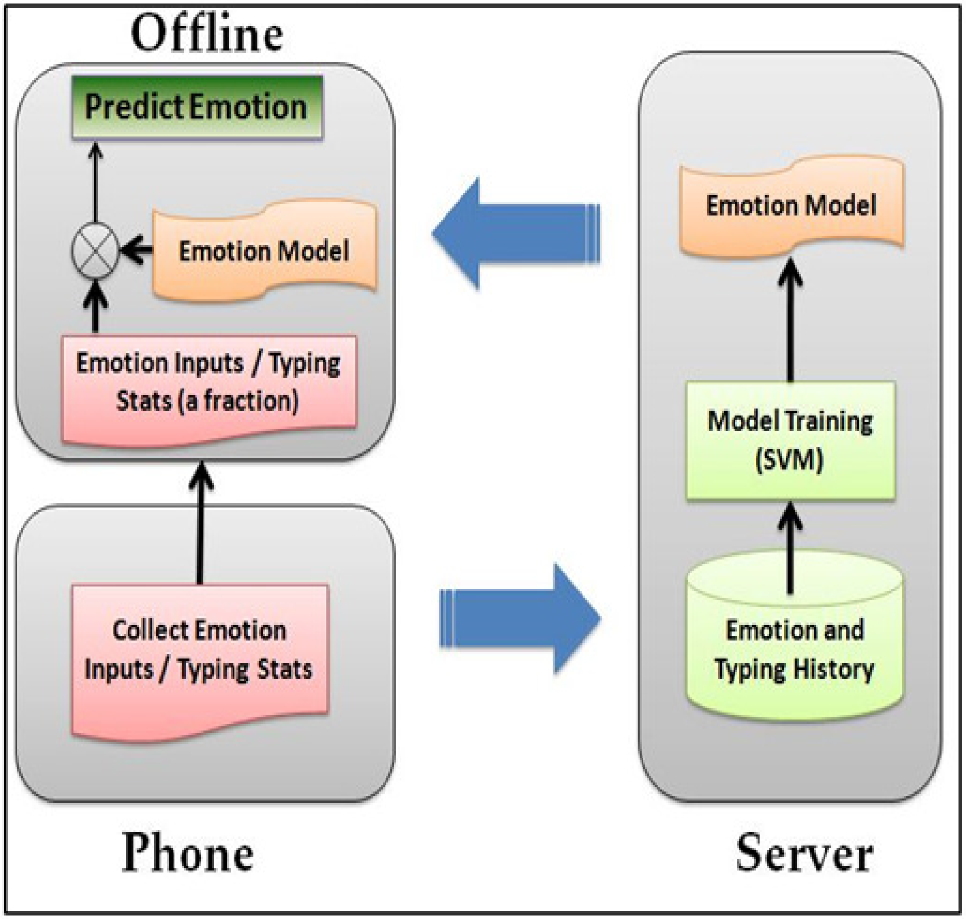
Objective

- A simple mechanism for emotion detection with different ESMs using smartphone
  - Non-intrusive (no additional device or sensor)
  - Energy-efficient (low power continuous channel)
  - Ensures privacy (won’t capture sensitive details)

- Tracking typing pattern of user satisfies all of these criteria

System Architecture

- SmartPhone
  - Tap Data collection
  - Communicate with server
- Background Server
  - Building the model based on tap data
- Features
  - ITD [Time elapsed between two consecutive tap events]



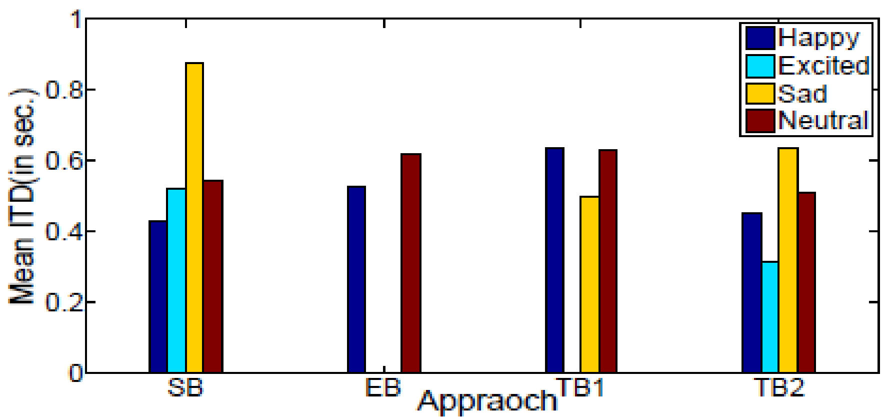
TapSense System Architecture

DataSet Generation

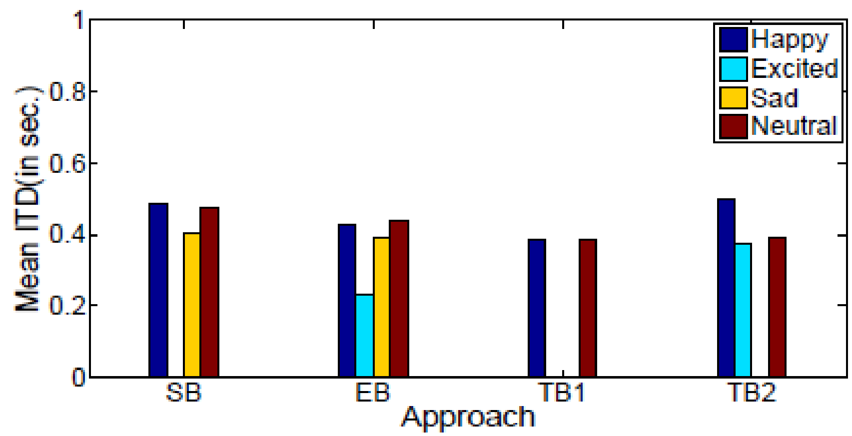
- Participants
  - 10 students aged between 19 – 24
  - Collected data for 16 days
  - Changed ESM (TB,EB,SB) after every 4 days
  - Noted every typing event and measured time elapsed between two typing events (ITD)

- ESM configurations
  - Signal-based (SB) [2 min idle period during typing]
  - Event-based (EB) [Change of application]
  - Time-based (TB1) [Periodicity 3 hr]
  - Time-based (TB2) [Periodicity 30 min]

Inter-Tap Distance (ITD) Distribution



(a) Mean ITD distribution – category I user

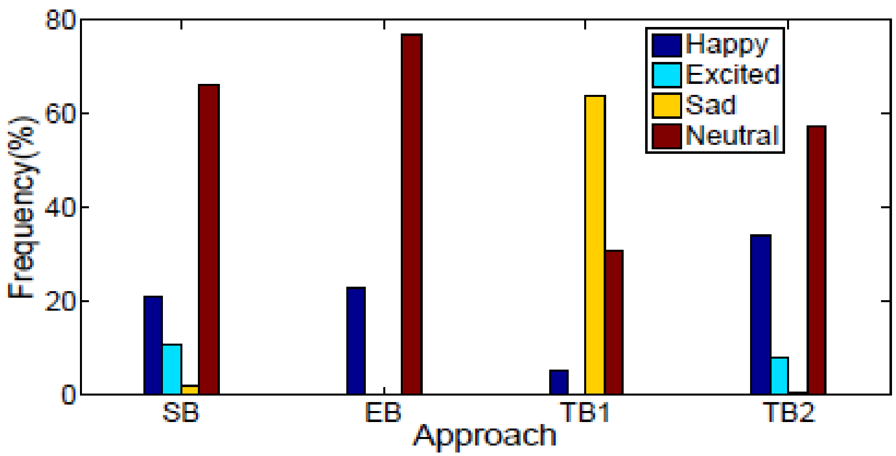


(b) Mean ITD distribution – category II user

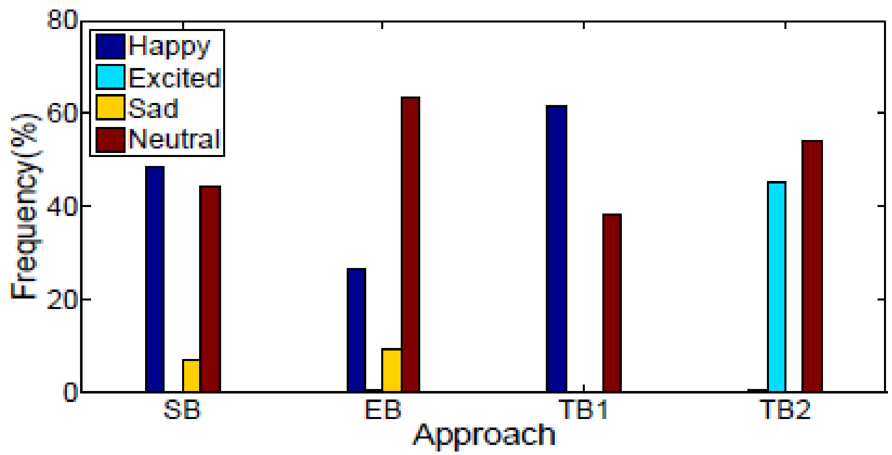
For category 1 : ***typing speed vary across emotion states***

For category 2 : ***typing speed does not vary significantly across emotion states***

Emotion Label Distribution



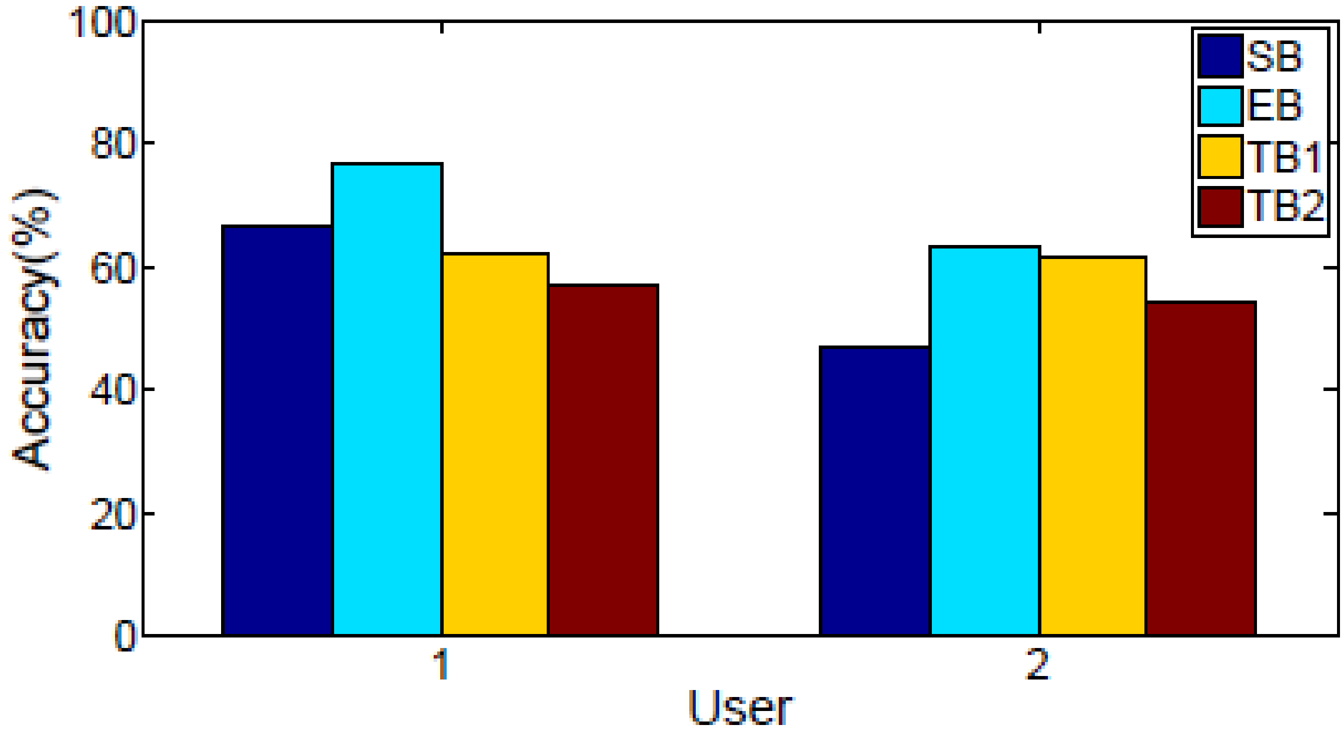
(a) Emotion state distribution – category I user



(b) Emotion state distribution – category II user

Emotion states labeled using different ESM approaches vary across both the users ➔ ESM can impact the user's response, assuming the trend of her emotion remains similar across the tests

Does ESM techniques influence accuracy of emotion detection?



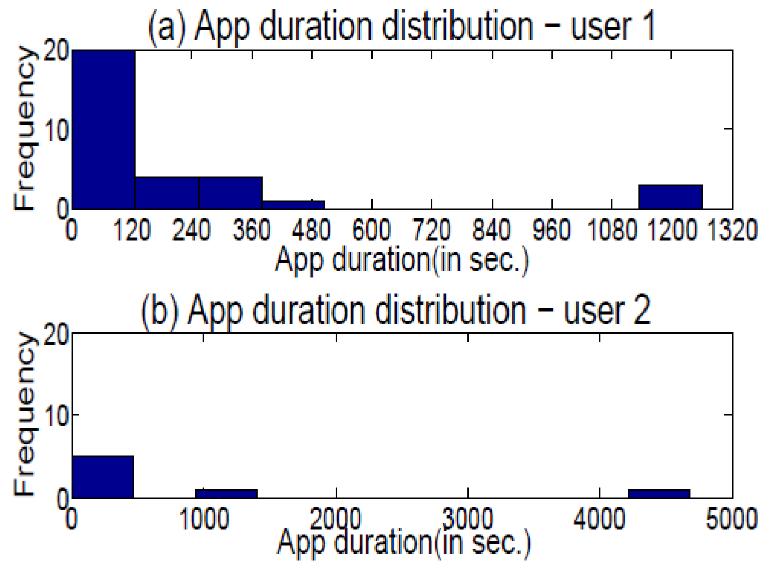
Classification accuracy – Feature; ITD only

Event-based ESM performs best for both type of users in both the models

How different are the ESM approaches?

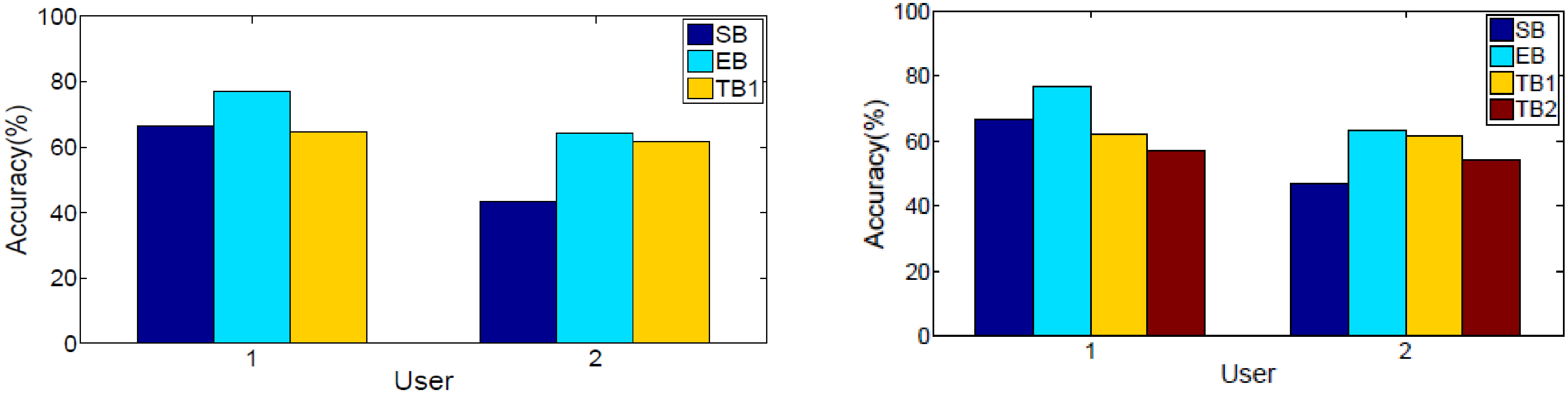
Table 1: Classification Accuracy in Cross-validation

User#	Train App	SB	EB	TB1
1	SB	66.50	77.05	31.0
1	EB	66.50	77.05	31.0
1	TB1	4.31	7.57	62.0
2	SB	46.82	27.41	60.87
2	EB	44.40	63.35	38.33
2	TB1	48.69	26.57	61.66



- Cross-training and testing → shows if two ESMs are identical in collected data quality
- For user 1, SB and EB performs identically, but not for user 2

How does ESM approaches depend on additional features?



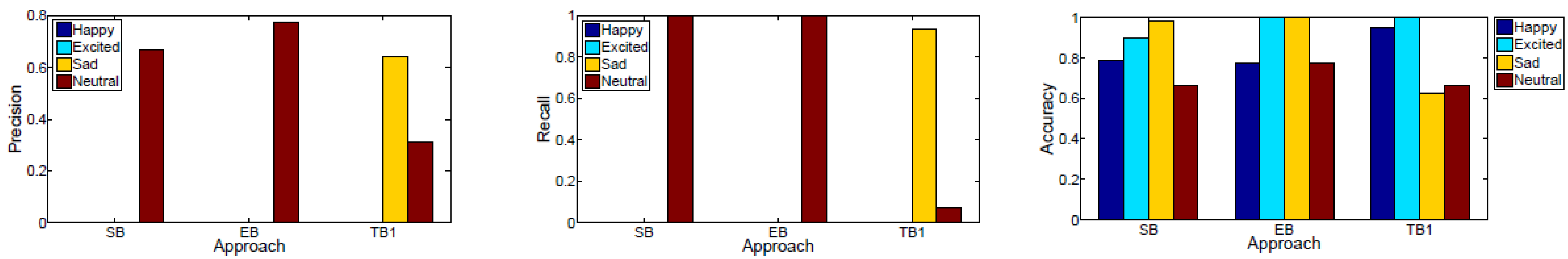
Classification accuracy

Features: ITD and App category

- Adding Application category does not improve accuracy much.
- Users tend to spend a significant proportion (80%) of time in IM apps, compared to texting or other apps

What is the role of sampling approaches on detecting individual emotion states?

User-1: significant variation in different typing speed across emotion states



(a) Precision (b) Recall (c) Accuracy

For SB, EB precision and recall for Neutral state is high

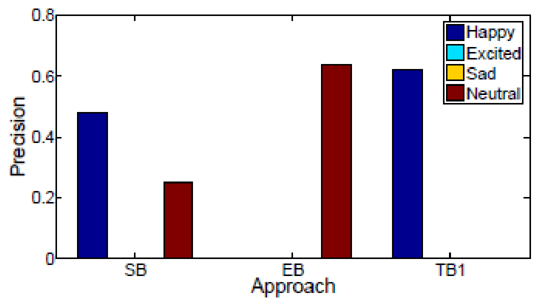
For TB1, recall is low for neutral state, but precision and recall are high for sad state

Neutral and happy states are detected with reasonable high accuracy across all approaches

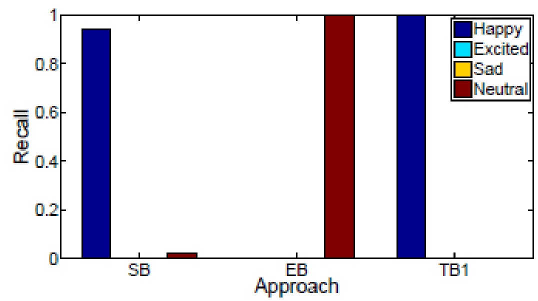
Sad and excited states are having higher accuracy (few sample points)

What is the role of sampling approaches on detecting individual emotion states?

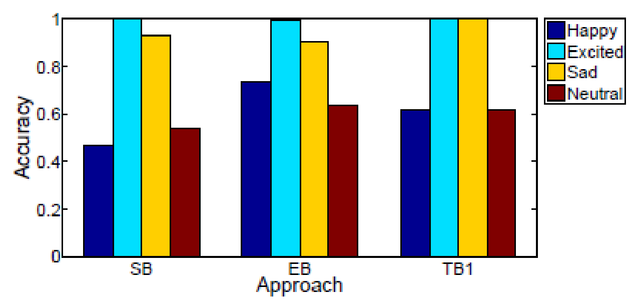
User-2: NOT significant variation in different typing speed across emotion states



(a) Precision



(b) Recall



(c) Accuracy

For EB precision and recall for Neutral state are high
For SB,TB1, precision and recall are high for Happy state

Accuracy of neutral and happy states are not high
Sad and excited states are having higher accuracy (few sample points)

Conclusion

- Tested prediction accuracy with different ESM
  - Results indicate that careful selection may help
- TapSense app
  - Careful ground truth collection may simplify the design of the classifier
- Open Question on designing emotion detection app
  - a simple ESM design, like periodic user feedback collection, coupled with a number of features for generating the model?
  - an ESM design that is adapted to the monitored feature, which may reduce the complexity of feature selection to build the model.

Future Work

- Explore a hybrid ESM technique
- Leverage ideas from anticipatory mobile computing
- Stronger validation of ITD as a feature using more participants

Reviewers' Comments

- Does it matter what the user is typing? Or their relative typing skill?
  - Yet to look at other features, like errors during typing, emoticons used, etc.

- Suggest authors to consider convenience/level of intrusiveness for users
  - Hybrid ESM, with a budget on number of times questionnaires can be fired, is a move in this direction

- A bit skeptical that users would use mobile phones when they are in a negative mood
  - Stress sensing has been shown to work
  - Usage pattern may reveal withdrawal → negative mood ?

Reviewers' Comments

- Brain sensors are becoming available and non-invasive, and they can be used for basic emotion detection
  - If these sensors become a everyday companion like smartphones, it may open up alternative modes for emotion sensing
- How this could complement to existing approaches(e.g., using audio or inertial sensors)
  - The audio and inertial sensors can provide more context, but can be turned on selectively to limit (i) power consumption (ii) privacy concerns


THANK YOU


?