

附件C：译文

从线上行为到线下销售

【译者】重庆大学 计算机学院 物联网一班 陈点

指导教师评定成绩

(五级制)：

指导教师签字：

Ping Luo[#], Su Yan^{#,b}, Zhiqiang Liu^s, Zhiyong Shen^s, Shengwen Yang^s, Qing He[#][#]Key Lab of Intelligent Information Processing of Chinese Academy of Sciences (CAS),
Institute of Computing Technology, CAS, Beijing, China.^bUniversity of Chinese Academy of Sciences, Beijing, China.^sBaidu, Inc., Beijing, China.

摘要

为了与穿着睡衣在线购物的安逸相抗衡，线下商场的经营者们越来越把目光放在通过确认来客偏好的方式，来提升顾客的满意度和回头率。但是，大多数这方面的研究都只是建立在线下消费的历史记录之上。我们也可以得益于互联网来获得顾客的在线行为，像是搜索记录、网页浏览记录或是在线购物记录什么的。有没有可能这些从不同的方式（即在线和离线）中看似毫不相干的信息，是存在某种相互关联的？我们怎样才能利用在线的行为和离线的行动，组合起来找到可以促进离线零售的方法呢？

在本次研究中，我们将这项任务构想为一个跨模态的推荐问题，并且通过制定一个概率性图形化模型来展示其解决方案，我们把这个模型叫做O2OTM。具体而言，这种方法明确地在线上线下的主题间建模，使在线和离线行为两者的似然性最大化。并且，作出的推荐仅基于在线与离线主题而定，记作 (t, l) ，如果一个在线主题 t ，与不包含此主题的平均响应数量相比，可以大大提高其相对应的离线主题 l 的响应，我们认作其具有较高的 $lift$ 值。此外，我们同时通过线上实测与回归实验对解决方案进行评估。这个模型实测部署于某个著名的购物商场在北京的周年纪念宣传活动中，与基准线相比，平均每条信息带来的购买行为数量上，我们的方案带来了29.75%的提升。此外，我们的模型还在线上搜索主题与线下品牌主题间，发现了一些有趣而具有解释性的联系。

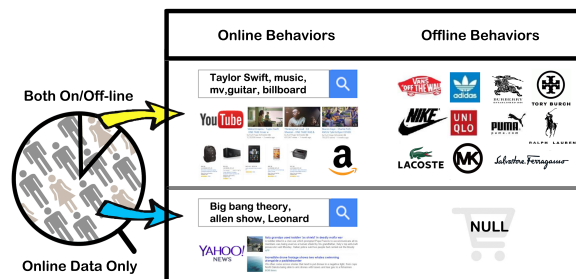
1. 绪论

推荐系统(RS)在近几十年来一直很流行，通过在庞大的候选集中促成选择以提升人们的生活质量。再此，我们关注推荐系统中的线下销售。虽然线上购物具有着蓬勃发展的趋势，线下销售在2014年依然分别覆盖了92%的美国销售市场和90%的中国销售市场。因此，在离线销售中任何微小的提升都可能带来巨大的收益。

与此同时，随着IT系统渗透到这个行业，为了更好地识别客户的喜好，客户的线下消费日志能够被记录下来，以便于提高满意度与留住客源。然而，以往的研究往往是把离线行为和在线的分开考虑，并仅仅根据离线（或在线）提出建议方案。

在本篇论文中，我们以一个联合的视角来研究线上与线下的行为。在线行为包括互联网上的全部活动，如网页浏览，网络搜索，在线视频观看，在线听音乐，网上购物等，离线行为包括现实生活中的所有活动，如在商场购物，旅游，外出就餐，看电影等。图像1展示的是我们这两种模式下数据中的一个例子。对于一小部分的用户，他们线上线下两者的活动我们都可以知道。在线上，我们知道他们在搜索引擎中搜了些什么，看了什么视频，以及在线购物了哪些物品。在线下，我们知道他们在一个线下的商场消费了哪些品牌。有了这些数据，我们可能会问：用户在YouTube上观看的视频，可以解释他们选择出行的旅游地点吗？会不会用户在网上阅读的新闻与他们在图书馆借什么书有关？如果我们将线上和线下的行为连接在一起，是否可以让这两种模式的数据在更具意义的推荐上激发出化学反应呢？

这项研究是这个方向的先驱。作为一个案例研究，在这里我们考虑的是在网上搜索日志的在线行为，线下购物记录的离线行为（注意，该模型可以应用到任何类型的线上线下活动）。最初的这两组数据分别来自互联网公司和本地商场。我们将展示如果互联网公司提供在线数据，和本地零售商提供的离线数据进行合作，将会是多么有利可图的双赢局面。



图像 1: 离线行为与在线行为

我们在这项研究中的基本假设为，客户的在线行为可能反映了他们的兴趣，从而影响他们的离线行为。从现实世界的的数据中，我们发现一些有趣的例子可以表明用户的搜索记录着实与他们线下的购物行为有关（尽管人们可能并不一定会确切地在网上搜索他们在线下购买的商品）。以服装品牌EVISU为例：读者可以通过其在店内的广告海报看出的这个服装品牌的风格，如图像2。同时，表格1展示了选择了这个品牌的客户查询的一些关键的词汇。这清晰地表明了购买EVISU的人们在日本动漫、电脑游戏上感兴趣，这说明他们是时尚的新一代。这也与EVISU的品牌精神相一致。



图像 2: EVISU 的一份店内海报

词汇	描述
火影	<i>Naruto</i> , a Japanese anime
海贼王	<i>One Piece</i> , Japanese manga series
大学	campus
动漫	animations
lol	<i>League of Legends</i> , online battle arena game
ios	iphone operating system
ps4	a video game console from Sony Entertainment

表格 1: EVISU 顾客搜索的关键词

在这个假设下，如果在线和离线行为之间的关系可以被自动的发掘出来，那将会给线下与线上双方的商场经营者带来极大的收益。一个典型的离线应用是商场的促销活动。为了增加销售收入，平均每年一次或两次，商场会进行发放优惠券到目标客户的推广活动。那么，应该发放哪种优惠券给某个特定的用户呢？撇开折扣的多少不谈，我们有理由假设，一个用户越喜欢某个品牌，那么他就会越有可能用这种（品牌的）优惠券。在此前，这种线下购物优惠券的推荐仅仅进行基于用户的线下购物的历史。随着在线行为的支持，更准确的建议会显著的增加预期收益。这一点也在北京著名的购物中心进行的实地试验中得到了验证（细节参见Section 5）。

对于在线的应用，当把线上线下之间的联系应用起来，可能会帮助当地的零售商进行更有效的在线对目标客户推广一个品牌的广告。举个例子，如果我们知道买EVISU的人们经常看日本动漫，那我们可以直接在他们观看影片之前直接把EVISU的优惠券投放给他们。与此同时，与主题相关的代表词，则正是品牌的负责人们需要为其搜索结果中投放的广告出价竞标的关键词。

受到这些应用程序的启发，我们在这里提出的概率图模型，用以描述线上线下的行为，以达到离线零售推荐的目的。与RS以前的研究相比，这项任务给我们带来了一些新的挑战。

- 多模态和形态遗漏。线上线下的活动，实际上形成了用户行为的两种模式。以往的研究采取使用的仅有一种模式（在线或离线）用于推荐，而忽略其他部分。另外，如图像 1 中，只有用户的一小部分同时具有在线和离线活动。如左边一样，大量商场（住在商场附近的）潜在客户用户，他们可能只留下了一些在线轨迹，但是，可能之前并没有在商场消费过。这意味着缺失他们的离线方式。因此，我们需要将这些线上的轨迹转换成的线下消费的意象与偏好。
- 可解释性。先前的建议方法，如矩阵分解，通常是不太可解释的。在他们的方法中，由于用户和项目被映射到潜在的向量中，很难解释每个维度表示着什么。但是，如果提供了销售推荐的理由，对于商场所有者来说将变得更加容易接受。此外，对于网络广告的应用，我们需要知道哪些是预测到在某品牌的消费上，所确切对应的关键词。在这项研究中，在线轨迹，尤其是用户搜索的询问，为我们提出的建议的解释与说明，提供了独特的机会。因此，我们不仅要做出的建议，同时也要对在线和离线行为之间的关系明确建模，使得离线推荐可以通过在线搜索词来进行解释。稍后，我们将举出更多有意义的例子来说明我们模型的可解释性。
- 可预测与不可预测的特征。同样值得强调的是，并不是所有的在线行为都将预计线下活动。例如，大多数人都使用搜

符号	描述
\mathcal{B}	the set of all the brands
\mathcal{V}	the vocabulary set of all the search words
\mathcal{U}	the customer set
U	the total number of customers
\mathcal{B}_u	the brands bought by customer u
\mathcal{W}_u	the words searched by customer u
\mathcal{P}_u	the brands which are not consumed by u
θ	distribution over online topics for customers
ϕ	distribution over search words for online topics
π	distribution over offline topics for online topics
φ	distribution over brands for offline topics
π	distribution over offline topics for online topics
$\alpha, \beta, \eta, \lambda$	Dirichlet prior
x, y	online topic variable
z	offline topic variable
w	word variable
b	brand variable
t	index for online topics
l	index for offline topics
v	index for search words
k	index for brands

表格 2: 符号变量

索引来搜索最新消息。直观地说，新闻相关的关键词，例如，当地的地名和突发新闻中人物的名字，可能无法预测或因果影响到任何线下消费。因此，提出的模型需要自动地从不可预测的在线特征中区分出可预测的，以便仅基于可预测的在线特征作出推荐。

在前十年来，许多研究者早已开始了对推荐算法的研究。对于这些推荐技术相关的作品的详细讨论将在Section 7 中有。在此我们认为，虽然以前的研究可能可以分别解决一些挑战，但仍然不存在一种模式可以解决所有这些问题，特别是对解释性和预测功能的问题上。在本文中，我们假设在线和离线的行为可以被分组为一些语义相关的主题。于是，我们提出一种基于lift的推荐方法，这种推荐仅基于对在线和离线主题，用 (t, l) 表示，当lift取较高值时，从而使得在线主题 t 的存在，比起不存在时大大增加了离线主题 l 的平均响应数量。为了评估lift的值，我们提出了一个概率图模型，称为Online to Offline Topic Model (O2OTM)，用来明确建立在线和离线主题之间的关系，从而使得生成所有在线和离线行为的似然性最大化。我们认为，这种基于lift的推荐方法将是现场推荐方案的更加实际有效。

最后，我们通过回溯和实地实验两种方法评估我们的模型，以显示其在几种基准下的优越性，也会给出更多的例子来说明它的可解释性。对于现场评估，我们在北京一家著名购物中心的销售活动中部署我们的策略。我们发现，在我们的模型中每一个促销的消息比起随机购买方法带来了29.75%的增幅，收入超过只使用离线行为的方法12.5%。因此，我们认为，通过结合线上和线下的行为，我们的解决方案对于离线销售的业务影响有着巨大的潜力。

2. 问题评估

在这里，我们以客户的线上和线下的行为作支持，为他们推荐线下销售的品牌。在这一节中，我们将描述我们掌握的数据并制定问题。所有相关的数学符号在表格 2 中进行了描述。

对于每一个推荐任务，用户与名目是不可或缺的两个元素。在我们的问题中，用户都看作是一个商场的潜在客户，而名目是指在商场出售的产品品牌。 \mathcal{B} 是所有品牌的集

合, \mathcal{U} 是所有客户的集合。同时, 我们可能知道这些客户在某个流行的搜索引擎发出的搜索查询。 \mathcal{V} 就是词汇集所有的搜索词。

对于每一个顾客 $u \in \mathcal{U}$, 他的线下行为表示为 $\mathcal{B}_u = \langle b_{u,i} | b_{u,i} \in \mathcal{B} \rangle$, 其中 $b_{u,i}$ 是顾客 u 购买的第 i 种品牌名。同时, 顾客的在线行为表示为 $\mathcal{W}_u = \langle w_{u,j} | w_{u,j} \in \mathcal{V} \rangle$ 其中 $w_{u,j}$ 是用户 u 搜索的第 j 个词标记。换句话说, 在线和离线的行为都以 bag-of-words 的方式进行表达。同时需要注意的是, \mathcal{B}_u 或者 \mathcal{W}_u 可以是空集。在所有这些数据的基础上, 我们把重点放在对品牌的推荐上。特别地, 对于每个用户 u 我们为其在 \mathcal{B} 中的每个品牌提供一个评级。

3. 线上到线下的话题模型 (O2OTM)

在我们详细的介绍基于提升量 (lift) 的推荐方法之前, 我们要提出一个线上到线下的话题模型 (O2OTM), 这种模型为线上与线下的话题之间的关系明确地建模, 基于从线上与线下两端获取的的语料数据 $\{\mathcal{W}_u, \mathcal{B}_u\}_1^U$

3.1 对线上线下行行为分别建模

首先, 基本的话题模型建模方法使用的是 LDA, 可分别应用于在线与离线行为的建模中。LDA 基于一个理念用以对文档内容建模, 即在一篇文档中词的概率分布, 可以表示为一系列话题对于词的概率分布的集合。

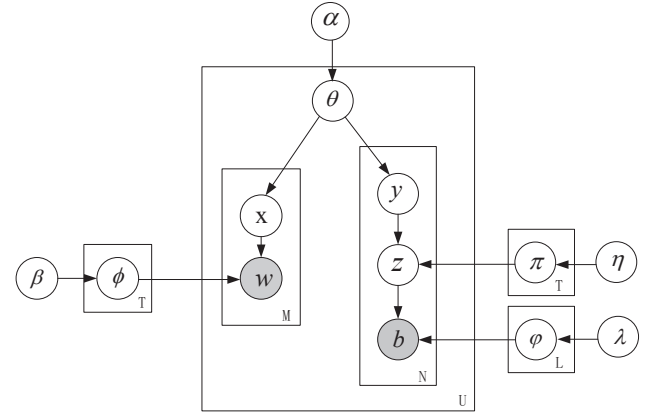
在这项研究中, 由于在线离线行为均由词包 (Bag-of-Words) 模型表示, 我们可以将每个 \mathcal{W}_u 或者 \mathcal{B}_u 看作一个文档来组成在线语料库 $\{\mathcal{W}_u\}_1^U$ 和离线语料库 $\{\mathcal{B}_u\}_1^U$ 。从而, LDA 可以分别应用到这两个语料库产生的在线主题 (由 x 表示) 和离线主题 (由 z 表示)。在此, 一个在线主题 $p(w|x)$ 是基于搜索词的概率分布, 同时, 离线主题 $p(b|z)$ 是基于品牌的概率分布。这种算法有助于将搜索词和产品品牌分组到语义主题中, 并且可以用以回答“搜索词相似的在线用户”和“购买品牌相似的线下顾客”的询问。然而, 这种算法不提供任何与在线离线主题间关系有关的任何信息, 因此不能从在线行为到离线行为进行模态迁移。

3.2 对线上线下行行为共同建模

提出的 O2OTM 模型吸取了概率图解法的优势, 通过考虑线上线话题之间的明确依赖为线上线下行行为建模。在 O2OTM 模型中, 每个用户 u 采取如下的概率过程来生成在线与离线行为, 表示为 $(\mathcal{W}_u, \mathcal{B}_u)$ 。这个过程可以分为两个部分。第一步, 对于每个搜索词 $w \in \mathcal{W}_u$, 从在线话题特征 u 的分布 (表示为 θ_u) 中选出一个在线话题 x , 然后从这个选定的在线话题中生成单词。第二步, 对于每个品牌 $b \in \mathcal{B}_u$, 再次从 θ_u 中选择一个在线话题 y 。在 y 的基础上, 从离线话题特征 y 的分布 π_y 中选择一个离线话题 z , 然后从这个选定的离线话题中生成品牌。

这个过程所对应的图形模型见图像 3。在图中, 左边的部分生成了在线行为, 同时右边的部分生成了离线行为。在这个模型中, 每个用户 u 与在线话题的分布 θ_u (借鉴于一个对称的 Dirichlet(α) 先验) 相关联。 u 的混合权重用来选择一个在线话题 x (或者 y)。并且, 每个基于 x 的搜索词通过一个 ϕ_x 分布 (借鉴于一个对称的 Dirichlet(β) 先验) 生成。与此同时, 每个基于 y 的线下品牌通过一个 π_y 分布 (借鉴于一个对称的 Dirichlet(η) 先验) 生成, 每个品牌通过一个 φ_z 分布 (借鉴于一个对称的 Dirichlet(λ) 先验) 生成。

在此, x 和 y 表示在线话题。我们用这两种不同的符号来从用以生成离线品牌的搜索词当中区分出用以生成在线话题的搜索词。此外, x 和 y 均由同一个 θ_u 分布中生成。因此, θ_u 依赖于线上线下行两者的行为。还值得一提的是, 线上于线下主题之间的依赖关系以 π 为蓝本。它能够用以从线上到线下的模态迁移。



图像 3: O2OTM 的图形化模型

3.3 O2OTM 的 Gibbs 采样算法

在本次研究中我们使用 Gibbs 采样来估计 O2OTM 的参数。通过应用 Gibbs 抽样, 我们构建了收敛到对潜在变量后验分布的马尔可夫链。马尔可夫链连续状态间的过渡, 归因于借鉴标准狄利克雷积分重复迭代总结出来的其他参数分布得来的 x, y, z 。特别的, 我们用以下三个算式来更新 x, y, z 。

首先, 我们有

$$p(x_i = t | \vec{x}_{-i}, \vec{y}, \vec{z}, \vec{w}, \vec{b}) = p(x_i = t | \vec{x}_{-i}, \vec{y}, w_i = v) \propto \frac{O_{t,-i}^v + \beta_v}{\sum_{v=1}^V O_{t,-i}^v + \beta_v} \cdot \frac{O_{u,-i}^t + \alpha_t}{\sum_{t=1}^T O_{u,-i}^t + \alpha_t}, \quad (1)$$

其中 $x_i = t$ 表示语料库中第 i 个词标记在语料库中对应的第 t 个在线话题, $w_i = v$ 代表第 i 个词标记对应着字词库中的第 v 个, \vec{x}_{-i} 表示除了第 i 个以外所有的在线话题表示。除此之外, $O_{t,-i}^v$ 表示第 v 个词在第 t 个在线话题中被观测到出现的次数, 不包括当前例。 $O_{u,-i}^t$ 表示第 t 个在线话题在用户 u (即拥有搜索词 w_i 的) 中被观测到出现的次数, 不包括当前例。在此, 在线话题 x 被分配给语料库中的每个词标记。

然后, 我们有

$$p(y_j = t | \vec{y}_{-j}, \vec{x}, \vec{z}, \vec{w}, \vec{b}) = p(y_j = t | \vec{y}_{-j}, \vec{x}, z_j = l) \propto \frac{O_{t,-j}^l + \eta_l}{\sum_{l=1}^L O_{t,-j}^l + \eta_l} \cdot \frac{O_{u,-j}^t + \alpha_t}{\sum_{t=1}^T O_{u,-j}^t + \alpha_t}, \quad (2)$$

其中 $y_j = t$ 表示第 t 个在线话题中对应的第 j 个品牌标记, $z_j = l$ 代表在第 j 个品牌标记中第 l 个离线话题中被观测到出现的次数, \vec{y}_{-j} 表示品牌标记中在第 t 个在线话题中被观测到出现的次数, 不包括当前例。此外, $O_{t,-j}^l$ 代表第 l 个线下品牌在第 t 个在线话题中被观测到出现的次数, 不包括当前例。 $O_{u,-j}^t$ 则代表第 t 个在线话题在用户 u (即购买品牌 b_j 的) 中被观测到出现的次数, 不包括当前例。在此, 在线话题 y 被分配给语料库中的每个品牌标记。

最终, 我们有

$$p(z_j = l | \vec{z}_{-j}, \vec{x}, \vec{y}, \vec{w}, \vec{b}) = p(z_j = l | \vec{z}_{-j}, y_j = t, b_j = k) \propto \frac{O_{l,-j}^k + \lambda_k}{\sum_{k=1}^B O_{l,-j}^k + \lambda_k} \cdot \frac{O_{t,-j}^l + \eta_l}{\sum_{l=1}^L O_{t,-j}^l + \eta_l} \quad (3)$$

其中 $z_j = l$ 表示第 l 个离线话题中对应的第 j 个品牌标记, $b_j = k$ 代表在观测到的第 j 个品牌标记是品牌词库里的第 k 个词, $y_j = t$ 表示观测到的第 j 个品牌标记是在线话题中的第 t 个。 \vec{z}_{-j} 代表除了第 j 个之外所有的离线话题所对应的品牌标记。除此之外, $O_{l,-j}^k$ 表现了观测到的在第 k 个品牌上出

现第 l 个离线话题的次数，不包括当前例。 $O_{t,\neg j}^l$ 表现了观测到的在第 l 个离线话题上出现第 t 个在线话题的次数。在此，离线话题 z 被分配给语料库的每一个品牌标记。

在马尔科夫链收敛之后，模型变量可以通过以下的式子进行估计

$$\theta_{ut} = \frac{O_{ut}^t + \alpha_t}{\sum_{t=1}^T O_{ut}^t + \alpha_t} \quad (4)$$

$$\phi_{tv} = \frac{O_{tv}^v + \beta_v}{\sum_{v=1}^V O_{tv}^v + \beta_v} \quad (5)$$

$$\pi_{tl} = \frac{O_{tl}^l + \eta_l}{\sum_{l=1}^L O_{tl}^l + \eta_l} \quad (6)$$

$$\varphi_{lk} = \frac{O_{lk}^k + \lambda_k}{\sum_{k=1}^B O_{lk}^k + \lambda_k} \quad (7)$$

在这里， T, L 分别是在线与离线话题的数量。 V, B 是词库和品牌库的大小。

在这项研究中，我们不估计超变量 $\alpha, \beta, \eta, \lambda$ 的值，而是通过分别将这些平滑参数固定为0.05, 0.01, 0.01, 0.01。该算法开始需要随机为在线主题分配搜索词，并为离线品牌随机分配在线和离线主题。该算法的每次迭代包括为语料库中每个词标记和品牌标记运用算式 (1),(2),(3)。这种采样过程至少需要执行2000轮直至收敛。

3.4 模型推理

在我们学习到了所有的变量 θ, π, φ 之后，我们可以推断一个品牌和在线主题之间的关系。特别地，对于一个在线话题 $y = t$ 和一个离线品牌 $b = k$ ，我们有

$$\begin{aligned} p(y = t | b = k) &= \sum_{l=1}^L p(y = t | z = l) p(z = l | b = k) \\ &\propto \sum_{l=1}^L p(y = t) \pi_{tl} \varphi_{lk} \end{aligned} \quad (8)$$

话题编号 #	词汇	描述
3	学校	campus
3	学院	school
3	大学	university
3	研究生	graduate students
3	兼职	part-time job
3	专业	majors in school
15	运势	horoscope
15	星座	12 zodiac signs
15	八字	Chinese horoscope according to the birthday
15	名字	name
15	周公解梦	understanding the dreams
44	皮肤	skin
44	头发	hair
44	发型	hair style
44	化妆品	cosmetics
44	排行榜	top charts

表格 3: 图像 4中出现的词的英文翻译

有了这个推导，在这里我们向您展示一个品牌和在线主题之间的关系的一个例子。*TEENIE WEENIE* 是一个



(a) Teenie Weenie的一份店内海报



(b) 三类高相关在线主题的标志词

图像 4: 品牌 *Teenie Weenie* 与它的相关词。这些词的解释详见 表格 3。

为15到25岁年轻女孩设计的韩国品牌。它因其学校式样的裙子而出名，见图像 4(a) 中的海报。图像 4(b) 展示了这个品牌在线话题的分布，表格 3 中也展示了Top3的代表词。第一个话题主要是关于校园生活，这预示着该品牌在某种程度上关系到大学生。第二个话题是关于星座。很多人尤其是年轻女性感兴趣的是他们的星座，它可以预测爱情，财富，事业等。第三话题是关于发型和化妆品，可以看出这个品牌与在乎自己的外表的女性团体相关。这三个话题显示出 *TEENIE WEENIE* 的顾客是年轻女性团体，尤其是为了爱情与未来而关心自己外观的大学女生。我们认为，这种关系可以显示出消费者对品牌的偏好，帮助设计师根据客户的兴趣和口味来调整商品。

4. 跨模态推荐算法

在这一节中，我们将展示跨模态推荐的算法。

4.1 实地评估的推荐算法

使用变量 θ, π, φ ，我们可以为顾客提供推荐。第 i 个用户购买第 k 个品牌的概率可以使用如下方法计算：

$$\begin{aligned} p(b = k | u = i, \theta, \pi, \varphi) &= \sum_{t=1}^T \sum_{l=1}^L p(b = k | z = l, \varphi) p(z = l | y = t, \pi) p(y = t | u = i, \theta) \\ &= \sum_{t=1}^T \sum_{l=1}^L \varphi_{lj} \pi_{tl} \theta_{it} \end{aligned} \quad (9)$$

通过对 θ_{it} 的估计，推荐将基于线上线下行为给出。由于这种方法可以方便地生成的品牌排名为用户，我们将进行实测评估。此方法与只使用离线行为的方法相比，是特意开发用来展示“在线行为可能会为建议任务提供额外的价值”的。

4.2 实际系统中的推荐

算式 (9) 中的推荐算法具有如下的缺点。其一，这种算法使用所有在线主题来提供线下推荐，但我们认为，如前面提到的，一些网上的话题可能无法预测或因果关系到离线消费。此外，虽然随着消费意向的主题淹没了这些不可预知的在线主题，但它们可能占据了用户的搜索内容的很大一部分。例如，从初孕女子的搜索内容中，比起她的工作和日常生活中进行比较，可能很大一部分不会包含关于孕期保健的事情。因此，对于该用户的在线主题分布中有关于孕期保健的话题处于较低值，基于公式 (9)，有关婴儿和孕妇的品牌可能不会稳居高位。

4.2.1 提升量

因此，我们需要一项措施来自动识别预测在线主题和其对应的离线主题的主题对，因为建议是仅由这些主题对为基础。

特别地，我们在线上话题 t 与线下话题 l 之间定义如下这种参量 $lift$ 。提醒一下我们提出的 O2OTM 模型输出的就是线上和线下的主题之间的关系。顾名思义，给定线上话题 t 时线下话题 l 的概率，通过 $\pi_{tl} = p(z = l | y = t)$ 分配使得所有数据的似然性最大。之后，基于 π 我们也可以计算出 $\pi_{\neg tl} = p(z = l | y = \neg t)$ ，给定其他所有在线话题包括 t ，度量离线话题 l 的可能性。那么， $lift$ 可以如下定义

$$lift_{tl} = \frac{p(z = l | y = t)}{p(z = l | y = \neg t)} = \frac{\pi_{tl}}{\pi_{\neg tl}} \quad (10)$$

这个提升值简单来说只是两个数值的比率：包含 t 的目标 l 的响应与不包含 t 的目标响应的比值。它实际上是用来衡量预测 l 的时候，具有 t 比起不具有 t 会有着多么大的提升。这个提升值越大，代表在线话题 t 与离线话题 l 之间的关联性越强。

4.2.2 提升量高的示例对

在表格 4 中，展示了高 $lift$ 值的一些例子，从我们的数据中选出的一些在线与离线的话题对。这些例子是通过筛选得到的，以便于覆盖不同群体的顾客。

- 对于第 I 组，搜索词和宠物医疗相关，并且这个在线话题到对应着为年轻女孩设计的时尚品牌具有 11.5 的提升值。由于这些品牌大多是从日韩紧扣的，他们比起常规的要更加昂贵，表明了这些女孩为了养宠物可以负担得起这些。

- 对于第 II 组，它对应着喜欢日漫和电脑游戏的年轻群体的（大多是男孩）他们大多喜欢穿一些休闲品牌，比如 VANS, ADIDAS ORIGINAL 等等。

- 对于第 III 组，它对应着孕妇和年轻妈妈们的群体。他们会搜索关于孕期护理、新生儿护理和产后恢复的知识。同时，他们也对应着两种提升度较高的线下话题。第一个品牌群面向新生儿，第二个品牌群面向年轻女孩的一些品牌，这表明新妈妈们依旧不会停止为自己购买些什么。

- 对于第 IV 组，对应的是拥有幸福生活的家庭主妇群体。在线上，他们搜索有关减肥和热门电视剧、网络小说等材料。在线下，他们买两种类型的品牌。一种是为了她们的孩子，另一种是为了他们的丈夫选择的户外运动装扮的品牌。

- 对于第 V 组，我们认为这是高端化妆品与珠宝品牌的消费者。我们惊讶地发现当他们搜索武侠故事、都市恋爱小说和文学的时候，提升值到达了 19。并且搜索整容和手机游戏的人们喜欢去购买这些昂贵的商品。

- 对于第 VI 组，我们认为这些是寻求股票，基金，证券和投资的人。这些是有着一些闲钱可以用来资产管理的富人们。在这个在线话题中具有最高提升值的离线话题，是和有着各种中国菜品的高档餐厅有关的。这表明，美食是这些人的重中之重。

值得一提的是，所有这些对具有高提升值的线上和线下主题对，都是我们提出的 O2OTM 模型自动发现的。他们是

与我们的在线和离线行为的基本认识是一致的。这些是与我们的在线和离线行为的基本认知相一致的。

4.2.3 基于提升量的推荐

那么，通过在提升值上加上一个人为界定的阈值 $\rho > 0$ ，我们可以找出在线和离线主题对中，提升值不小于 ρ 的。顾名思义，我们有

$$\mathcal{R} = \{(t, l) | lift_{tl} \geq \rho\} \quad (11)$$

于是，对于有着高提升值的话题对 \mathcal{R} 我们提出了以下的推荐算法。以周末的购物的推荐方案作考虑。在最近的工作日的搜索词收集推断他们的在线主题。当一个在线话题 t 出现了至少 q 次的时候，它就会与 \mathcal{R} 进行比对以检查 t 是否出现于 $\mathcal{T} = \{t | (t, l) \in \mathcal{R}\}$ 当中。如果是的话，所对应的离线话题中最高品牌将被选中作推荐。需要一些后续的处理步骤，通过考虑用户最近的购物历史和推荐日志，以避免重复的建议。在这次研究中，两个阈值被设定为 $\rho = 4.5$ 以及 $q = 10$ 。

在这个算法中，推荐是通过一个触发方式做出的。它代表着推荐只会当在线话题具有一个很高的值，并持续了足够的时间且对应的推荐被认作足够可信赖的时候才会做出。它可以自动判断什么时候提出建议，即使在许多不可预知的搜索词出现的时候。如同表格 4 中所示的主题对清晰地反映了离线的消费意向，我们相信通过这种基于提升值算法的推荐将会对潜在客户具有更多的吸引力，同时也可以避免对用户不必要的打扰。

5. 实地评估

我们将提出的提升度模型部署在真实世界中，选择了北京一个著名商场的销售活动中推行。为了庆祝周年庆，商场打算选择各种品牌送出大量优惠券来吸引顾客。这些优惠券在周年庆的前三天通过短信发给用户，每个短信中包含 3 种不同品牌的优惠券。对于这家商场的 VIP 用户，我们知道他们在线搜索查询记录和线下消费历史的信息。我们将这些用户随机分为三组，其中每组包含大约 10,000 个用户。由于数量较大，我们假定这三组用户具有相同的人口统计信息（如性别，教育程度，消费能力等）。

然后，我们评估三种方法下推荐算法的效果，它们分别用于这三组用户的优惠券建议。这三种方法包括：1) RAN，表示随机建议；2) LDA，仅基于离线消费历史的方法。特别的，LDA 算法应用于 $\{\mathcal{B}_u\}_{u=1}^U$ ，并且，可以通过 $p(b|u) = \sum_z p(b|z)p(z|u)$ 来进行离线销售的推荐；3) LIFT，表示同时基于在线与离线数据两者的基于提升度的算法。这三种方法是被故意设计来回答，与随即方法相比，基于历史数据的建议是否有效的问题，以及只使用离线数据相比，同时使用在线行为这一提议是否具有额外的价值。

算法	IPR	APPC 值 (元)	OCPPUP 值 (元)
RAN	2.34%	1656.64	38.76
LDA	2.39%	1870.30	44.70
LIFT	2.59%	1942.10	50.30

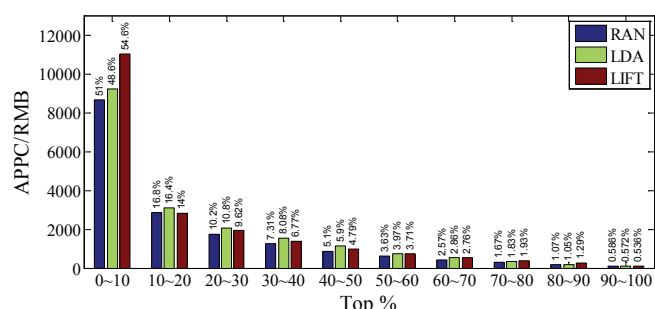
表格 5: 实地评估结果

在这次实地测试中，我们记录了这三种算法在周年庆时对应的三个用户群在商场的消费情况：

- IPR, 存货购买率 指的是人们在纪念日人们在商场消费的比率。
- APPC, 用户平均购买量。
- OCPPUP, 单位促顾客购买量 其值等于 $IPR \times APPC$ 意味着每个推荐信息能带个商场的购买量。

结果详见表格 5。可以清晰的看出，LIFT组的消费情况比三种算法的其他两组要好。尽管IPR的提升比较小，但APPC的值提升了很多。它带来的50.3的OCPPUP值，表示每一条带有使用LIFT推荐的优惠券的短信，平均会带来50.3人民币的收益。对于另外两组用户，OCPPUP值分别为44.7和38.76。因此，LIFT组在OCPPUP上比起RAN组有着29.75%的提升，比起LDA组有着12.5%的提升。

由于这三组用户的人口统计数据是相同的，我们认为使用LIFT能更好地增加优惠券推荐的购买建议。这表明，使用同时与在线和离线行为的O2OTM模型的用户，他们的喜好能更准确地得到更具吸引力的建议。这清楚地表明，如果基于在线和离线行为的建议被引入到离线零售业，将会具有巨大潜力的业务影响。



图像 5: 使用APPC进行比较的排名

此外，我们想要知道用户的消费增长从何而来。具体地，对于每个用户组，我们对在他们所购买的品名进行排序。然后，我们把在这次销售活动中购买了物品的用户，根据他们的评级分成10个子分组。换句话说，前10%的用户在第一子分组，前10%-20%的用户在第二子分组，以此类推。然后，对于每个子组，我们计算其APPC值。结果见图像 5，其中每个条上的数字代表这个子分组的用户比起所有用户的总购买量所占的购买比例。

我们进行了以下观察。首先，消费排名为前10%的用户贡献的总收益高于LIFT用户群总收益的54.6%。其次，LIFT中APPC增长主要来源于这些高排名的高质量客户。这表明这些高价值客户对于优惠券推荐的表现更敏感。提供高价值客户更好的建议，它会带来更多的购买。

6. 回馈评价

在本节中，我们比较了在回馈评价中推荐方法的性能。

我们使用的数据集每个都是由在线部分和离线部分组成。在线部分包含用户的搜索查询内容，这些内容都是来自一个在中国会被广泛使用的搜索引擎。每个搜索记录有三个属性，分别是用户ID，查询语句和日期。注意，一条查询语句可以被解析成一系列单词的集合。离线部分包含了顾客在当地商场的消费历史记录。每次消费记录由用户ID，所买的品牌和日期构成。

在回馈评价中，我们使用了两个数据集，这两个集合分别来自中国的两家不同的商场，分别标记成商场A和商场B。总计，我们从商场A的到了6539条用户记录，732个线下品牌和30097条搜索文字。对于商场B，则有6961条用户记录，789个线下品牌以及103597条搜索文字。对于每个数据集，我们使用了2013/01/01到2014/12/31期间的数据作为训练集，用2015/01/01到2016/06/01的数据作为测试集。

很多方法都能够用来评估推荐系统的性能，像是绝对平均误差（MAE），均方根误差（RMSE），精确度，召回率等等。因为这个问题和评分预测无关的，所以采用MAE和RMSE作为评估方法是不合适的。我们也不能使用召回率进行评估，

因为对于一个用户会买的所有品牌的参考标准我们是不知道的。我们只能计算，在测试集中，一个用户最终买下了被推荐的品牌的百分比。在的到了针对每个用户的精确度后，我们会计算它们的平均值并得出针对整个用户群体的精确度 $precision@K$ 。

在这个评估中，我们关注的是对用户之前没有消费过的产品进行推荐，因为这些没有消费的产品可能更加精确的反映了用户的偏好。我们知道，之前有针对会被重复消费的商品的研究，这些会被重复消费的商品，很多是由商品的使用经历，商品的属性（比如可重复使用或者一次性），和一些用户的心理问题（比如，无聊了）所决定的。然而，对会被重复消费的商品推荐已经超出本次研究的范围。

这里我们有两个基准方法，RAN和LDA，在实地评价中它们是一样的。RAN方法是指随机推荐方法，而LDA是只使用了离线行为的方法。对于RAN方法，它的 $precision@K$ 值能够通过以下公式进行估计，

$$precision@K = \frac{1}{|U|} \sum_{u \in U} \frac{|\mathcal{T}_u - \mathcal{B}_u|}{|\mathcal{B} - \mathcal{B}_u|}, \quad (12)$$

其中 $(\mathcal{B} - \mathcal{B}_u)$ 指的是所有品牌的未消费顾客 u 的集合， $(\mathcal{T}_u - \mathcal{B}_u)$ 包括了所有测试集中用户 u 进行了购买而未消费的品牌。需要注意对于RAN这个量是独立于 K 的。

在这次的回馈评价中，我们通过公式 (9) 比较了两种基准算法与推荐算法O2OTM。这种算法可以生成未消费品牌的全评价排名。但是，基于提升值的算法在线上话题没有达到高提升值时，可能不会推荐任何东西。所以，我们在回馈评价中没有包括这一部分。

需要注意的是依然有着一些强大的基准算法，比如可以为用户和物品嵌入特征（比如，所有搜索词）的分解机推荐模型。但是，这些算法不能提供任何推荐的解释性。此外，回馈评价的目的只是为了展示在线行为可以为离线销售提供额外价值。所以，我们的评估中不包括这些算法。

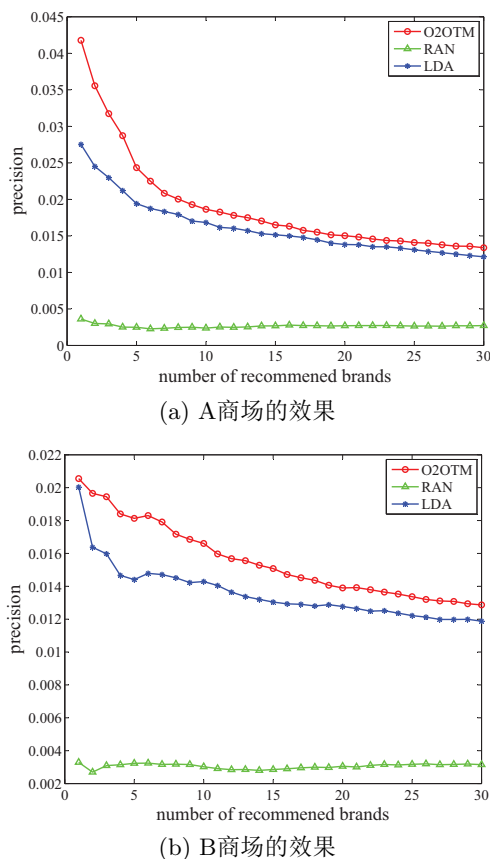
在一些初始化实验之后，为了回馈评价，我们将在线话题的数量设为 $T = 50$ ，将离线话题的数量设为 $L = 30$ 。两个商场 $precision@K$ 的结果在图像 6 中进行了展示。它清晰的展现出O2OTM显著地比LDA更好。在A商场，当 K 较小时O2OTM算法的提升更加明显，但是在 K 到达15时提升会变小。在B商场，当 $K = 1$ 时O2OTM和LDA的表现是相似的，但是当 $K > 1$ 是O2OTM的提升就变大了。同时，RAM算法的线是扁平的，和我们的分析相一致。

7. 相关工作

我们的工作和三个领域有关，i) 推荐方法ii) 话题建模方法和iii) 多模态学习

推荐 很多类型的物品，像是书籍，音乐，移动应用，甚至是兴趣点，都已经被开发过相关的推荐算法服务。协同过滤是最出名的推荐方法，其核心思想是，有着相似兴趣的用户可能会在相似的商品表现出相似的评分。尤其是，如果一个推荐是基于相似的用户（那些对相同的物品上有着类似评分的目标用户）做出来的，它就被称为基于用户的协同过滤。另一方面，如果一个推荐是基于相似的物品（那些被相同集合中的人打出了相近分数的物品）做出来的，那么它就被称为基于物品的协同过滤。除了这些最相近的科技，矩阵分解是在假设“用户和物品是与潜在的因素联系起来的，并且观察到的用户—物品评分矩阵是2的倍数”下开发出来的。这些潜在因素可以通过最小化一个loss函数学习到（loss函数被定义成评分的观察值和预测值之间距离的平方和），或者通过最大化回归残差的似然性学习（假设回归残差遵循标准高斯分布）。

为了解决数据稀疏的问题，很多不同的方法被提出，以求让推荐包含用户和物品的各种各样的特性。例如，社交网络被引入来约束分解的过程，在这个方法下，有着社会联系



图像 6: 在两个商场中，三种算法下precision@K 的值

的人们会倾向于有相似的用户特征向量。对于有着丰富文本信息的物品，提出了结合矩阵分解和LDA的fLDA模型。为了推荐科技类文章，开发了CTR模型，其假设“物品的潜在因素是它的文档主题分布和潜在的偏移矢量之和”。Yin等人开发了一个模型来解释无声的观看行为，将其转换成“伪投票”来帮助推送文章。另外，分解机的一般模型也被提出，来引进物品和用户的各种类型的特征，以得到有着可观性能的推荐效果。

然而，这些相关工作全都没有考虑在多模态角度下的推荐方法，尤其是在两个模式只是有着微弱的联系的时候。对于跨模态下的推荐任务，我们明确地考虑了两个模式之间的关系，并同时在两个模式上建立了话题模型。这些联系不仅清晰地跨模态迁移提供了可解释性的帮助，还有助于识别具有高提升值的联系。

话题建模 PLSA 和LDA 是比较流行的在共生矩阵上获取话题的方法。不同的话题建模方法的变体被提出，用以描述数据的额外特性。例如，作者话题模型就是在文章上考虑作者信息所提出的模型。TOT模型是为了用话题的流行时间来描述话题所提出的模型。这些之前的工作提供了“如何描述我们拥有数据的概率生成过程”的基本思想，但是它们没有一个被应用在跨模态数据上。

多模态学习 随着多模态下数据的快速积累，多模态学习最近在近年收到了广泛的关注。大多数之前的工作都是关注于用多模态的特征来监督或者半监督问题学习的过程。协同训练是解决这个任务的最出名的方法。最近，不同模态的重要性也被考虑进来，用于从强模态的监督中提取弱模态的鉴别特征。注意所有的这些模态都是在特征空间中的。关于在

这项研究里的跨模态推荐任务，在线模态被用作特征空间，而离线模态则被认为是预测标签。

8. 结论

在本篇论文中，我们基于用户的在线行为来研究离线销售的推荐问题。我们提出了概率模型O2OTM，它展现了线上线下两种活动的话题模型（Topic Model），同时建立了在线和离线话题之间的关系。这些产生的关系可以直接对跨模式的建议进行支持，并为建议提供解释。此外，提升度（Lift Measure）用以标识在线与离线的主题间，具有很强甚至是因果关系程度的联系。这次实地的测试表明了这种跨模态的建议中，与藏着巨大的商业影响力。

我们强调，这种跨模态的推荐模型有着广泛的应用前景。举例说明，基于搜索日志，我们就可以对餐馆、移动应用、旅行网站等做出推荐。此外，我们将会使得该模型可以扩展为大规模数据所用，并构建跨模态推荐的普适系统。

译文原文出处：

[1] P. Luo, S. Yan, Z. Liu, Z. Shen, S. Yang and Q. He. From online behaviors to offline retailing. In KDD, 2016.

	在线词(英文描述)	提升值	离线品牌(英文描述)
I	1. 狗狗 (dogs, puppies) 2. 泰迪 (<i>Teddy Bear Puppy</i> , a breed of small lovely dog) 3. 宠物 (pet) 4. 狗粮 (dog food) 5. 绝育 (sterilization for pet) 6. 猫咪 (cats) 7. 宠物医院 (pet hospital) 8. 疫苗 (Pet vaccine) 9. 金毛 (<i>Golden Retriever</i> , a breed of large dog with thick yellow hair) 10. 拉布拉多 (<i>Labrador</i> , a breed of large dog)	11.5	1. TUTUANNA (a Japanese brand famous for lovely girl underwear) 2. PEACH JOHN (girl underwear brand of hot sexy and cute style) 3. INNISFREE (a Korean cosmetics brand) 4. SIXTY SIGHT (women fashion brand, famous for women's lingerie, dresses) 5. SNIDEL (a Japanese girl clothing brand) 6. LILY BROWN (a Japanese girl clothing brand) 7. EARTH MUSIC (a Japanese girl clothing brand of simple and natural style) 8. ORBIS (a Japanese cosmetics brand) 9. The Body Shop (natural skin care, make-up brand) 10. Lowrys Farm (a Japanese girl clothing brand)
II	1. 游戏 (games) 2. 漫画 (comic series) 3. 怪物猎人 (<i>Monster Hunter</i> , a PSP game) 4. 盗墓笔记 (<i>The Secret of Grave Robber</i> , a Chinese net-novel) 5. 食尸鬼 (<i>Tokyo Ghoul</i> , a Japanese anime series) 6. 火影忍者 (<i>Naruto</i> , a Japanese anime series) 7. 最终幻想 (<i>Final Fantasy</i> , a role-playing video game) 8. 暴走 (<i>Rage Comic</i> , a comic series) 9. 战争 (wars) 10. 上古卷轴 (<i>The Elder Scrolls</i> , an action role-playing video game)	7.5	1. ADIDAS (a famous brand for sportswear) 2. NIKE (a famous brand for sportswear) 3. VANS (a brand for skate shoes and clothing for the young) 4. CONVERSE (a brand for sportswear and lifestyle footwear) 5. NIKE 360 (a classical sportswear series of NIKE) 6. ADIDAS NEO (a sports and leisure series of ADIDAS) 7. ADIDAS ORIGINAL (a classical sportswear series of ADIDAS) 8. LEVI'S (a famous brand for jeans and jackets) 9. PAUL FRANK (a fashion and lifestyle brand known for its bright color palette and stylish products) 10. MLB (a brand for baseball clothing)
III	1. 新生儿 (newborn baby) 2. 哺乳期 (suckling period) 3. 产后 (postpartum recovery) 4. 奶粉 (milk powder) 5. 母乳 (breast milk) 6. 睡觉 (sleep) 7. 大便 (baby stool) 8. 月子 (confinement in childbirth) 9. 剖腹产 (caesarean birth) 10. 满月 (babies' completion of its first month of life)	6.0 4.6	1. 乐友 (a brand special for newborn children related products) 2. ORCHESTRA (a children clothing brand) 3. TOYS R US (a kids store for all toys, video games, dolls, etc.) 4. MOTHERCARE (a brand for selected nursery accessories) 5. BABYLAND (a brand for children indoor interactive playground) 1. SIXTY SIGHT (a women fashion brand, famous for young women's lingerie, dresses) 2. TEENIE WEENIE (a South Korean clothing brand special for the young women) 3. E-LAND (a South Korean clothing brand special for the young women) 4. 阿童木 (a women and kids clothing brand) 5. Walker One (a brand for young women's shoes)
IV	1. 郑多燕 (a South Korean diet writer, famous for her fitness exercise) 2. 柔术 (human body flexibility exercise) 3. 天猫 (<i>TMall</i> , an online shopping site in China) 4. 奶粉 (milk powder) 5. 奇热 (a website for online TV series) 6. 钟汉良 (a famous TV series actor) 7. 新浪网 (<i>Sina</i> , a Chinese online media website) 8. 慢慢 (a shopping guide website) 9. 仙途 (a netnovel) 10. 生肖 (the Chinese zodiac)	7.6 4.5	1. 巴拉巴拉 (a brand special for children related products like children's shoes, toys and clothing) 2. OSHKOSH (a well-known American brand for children's clothing) 3. 爱法贝 (a children clothing brand) 4. 法米尼 (a children clothing brand) 5. 斯乃纳 (a brand for pregnant women's clothing and care) 1. Columbia (an American outdoor sportswear brand) 2. Jack Wolfskin (a Germany outdoor sportswear brand) 3. 探路者 (a Chinese outdoor sportswear brand) 4. The North Face (an American outdoor sportswear brand, specializing in outerwear and equipment) 5. KAILAS (a brand in the manufacturing of technical climbing gear and apparel)
V	1. 武侠 (Chinese gongfu story series) 2. 文学 (literature) 3. 动漫 (comics and animations) 4. 激情 (emotional love novels) 5. 都市 (urban novels) 1. 幼儿园 (kindergartens) 2. 双眼皮 (double fold eyelid) 3. 保卫萝卜 (<i>Carrot Fantasy</i> , an funny tower defence game on the phone) 4. 玻尿酸 (Hyaluronic Acid, used to keep skin elasticity and tension) 5. 整形 (plastic surgery)	19.0 5.8	1. Dior (a French fashion brand famous for cosmetics, skin care, makeup and fragrances) 2. Chanel (a French fashion brand for perfume and bags) 3. BIOHERM (a brand of high-grade skin care products) 4. FANCL (a nature cosmetics brand from Japan) 5. DHC (a skincare brand from Japan) 6. Estee Lauder (An American high-performance skincare and makeup brand) 7. 周大福 (a Hongkong jewellery brand) 8. 安莉芳 (a Hongkong underwear brand) 9. L'Oreal (a famous French cosmetics, haircare brand) 10. Aupres (a cosmetics brand from Japan)
VI	1. 股票 (stocks) 2. 基金 (funds) 3. 投资 (investment) 4. 股份 (stock shares) 5. 证券 (securities) 6. 阿里巴巴 (<i>Alibaba</i> , the largest e-commerce company of China) 7. 金融 (finance) 8. 上市 (IPO) 9. 黄金 (gold) 10. 股价 (stock price)	9.7	1. 唐宫 (a fancy Chinese chain restaurant brand for Cantonese cuisines) 2. 熔宴 (a Chinese restaurant famous for Hunan cuisines) 3. 便宜坊 (a famous roast duck restaurant in Beijing) 4. 论道茶叶 (a Chinese tea brand) 5. 苗乡楼 (a Chinese restaurant famous for Guizhou cuisines) 6. 鹿港小镇 (a Chinese restaurant famous for Taiwan cuisines) 7. 蕉叶 (a restaurant famous for Thai food) 8. 恒香栈 (a Hongkong tea restaurant) 9. 釜山料理 (a Korean barbecue chain restaurant in China) 10. 至爱 (a Chinese restaurant for mixed cuisines)

表格 4: 具有高提升值(lift)的在线/离线话题对的示例