Vrbo™

Travel better together

A Simple Deep Personalized Recommendation System

Pavlos Mitsoulis-Ntompos, Meisam Hejazinia, Serena Zhang, Travis Brady

Vrbo, Part of Expedia Group, Austin, Texas and London, UK

Overview

- Introduction
- Business context
- Vrbo use case within Expedia Group
- Position of this research
- Model
- Conclusion

Travel better together

Largest Travel Platform in the World ...

High Volume and diversity of **Customers**

750M+ monthly visits[1]

2M+ active corporate travelers

Powering **~100K** offline travel agents[2]

35K+ B2B partners leveraging our platform

50M+ contacts handled annually[3]

Broad and diversified **Supply Partners**

1.1M+ properties on core lodging platform[4]

2M+ Vrbo online bookable listings

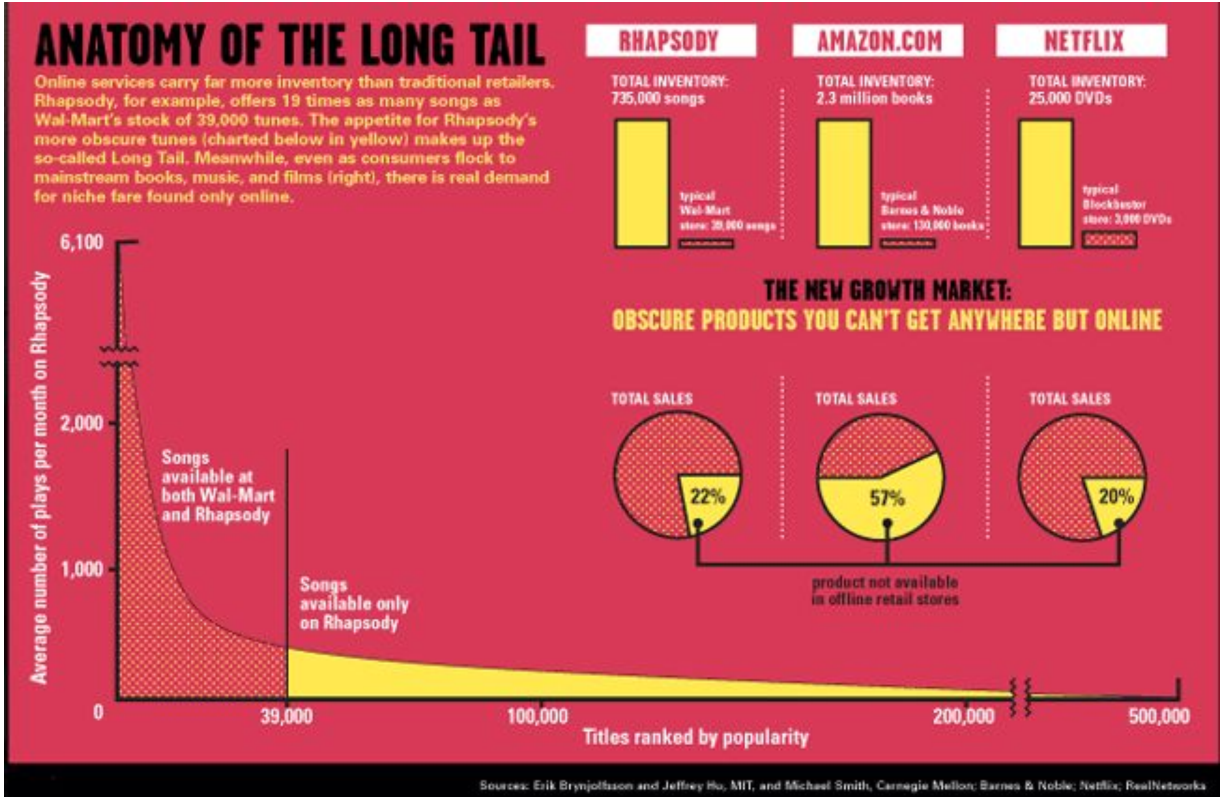
500+ airlines

175+ car rental companies

35K+ unique activities

Notes: Expedia Group data shown as of 3/31/19, unless otherwise noted. 1. Monthly visits based on data for Brand Expedia, Hotels.com, Orbitz, Travelocity, Wotif, Vrbo, trivago and Hotwire combined during 2018. 2. Offline travel agents based on number of sales agents in Global Customer Operations, Expedia Partner Solutions (EPS), Vrbo, Classic Vacations, CruiseShipCenters, Travel Agent Affiliate Program (TAAP). 3. Contacts handled annually include calls, emails, chats and social media. 4. Includes more than 460,000 integrated Vrbo listings. This information is based on Q2 19' Earning Report.

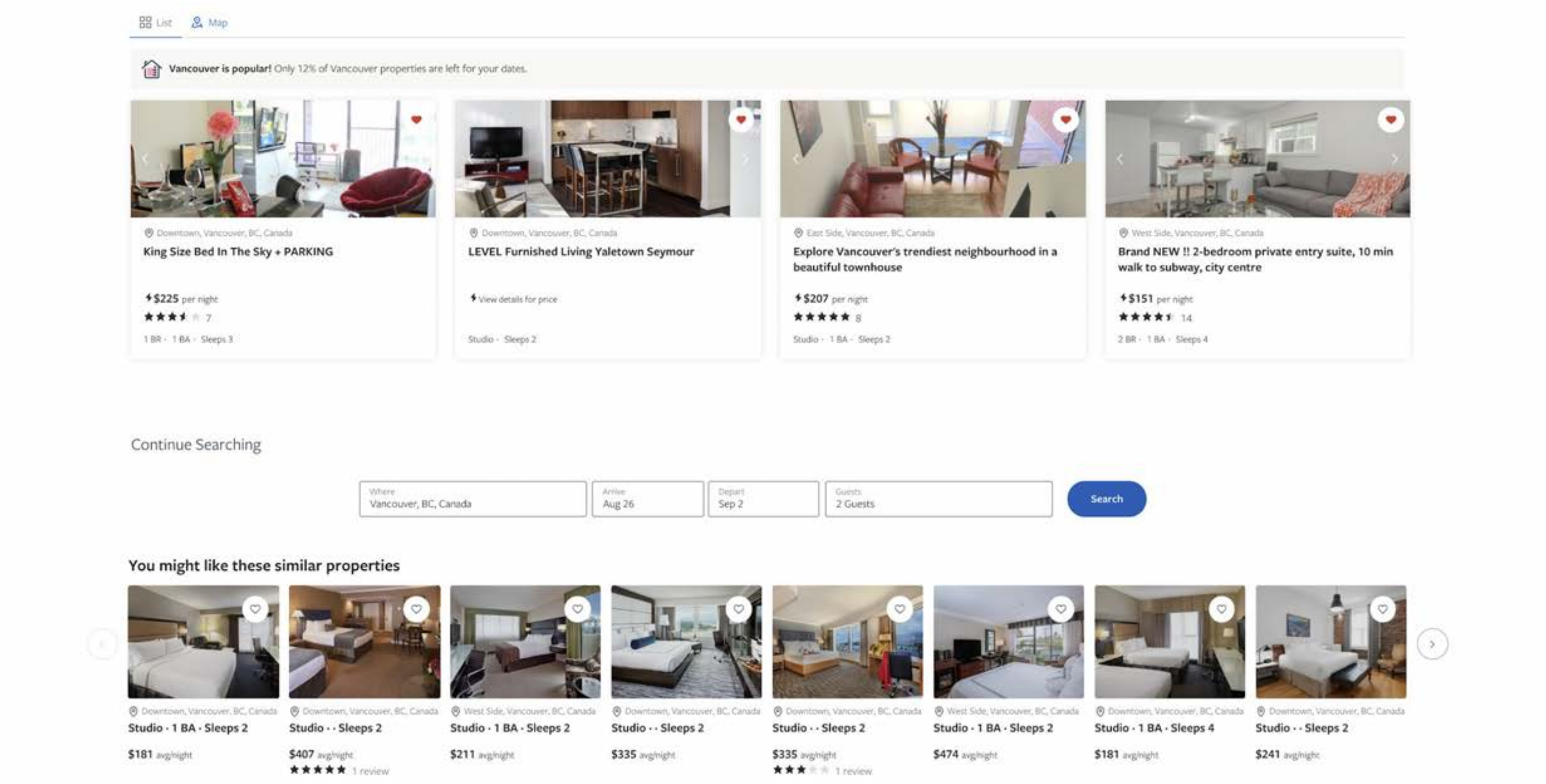
Recommender Systems and the Long Tail



Source: Erik Bryjolfsson, Yu Jeffrey Hu and Michael D. Smith, MIT Sloan,"From niches to riches: Anatomy of the long tail." Sloan Management Review 47.4 (2006): 67-71.

Two-sided Vacation Rental Marketplace Platforms

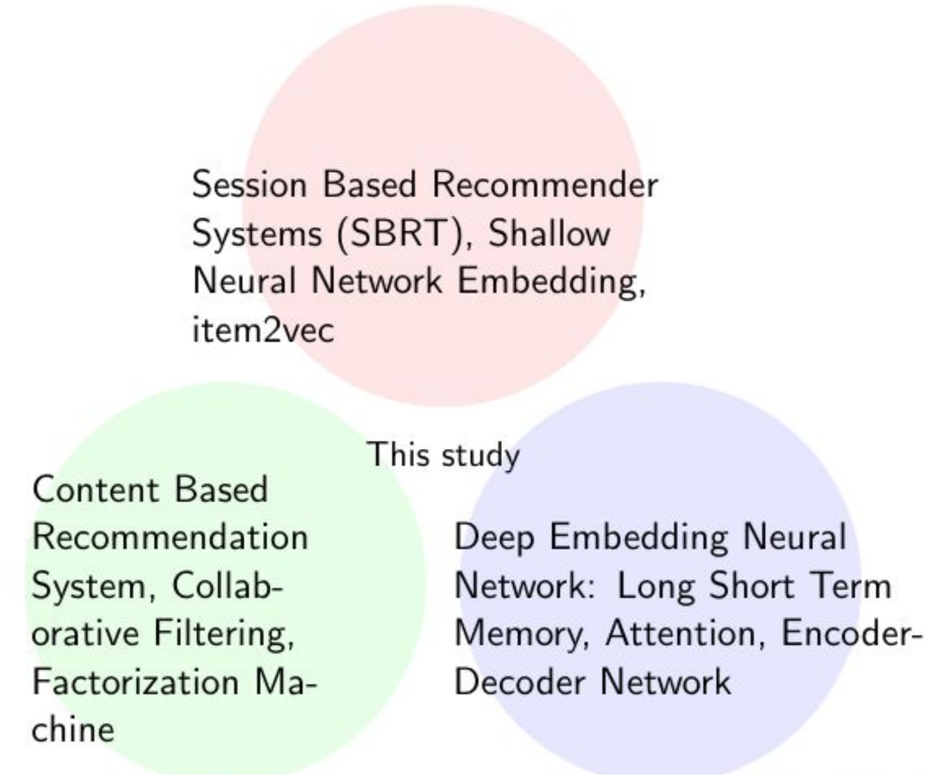
Recommendation use case in Vrbo



Research questions

- How to personalize the item recommendations based on trip context?
- How to make personalized recommendation scalable (millions of listings and travelers)?
- How to incorporate both seasonality and trip type (e.g. with family to mountain or beach in summer) in personalizing recommendations?
- How to capture the non-linear and complex search behavior of the travelers to personalize recommendations
- How to serve personalized recommendations to all the heterogeneous use cases from discovery to final booking on home page, landing page, property detail page, trip board, email, etc. in production environment?

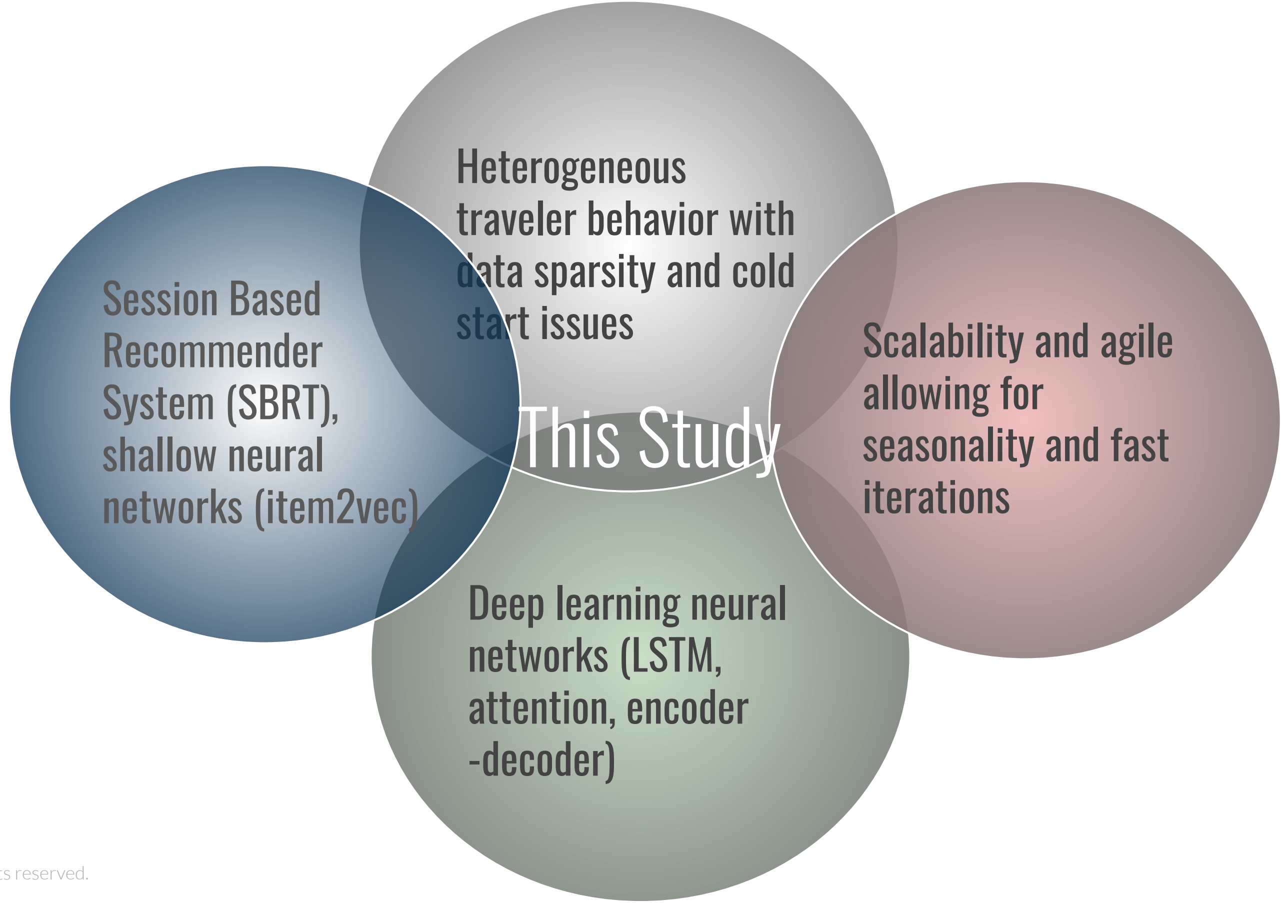
Position of this research in literature



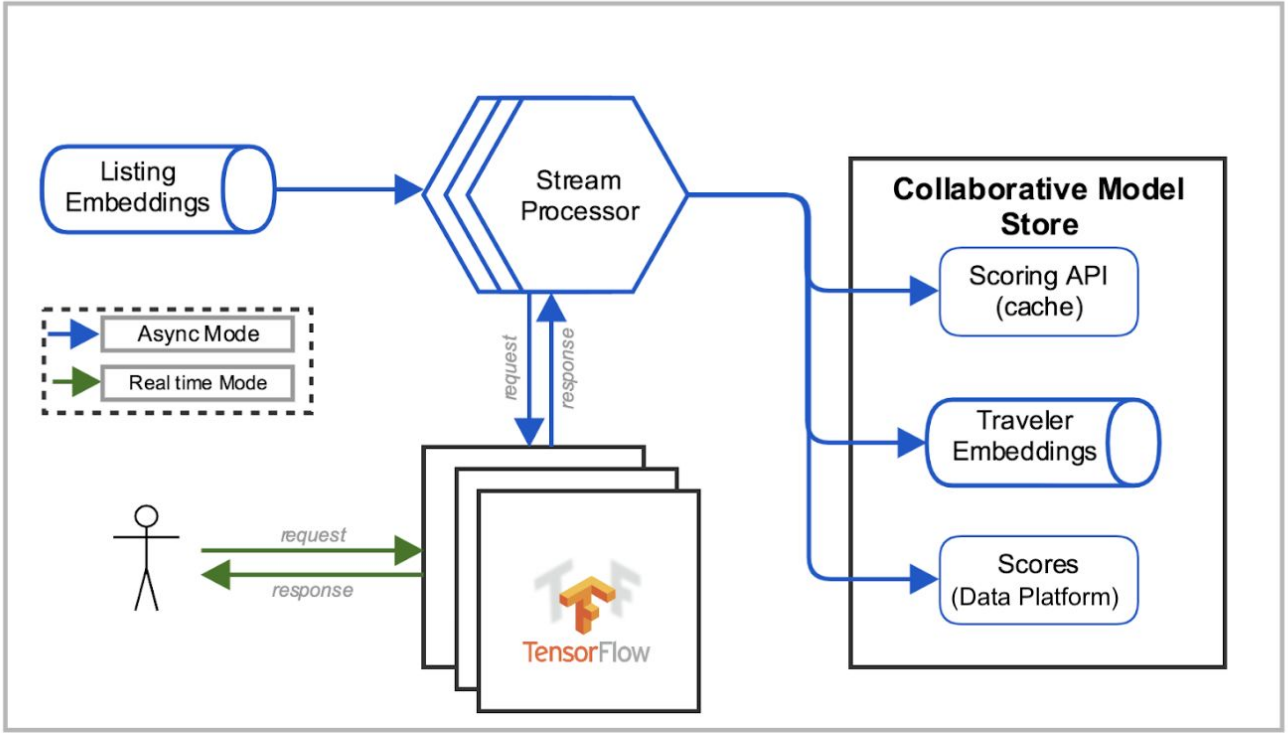
Gap in the literature and this study

- Personalization of recommendation
- Incorporating for trip context
- Importance of Scalability
- Importance of Computational Efficiency to fit in agile framework
- Incorporating non-linearity
- Architecture guided by theory: traveler bounded rationality, traveler's memory and item encoding and decoding, traveler's attention

Position of this research



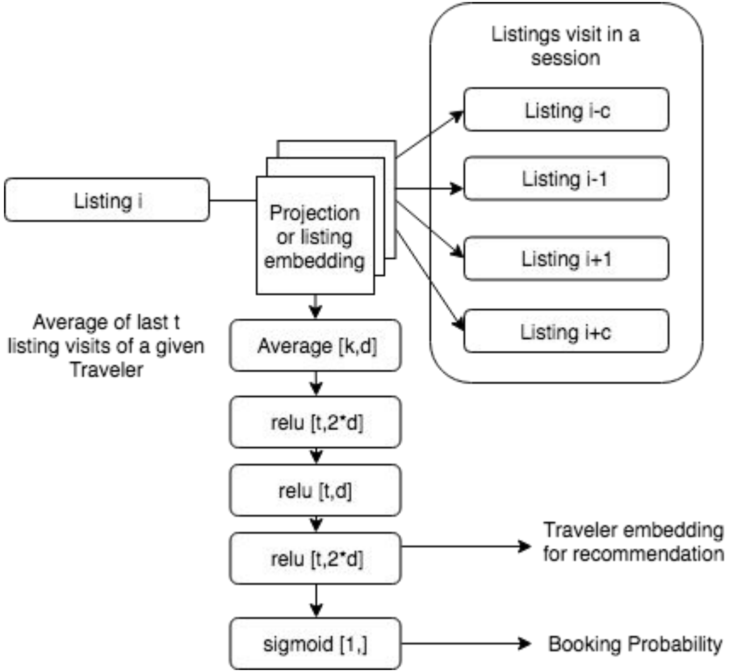
Overview of the System



Overview of the Methodology

- Internal anonymized clickstream data is collected for millions of users from two different seven-day periods:
  - The 1st dataset was used to generate embeddings using Deep Average Network and the LSTM with Attention
  - The 2nd dataset was used to evaluate the learned embeddings on the Traveler Booking Intent Model
- Three baselines methods:
  - Random Listing Embedding Selection
  - Averaging Embeddings
  - LSTM with Attention

Traveler Embedding Model (1)



$$\frac{1}{S}\sum_{s=1}^{S}\sum_{-c\le i\le c, i\neq 0}\log p(l_{i+j}|l_i), \text{ where } p(l_{i+j}|l_i) = \frac{\exp(\nu_{l_{i+j}}^T \nu_{l_i})}{\sum_{l=1}^{L}\exp(\nu_l^T \nu_{l_i})}$$

$$\text{However, by negative sampling } p(l_{i+j}|l_i) = \frac{exp(\nu_{l_{i+j}}^T \nu_{l_i})}{1+\exp(\nu_{l_{i+j}}^T \nu_{l_i})}$$

Traveler Embedding Model (2)

- Target to predict: $\mathrm{P}(Y_j|\mathrm{S}_j, C_j) = \mathrm{sigmoid}(f(\nu_{j.}))$
- Deep Average Network (DAN):

$$f(\nu_{j.}) = \mathrm{relu}(\omega_1 \cdot h_2(\nu_{j.}) + \beta_1)$$
$$h_1(\nu_{j.}) = \mathrm{relu}(\omega_2 \cdot h_1(\nu_{j.}) + \beta_2)$$
$$h_2(\nu_{j.}) = \mathrm{relu}(\omega_3 \cdot \frac{1}{k}\sum_{i=1}^{t} \nu_{ji}) + \beta_3)$$

LSTM Architecture: $f(\nu_j^t) = \mathrm{sigmoid}(\omega_f[h_t, \nu_j^t] + \beta_f) \cdot f(\nu_{j.}^{t-1})$
$+ \mathrm{sigmoid}(\omega_i[h_t, \nu_j^t] + \beta_i) \cdot \tanh(\omega_c[h_{t-1}, \nu_j^t] + \beta_c)$

- LSTM + Attention: $f(\nu_j) = \mathrm{softmax}(\omega^T \cdot h_T)\tanh(h_T)$

Traveler Booking Intent (TBI) Model

- A scalable, end-to-end machine learning system in production that predicts conversion, aka booking intent probability
- XGBoost framework is used for advantages in performance and efficiency

$$L^{(t)} = \sum_{j} l(y_j, \hat{y}_j^{(t-1)} + f_t(\mathbf{x}_j)) + \Omega(f_t)$$

- Handcrafted features using session data that includes user onsite interactions, e.g. search, property listing page, check out flow and contacts (inquiries and messages).

Experiments and Results (1)

Table: Comparison between Model Settings

	Performance Metrics			
Algorithm	AUC	Precision	Recall	F-Score
Random	0.973	0.821	**0.633**	0.715
Averaging Embeddings	0.971	0.816	0.628	0.71
LSTM + Attention	0.976	0.877	0.62	0.727
DAN	**0.978**	**0.888**	0.628	**0.735**

Experiments and Results (2)

Table: Performance Uplift to TBI Model

	Performance Metrics			
Settings	AUC	Precision	Recall	F-Score
Only Hand-Crafted Feat.	0.975	0.817	**0.651**	0.724
Hand-Crafted + DAN Feat.	**0.978**	**0.888**	0.628	**0.735**

Conclusions

- Combined shallow and deep neural network embedding for listing and traveler embedding and tree based boosting methods
- Deployed end-to-end manner and traveler embeddings are made available to any team in the company
- Extension by incorporating more contextual spatio-temporal information in the model
- Currently working on the creation of a scoring layer that combines listing and traveler embeddings to personalize recommendations and ranking of search results

Q&A

Thank You for the Opportunity!

We're hiring!

Visit our <u>Careers Page</u> to join our team!

https://lifeatexpedia.com/jobs