Large-scale Automatic Depression Screening Using Smartphone Data

Dr. Bing Wang

Computer Science & Engineering Department

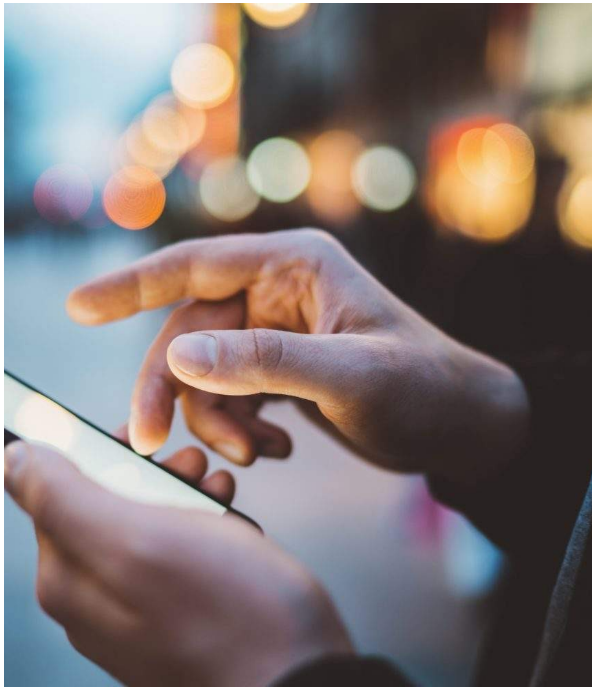
UCONN | UNIVERSITY OF CONNECTICUT

Introduction

- Depression is a serious illness
  - Significant effects on both physical and mental health
  - Higher medical costs, mortality
- Current diagnosis methods
  - Rely on physician-administered or patient self-administered surveys
  - Subjective, burdensome, and recall bias
  - Lack of trained professionals (14.5 psychiatrists per 100,000 in United States)

Urgent need for an accurate, objective and easily-accessible depression screening tool for mass usage

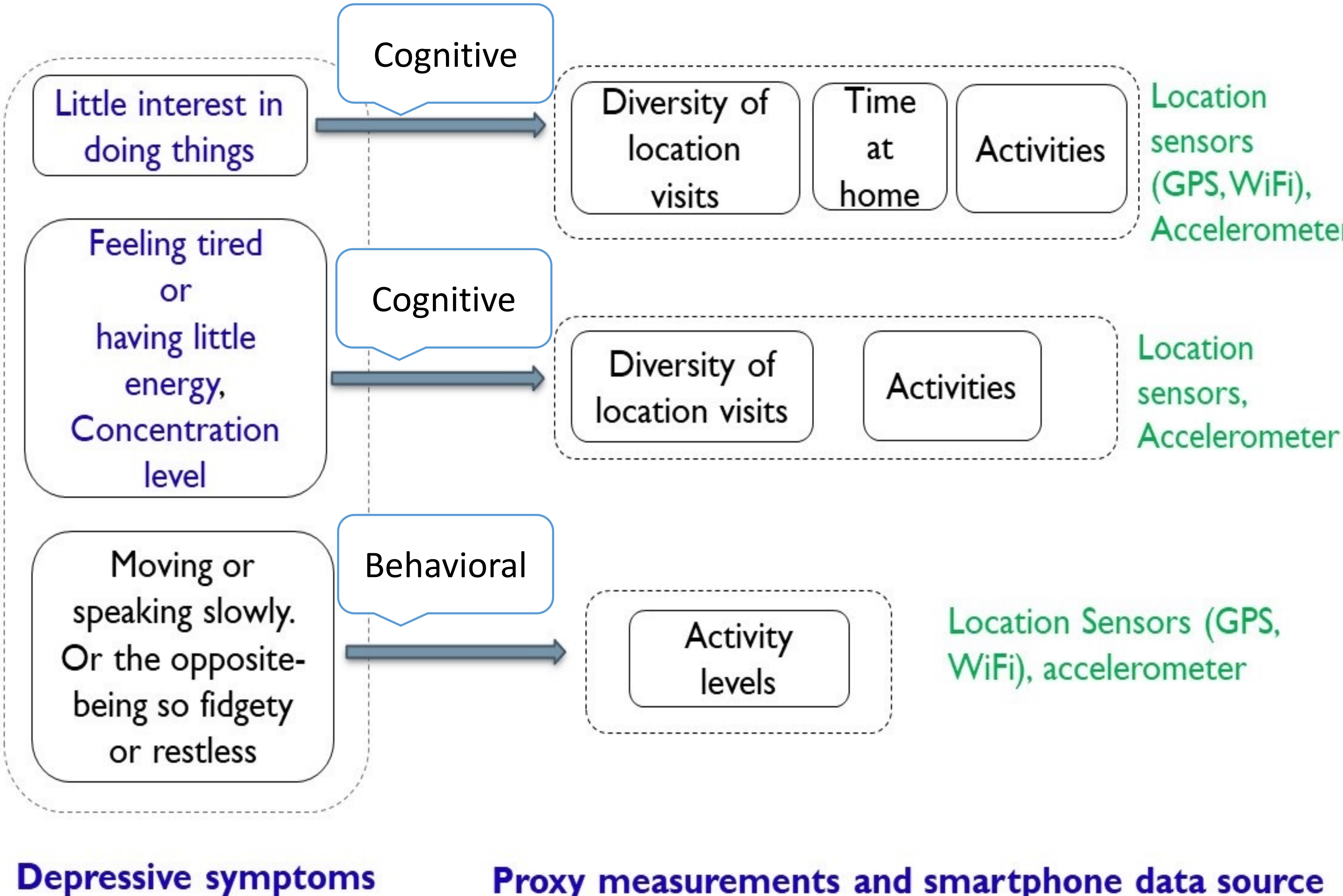
New Opportunities



- Automatic depression screening using smartphone sensing data
  - Ubiquitous adoption of smartphones
  - Rich set of sensors reflect a user's behavior

- Advantages
  - Objective, data automatically collected
  - Ideal for continuous monitoring

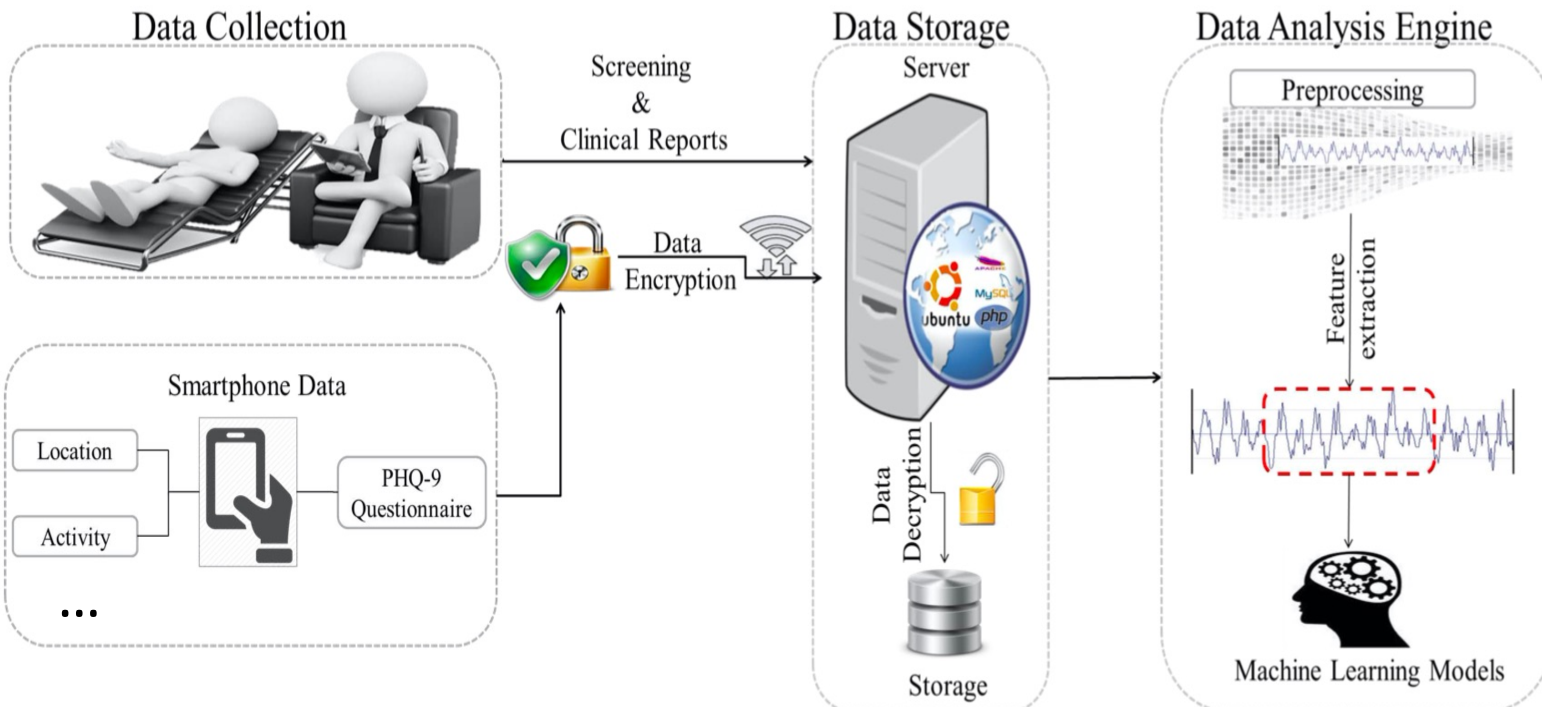
Why it might work?



Challenges

- Human behaviors extremely stochastic
- Correlations between behavior and depression are complex
- Opportunistic data capture
- Identifying meaningful features

Our Experience: LifeRhythm Project



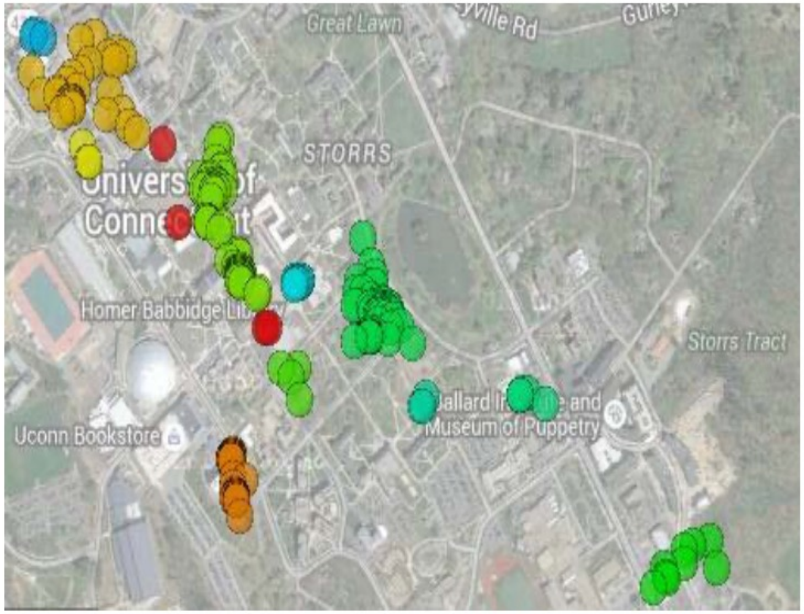
LifeRhythm: a system for automatic and pervasive depression screening using smartphone data

Outline

- Binary prediction: depressed or not
  - Case study: using location data
- Other types of prediction
  - self-report depression scores
  - depression severity
  - depressive symptoms
- Using other types of sensing data
  - Activity data
  - Meta-data of Internet traffic
  - Social interaction (statistics of SMS, phone)
- Conclusion and future work

Depression Prediction Using Location Data

- Location data
  - behavioral characteristics

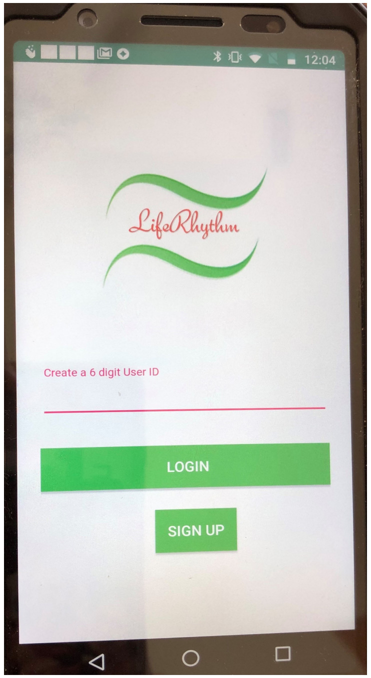


- Two settings: passive data collection, requiring no user interaction
  - Running app to collect location data on individual phones
  - Does not require running app on phones

Depression Prediction Using Location Data Collected on Phones

Data Collection

- **Location data: passively collected**
  - LifeRhythm app for Android and iOS
  - GPS: longitude, latitude
    - Andriod: periodic every 10 min; iOS: event-based
  - WiFi association events



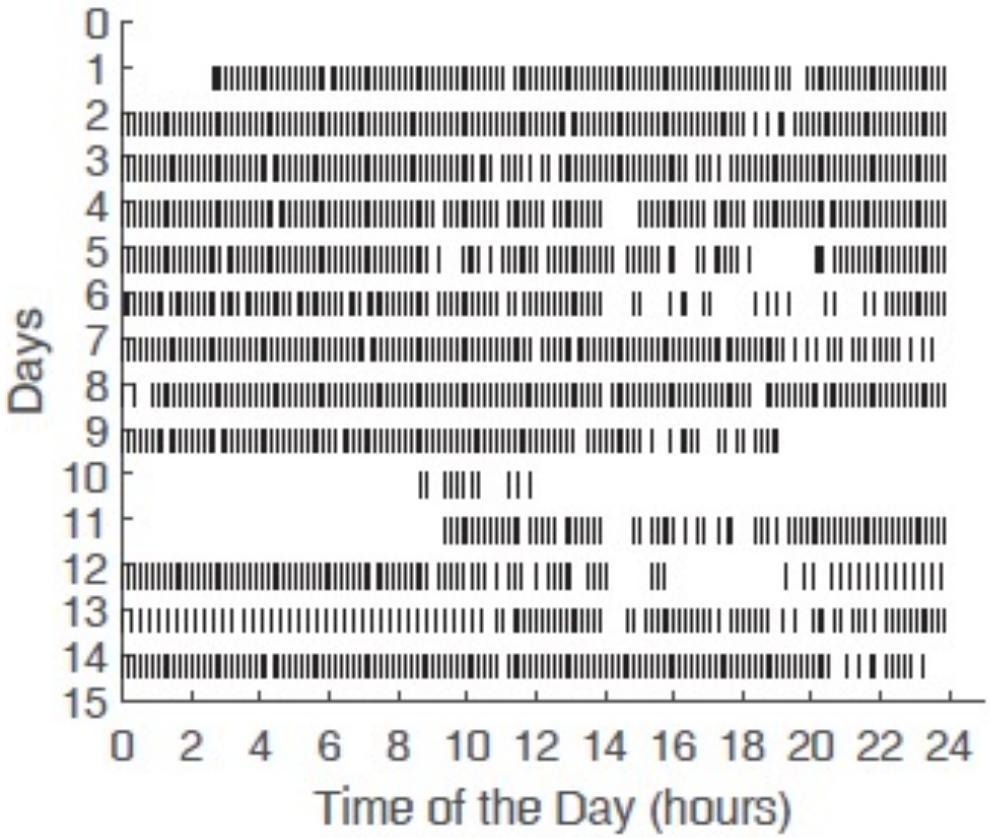
- **Questionnaire responses**
  - Self-reported depressive symptoms
  - Collected by LifeRhythm app
  - Weekly or biweekly, using notification on phones

- **Clinical assessment**
  - Initial screening: depressed or non-depressed
  - Follow-up meetings with depressed participants (once or twice per month)

Participant Recruitment

- Two-phase study at University of Connecticut
  - Participants: full-time students, aged 18-25
- Phase 1: 10/2015 – 5/2016
  - 79 participants (19 depressed; 60 non-depressed)
  - Participation duration: up to 8 months
  - Self-report questionnaire: PHQ-9 (Patient Health Questionnaire) 9-item, biweekly
- Phase 2: 2/2017 – 12/2027
  - 103 participants (39 depressed; 64 non-depressed)
  - Self-report questionnaire: QIDS (Quick Inventory of Depressive Symptomatology), 16-item, weekly
- Participants use own phones: Android (variety of manufactures) or iPhone

Large Amount of Missing Data



GPS location data for an Android user over 14 days

Reasons: low battery level, poor GPS signal, or user turns off GPS

How to deal with missing data?

- Common approach: remove time periods with substantial missing data
  - All the data (even high-quality data) in the time period are removed
- Existing data imputation approaches do not work well
  - Matrix Factorization, Multiple Imputation by Chained Equations, and Nuclear Norm Minimization
- Our approach: data fusion

Fusing GPS and WiFi Association Data

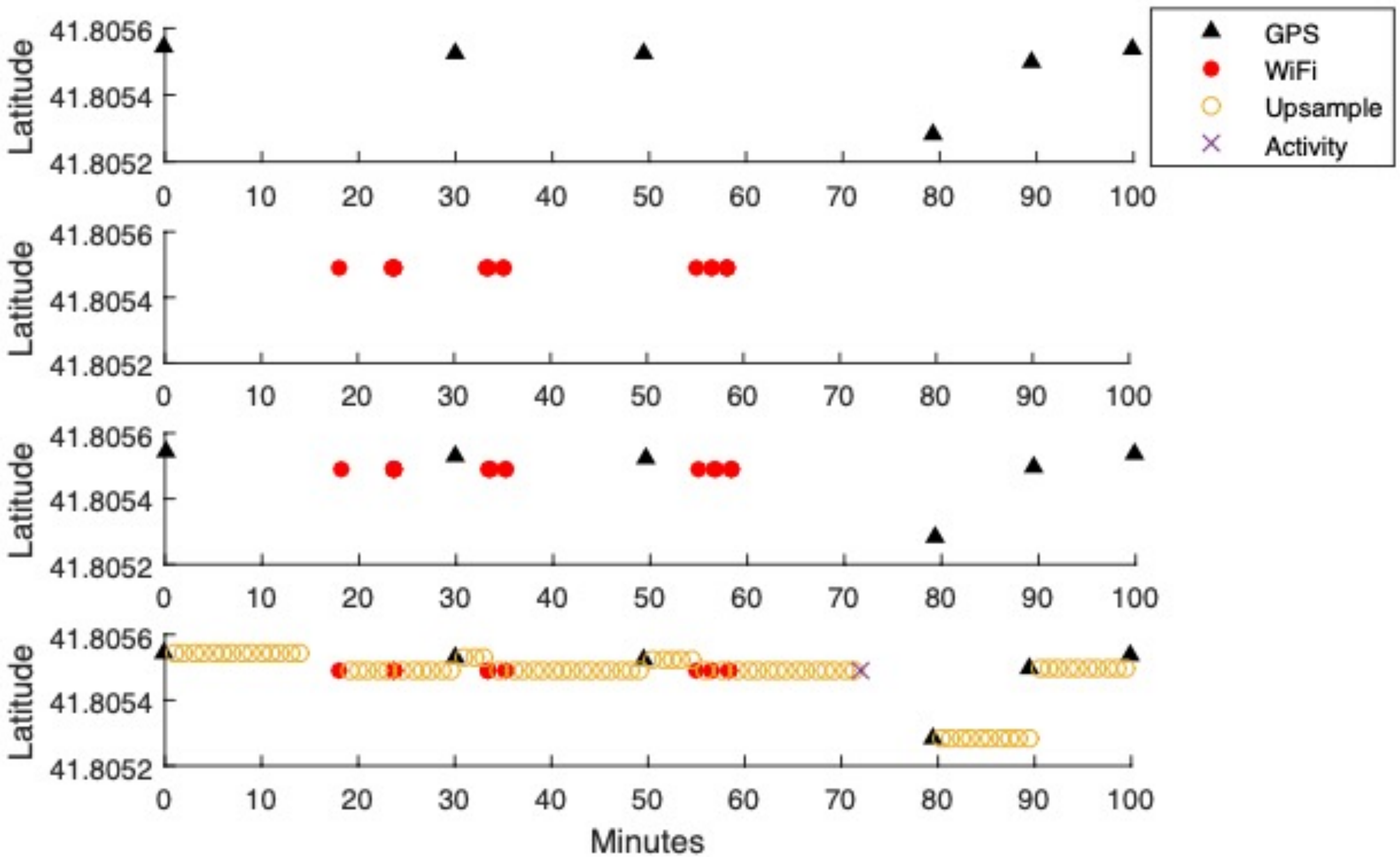
- Both collected on phones
- Both provide location information
  - GPS: longitude, latitude with accuracy
  - WiFi association: use location of Access Points (APs) to approximate phone locations
- GPS: fine-grained
  - but high energy consumption and does work well indoors
- WiFi association: coarse-grained
  - but low energy consumption and works well indoors

Question: Can we use GPS and WiFi data to get more complete data? Can it improve depression prediction?

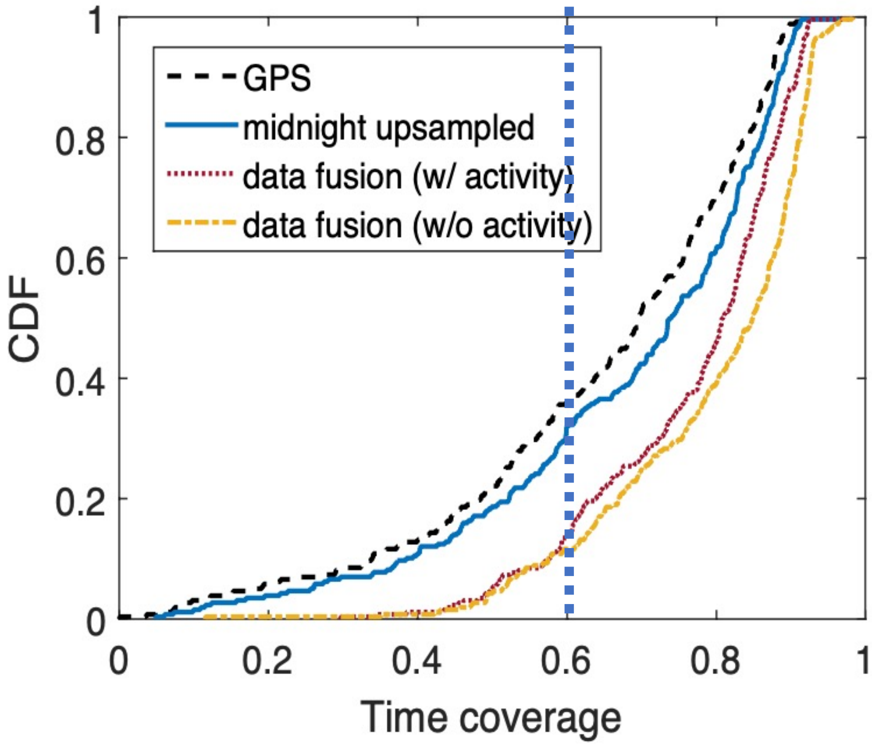
Data Fusion: High-level Approach

- Automatically determine (longitude, latitude) of WiFi APs
  - If phone associates with an AP at time *t*, find GPS records close to *t* to approximate the AP location
  - Use GPS records collected from multiple users and points of time to improve approximation
- Merge GPS and WiFi data
- Obtain location estimates in continuous time

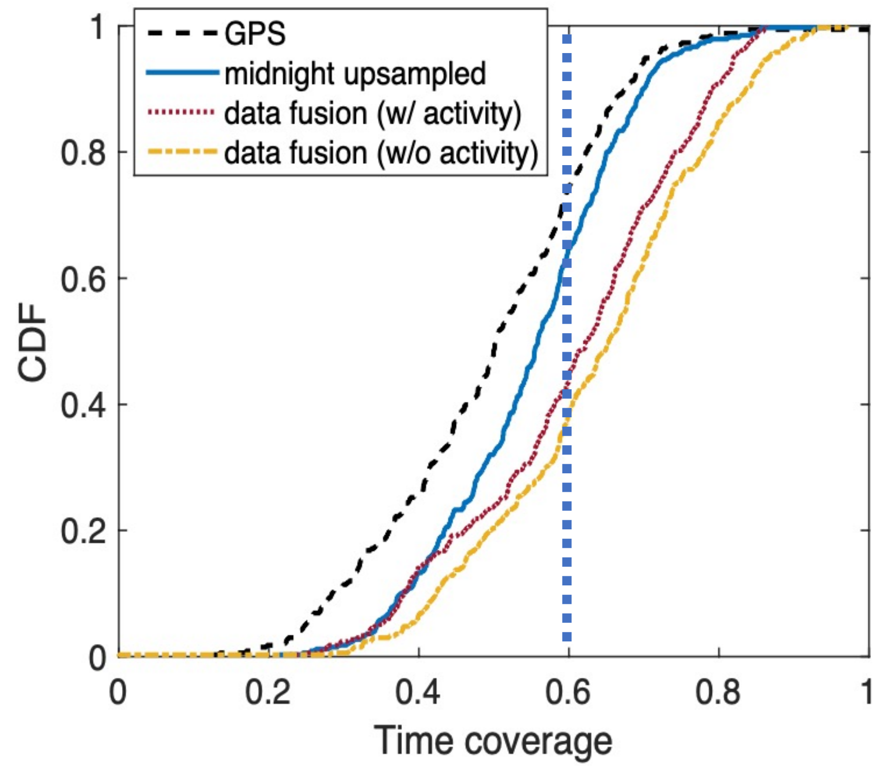
Data Fusion: Illustration



Data Fusion Results



Android data



iOS data

Significantly better time coverage after data fusion

Data Analysis and Model Training

- Phase 1 study
  - Analysis based on PHQ-9 intervals
    - Day when filling in a PHQ-9 + previous 14 days
  - Extract features from location data in PHQ-9 interval
  - Correlate features with PHQ-9 score
  - Use the features to train classification models to predict depression status
    - depressed or non-depressed: based on clinician assessment
- Phase 2 study
  - Analysis based on QIDS intervals
    - Day when filling in a QIDS + previous 7 days

Feature Extraction

- Based on raw location data
  - Location variance: variability in participant's location
  - Time spent in moving (i.e., speed > 1km/hr)
  - Total distance
  - Average moving speed
- Based on location clusters (using DBSCAN)
  - # of unique locations/clusters
  - Entropy
    - Variability of time that a participant spends at different locations
  - Normalized Entropy (Entropy normalized by # of clusters)
  - Time spend at home
    - Home cluster: participant most frequently found between [0, 6]am

Correlation Results

- Pearson's correlation (significance level α = 0.05)

Phase I Android dataset

	No data fusion		Data fusion	
Features	R-value	P-value	R-value	P-value
Loc_{var}	-0.15	0.07	-0.24	0.001
Distance	-0.13	0.11	-0.04	0.61
AMS	-0.09	0.28	-0.04	0.62
Move	0.06	0.43	-0.11	0.11
Entropy	-0.16	0.05	-0.28	10^{-4}
NEntropy	-0.21	0.01	-0.26	10^{-4}
Home	0.18	0.03	0.23	0.003
N_{loc}	-0.09	0.28	-0.16	0.03

- Correlation after data fusion is significantly better

Classification Results

- Use Support Vector Machine with Radial Basis Function kernel

Phase-I dataset

	F1 score	Precision	Recall	Specificity
Android (no data fusion)	0.50	0.64	0.41	0.71
Android (w/ data fusion)	0.66	0.83	0.56	0.88
iOS (no data fusion)	0.50	0.59	0.42	0.77
iOS (w/ data fusion)	0.76	0.77	0.76	0.77

- leave-one-user-out cross validation
- much better classification results after data fusion

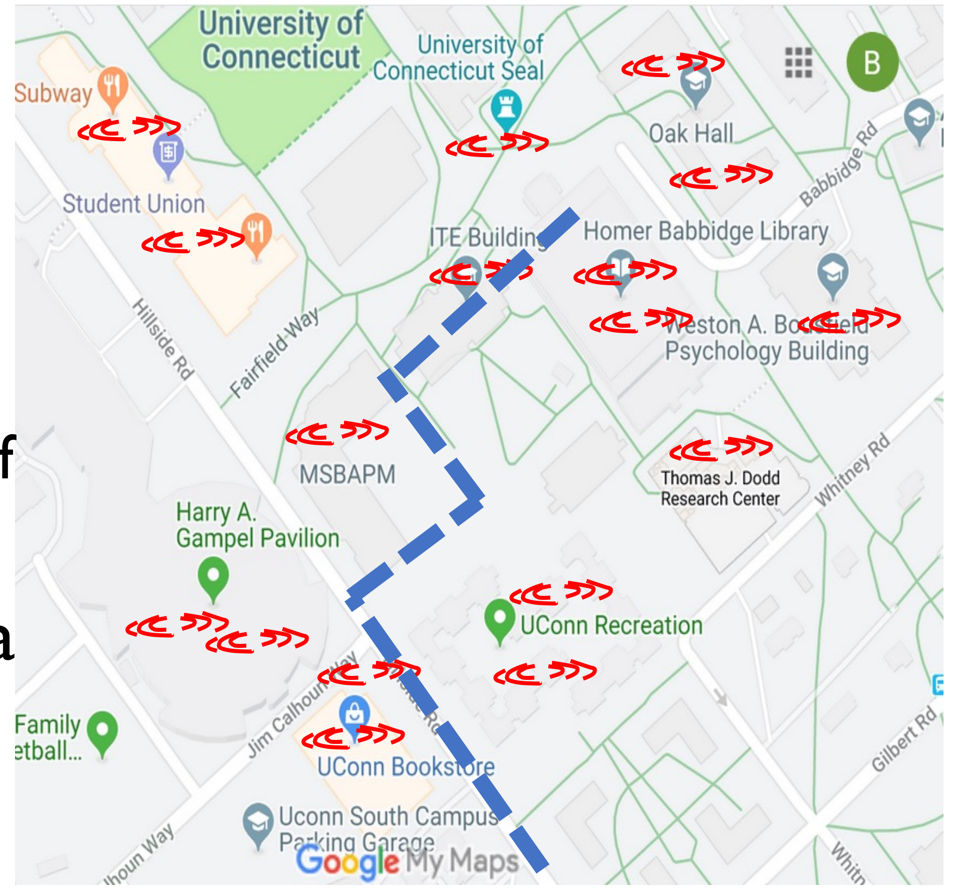
Depression Prediction: Other Approaches?

- So far, using location data directly captured on phones

Question: Can we get by without direct data capture (running apps) on phones?

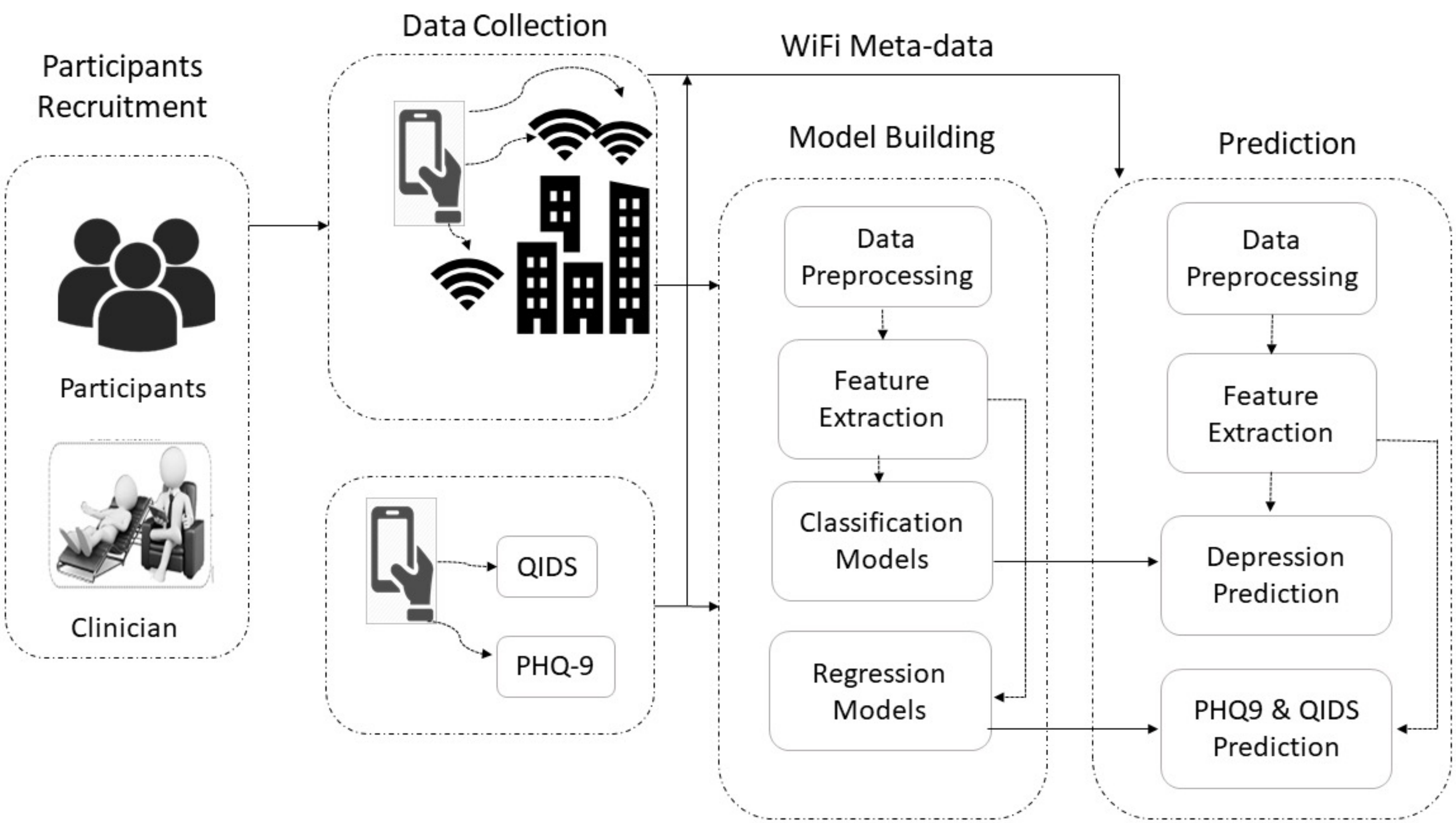
Using Meta-data from WiFi Infrastructure

- WiFi widely deployed in institutions
- Phones associate with nearby access points (APs)
- Locations of phones/users
  - Approximated by locations of APs associated with
- Obtain WiFi association data directly from infrastructure
  - Standard protocols



Large-scale Depression Prediction Using Meta-data from Institution WiFi Infrastructure

High-level Approach



- Large-scale automatic depression screening in an institution
- No need to run any app on individual phones

Privacy and Responsible Usage of Data

- Clearly very important!
- Need careful mechanisms for user consent and privacy
- Usage scenarios
  - Depression screening at population level
    - e.g., estimate the rate of depression in an institution
    - No need of user identity
  - Depression screening at individual user level
    - Users elect to use the service
    - Need certain user identity info
    - User privacy can be preserved via cryptographic techniques

Our Focus

- Is it feasible to learn accurate prediction models for depression screening from WiFi meta-data?
- Challenges:
  - WiFi location data is of lower resolution
  - Location collection is opportunistic
    - can only be captured when a phone is connected to institution WiFi network
- There is hope
  - Users prefer to connect to institution WiFi network
  - Use data collected over a period of time

Two Settings

- 24-hour monitoring
  - Applicable to students living on campus
  - 24-hour data provides more insights into a user's behavior
- Daytime (8am-6pm) monitoring
  - Applicable to commuting scenario
  - Does daytime data alone already provide sufficient insights?

Data Analysis at Three Levels

- AP level analysis (serves as baseline)
  - Treat each AP as a unique location
- Building level analysis
  - Treat each building as a unique location
  - Obtain the building where an AP is located beforehand
- Enhanced building level analysis
  - Building categories: entertainment, sports, library, …

Feature Extraction

- Number of unique locations
- Entropy and normalized entropy
- Time spent at home (for 24-hour monitoring only)
  - Home: most frequently found location between [0,6]am
- Circadian movement
  - Extent of a user's locations following circadian rhythm
- Number of significant locations visited
  - Number of locations visited among top-10 locations spent time
- Routine Index
  - Difference of locations visited across days
- Additional building category features
  - Duration, # of days visiting different categories of buildings

Correlation Results

		All		Depressed		Non-depressed	
	Features	**r-value**	**p-value**	**r-value**	**p-value**	**r-value**	**p-value**
Phase I 24-hour monitoring	Entropy	-0.36	0.00	-0.40	0.01	-0.24	0.009
	Entropy_N	-0.33	0.00	-0.44	5×10^{-3}	-0.06	0.51
	Home	0.22	6×10^{-3}	0.40	0.01	-0.20	0.03
	N_{loc}	-0.26	9×10^{-4}	-0.22	0.10	-0.36	10^{-4}
	CMove	0.01	0.86	-0.20	0.22	0.12	0.19
	N_{sig}	-0.13	0.12	-0.16	0.36	0.008	0.93
	RIndex	0.32	10^{-4}	0.34	0.03	0.05	0.60
Phase I Daytime monitoring	Entropy	-0.37	10^{-4}	-0.33	0.02	-0.39	3×10^{-4}
	Entropy_N	-0.36	10^{-4}	-0.41	0.02	-0.22	0.04
	N_{loc}	-0.24	0.04	-0.09	0.60	-0.41	10^{-4}
	CMove	-0.19	0.04	-0.23	0.24	-0.11	0.32
	N_{sig}	-0.12	0.21	-0.10	0.60	-0.02	0.84
	RIndex	0.46	0.00	0.37	0.05	0.51	0.00

- Stronger correlation for depressed participants
- Significant correlation even for daytime monitoring

Classification Results

Phase-I data, 24-hour monitoring

	F_1 score	Precision	Recall	Specificity
AP level	0.66	0.63	0.70	0.58
Building level	0.84	0.90	0.77	0.90
Enhanced building level	0.83	0.78	0.89	0.75

Phase-I data, daytime monitoring

	F_1 score	Precision	Recall	Specificity
AP level	0.74	0.74	0.75	0.72
Building level	0.75	0.67	0.80	0.62
Enhanced building level	0.74	0.71	0.78	0.69

- Even day-time monitoring provides accurate prediction
- Comparable results a those when using location data directed collected on phones

Other Types of Prediction

- So far, using location data for binary prediction

- Predicting self-report depression score
  - PHQ-9 scores (numerical value: 0-27, Phase I)
  - QIDS scores (numerical value: 0-27, Phase 2)
- Predicting depression severity
  - Severity level: stable, mild, moderate, severe; level 0-3
- Predicting depressive symptoms
  - Whether a depressive symptom is present or not

Automatic Prediction of Depressive Symptoms

Depressive Symptoms

- 8 major categories
  - Behavioral: appetite, energy, psychomotor, sleep disturbance
  - Cognitive: concentration, interest, self-criticism, feeling sad

- Prediction approach
  - Train binary classification model for each symptom
  - Ground truth: Symptoms' scores from questionnaire
    - 0: no symptom, 1-3: occasional to persistent symptoms

Prediction of Depression Symptoms (Phone Data)

Phase I
iOS dataset
Depressed:

	F_1 score	Precision	Recall	Specificity
Appetite	**0.79**	**0.75**	**0.84**	**0.67**
Concentration	**0.70**	**0.65**	**0.75**	**0.50**
Energy	**0.74**	**0.64**	**0.87**	**0.57**
Feeling-depressed	**0.83**	**0.90**	**0.76**	**0.88**
Interest	**0.81**	**0.78**	**0.84**	**0.82**
Self-criticism	**0.81**	**0.80**	**0.82**	**0.78**
Sleep	**0.68**	**0.68**	**0.69**	**0.70**

- Both behavioral and cognitive depressive symptoms can be predicted accurately
- Overall better results for depressed population

Prediction of Depression Symptoms (WiFi Meta-data)

Phase I
24-hour monitoring
Depressed:

	F_1 score	Precision	Recall	Specificity
Appetite	**0.76**	**0.68**	**0.85**	**0.50**
Concentration	**0.63**	**0.57**	**0.71**	**0.53**
Energy	**0.80**	**0.73**	**0.89**	**0.67**
Feeling-depressed	**0.85**	**0.83**	**0.78**	**0.84**
Interest	**0.86**	**0.79**	**0.85**	**0.56**
Self-criticism	**0.86**	**0.84**	**0.89**	**0.83**
Sleep	**0.70**	**0.65**	**0.76**	**0.54**

- Similar results as those when using smartphone sensing data directed collected on phones

Using Other Types of Sensing Data

- So far, consider location data only

- Other types of sensing data
  - Collected by LifeRhythm app
  - Internet traffic
    - Traffic that destines to or originates from phones
    - meta-data: timing, source & destination IP addresses, size
  - Social interaction
    - SMS logs and/or phone logs
    - Statistics: # of incoming/outgoing messages/calls, …

Depression Prediction Using Meta-data of Internet Traffic

Using Internet Traffic

- Advantages of using Internet traffic
  - Easily collected
  - Much less energy consuming than collecting GPS data
  - Not sensitive to phone hardware
- For user privacy, capture only meta-data of each packet on phone; no payload
  - (timestamp, src IP, dest IP, size)

Question: How to extract behavioral features from packet-level meta-data?

Data Preprocessing

- Identify usage sessions (when accessing Internet)
  - remove keep-alive packets
  - handle background traffic
  - use heuristic to aggregate traffic to identify usage sessions

- Identify application categories
  - Based on destination IP address
  - Use public database to look up information on destination IP address
  - match keywords to determine application categories (mail, social, game, shopping, …)

Feature Extraction

- Aggregate usage features
  - Total duration online
  - Total # of sessions
  - Total off-duration
  - Internet usage in each time period (morning, afternoon, night, midnight)
- Category-based usage features
  - Duration of each application category
  - # of sessions of each application category
- Volume-based features
  - Total volume
  - Volume in each time period

Classification Results

❖ Binary classification: depressed or not

Phase-I, iOS dataset

	F_1 score	Precision	Recall	Specificity
Aggregate features	0.63	0.67	0.60	0.59
Category-based features	0.65	0.62	0.68	0.61
Aggregate + category-based features	0.71	0.71	0.71	0.63

- Using aggregate or category-based features alone can already lead to good prediction
- Aggregate + category-based features leads to even better prediction
- Usage based features generally leads to better prediction than volume-based features

Conclusion

- Using smartphone data for depression prediction is promising
  - Wide range of prediction
    - depression status, depressive symptoms, depression questionnaire scores, depression severity level
  - Wide variety of sensing data
    - Location, Internet traffic, activity, social interaction
    - Different tradeoffs in ease of data capture, energy consumption, scalability
- Further validation
  - Larger sample size
  - More diverse demographic groups

Future Work

- Depression prediction
  - Techniques to deal with missing data
  - Better features
  - Handling data from different platforms
  - Stable data collection
  - Large dataset, deep learning
  - …
- Personalized treatment of depression
  - Using smartphone data to predict response to treatment → assist clinician decision making
  - …

Related Papers

- Group website: https://nlab.engr.uconn.edu/

- C. Yue, S. Ware, R. Morillo, J. Lu, C. Shang, J. Bi, J. Kamath, A. Russell, A. Bamis, and B. Wang. "Fusing Location Data for Depression Prediction," IEEE Transactions of Big Data, Vol. 7, No. 2, June 2021.

- S. Ware, C. Yue, R. Morillo, J. Lu, C. Shang, J. Kamath, A. Bamis, J. Bi, A. Russell, and B. Wang. "Large-scale Automatic Depression Screening Using Meta-data from WiFi Infrastructure," Ubicomp, October 2019.

- S. Ware, C. Yue, R. Morillo, J. Lu, C. Shang, J. Bi, J. Kamath, A. Russell, A. Bamis, and B. Wang. "Predicting Depressive Symptoms Using Smartphone Data," Vol. 15, Smart Health Journal, March 2020.

- C. Yue, S. Ware, R. Morillo, J. Lu, C. Shang, J. Bi, J. Kamath, A. Russell, A. Bamis, and B. Wang. Automatic depression prediction using Internet traffic characteristics on smartphones. Smart Health Journal. Vol 18. November 2020.