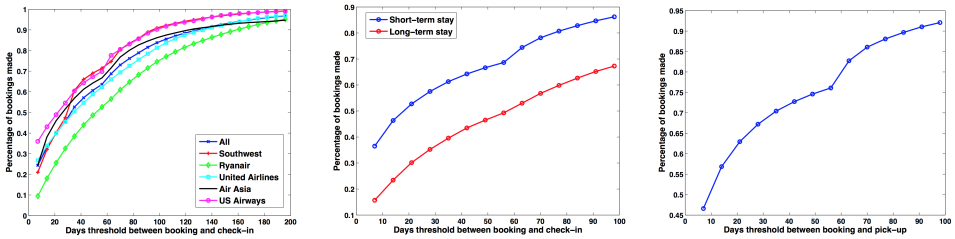
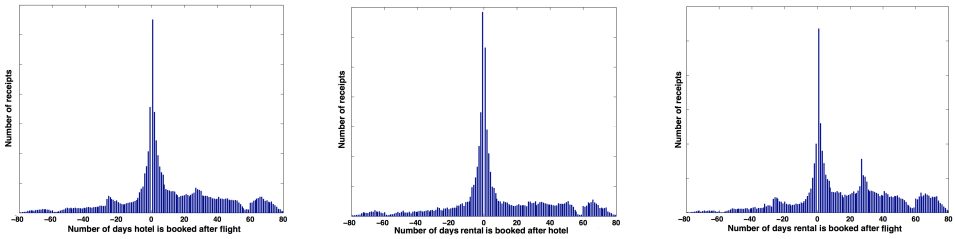


Figure 2: Cumulative distribution of receipts with respect to days booked in advance: a) flights; b) hotels; c) rentals



(a) Hotel vs. flight (b) Rental vs. hotel (c) Rental vs. flight

Figure 3: Number of days between coupled bookings from different channels

all bookings are made within one week in advance, while for long-term stay the percentage drops to 15%. In other words, customers are planning more in advance for longer, more expensive stays. We also observe jumps in the number of bookings 60 days in advance, which could be explained by black-out windows imposed by some loyalty programs.

Lastly, we see that more than 45% of rentals are booked within 7 days from car pick-up date. This grows to 70% one month prior to pick-up, indicating that rentals are planned much less ahead than either flights or hotels.

2.2 Analysis of cross-channel correlation

Next we analyzed time difference between bookings from different channels made for the same trip, where the bookings were deemed part of the same trip if the check-in times are within 24 hours. We only considered trips where the difference between booking dates and check-in date is longer than 30 days, in order to avoid the effect of late purchases which would bias the bookings to be temporally close. The results are given in Figure 3, where x-axis shows number of days between bookings from two different channels.

By considering the three subfigures, we can see that the majority of two-channel bookings are made within 10 days from each other. Moreover, we can conclude that when users are booking more than one channel for a single trip (e.g., booking a flight in addition to hotel, or rental in addition to flight), they are mostly purchasing them in the order "flight → hotel → rental", going from higher demand toward smaller demand channels. Interestingly, this conclusion is also confirmed by Figure 2, where 50% of flights are booked around 30 days in advance, 50% of short-term hotel stays around 20 days in advance, while half of car rentals are booked 10 days in advance. These insights are of great importance to booking agencies (e.g., they provide guidance on how and when to target users that already booked a flight).

3. PREDICTING BOOKING BEHAVIOR

In this section we turn our attention to behavior prediction task. In particular, we predict when a user will book a rental given they already booked a flight. We cast the problem as a classification task and predict labels "<5 days", "6-10", "11-20", "21-40", and ">41 or never", while we used age, gender, and past bookings in each channel as features. We trained logistic regression (we split the data into equally-sized training and test sets), and compared to baselines that predict most frequent label and most frequent label in age-gender groups (age buckets were set to "<26", "26-40", ">40"). The results are given in Table 1, where we see that the proposed approach outperformed the baselines.

Table 1: Predicting rental time given flight booking

Method	Rel. accuracy improv.
Most frequent	0%
Age-gender bucket	14%
Logistic regression	23%

4. REFERENCES

[1] K. Matzler and M. Waiguny. Consequences of customer confusion in online hotel booking. *Information and communication technologies in tourism 2005*, pages 306–317, 2005.

[2] H. Min, H. Min, and A. Emam. A data mining approach to developing the profiles of hotel customers. *International Journal of Contemporary Hospitality Management*, 14(6):274–285, 2002.

[3] M. Oppermann. A model of travel itineraries. *Journal of Travel Research*, 33(4):57–61, 1995.

[4] Å. Rudström and P. Fagerberg. Socially enhanced travel booking: a case study. *Information Technology & Tourism*, 6(3):211–221, 2003.

[5] E. Sezgin and M. Yolal. *Golden Age of Mass Tourism: Its History and Development*. InTech, 2012.

[6] B. A. Sparks and V. Browning. The impact of online reviews on hotel booking intentions and perception of trust. *Tourism Management*, 32(6):1310–1323, 2011.

[7] World Travel & Tourism Council. Economic impact of travel & tourism. Technical report, 2015.

[8] Y.-T. Zheng, Z.-J. Zha, and T.-S. Chua. Mining travel patterns from geotagged photos. *ACM Trans. on Intelligent Systems and Technology*, 3(3):56, 2012.