Towards Ambient Intelligence in Smart Healthcare: A CPS Perspective

John A. Stankovic
BP America Professor
Director, Link Lab
Department of Computer Science
University of Virginia

ACM Transactions on Computing for Healthcare (HEALTH)

Co-Editors-in-Chief

Insup Lee
University of Pennsylvania

John A. Stankovic
University of Virginia

ACM Transactions on Computing for Healthcare (HEALTH) is a multidisciplinary journal publishing high-quality, original research papers, survey papers, and challenge papers at the intersection of computer science, electrical and computer engineering, mechanical engineering, bio-medical engineering, behavioral and social science, psychology, and the health field, in general.

https://health.acm.org

University of Virginia

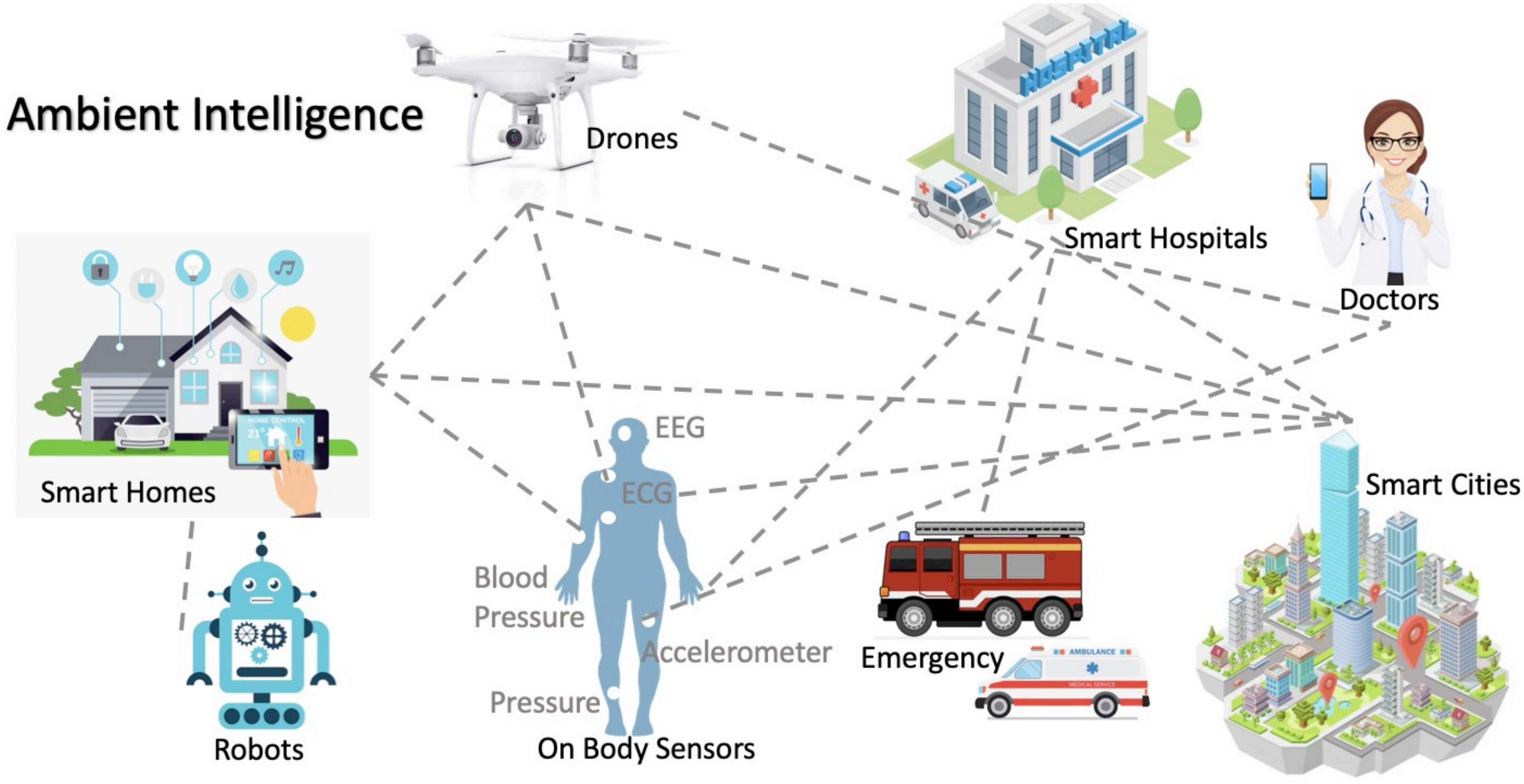
Rising (PhD) Stars in CPS

- Held at UVA in 2022; again in May 2023
- 124 applicants
- 33 awarded
  - Diversity 16
  - Different Universities represented - 22
- Supported by NSF

Ambient Intelligence(AmI)

- Definition: we envision AmI as a seamless infrastructure that integrates in-body, on-body, in-situ sensors, actuators, and *interacting cognitive assistants*.

Ambient Intelligence



IOT Infrastructure

University of Virginia

Vision

An ambient healthcare intelligence

WebMD
Big Data Collections
ML/Analytics
Data Mining

General Population

Nano-pills
Pacemaker

Actuations

Holistic

Wearables

Sensors
Actuators

University of Virginia

Outline

- A Few CPS Challenges
- Wearables/Cognitive Assistance
- Acoustics
- Robust Models: Properties/Uncertainties
- Lessons Learned from a Real Deployment

CPS Challenges

Physical World

- In the WILD
  - Sensing
    - Noise, missing data, multi-modal, …
    - Physiology, psychology, environmental context, …
    - Person ID
  - Actuation
    - Expanded concept

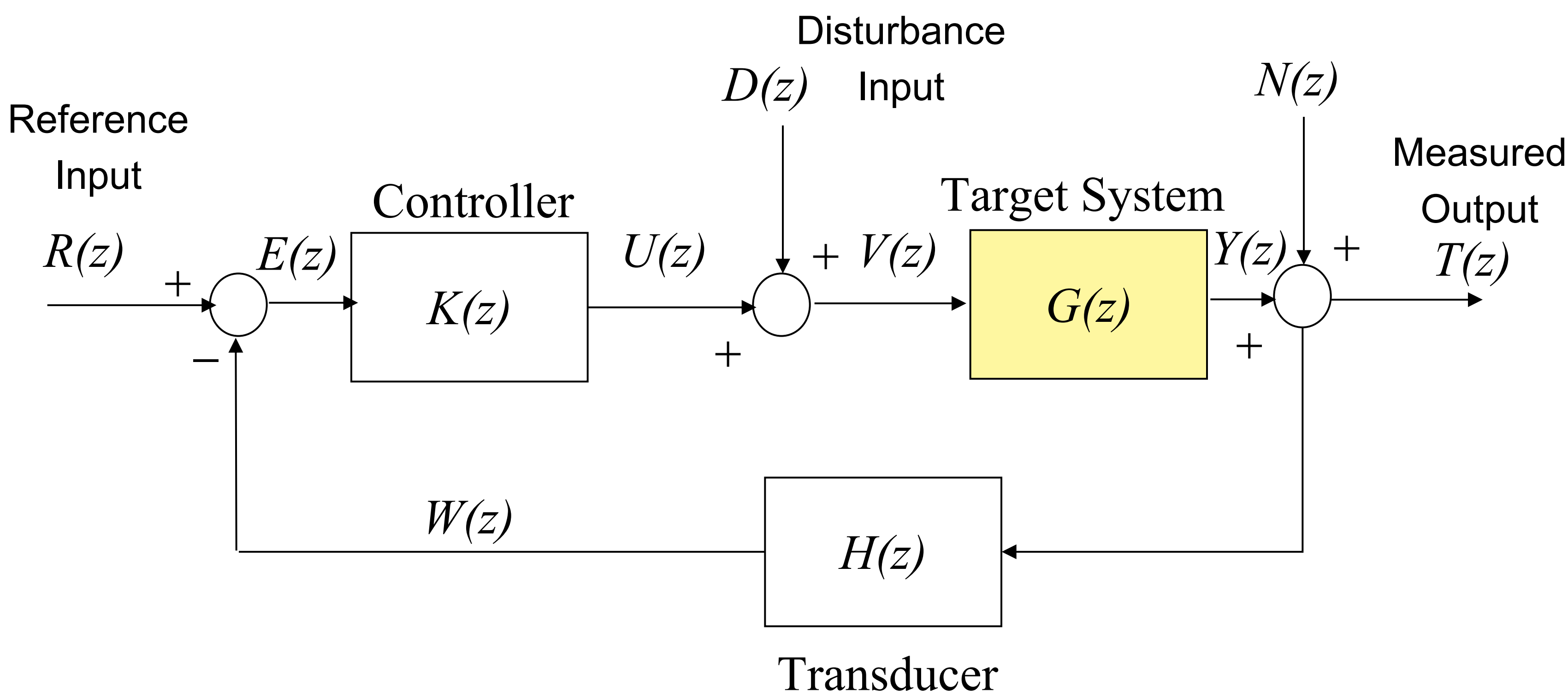
Cyber World

- Realities of physical world
- Realities of human health and behavior
- Understand context
- Highly diverse and dynamic environments
- Comprehensiveness
- Safety, Guarantees
- Continual Learning
- Integrate control, AI, ML, cognitive assistance, ….
- …

Two Challenges

- Models
- Interactions of Cognitive Assistants

Classical Feedback Loop

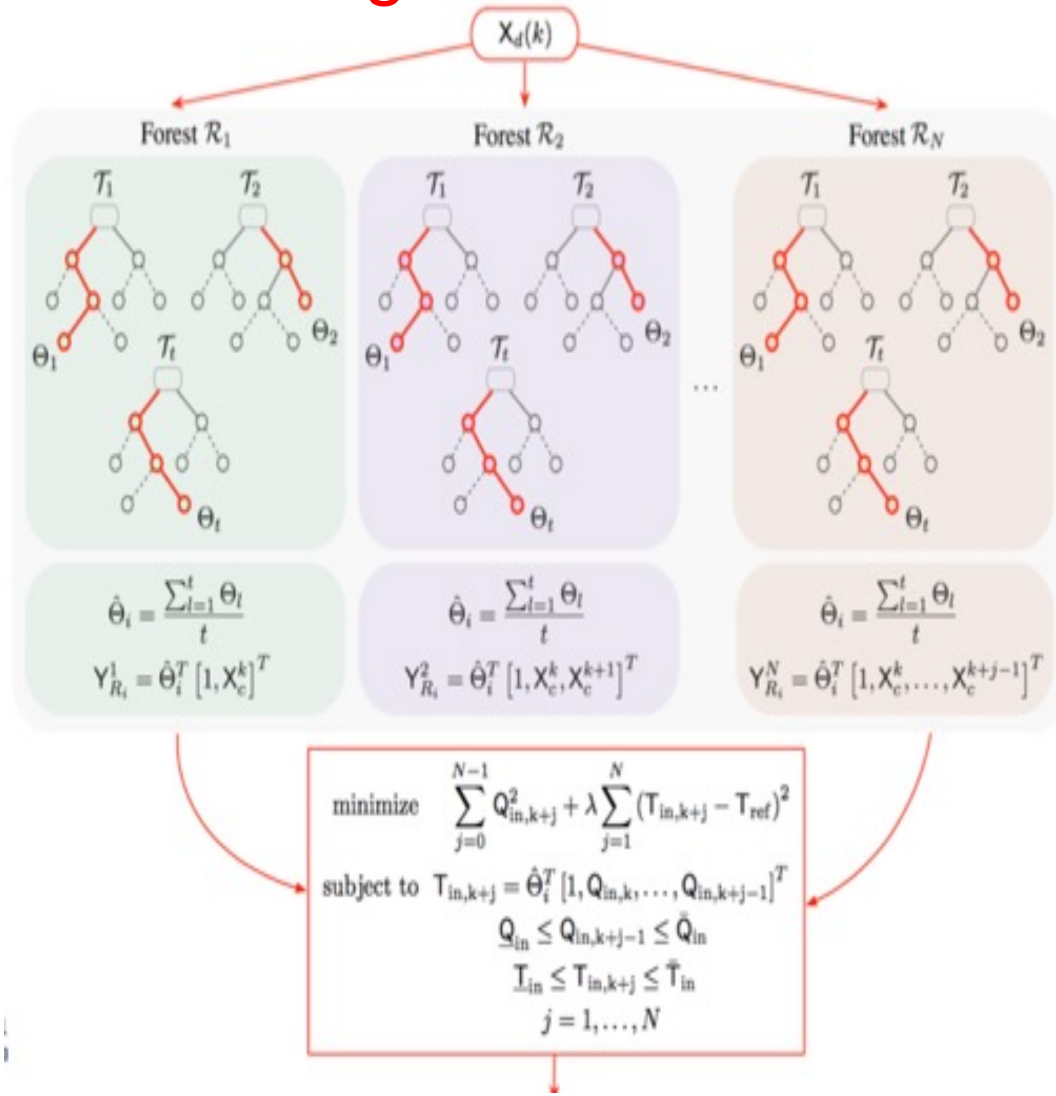


Central Challenge: Models

- Data Driven Alone – brittle, no guarantees
- Use Formal Methods and Knowledge Graphs to integrate properties into ML models
- Address Uncertainties into ML Model Predictions

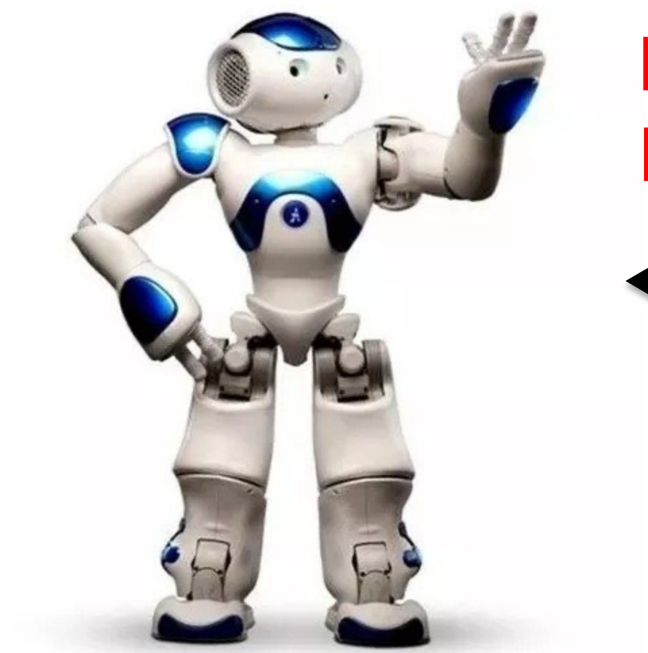
ML and FC Modeling

Decision Trees and Regression Model

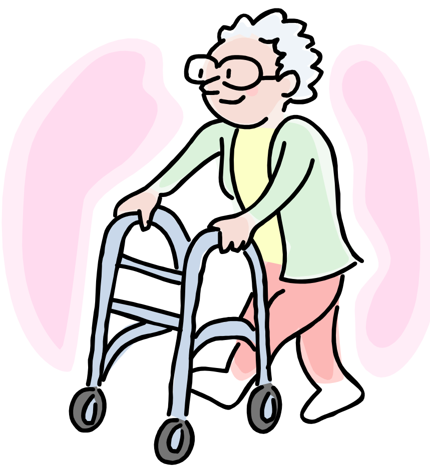


Madhur Behl
et al

Interacting Agents



Exchange
Intelligence



Personal Context
Environmental Context
Objectives
Decision making
N control loops

Exchange Intelligence

- Example
- Rather than passing (raw) data such as mood, physiology measures, current medicine, …
- Pass assessments such as
  – John is experiencing a medical problem and needs medication and stress reducing actions

AmI - Hype or Revolution

- Smart Watches
  - More and more sold; more and more sensors
- Smart Skin
- Smart Textiles
- …

- In-situ

Today's Main Themes

- Wearables (Smart Watches)
- Cognitive Assistance (ML and NLP)

Towards Ambient Healthcare Intelligence

- Acoustics (ML)
  - In situ => mood at distance

Powerful Modality

UVA ENGINEERING
LINK LAB

University of Virginia

iAdhere – verbal medication and exercise reminder system

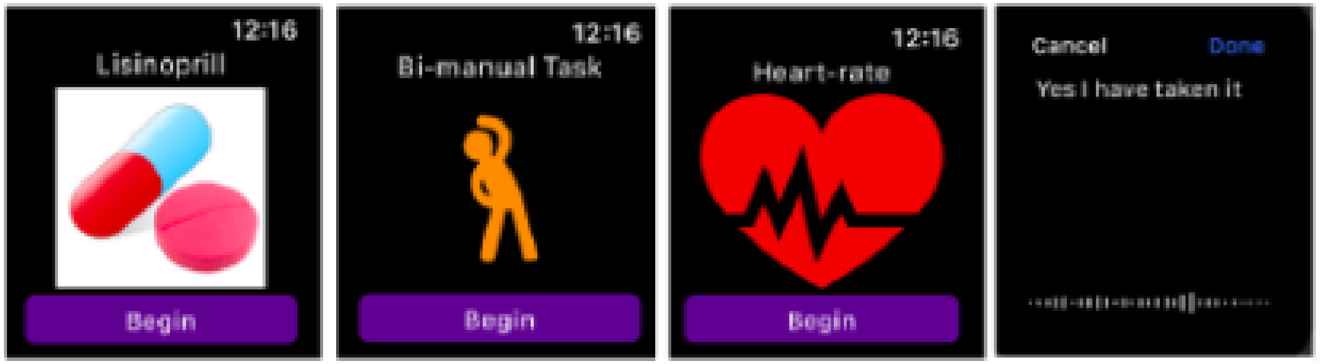


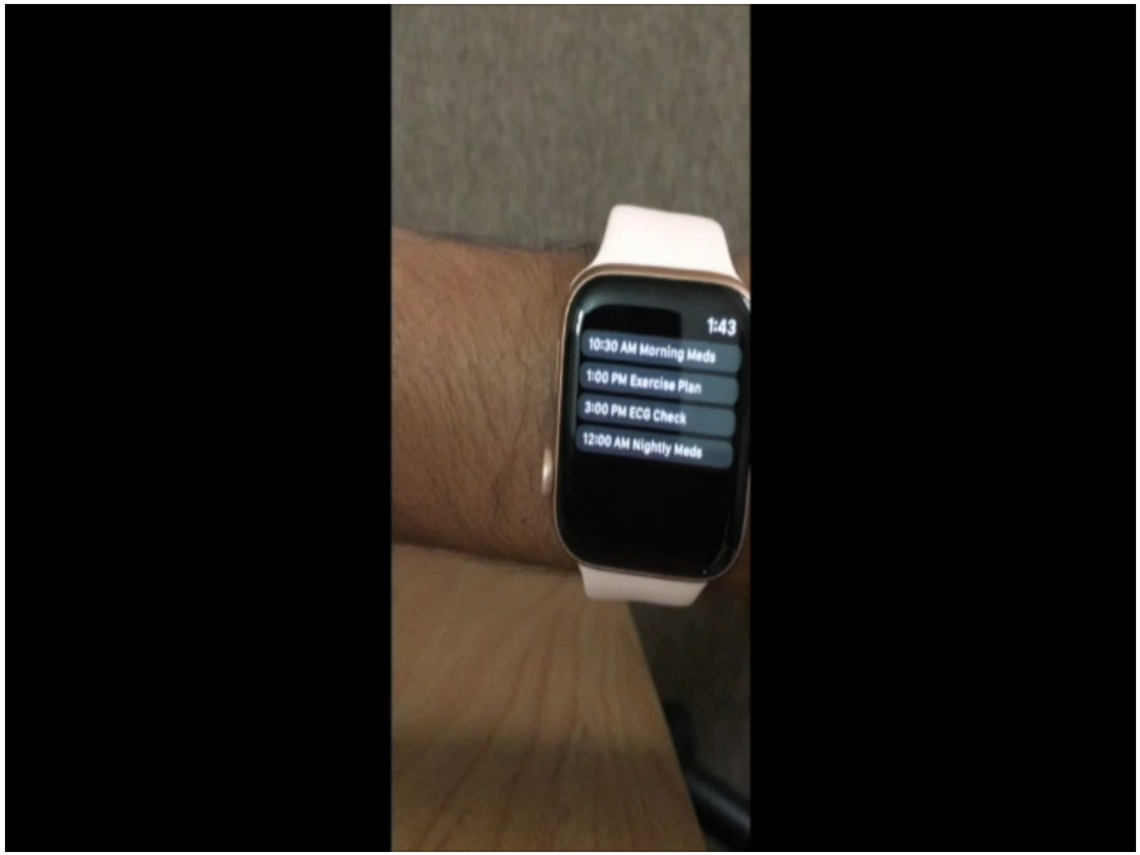
Figure 1: User Interfaces of the reminders to and the response from the users

For stroke patients

Using Apple Watch – with microphone and speaker

Applying in a Telemedicine setting

Demo – A Few Features



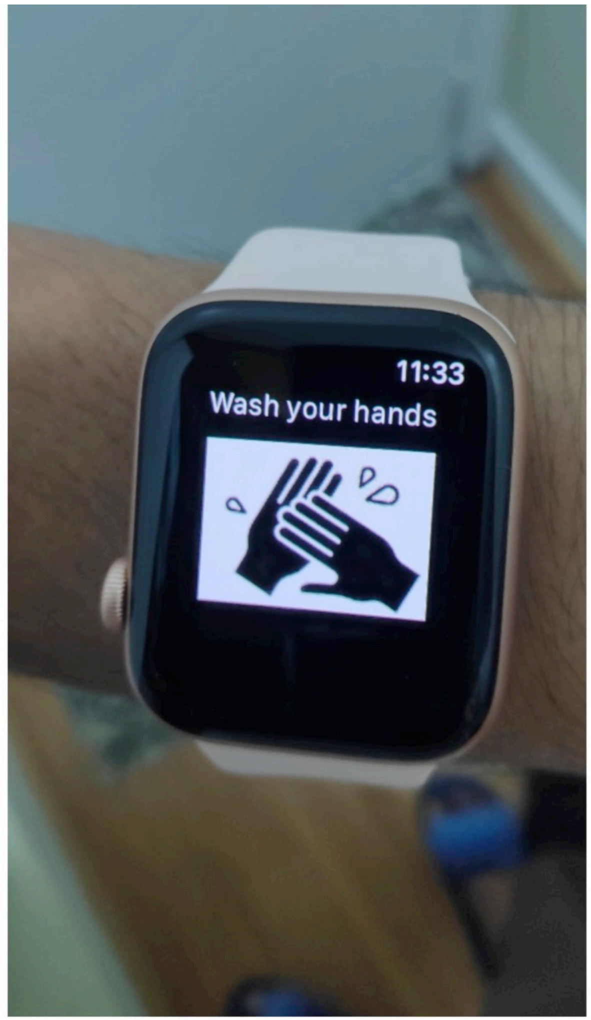
Earlier version called Medrem

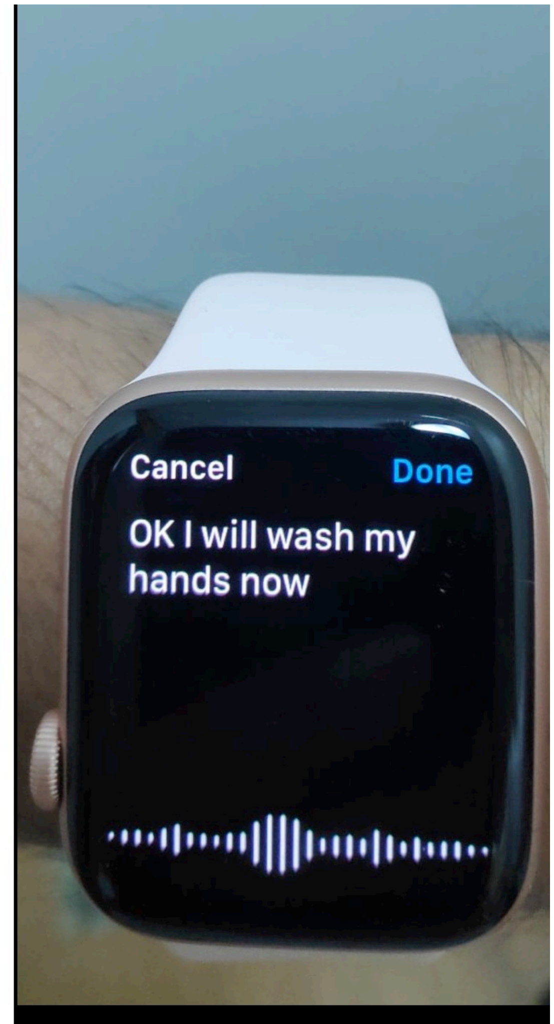
Services Expanded for Pandemics

- Collections of services on a smart watch
  - Handwashing (or general hygiene/elderly)
    - Quality of …
  - Mood/Depression/Anxiety/Loneliness
  - Voice based conversations
    - Pandemic info
    - Reminders/Alerts/Advice
  - Physiological parameters and more
    - Symptoms (coughing, fever, …)

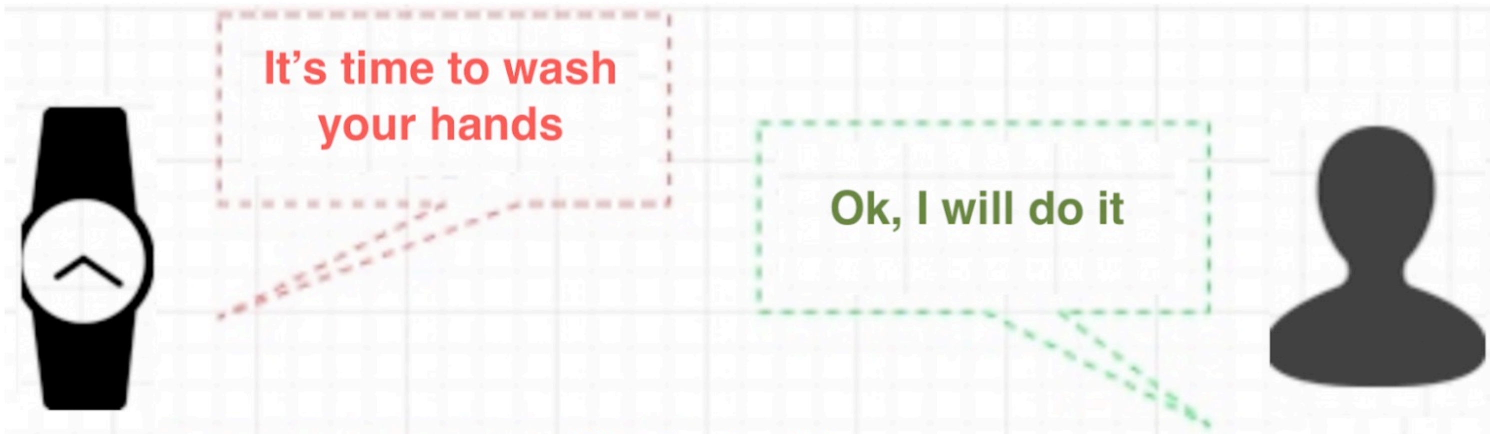
  - VoiceCare: published in EMBC

Smartwatch App

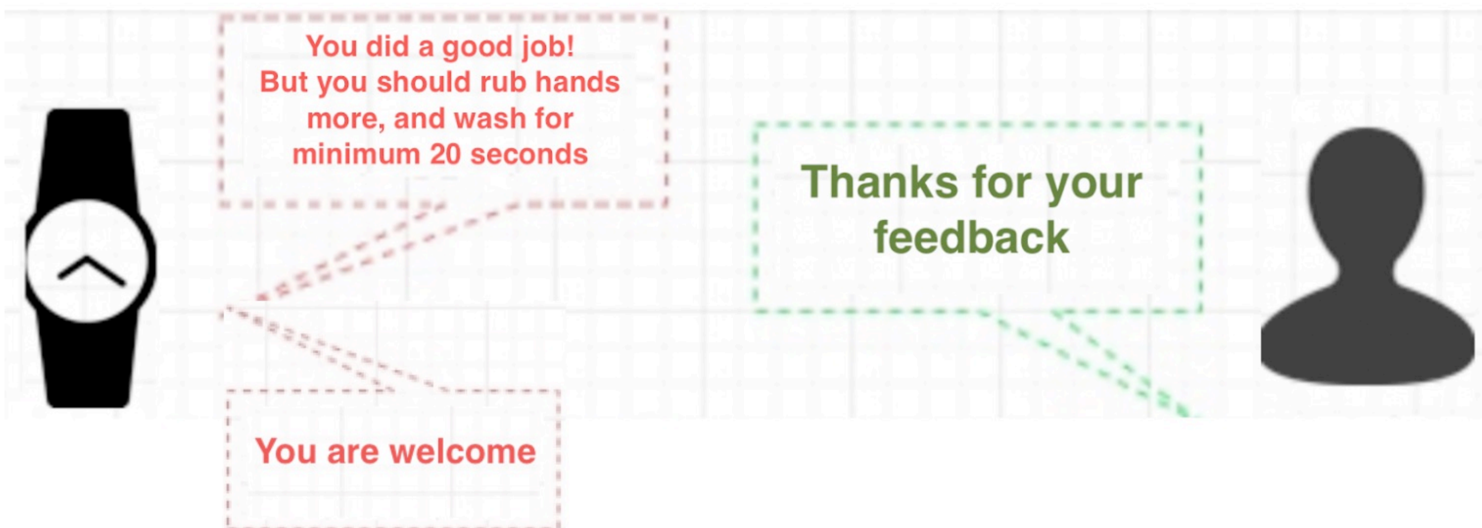




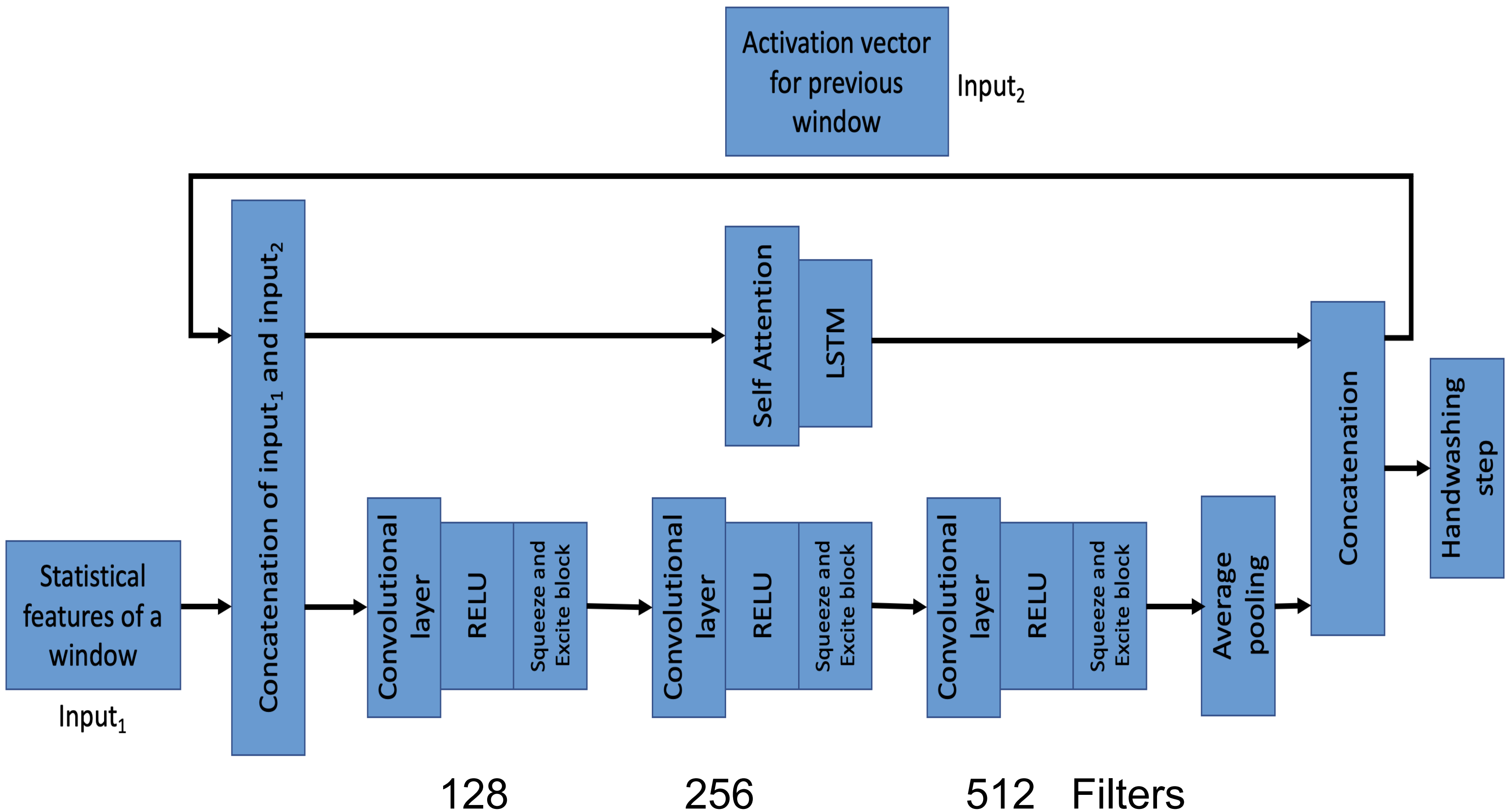
Dialogue



After handwashing..



Solution – A Hybrid DNN



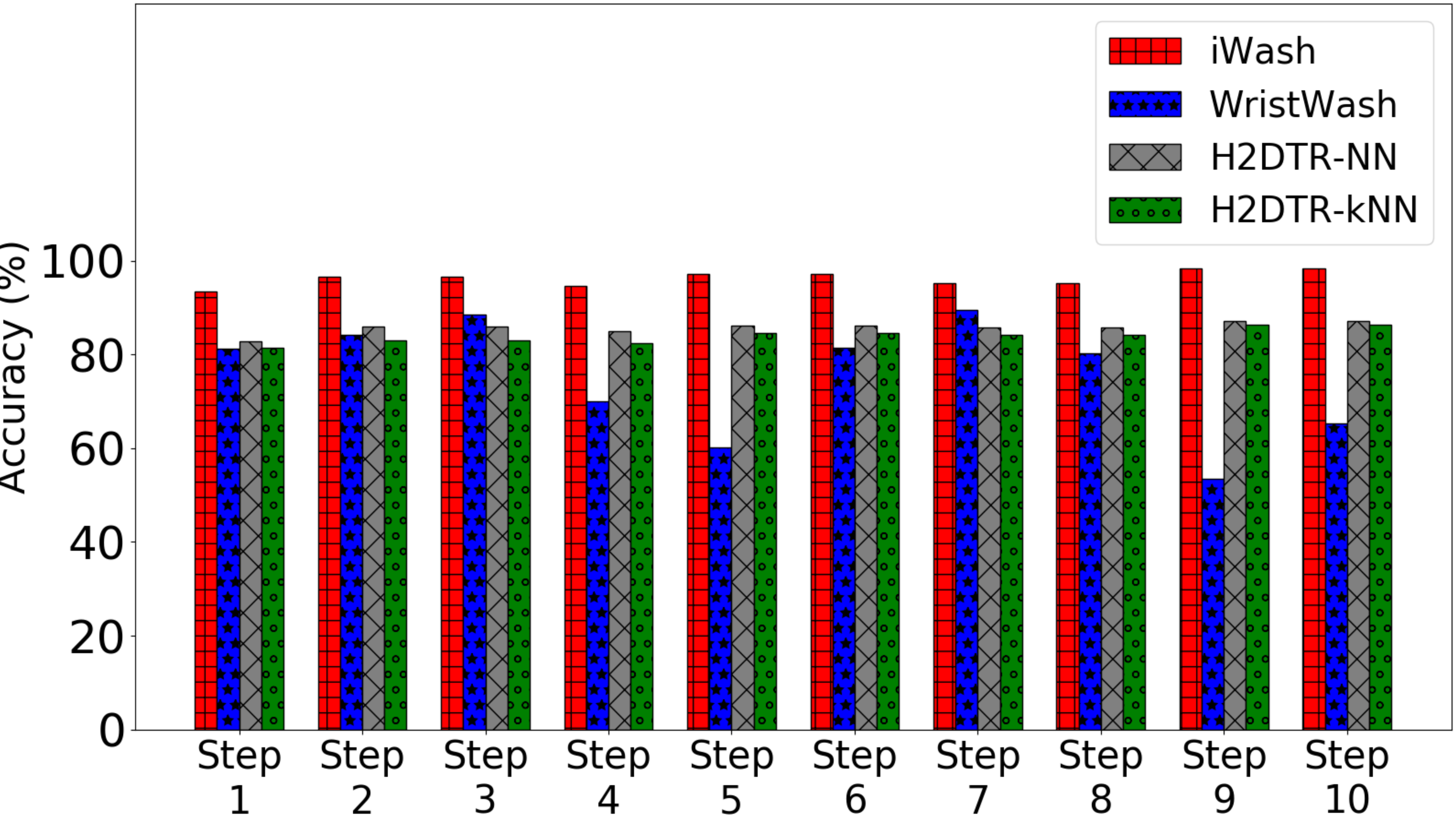
S. Samyoun, S. Shubba, A. Mondol, and J. Stankovic, iWash, A Smart Handwashing Quality Assessment and Reminder System with Real-Time Feedback in the Context of Infectious Diseases, *CHASE*, Dec. 2020.

University of Virginia

Evaluation

- Our own dataset
- 14 participants
  - Each 19 HW sessions
- 3 practice runs
- Video for Ground Truth

Evaluation


iWash
WristWash
H2DTR-NN
H2DTR-kNN
Accuracy (%)
0
20
40
60
80
100
Step 1
Step 2
Step 3
Step 4
Step 5
Step 6
Step 7
Step 8
Step 9
Step 10

Acoustics: Exploiting Speech

- Distance Emotion Recognition
  - Happy, sad, angry, neutral
- Anxiety and Depression
- …

Mental Health

UVA ENGINEERING
LINK LAB

University of Virginia

Distance Emotion Recognition

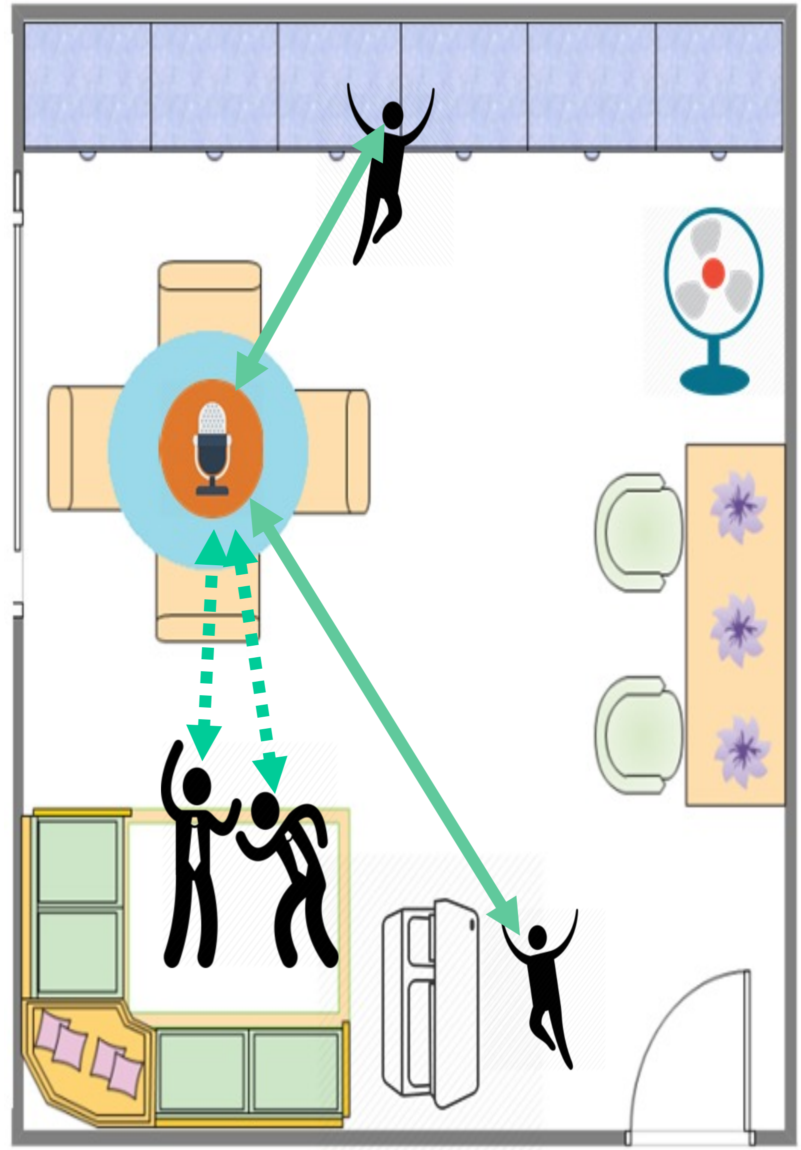
Close to microphone

Fixed distance

iia

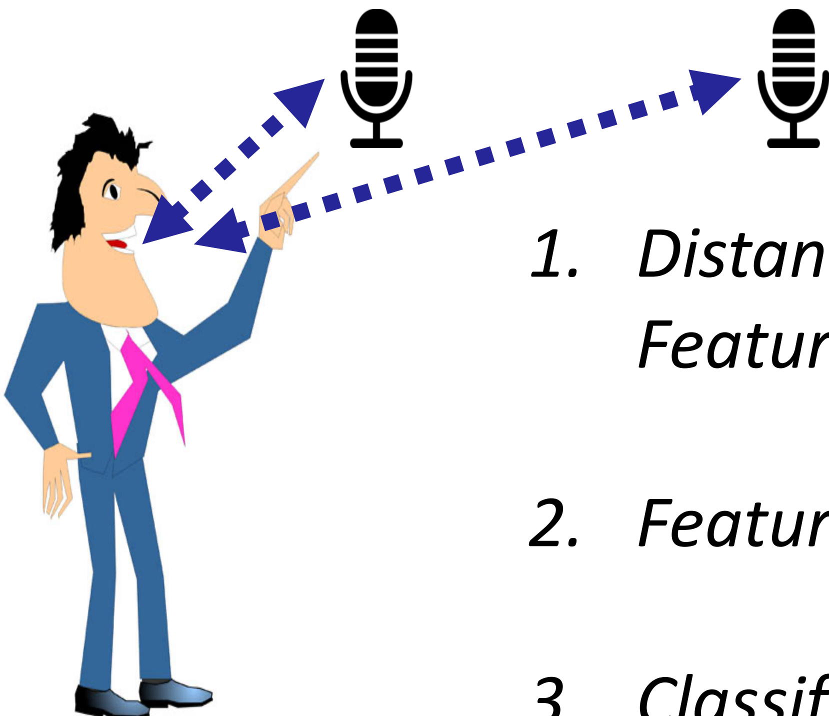
24/7

A realistic indoor speech emotion recognition system



- *Reverberation*
- *Ambient noise*
- *De-amplification of speech*
- *Overlapping of speech*

Solution



1. *Distance Agnostic Features/code words*

2. *Feature Modeling: Emo2vec*

3. *Classifier: LSTM*

Select Robust Features

Consider 231 LLD features

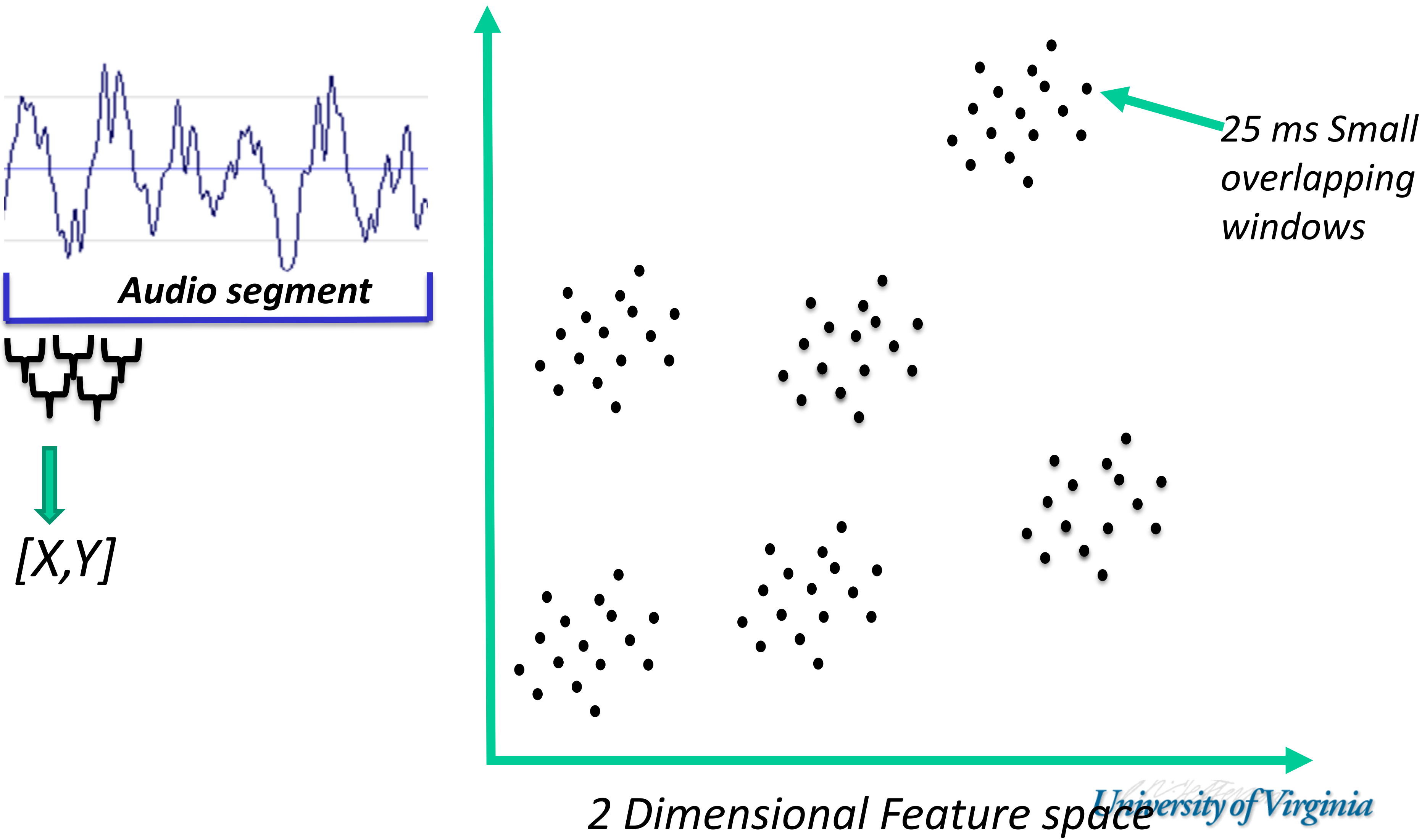
Feature	Count
Mel-Frequency cepstral coefficients (MFCC) 1-25	25
Root-mean-square signal frame energy	1
The voicing probability computed from the ACF	1
The fundamental frequency computed from the Cepstrum	1
Pitch	1
Harmonics to noise ratio (HNR)	1
Zero-crossing rate of time signal	1
PLP cepstral coefficients compute from 26 Mel-frequency bands	6
The 8 line spectral pair frequencies computed from 8 LPC coefficients	8
Logarithmic power of Mel-frequency bands 0 - 7	32

Select 48 LLD features:

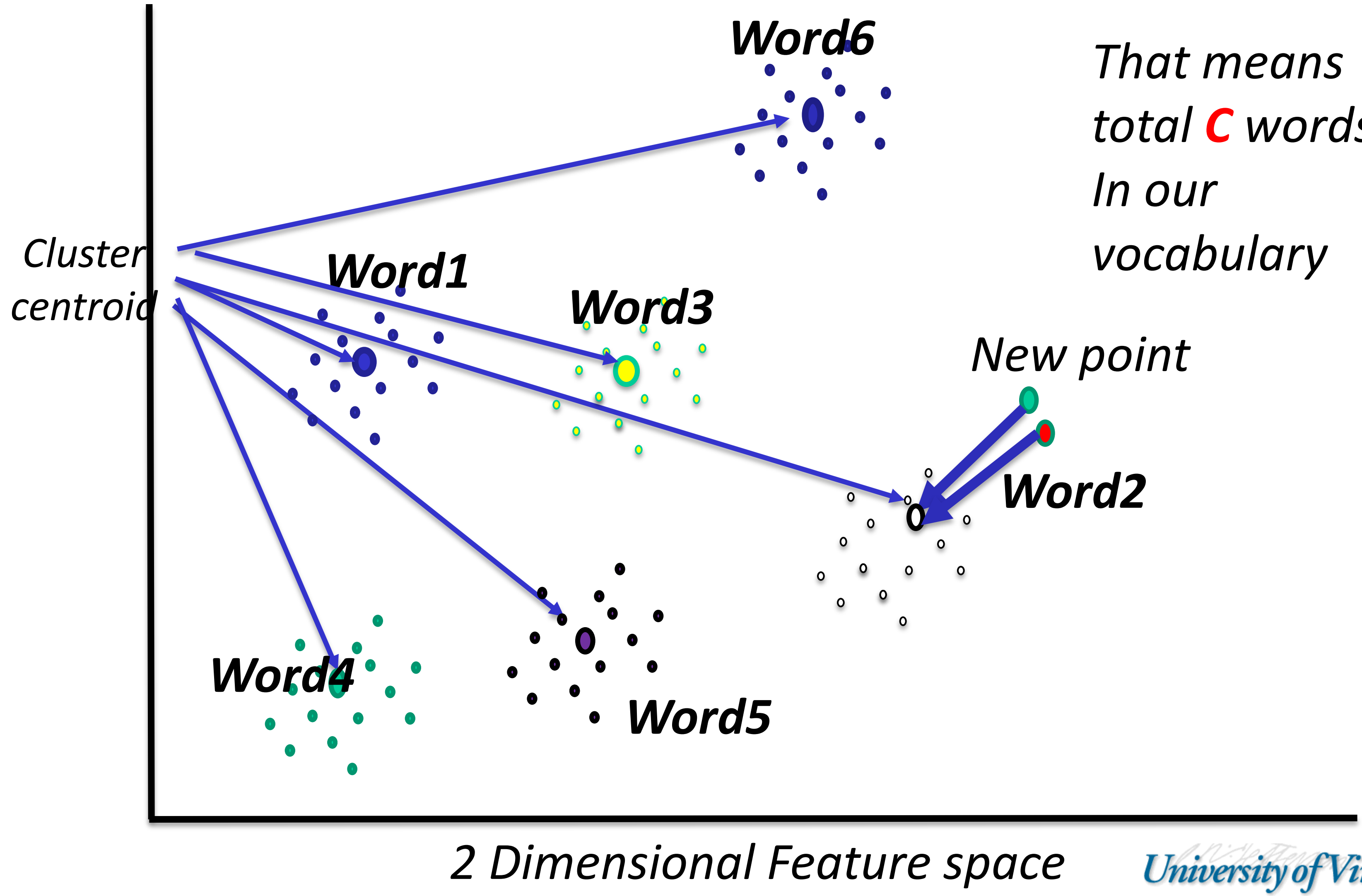
- *5 MFCC*
- *Voice probability*
- *Fundamental frequency*
- *Zero crossing rate*
- *8 line spectral pair frequencies*
- *32 logarithmic power of Mel-frequency bands*

Delta and delta-delta of these 77 features

Audio (Code) Words



Audio Word



C clusters

That means total C words In our vocabulary

Code Book Sizes

- Tested 500 to 2500 in increments of 500
  - K-means clustering
- Interesting Result: Different code book sizes for different emotions

Adaptation of Word2Vec : Emo2vec

- Convert audio words into vectors
- Words which occur in similar context (that means with similar neighbor words), for **a specific emotions** have similar vector representations.

Words A and C in similar Context but Not for Happy

(word,{Neighbour set})

⋮

(A , {P,Q,R,S,T,U,V,W,M})
(A , {P,R,Q,S,T,U,V,W,N})

⋮

(B , {O,P,Q,R,S,T,U,V,W})
(B , {P,Q,R,S,T,N,U,V,W})
(B , {P,R,S,T,U,V,W,M,Q})

⋮

(D , {E,F,E,G,H,E,B,C})
(D , {G,H,F,E,J,I,GW})
(D , {F,O,X,D,K,M,N,J})

Input corpus of happy D_H

(word,{Neighbour set})

⋮

(C , {P,Q,R,S,T,U,V,W,X})
(C , {P,R,Q,S,T,U,V,W,N})
(C , {P,R,Q,S,T,U,V,M,N})

⋮

(E , {F,X,P,Y,Z,S,T,W})
(F , {A,P,E,G,H,H,J,J})
(J , {E,F,J,M,M,K,N,P})

⋮

Input corpus of Not happy D_N

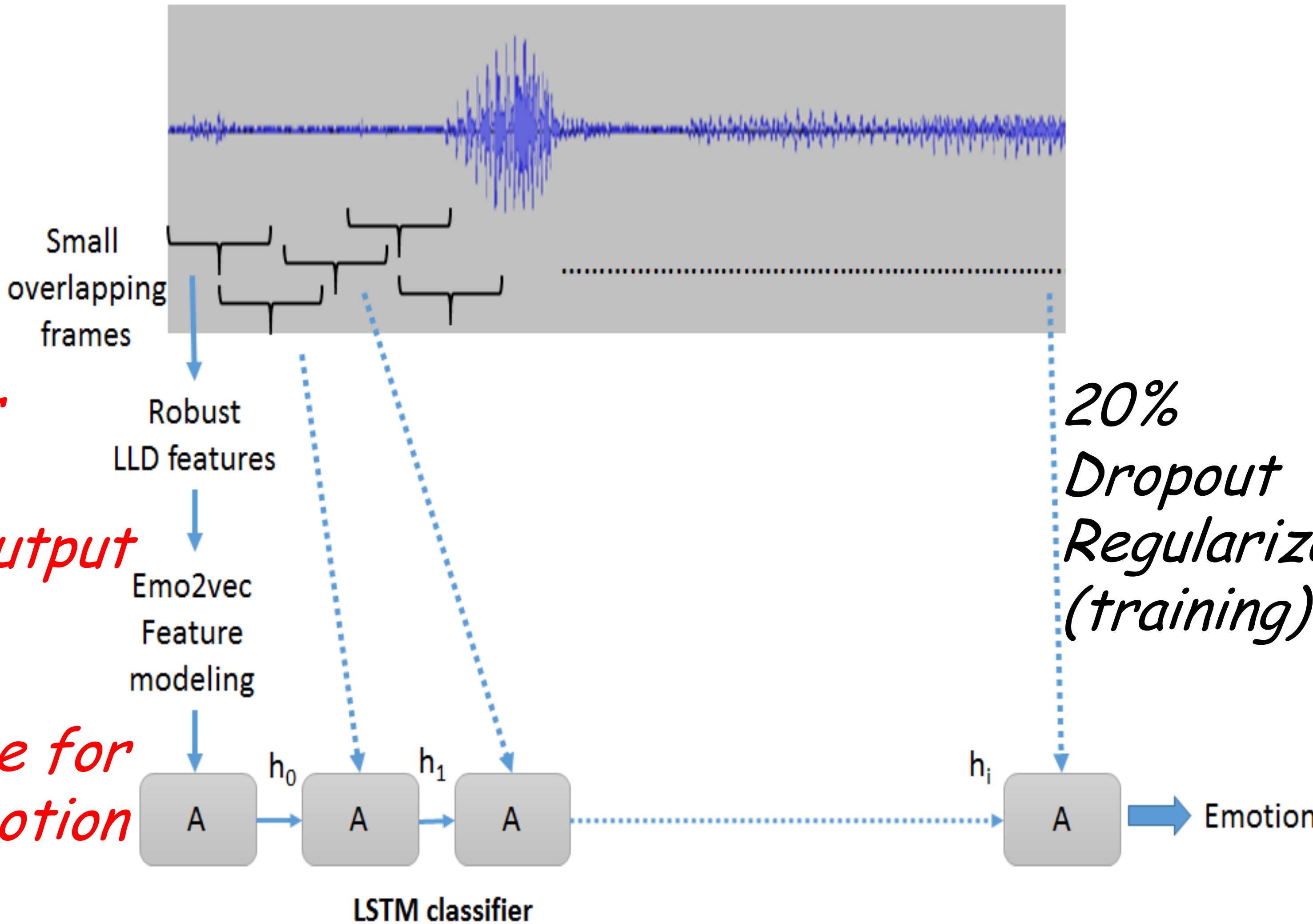
Words A & B, appear In similar context (with similar neighbors) for Emotion Happy

LSTM Classifier

100 Neurons

Dense output Layer

Separate for each emotion



20% Dropout Regularization (training)

Evaluation

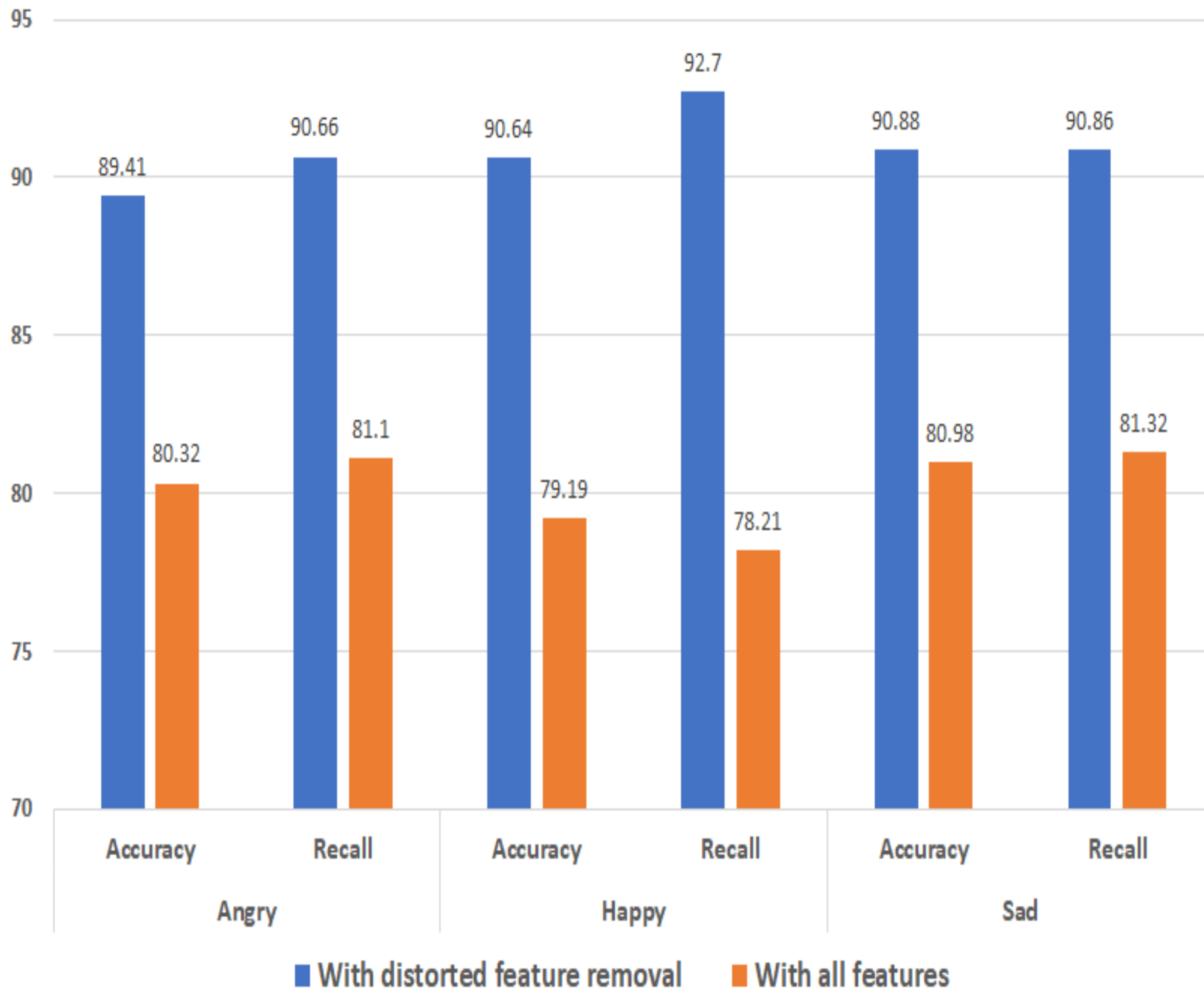
- 2 literature datasets
- Our own family-discussion experiments
  - 12 families; 28 people
  - Spontaneous discussions
  - Similar performance to literature datasets
- 4 Baselines
  - Approximately 16% better than best baseline

A. Salekin, Z. Chen, M. Ahmed, J. Lach, D. Metz, K. de la Haye, B. Bell, and J. Stankovic, Distance Emotion Recognition, *ACM Interactive, Mobile, Wearable, and Ubiquitous Technologies*, Vol. 1, Issue 3, Sept. 2017, 96:1-96:24.

Observation

- In the past – Happy versus Angry difficult (acted datasets)
  - Why? - Use of Energy based features
  - In real setting: Laughter helps discriminate

Elimination of distorted features helpful?



versity of Virginia

Effect of Distance

- As we move from mic to 6m away from mic, drop in accuracy is about 5%

- State of art: the drop is about 12%

Acoustics – Open Q

- Do we have Ground Truth?
- Noise, Reverberation, and Distance (in the wild)
- Similar Voices (father/son, …)
- Overlapped speech
- Speech changes with have a cold, tired, drinking alcohol, taking medicine, …
- What affect and multiple affects (dynamic)
  - People have more than 1 emotion at the same time
- …

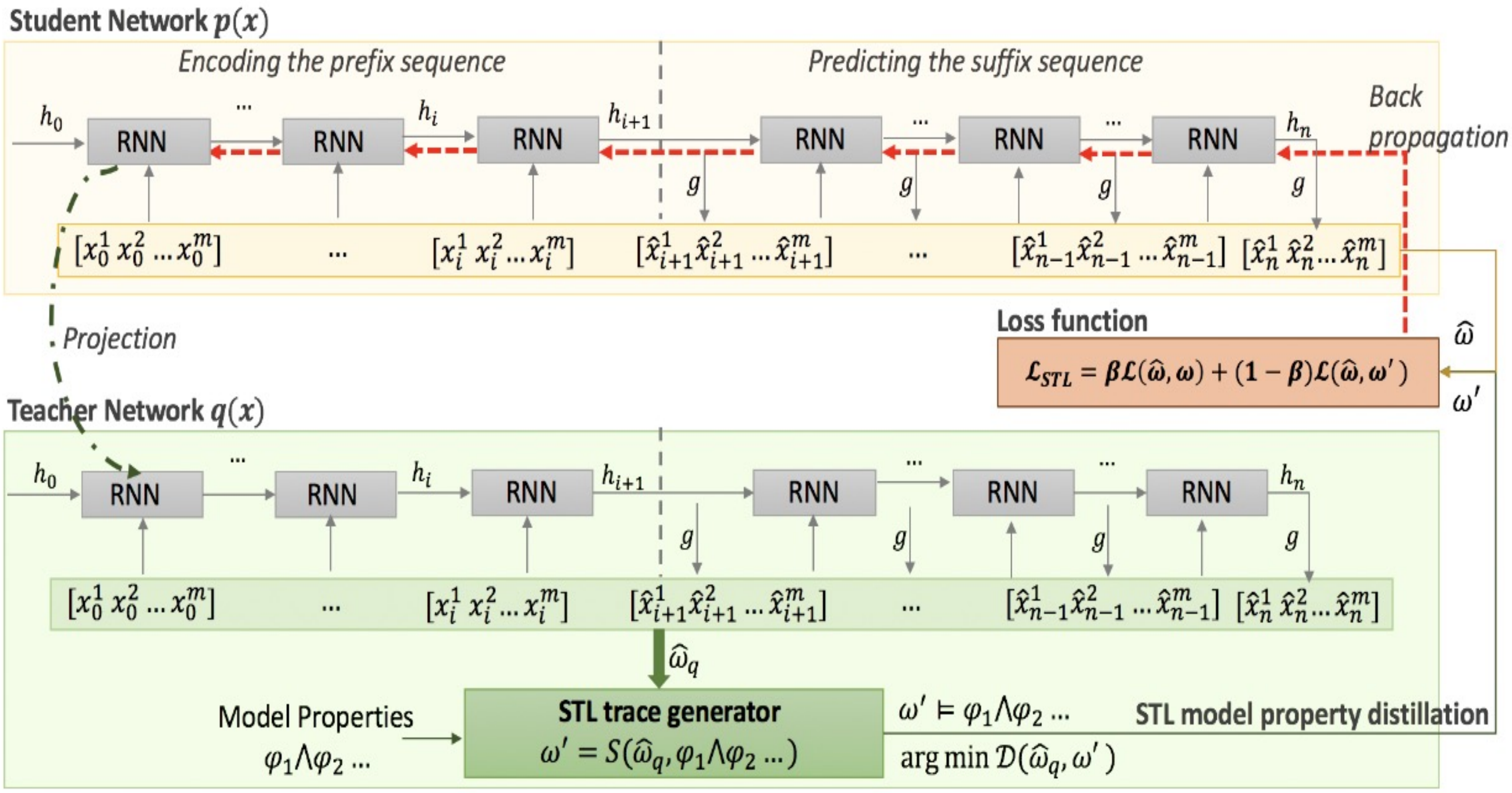
Models are Central to AmI

- Robustness
- Guarantees
- Uncertainties
  - Softmax is too limiting
- Dynamic Learning
  - Meta learning, Knowledge Graphs, GNN, …

Robustness

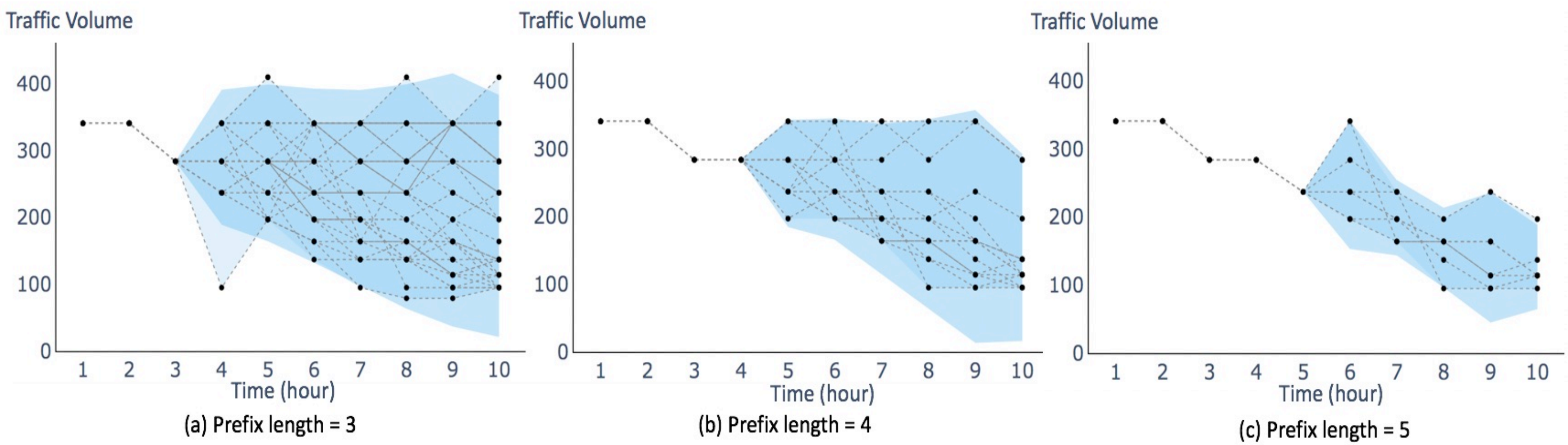
- Integrate properties into models
  - Physics (a lot of work here)
    - Typical: Model Equations
  - World Knowledge (Knowledge Graphs)
    - City has 5 major hospitals
    - Parents had dementia
  - Use Formal Methods

Properties into Models via FM



STLnet

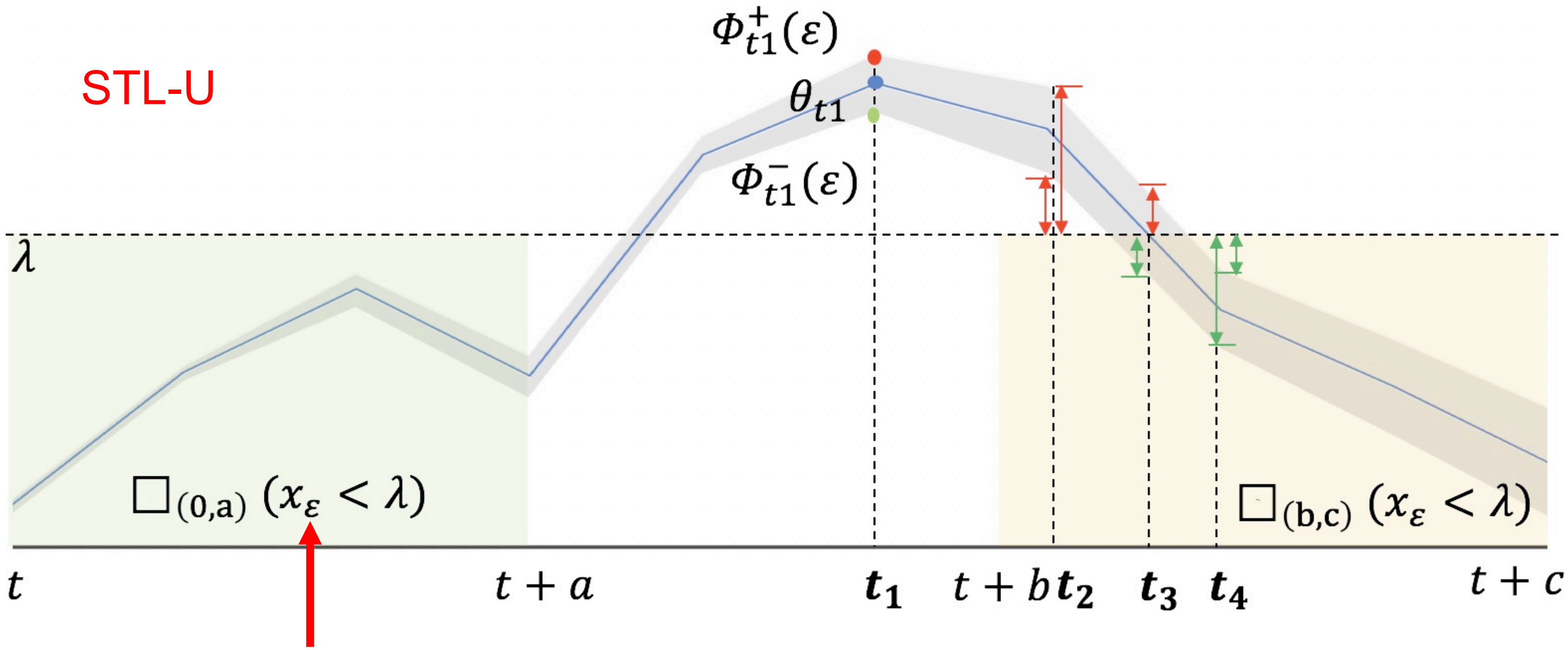
Uncertainty Exists
Smart City



Smart City Traffic Volume

STL-U: Uncertainty

STL-U



FLOWPIPE

AQI (Pollution)

STLnet and STL-U

- Demonstrates the feasibility of integrating formal methods and Bayesian deep learning toward accurate estimation of uncertainty.
- Incorporate real-world requirements to guide the learning process.

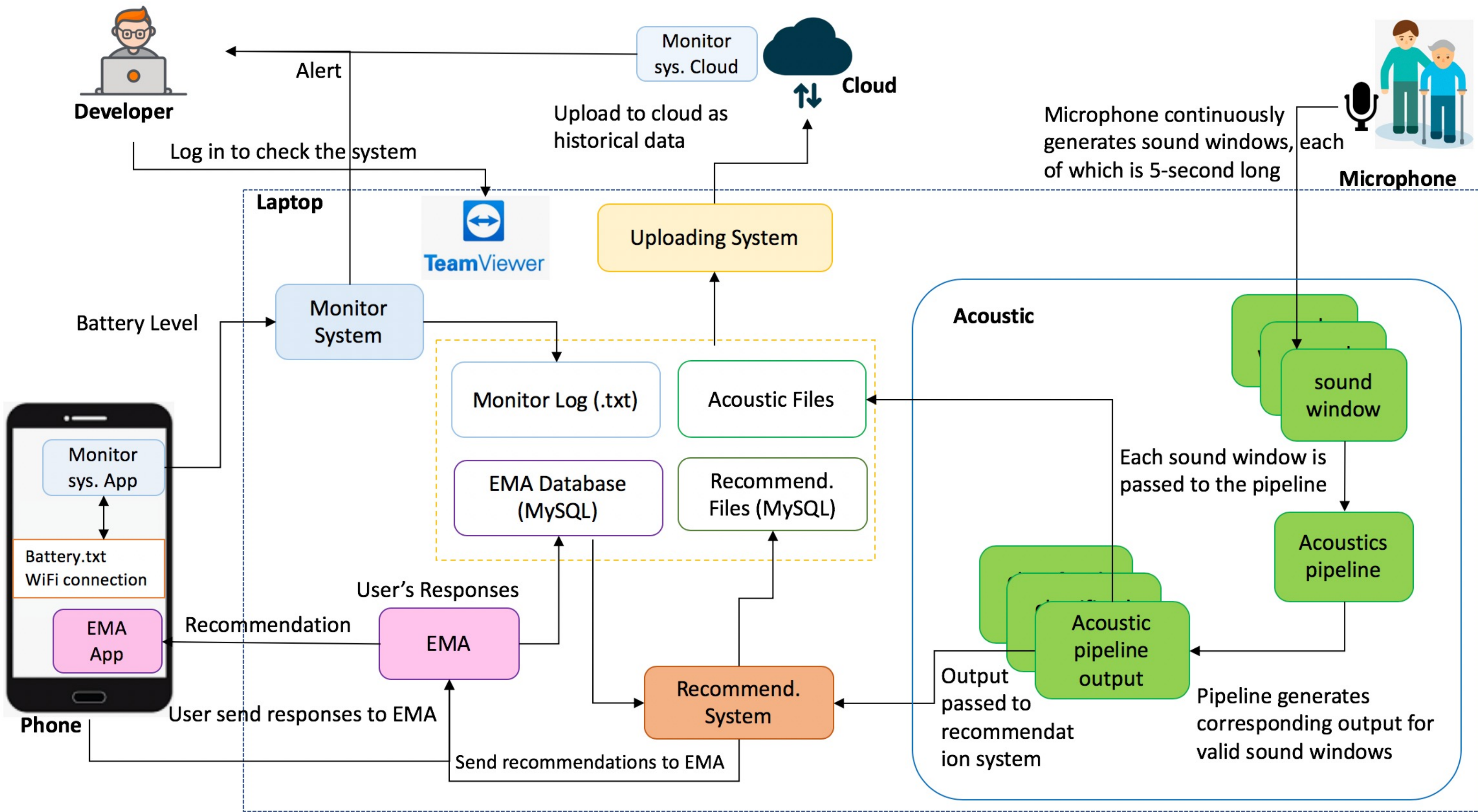
M. Ma, J. Stankovic, L. Feng, Predictive Monitoring with Logic-Calibrated Uncertainty for Cyber Physical Systems, *EMSOFT*, Oct. 2021.

M. Ma, L. Feng, and J. Stankovic, STLnet: Signal Temporal Logic Enforced Multivariate Recurrent Neural Networks, *NeurIP*, 2021.

Real Deployments

- Family Eating Dynamics for Obese Families - USC – 23 families
- Alzheimer's Patients and Incontinence, 13 families
- Alzheimer's Patient – Caregiver Interaction – Ohio State and Univ. of Tennessee, 6 four-month deployments
- Smart Watch Handwashing, UVA PICU and 15 homes

Alzheimer's – PCR System



6 Real Deployments
4 Months Each

University of Virginia

Lessons Learned – Expect the Unexpected

- Pulled wires
- Internet down (a lot)
- Moved system to another room
- Went on vacation
- Got covid
- Same exact context – different human choice
- Similar voices
- A lot of TV sounds
- Construction noise
- Yelling across rooms detected as anger

Sounds Encountered

Physiological: Sneezing, nose blowing, sniffling, clearing throat, hiccup, eating, burp, humming, laughter, drinking, snoring

Objects: phone vibrating or ringing, typing, mouse wheel, unwrapping food, papers rustling, clothes rustling, television, piano, moving furniture, doors opening and closing, objects dropping or moving, footsteps, pouring liquid, coffee percolation, dishwasher, cleaning sounds

Ambient: truck backing up, siren, birds chirping, passing airplane, traffic, motorized tools (lawnmower, etc)

Better Methodology

- Develop ML solutions with datasets
- Test in lab
- Test in trial homes for short periods of times (but w/o real patients)
- Real Deployment

- Open Q: What to do at development time?
  - To better ensure success at deployment time

Summary

- Future of AmI for Healthcare (a long way off)
- AI/ML and CPS
- Key modality: Acoustics/Speech (Privacy)
- Robust and Dynamic Learning Models

Thanks

- Abu Mondol, now at Amazon
- Sirat Samyoun, at UVA
- Asif Salekin, now faculty at Syracuse
- Meiyi Ma, now faculty at Vanderbilt
- Sarah Preum, now faculty at Dartmouth

University of Virginia

ity of Virginia