GeoLifecycle: User Engagement in Geographical Change and Churn Prediction in LBSNs

Presenter: **Young D. Kwon**
ydkwon@cse.ust.hk

2019. 09

HKUST
Department of Computer Science and Engineering
System and Media Lab
Young D. Kwon, Dimitris Chatzopoulos, Ehsan Ul Haq,
Raymond Chi-Wing Wong, and Pan Hui

香港科技大學
THE HONG KONG
UNIVERSITY OF SCIENCE
AND TECHNOLOGY

Why User Engagement and Churn Prediction?

❖ Proliferation of Location-Based Social Networks (LBSNs)

- Heavily rely on User Generated Content (e.g., reviews)
- Users can stop contributing at any time

Main Goal: Maintainability → Analysis: User Engagement → Application: Churn Prediction

Leave or Stay?

✓ 20 important stats and facts, March 2018. URL: https://expandedramblings.com/index.php/by-the-numbers-interesting-foursquare-user-stats/
✓ Yelp Factsheet, August 2018. URL: https://www.yelp.com/factsheet
✓ 2002. Expert Systems with Applications. Turning telecommunications call details to churn prediction: a data mining approach. Wei and Chiu.
✓ 2012. WWW. Churn Prediction in New Users of Yahoo! Answers. Dror et al.

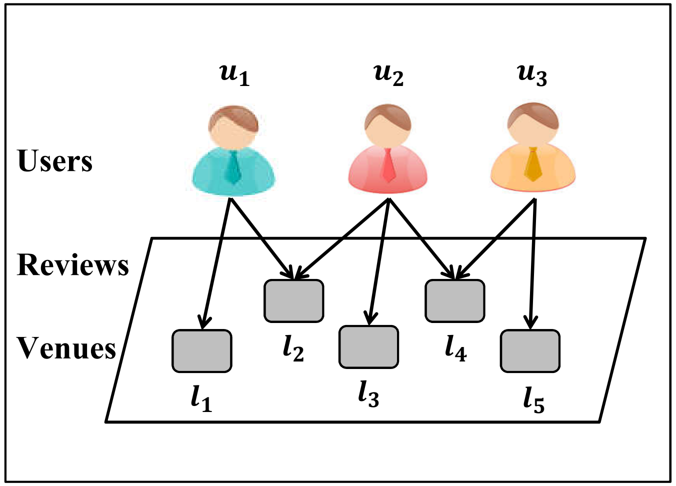
Challenges

❖ Limitations

- **Unclear how users engage with LBSNs**

  - LBSNs can capture online and offline experiences of users

Overview of LBSNs



- **Effects of various aspects (e.g., temporal, social, linguistics) are not fully studied**

  - Novel Offline Feature : Geographic, Venue-specific features
  - New Platform (Not studied yet)

✓ 2018. IMWUT. Revisitation in Urban Space vs. Online: A Comparison across POIs, Websites, and Smartphone Apps. H Cao et al.
✓ 2013. WWW. No Country for Old Members: User Lifecycle and Linguistic Change in Online Communities. Danescu-Niculescu-Mizil et al.
✓ 2015. WWW. All Who Wander: On the Prevalence and Characteristics of Multi-community Engagement. Tan and Lee.

Challenges

❖ Limitations

- Less Attention on churning of highly active producer-type users who contribute a majority of reviews

❖ Focus & Scope

- Focus on highly active producer-type users
- Limit the scope of user engagement to reviewing behaviors

✓ 2018. IMWUT. Revisitation in Urban Space vs. Online: A Comparison across POIs, Websites, and Smartphone Apps. H Cao et al.
✓ 2013. WWW. No Country for Old Members: User Lifecycle and Linguistic Change in Online Communities. Danescu-Niculescu-Mizil et al.
✓ 2015. WWW. All Who Wander: On the Prevalence and Characteristics of Multi-community Engagement. Tan and Lee.

Examined Research Questions

RQ1: How do highly active producer-type users engage in the services of LBSNs in terms of **geographical** exploration?

RQ2: How do **engagement patterns** of highly active producer-type users manifest themselves in various aspects?

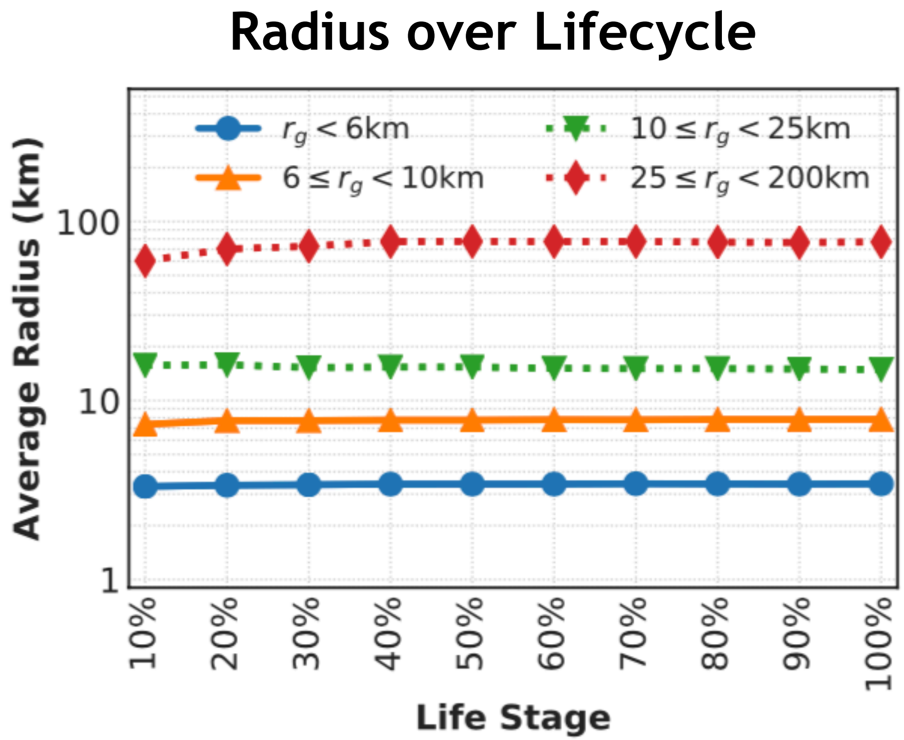
RQ3: To what extent **can we predict** the churning of users with significant contributions within a given period of time?

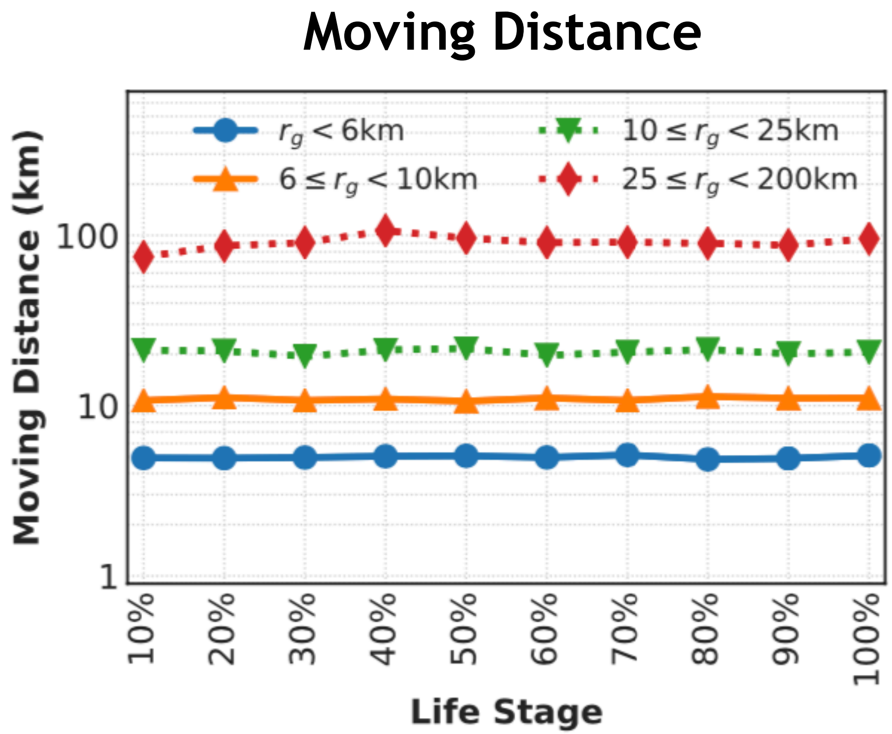
✓ Yelp dataset: https://www.yelp.com/dataset

✓ Foursquare dataset: Y. Chen, et al. 2018. Measurement and Analysis of the Swarm Social Network With Tens of Millions of Nodes. IEEE Access

RQ1

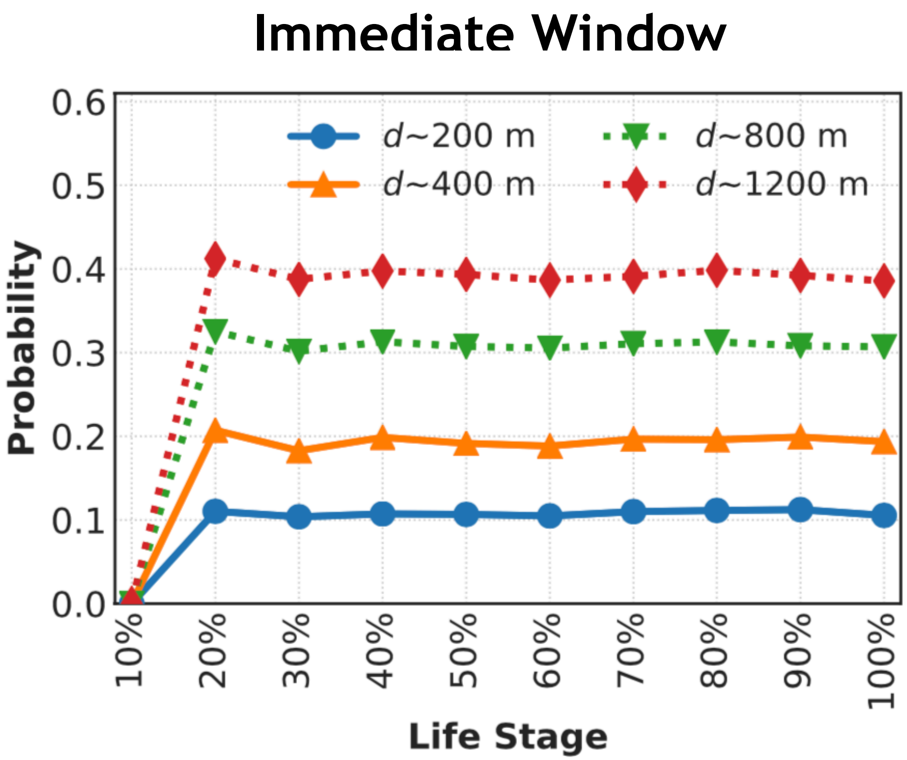
- **Geographical engagement patterns:** How do highly active producer-type users engage in the services of LBSNs in terms of geographical exploration?
  - $r_g(t)$: The **average radius** using a user's trajectory up to t^{th} reviews

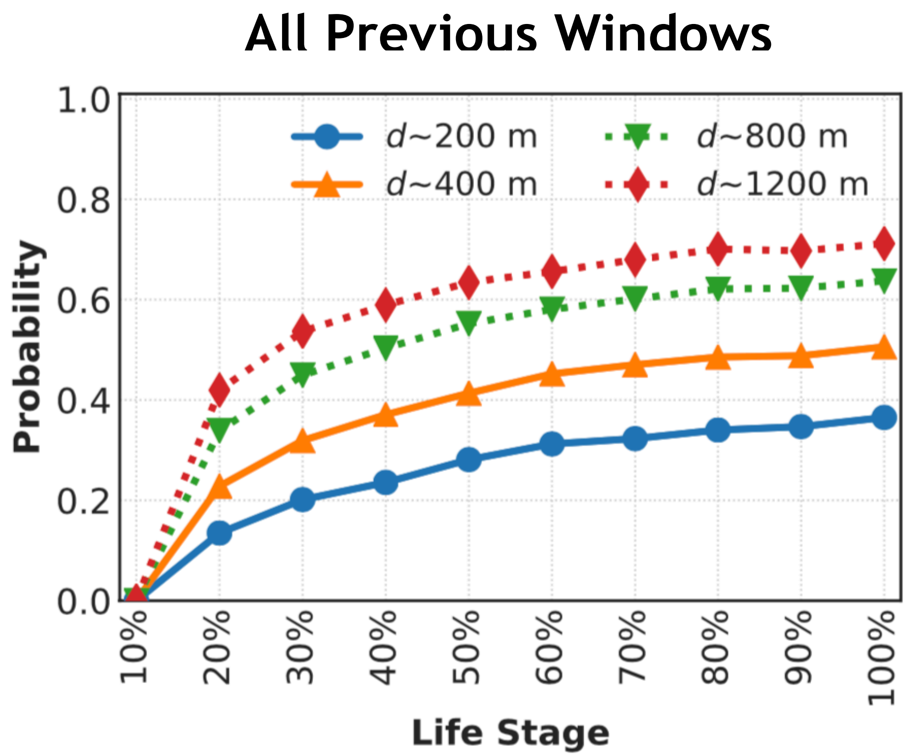




RQ1

- **Human life course**: will users settle down or keep exploring geographically?
  - d : Distance to define neighborhoods

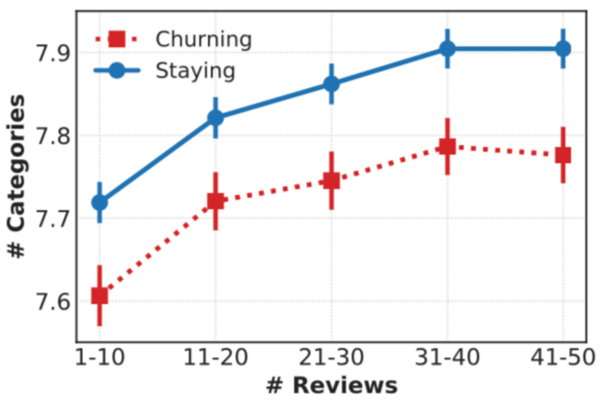




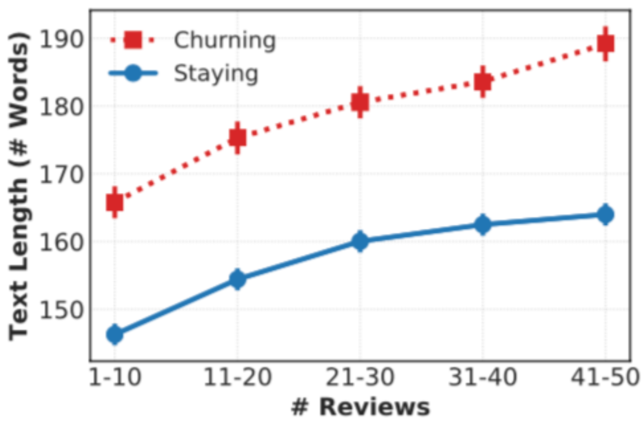
RQ2

- How do **engagement patterns** of highly active producer-type users manifest themselves in various aspects?

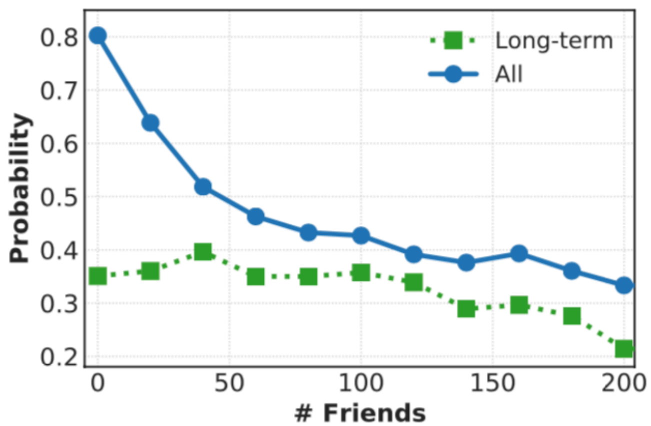
Venue-specific Aspect



Linguistic Aspect



Social Aspect



RQ3

- **Churn Prediction Task:** To what extent **can we predict churning** of users with significant contributions within a given period of time?

Classifiers

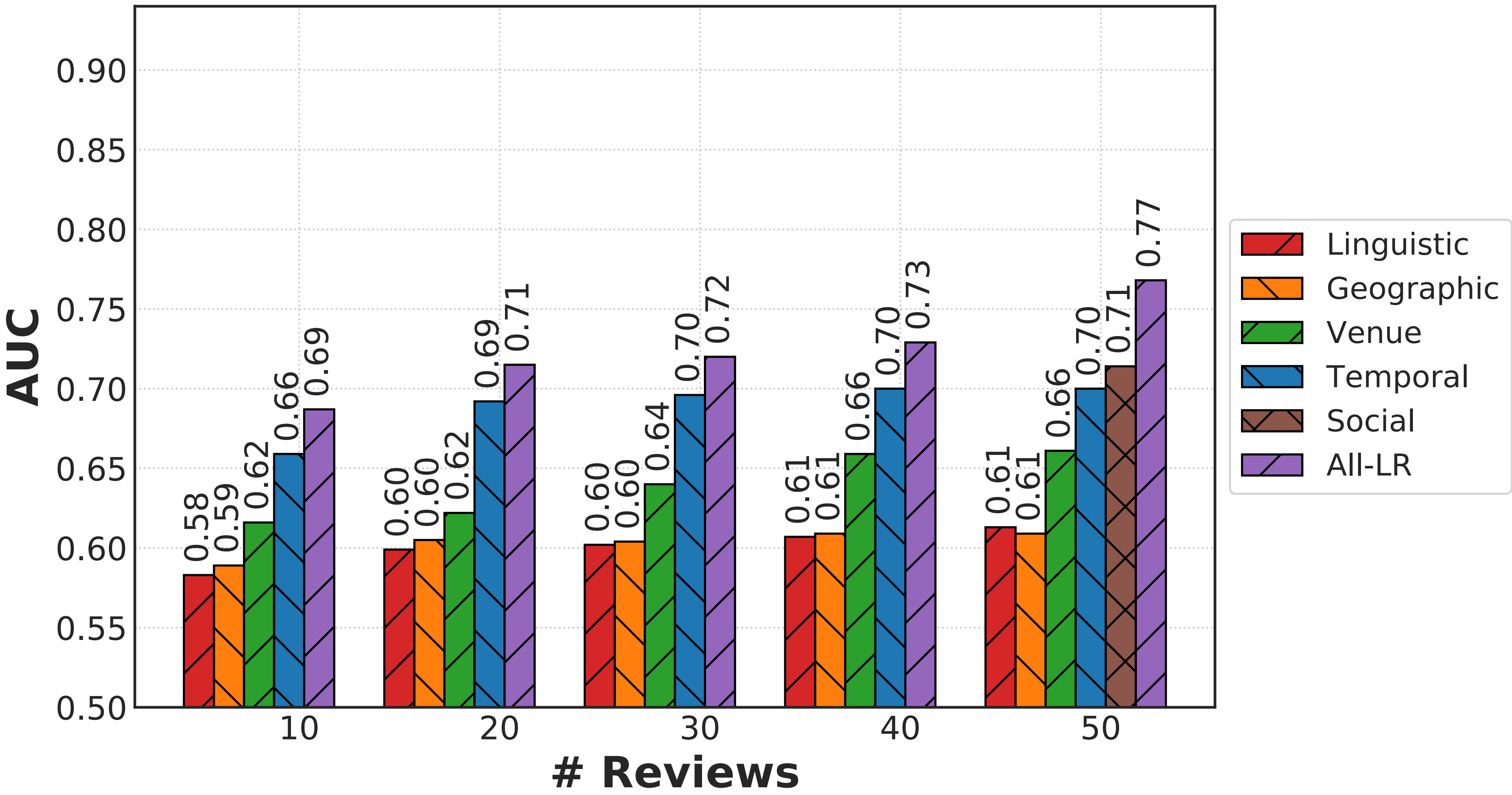
1. Logistic Regression (LR) with L2-Regularization

2. Stacked LSTMs

Models

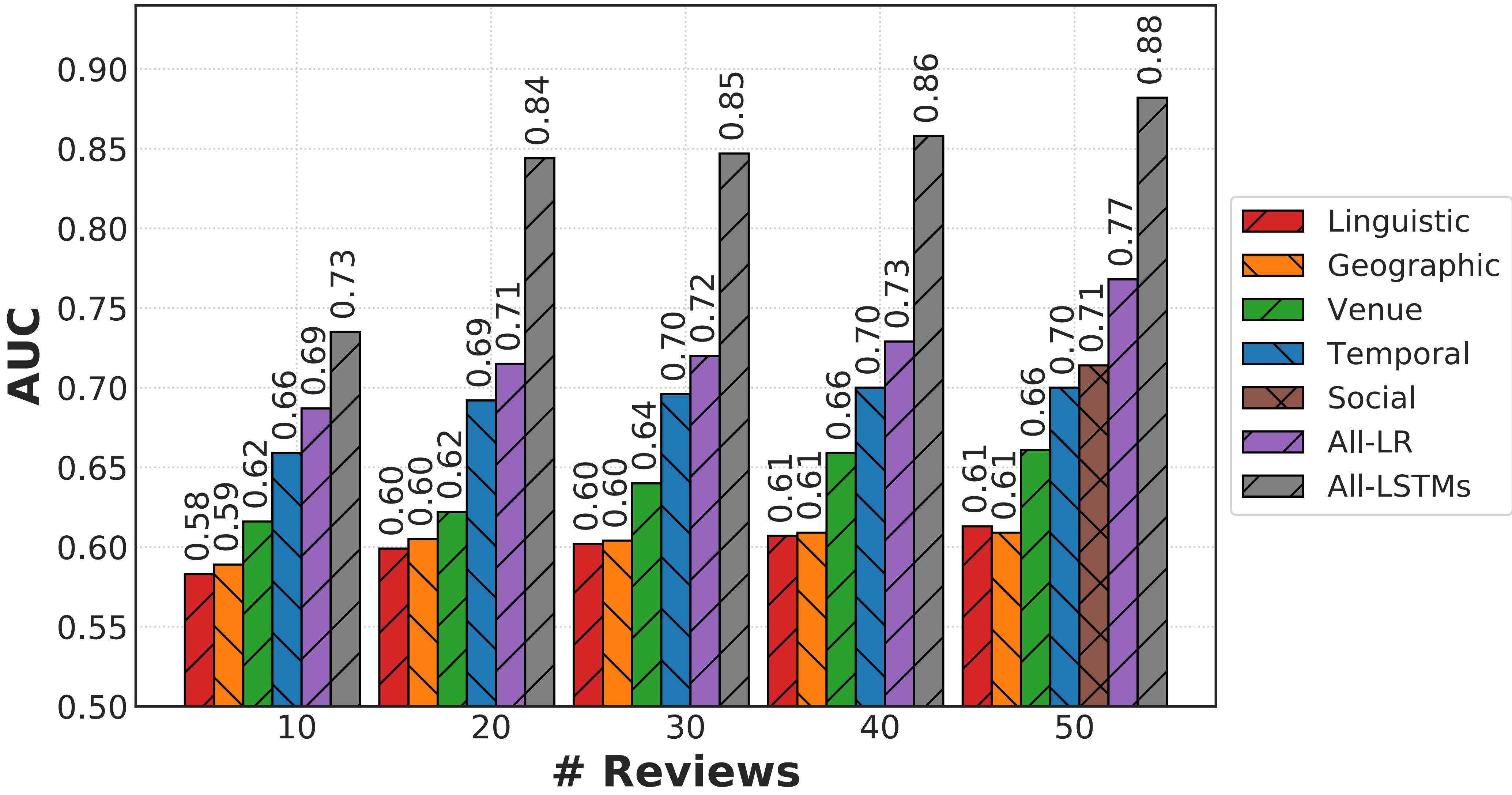
(F1) Temporal feature (Baseline)
(F2) Geographic feature
(F3) Venue property
(F4) Social feature
(F5) Linguistic feature
(F6) Top2 (based on feature importance)
(F7) Top2+Geo2
(F8) All
(F9:F15) Leave-one-out

RQ3



AUC
0.90
0.85
0.80
0.75
0.70
0.65
0.60
0.55
0.50
0.58
0.59
0.62
0.66
0.69
0.60
0.60
0.62
0.69
0.71
0.60
0.60
0.64
0.70
0.72
0.61
0.61
0.66
0.70
0.73
0.61
0.61
0.66
0.70
0.71
0.77
10
20
30
40
50
Reviews
Linguistic
Geographic
Venue
Temporal
Social
All-LR

RQ3



AUC
Reviews
0.50
0.55
0.60
0.65
0.70
0.75
0.80
0.85
0.90
10
20
30
40
50
0.58
0.59
0.62
0.66
0.69
0.73
0.60
0.60
0.62
0.69
0.71
0.84
0.60
0.60
0.64
0.70
0.72
0.85
0.61
0.61
0.66
0.70
0.73
0.86
0.61
0.61
0.66
0.70
0.71
0.77
0.88
Linguistic
Geographic
Venue
Temporal
Social
All-LR
All-LSTMs

Contributions

- Users **constantly wander** around diverse offline places
- The behavioral differences between churners and stayers are significant and are exhibited with their **first 10 reviews**
- LR models based on our findings **significantly improve the performance** over the baseline on the churn prediction task
- We achieve even higher performance in the task by employing a deep learning model

Take-Home Messages

- The average radii and moving distance of users are determined within 5-10 reviews and stable over their lifecycle
  - ➢ **More personalized services based on a user's average radius and moving distance**

- Users constantly write reviews to diverse locations
  - ➢ **Recommend to a user different venues located in geographically different neighborhoods that the user have not reviewed yet**

- We can accurately predict churning users
  - ➢ **Gamification techniques such as badges and rewards could be used to increase engagement levels of users**

Thank you!

Any questions?

You can find me at:

ydkwon@cse.ust.hk

http://www.youngkwon.org/

Young D. Kwon, Dimitris Chatzopoulos, Ehsan Ul Haq,
Raymond Chi-Wing Wong, and Pan Hui
System and Media Lab, Dept. of CSE, HKUST

香港科技大學
THE HONG KONG
UNIVERSITY OF SCIENCE
AND TECHNOLOGY