Continuous Experience-aware Language Model

Subhabrata Mukherjee[†], Stephan Günnemann[‡], Gerhard Weikum[†]

[†]Max Planck Institute for Informatik, [‡]Technical University of Munich

22nd ACM SIGKDD Conference on
Knowledge Discovery and Data Mining, San Francisco, 2016

Outline

- Motivation
- Prior Work
- Current Work
  - Modeling continuous experience evolution
  - Modeling continuous language model evolution
  - Inference
- Experiments
- Conclusions

Motivation

- Online communities are dynamic
  - Users join and leave
  - Adopt vocabulary, adapt to evolving trends
  - Mature over time
- How to capture evolving user maturity?

Example 1

- Consider following camera reviews by John:

  - *My first DSLR. Excellent camera, takes great pictures with high definition, without a doubt it makes honor to its name.* [Aug, 1997]

  - *The EF 75-300 mm lens is only good to be used outside. The 2.2X HD lens can only be used for specific items; filters are useless if ISO, AP,... . The short 18-55mm lens is cheap and should have a hood to keep light off lens.* [Oct, 2012]

Example 2

- Beer: Moosehead Lager

  - The smell of grains a malts on the nose with the slight hop aroma in there. The taste of the beer is crisp ....

  - The beer tastes absolutely terrible ...

Example

- Consider following camera reviews by John:

My first DSLR. Excellent camera, take great

RQ1: How to quantify this change in user maturity (referred to as experience in our work) ?

RQ2: How to model this evolution or progression in maturity?

r

specific items; filters are useless if ISO, AP,... . The short 18-55mm lens is cheap and should have a hood to keep light off lens.

[Oct, 2012]

Use-cases

- Recommend item to a user based on her maturity to consume it
  - Maturity / Experience evolves over time
- Identify experienced users
  - E.g.: Medical professionals in Health communities
- Crowd-sourcing applications / aggregation / community Q&A

Outline

- Motivation
- Prior Work
- Current Work
  - Modeling continuous experience evolution
  - Modeling continuous language model evolution
  - Inference
- Experiments
- Conclusions

Prior Work: Experience Evolution

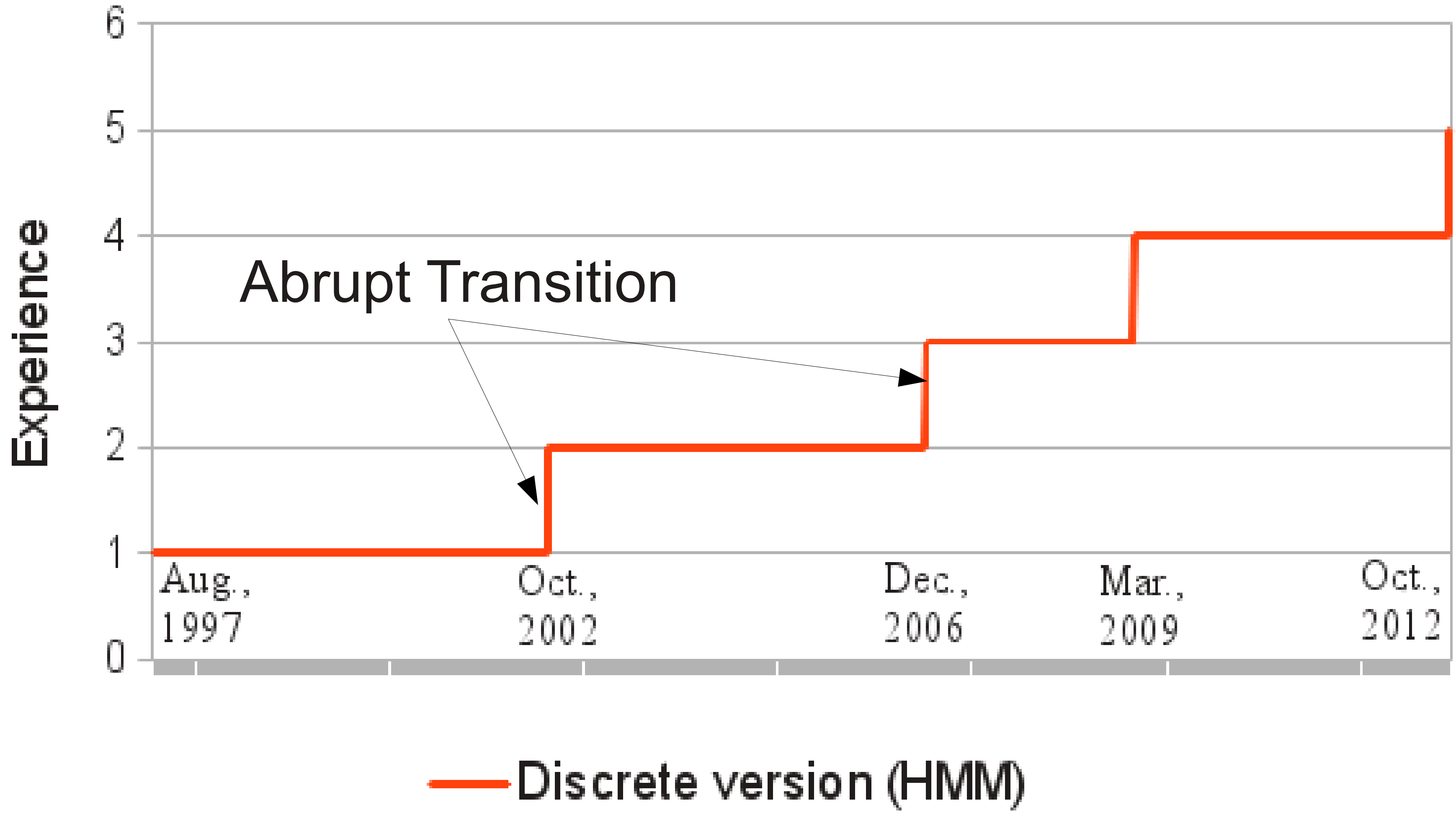
(1) Users at similar levels of experience have similar rating behavior and facet preferences

[J. McAuley, J. Leskovec, WWW '13]

(2) Additionally, users at similar levels of experience have similar writing style

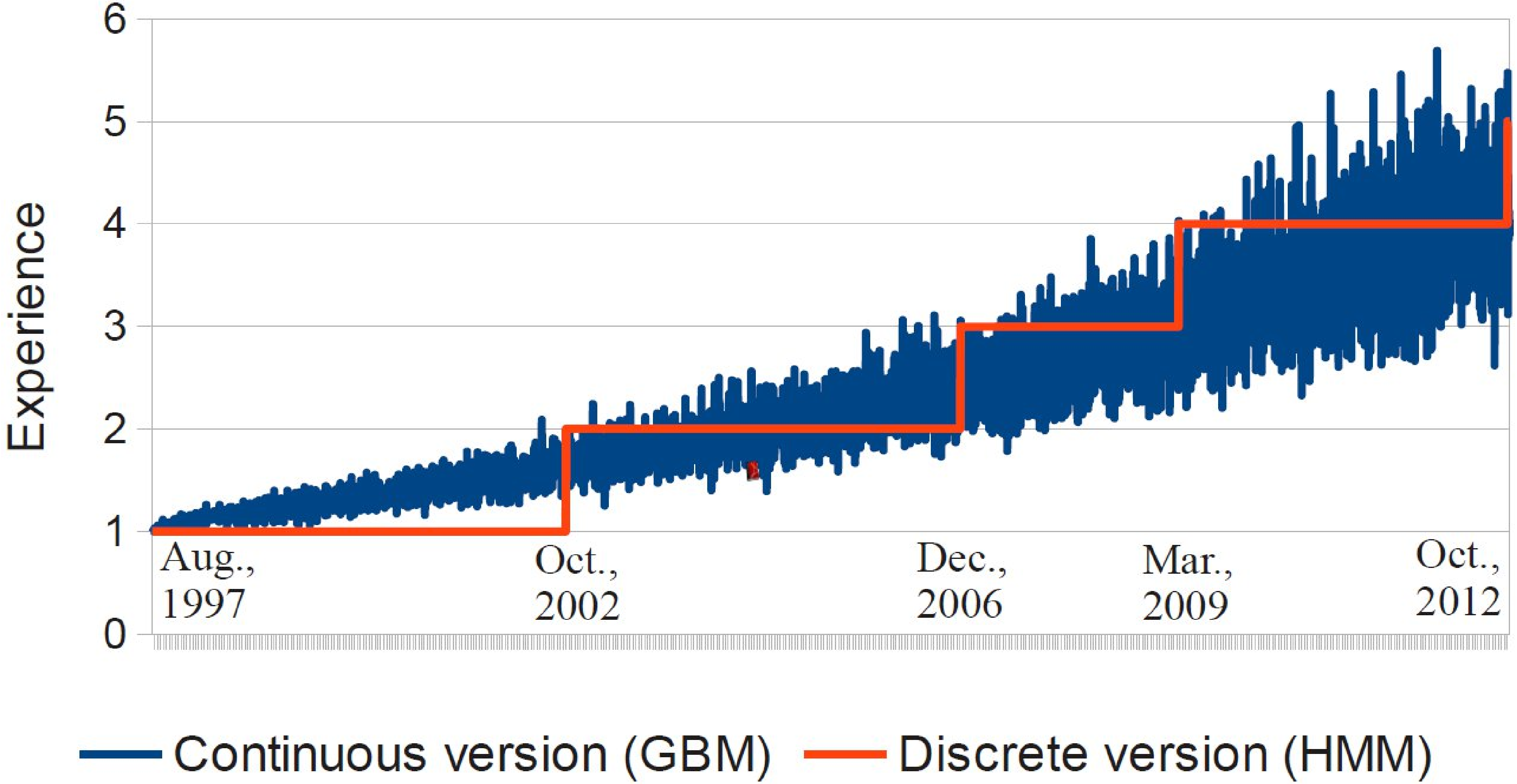
[S. Mukherjee, H. Lamba, G, Weikum, ICDM '15]

Prior Works: Discrete Experience

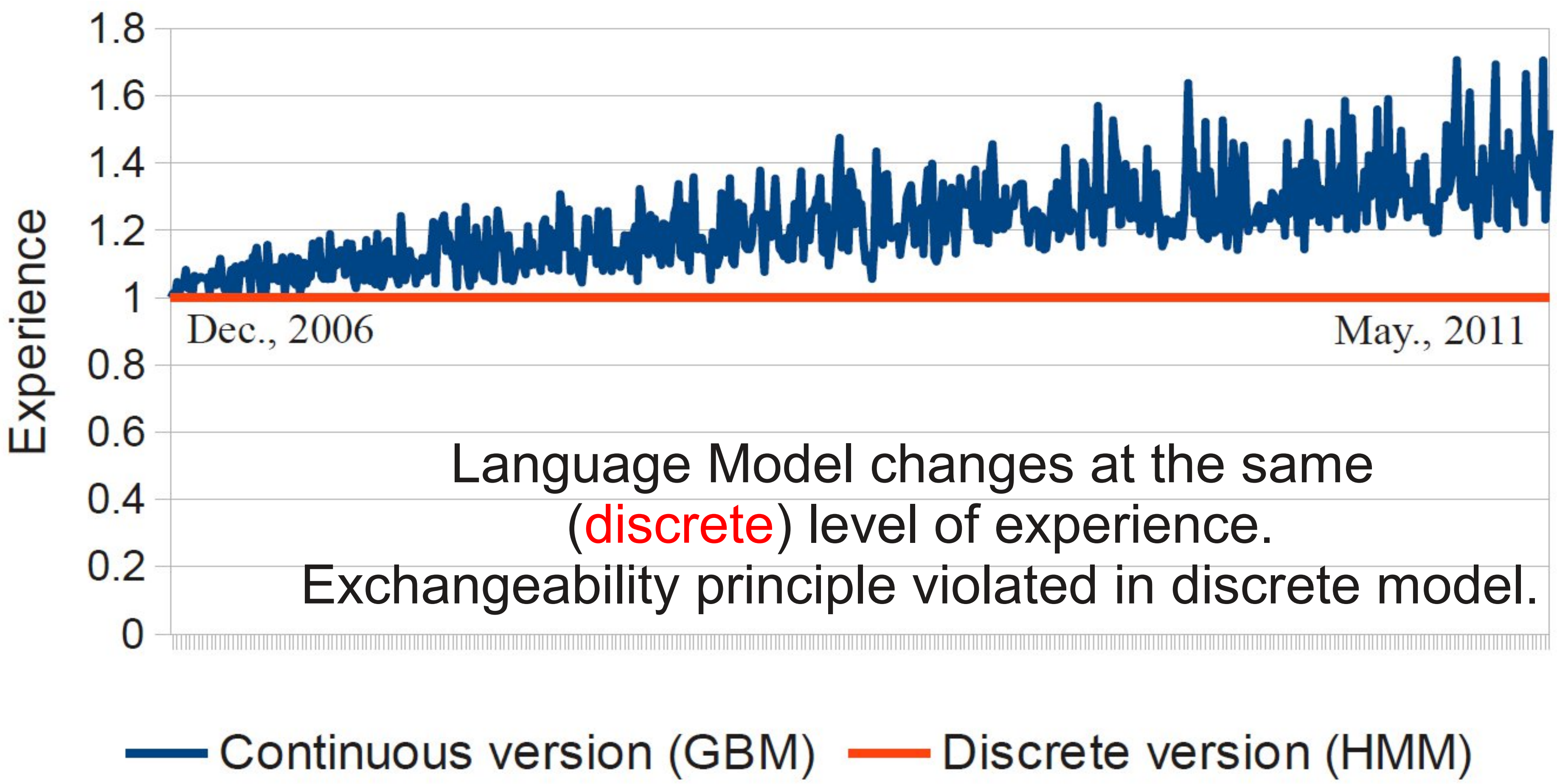


Assumption: At each timepoint (of writing a review) a user remains at the same level of experience or moves to the next level

Current Work: Continuous Experience



Effect of Discrete Evolution on Language Model



Outline

- Motivation
- Prior Work
- Current Work
  - Modeling continuous experience evolution
  - Modeling continuous language model evolution
  - Inference
- Experiments
- Conclusions

Continuous Experience Evolution: Assumptions

- Continuous-time process, always positive
- Markovian assumption: Experience at current time depends only on the latest observed experience
- Drift: Overall trend to increase over time
- Volatility: Progression may not be smooth with occasional volatility. E.g.: series of expert reviews followed by a sloppy one

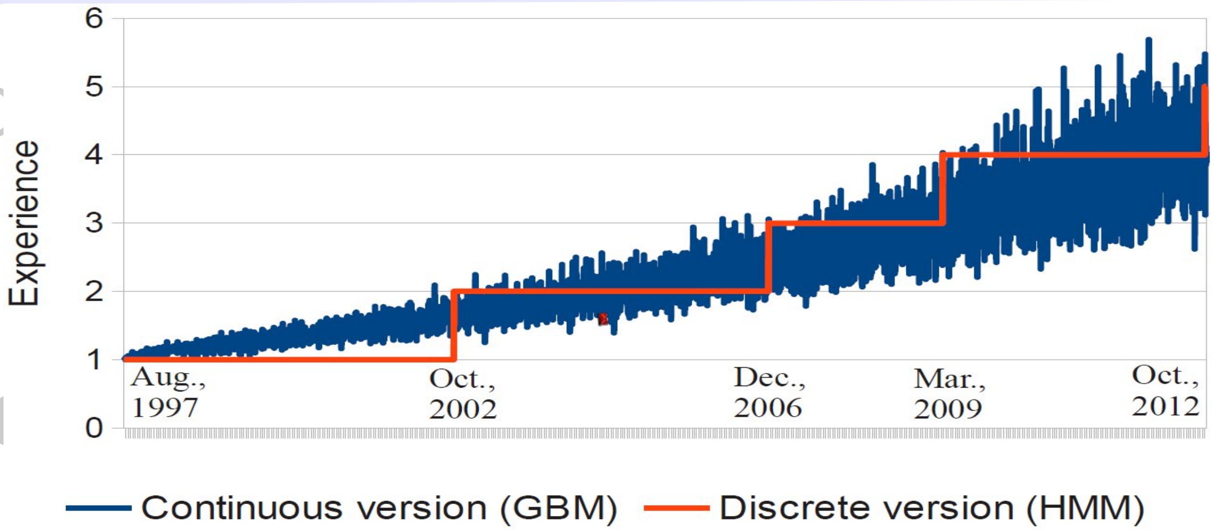
Continuous Experience Evolution: Assumptions

- Continuous-time process, always positive

We show these properties to be satisfied by the continuous-time stochastic process:

Geometric Brownian Motion

- M
  d
- Drift: C
- Volatili occasional
  volatilit d by a
  sloppy



Geometric Brownian Motion

- Stochastic process to model population growth, financial processes like stock price behavior with random noise
- Natural continuous-state alternative to discrete-state space Hidden Markov Model
- Continuous time stochastic process, where $\log(E_t)$ follows Brownian Motion with volatility and drift

Geometric Brownian Motion

- Stochastic Differential Equation: $dE_t = \underbrace{\mu E_t dt}_{\textbf{deterministic trend}} + \underbrace{\sigma E_t dW_t}_{\textbf{unpredictable volatility}}$

  $E_t \rightarrow$ Experience at time 't'

- Analytic solution: $E_t = E_0 \exp((\mu - \sigma^2 / 2) t) + \sigma W_t)$

$t \rightarrow$ time

Starting experience

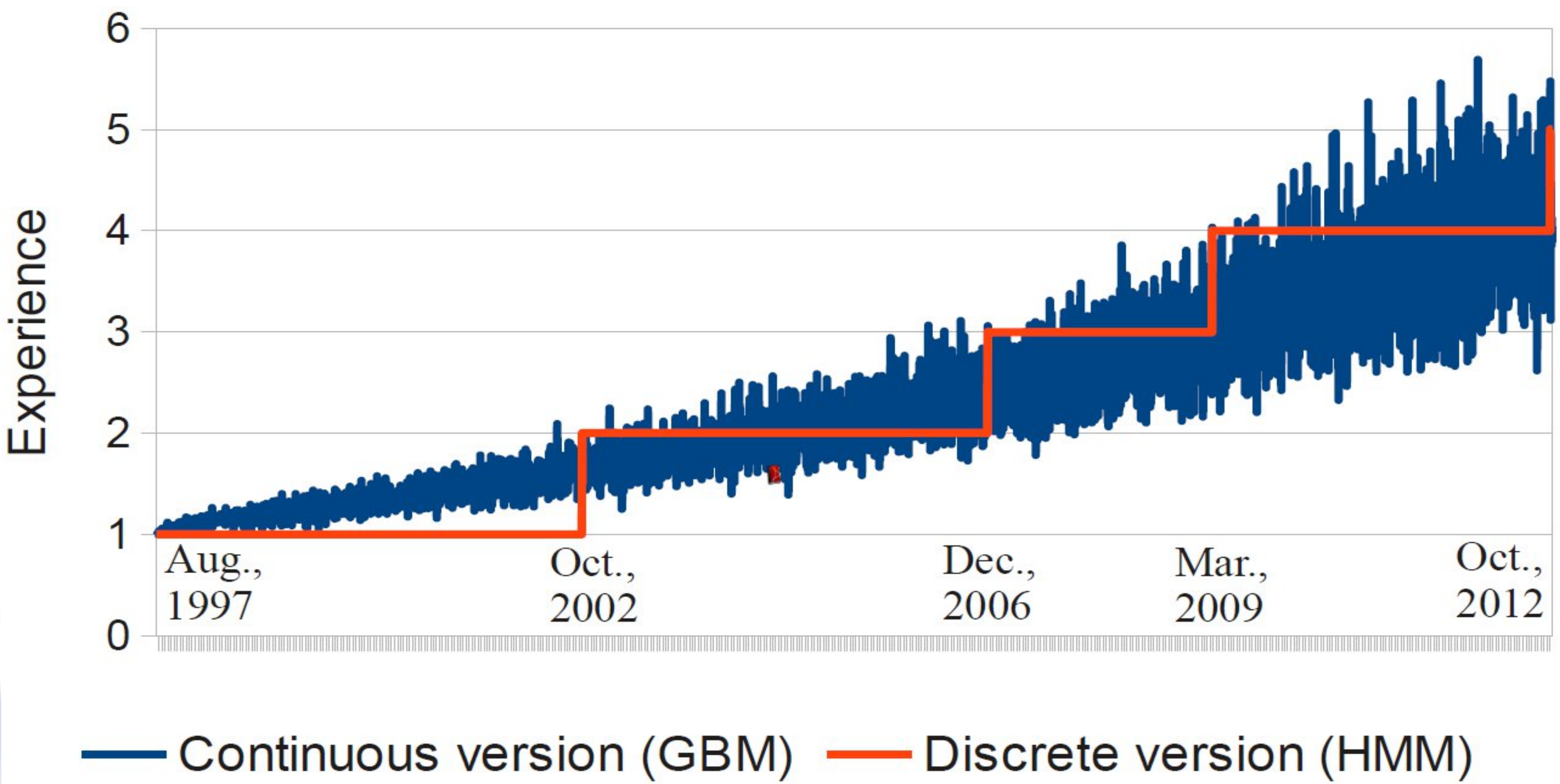
trend

volatility

Wiener Process / Standard Brownian Motion

Geometric Brownian Motion

- Stochastic Differential Equation: $dE_t = \mu E_t dt + \sigma E_t dW_t$



Outline

- Motivation
- Prior Work
- Current Work
  - Modeling continuous experience evolution
  - Modeling continuous language model evolution
  - Inference
- Experiments
- Conclusions

Continuous Language Model (LM) Evolution: Assumptions

- Users' LM also evolves with experience evolution
- Smoothly evolve over time preserving Markov property of experience evolution
- Variance changes with experience change between timepoints
  - If user's experience does not change between successive timepoints, LM remains almost same

Continuous Language Model

$\beta_{t,z,w}$ = Probability of observing word 'w' for facet 'z' at time 't'

$$\beta_{t,z,w} \sim \quad (\beta_{t-1,z,w} , \sigma \cdot |e_t - e_{t-1}|)$$

LM at time 't'   LM at time 't-1'   Experience change

The smell of grains a malts on the nose with the slight hop aroma....

The taste of the beer is crisp

Continuous Language Model

Following principles of standard dynamic systems with Gaussian noise

$$\beta_{t,z,w} \sim \text{Normal}\left(\beta_{t-1,z,w} \,,\, \sigma \cdot |e_t - e_{t-1}| \right)$$

LM at time 't'

LM at time 't-1'

Experience change

The smell of grains a malts on the nose with the slight hop aroma....
The taste of the beer is crisp

Continuous Language Model (LM): Challenges

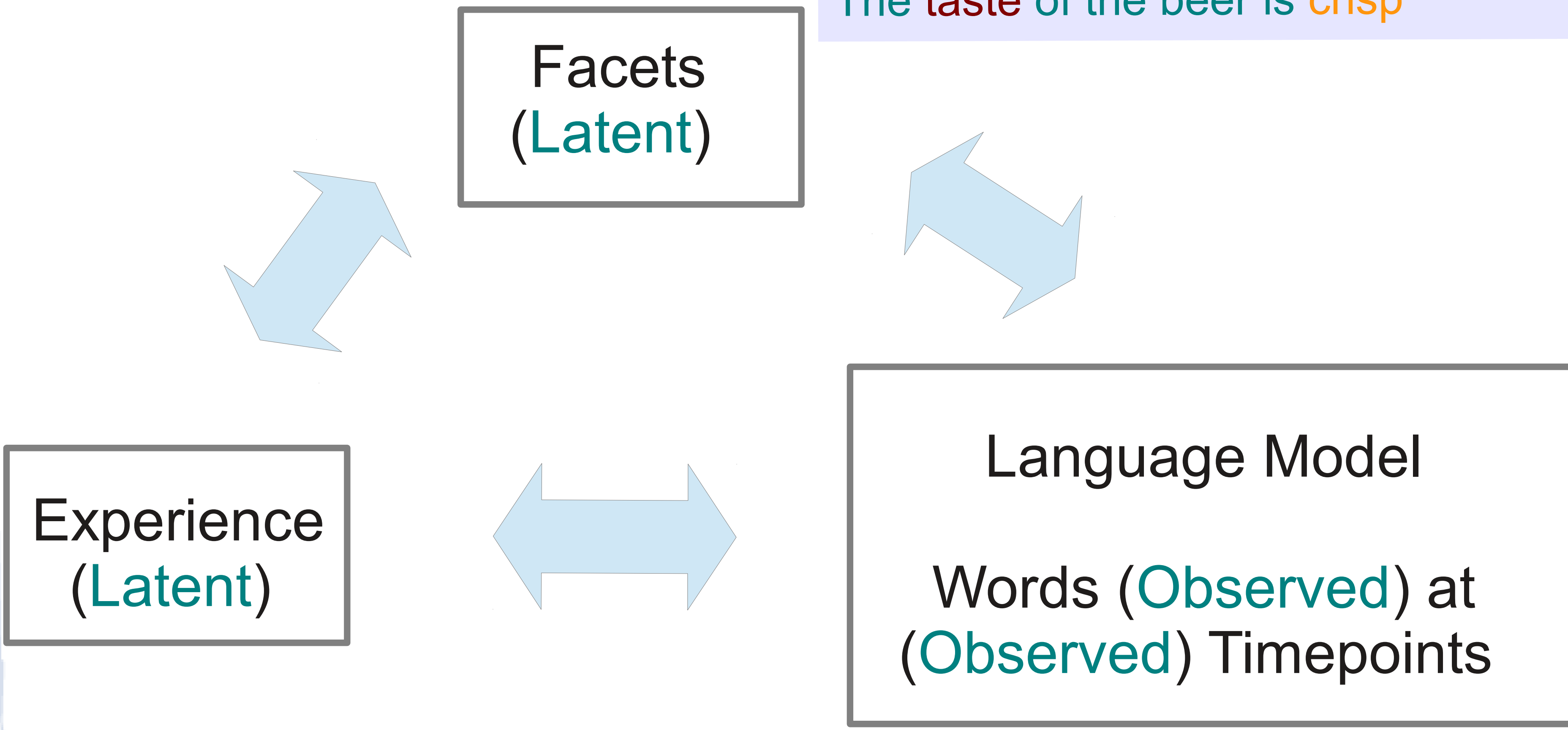
- Experience and LM are continuous distributions, but words in documents have to be generated from discrete distribution
- Temporal granularity: LM does not evolve at the same resolution as experience does

Outline

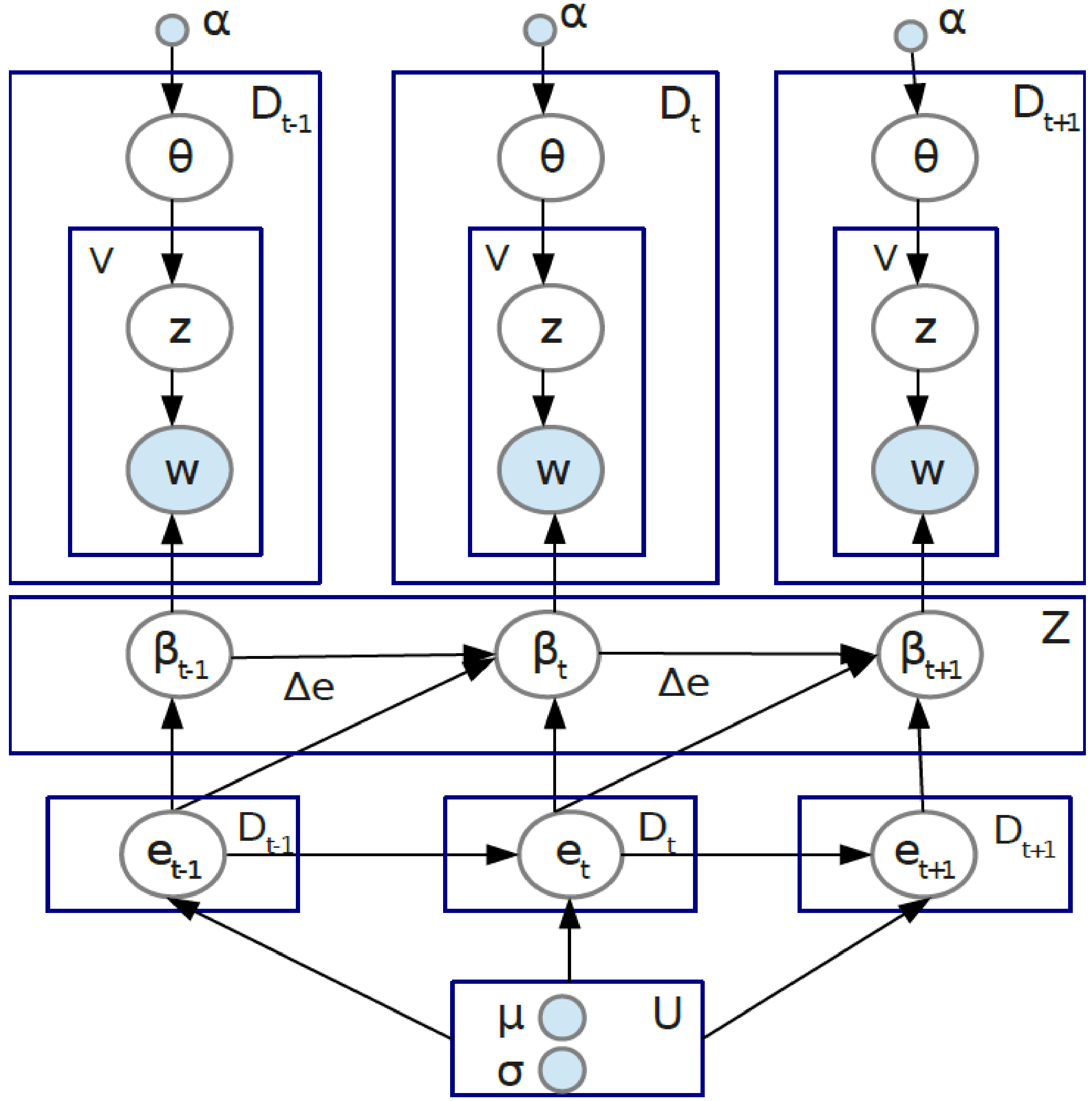
- Motivation
- Prior Work
- Current Work
  - Modeling continuous experience evolution
  - Modeling continuous language model evolution
  - Inference
- Experiments
- Conclusions

Inference

E.g.: The smell of grains a malts
on the nose with the slight hop aroma....
The taste of the beer is crisp

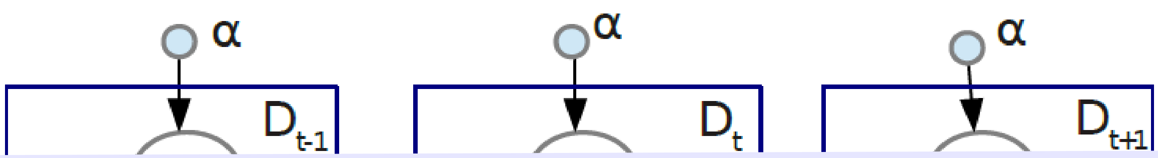


Inference



α
α
α
D_{t-1}
D_t
D_{t+1}
θ
θ
θ
V
V
V
z
z
z
w
w
w
Z
β_{t-1}
β_t
β_{t+1}
Δe
Δe
D_{t-1}
D_t
D_{t+1}
e_{t-1}
e_t
e_{t+1}
μ
σ
U

Inference



Topic (or, Facet) Model (Blei et al., JMLR '03)

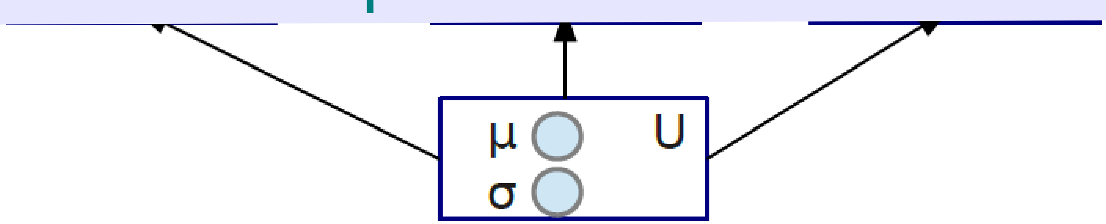
+ Users

    (Author-topic model, Rosen-Zvi et al., UAI '04)

+ Continuous Time

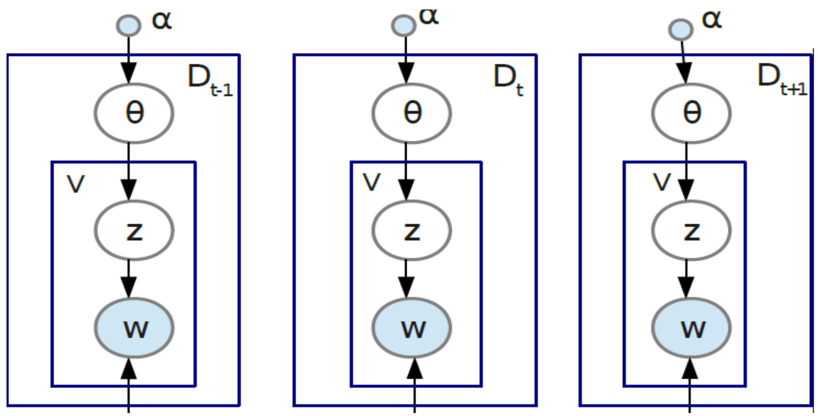
    (Continuous time dynamic topic model, Wang et al., UAI '08)

+ Continuous Experience



Inference: Estimate Facets

Facets (Latent)

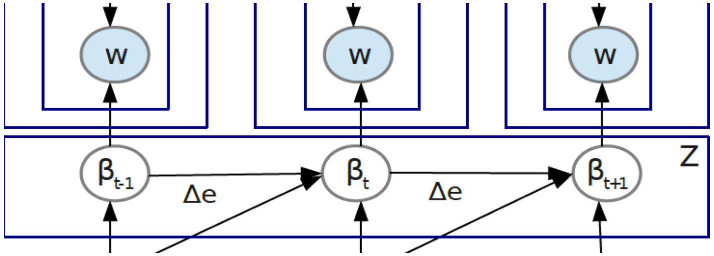


Estimate facets by Gibbs sampling as in standard Latent Dirichlet Allocation keeping LM and experience unchanged

Inference: Estimate LM

Estimate the following state transition model:

$$\beta_t \sim \text{Normal}\,(\beta_{t-1}\,,\ \sigma \cdot \Delta e_t)$$
$$w_n \sim \text{Multinomial}\,(\ \boldsymbol{f}\,(\beta_t)\,)$$



with Kalman Filter and previously *inferred* latent facets

Language Model (Latent)

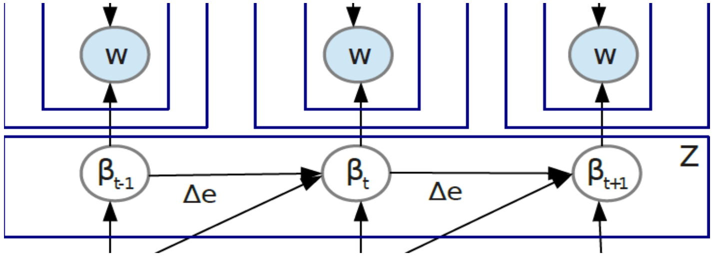
Words (Observed) at (Observed) Timepoints

Inference: Estimate LM

Estimate the following state transition model:

$\beta_t \sim \text{Normal}(\beta_{t-1}, \sigma \cdot \Delta e_t)$

$w_n \sim \text{Multinomial}(\boldsymbol{f}(\beta_t))$



with Kalman Filter and previously *inferred* latent facets

- Models sequential LM evolution
- Continuous-state analog to discrete HMM

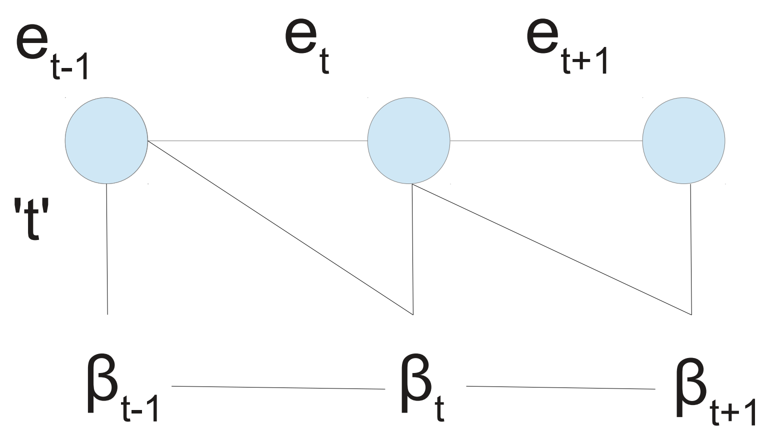
Language Model (Latent)

Words (Observed) at (Observed) Timepoints

Inference: Estimate Experience

e_t → experience at time 't'

β_t → LM at time 't'



$$\beta_t \sim Normal\ (\ \beta_{t-1}\ ,\ \sigma\ |e_t - e_{t-1}|)$$

Experience (Latent)

Change in experience at time 't' affects language models at time 't' and 't+1'

Exploit this to derive proposal distribution for Metropolis Hastings (MCMC sampling)

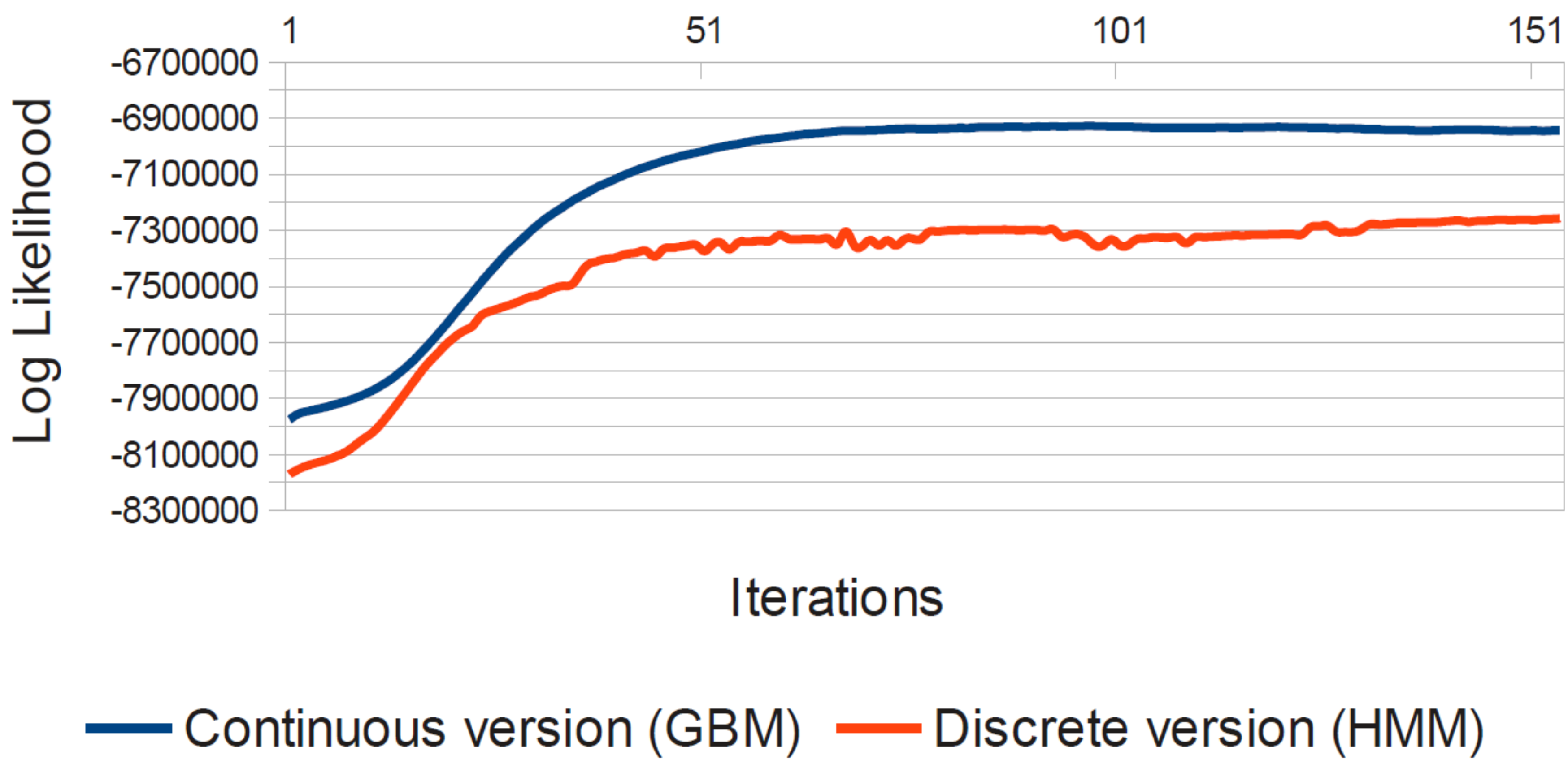
Outline

- Motivation
- Prior Work
- Current Work
  - Modeling continuous experience evolution
  - Modeling continuous language model evolution
  - Inference
- Experiments
- Conclusions

Dataset Statistics

Dataset	#Users	#Items	#Ratings	#Time (Years)
Beer (BeerAdvocate)	33,387	66,051	1,586,259	16
Beer (RateBeer)	40,213	110,419	2,924,127	13
Movies (Amazon)	759,899	267,320	7,911,684	16
Food (Yelp)	45,981	11,537	229,907	11
Media (NewsTrust)	6,180	62,108	89,167	9
TOTAL	885,660	517,435	12,741,144	-

Loglikelihood, Smoothness and Convergence

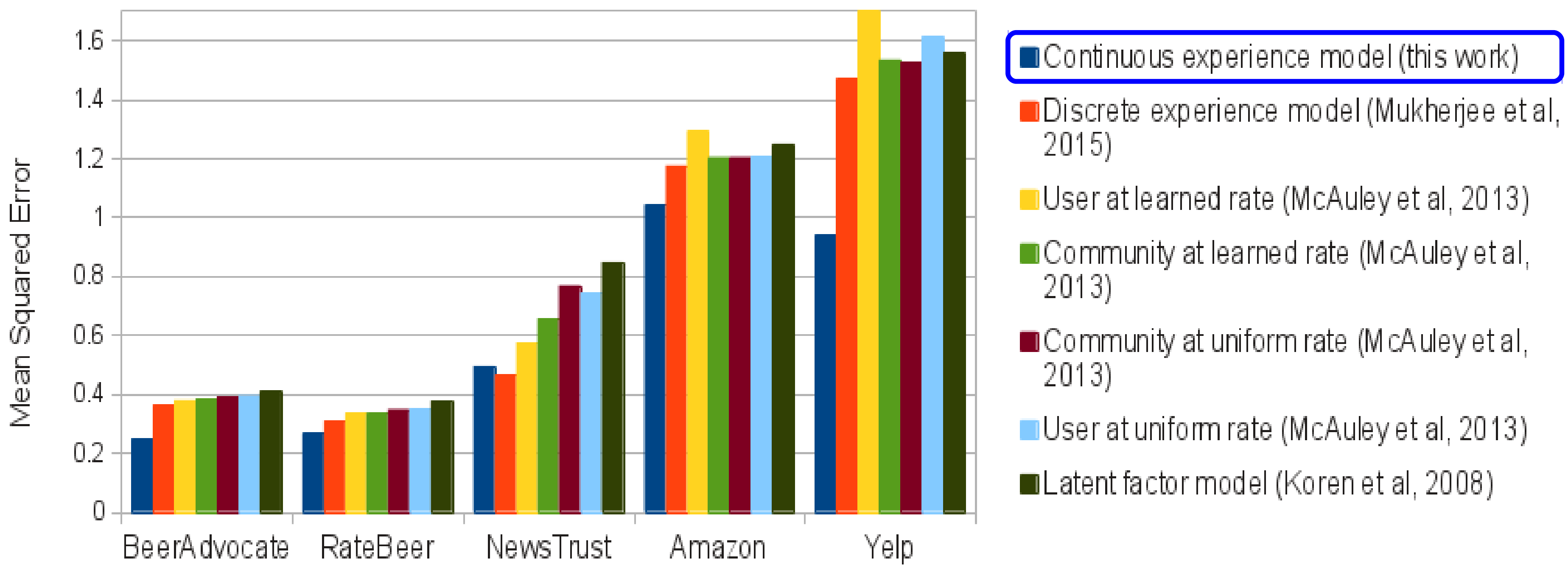


Mean Squared Error: Item Rating Prediction

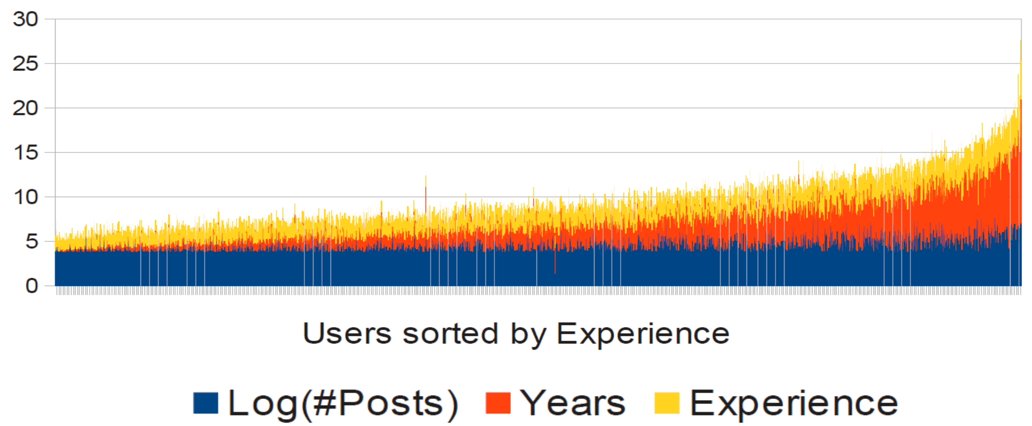
RQ3: Can we recommend items better, if we consider user experience ?

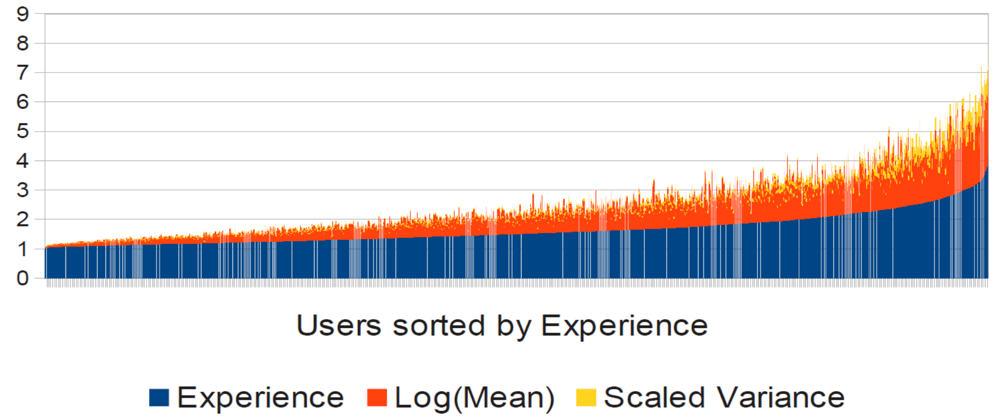
Mean Squared Error: Item Rating Prediction

RQ3: Can we recommend items better if we consider user experience ?



Experience Progression: Insights



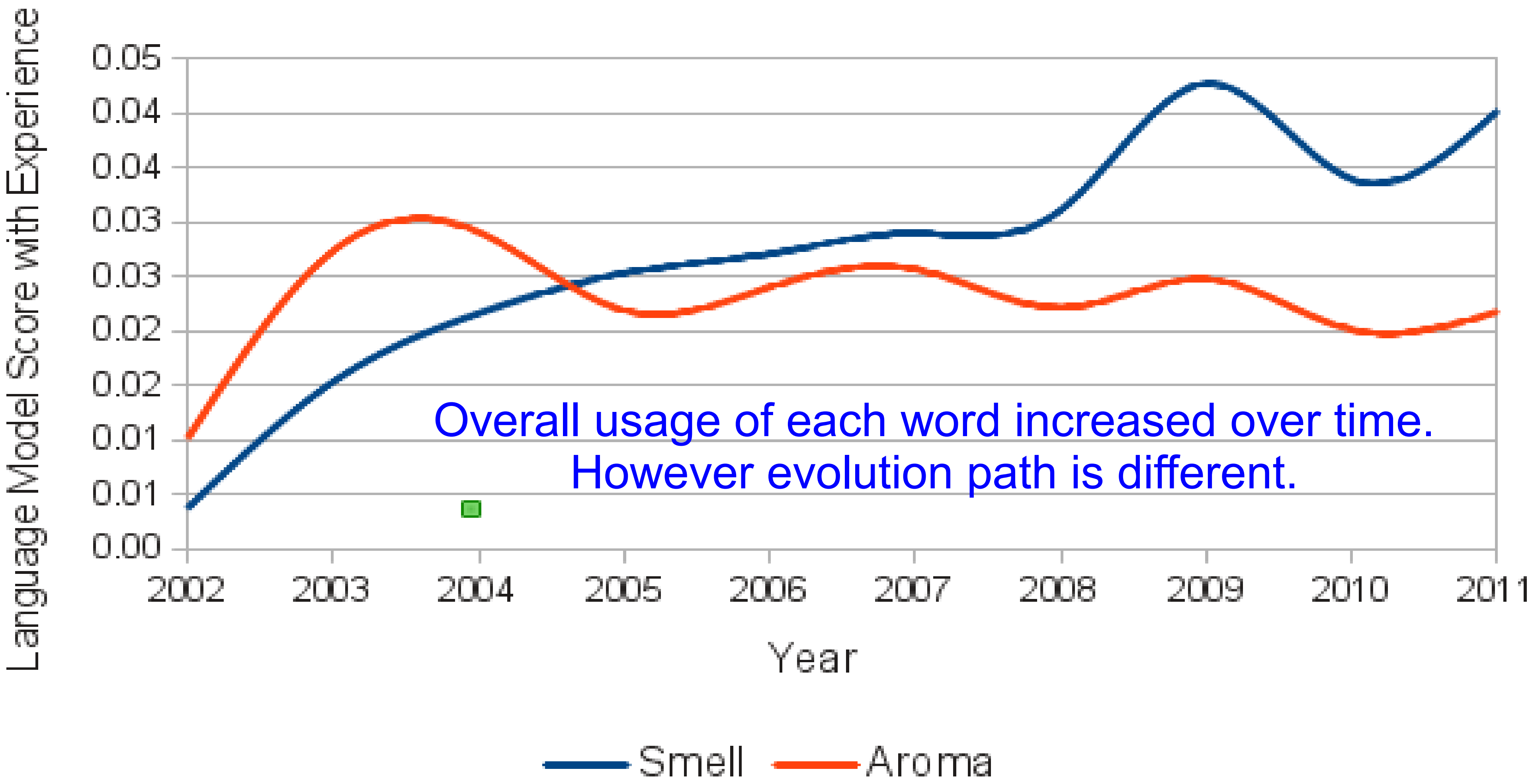


Experience Progression: Insights

- Experienced users mature faster than amateurs, exhibit a higher variance

- Progression depends more on time spent in community than on activity

Trace Evolving Norms with Time and Experience



Interpretability: Top Words by Experienced Users

	Most Experience	Least Experience
BeerAdvocate	chestnut_hued near_viscous cherry_wood sweet_burning faint_vanilla woody_herbal citrus_hops mouthfeel	originally flavor color poured pleasant bad bitter sweet
Amazon	aficionados minimalist underwritten theatrically unbridled seamless retrospect overdramatic	viewer entertainment battle actress tells emotional supporting
Yelp	smoked marinated savory signature contemporary selections delicate texture	mexican chicken salad love better eat atmosphere sandwich
NewsTrust	health actions cuts medicare oil climate spending unemployment	bad god religion iraq responsibility questions clear powerful

Interpretability: Top Words by Experienced Users

	Most Experience	Least Experience
BeerAdvocate	chestnut_hued near_viscous cherry_wood sweet_burning faint_vanilla woody_herbal citrus_hops mouthfeel	originally flavor color poured pleasant bad bitter sweet
Amazon	aficionados minimalist underwritten theatrically unbridled seamless retrospect overdramatic	viewer entertainment battle actress tells emotional supporting
Yelp	smoked marinated savory signature contemporary selections delicate texture	mexican chicken salad love better eat atmosphere sandwich
NewsTrust	health actions cuts medicare oil climate spending unemployment	bad god religion iraq responsibility questions clear powerful

Conclusions

- Users' experience evolve continuously in nature, along with their language usage
- Experienced users have distinctive writing style, maturing over time
- Recommendation models can be improved by explicitly considering user experience
- Finally, we proposed a Brownian Motion based stochastic model to capture the above phenomena