# Hidden semi-Markov model-based reputation management system for online to offline (O2O) e-commerce markets

Shengsheng Xiao, Ming Dong *

*Antai College of Economics & Management, Shanghai Jiao Tong University, 535 Fahua Zhen Road, Shanghai, 200052, PR China*

A B S T R A C T

The rapid development of information technology enables an increasing number of consumers to search and book products/services online first and then to consume them in brick-and-mortar stores. This new e-commerce model is called online to offline (O2O) e-commerce and has received significant managerial and academic attention. Compared with many extant e-commerce models (i.e., B2B, B2C and C2C), reputation management in this emerging model needs some improvement. It has to collect more raw reputation-related data, consider more reputation-related factors and show more comprehensive reputation evaluation results. As a stepping-stone in the research in O2O e-commerce, a new reputation management system (HSMM-RMS) has been developed based on a probabilistic model called the hidden semi-Markov model. By combining observable online and offline raw reputation information, the proposed system can accurately, promptly and dynamically provide O2O e-commerce participants with offline merchants' historical and predictive reputation information. Our Monte-Carlo simulation experiments indicate that the proposed system performs significantly better than the extant hidden Markov model-based reputation management system. A case study based on a real O2O e-commerce platform demonstrates the real application of HSMM-RMS. It also shows that the proposed system can provide a realistic solution for reputation management in the O2O e-commerce market.

© 2015 Elsevier B.V. All rights reserved.

## 1. Introduction

The rapid development of information technology enables increasing numbers of consumers to seek and book products/services online first and then consume them in brick-and-mortar stores. This emerging trend is called "online-to-offline (O2O) e-commerce" and can be exemplified by services such as Groupon, Yelp and Dianping in the restaurant business, TripAdvisor and Hipmunk in the travel business, and Zipcar and Uber in the transportation business. According to a recent report from the IIMedia Consultation Group, the market size of O2O e-commerce in China alone amounted to 98.68 billion RMB at the end of 2012 and will increase to 418.85 billion RMB by the end of 2015 [1]. Attracted by this huge profit potential, an increasing number of companies will enter into this new market. However, the latest feedback from this market shows that the flourishing prospect of O2O e-commerce largely depends on whether offline merchants can provide consumers with consistent products/services as claimed online. That is to say, offline merchants' reputation plays a vital role in the further development of O2O e-commerce. In fact, in the e-business market, merchants' reputation can help consumers reduce information asymmetry and increase their acceptance of e-commerce [2]. For many e-commerce models (e.g., B2C, C2C and B2B), reputation

management has become one of the critical factors that restrict their development [2–4], and this is also no exception to the O2O e-commerce model.

Because of the importance of reputation management, a variety of reputation management systems/models that can be used in many extant e-commerce markets have been developed during the past several years [5–8]. However, reputation management in the emerging O2O e-commerce needs some improvements. First, in the O2O e-commerce market, reputation management has to address participants from both online and offline settings (e.g., online consumers and brick-and-mortar restaurants on Yelp). To provide precise reputation assessment results, the reputation management in this scenario has to collect both online and offline raw reputation-related information,[1] whereas extant reputation management systems/models mainly focus on the information collected from either an online or offline setting alone. Second, reputation management in O2O e-commerce has to take the fluctuant demand information and offline merchants' real service capabilities into account when evaluating their reputations. Different from these traditional e-commerce models, customers in the O2O e-commerce

---

* Corresponding author. Tel./fax: +86 21 52301193.
   *E-mail address:* mdong@sjtu.edu.cn (M. Dong).

[1] For example, in order to provide precise reputation evaluation results for restaurants on Yelp, reputation information managers should use not only the online information such as online reviews but also some offline information about brick-and-mortar restaurants (e.g., the restaurants' reputation-related records such as real regulatory compliance records, industry standard compliance records and health inspection records).

market consume the online-reserved products/services in brick-and-mortar stores. Each offline merchant has limited service capacity; if the number of consumers who visit a merchant exceeds its maximum capacity in a certain period of time, its product/service quality may be affected. This implies that an offline merchant may not be able to provide consumers with products/services that are consistent with those claimed online. As a result, to check offline merchants' overload operation behavior, reputation management in the O2O e-commerce should consider their fluctuant demands and their real service capacities. Third, the predictive reputation information about offline merchants can help online consumers choose appropriate product/service providers and determine their consumption time in advance. The extraction of predictive reputation information, while feasible in the O2O scenario, is not considered by the extant reputation management systems.

To leverage these opportunities, we propose a new reputation management system (HSMM-RMS) based on a probabilistic model called the hidden semi-Markov model. In this system, offline merchants are modeled by HSMMs (i.e., HSMM with different parameter values). Their observable reputation-related information and hidden reputation states are denoted by observation variables and latent states in HSMMs. By collecting and processing raw reputation-related data from both online and offline settings, offline merchants' hidden reputation sequence $Q^*$ and the "optimal" reputation state structure HSMM $\lambda^*$ (which perfectly describes offline merchants' reputation behavior given the observable data) will be quickly derived and continually updated through a modified forward and backward algorithm. The proposed system is able to improve reputation management in the O2O e-commerce.

The main contributions of our study are threefold: First, we extend traditional reputation management studies that use online reputation-related information alone. We explore reputation management issues in the emerging O2O e-commerce setting by combining reputation-related information from online and offline channels. Second, we establish a systematic framework for a reputation management system in the O2O setting and illustrate how the raw information can be acquired, processed and distributed within this system. Third, our study also enriches extant reputation management models from the methodological perspective. We develop a hidden semi-Markov model to describe offline merchants' reputation behavior. It relaxes many extant assumptions on reputation behavior and performs better than the HMM-based reputation management model.

The rest of this paper is organized as follows: Section 2 reviews the related literature. Section 3 provides a detailed description of our HSMM-based reputation management system (HSMM-RMS) and derives the HSMM-based reputation management model. Section 4 evaluates the performance of our model through Monte-Carlo simulations. In Section 5, a case study based on a real O2O e-commerce platform is conducted to illustrate the application of HSMM-RMS. Section 6 draws conclusions and provides future research directions.

## 2. Literature review

In this section, relevant reputation management studies are reviewed, and some probabilistic reputation management models are summarized.

### 2.1. Reputation management-related studies

There has been a steady stream of studies that address reputation management problems over the past several decades. By using a classification method similar to that of Lee et al. [9], we broadly divide the extant studies into four types (see Fig. 1). Works of the first type focus on offline reputation management issues alone, and they do not directly discuss e-commerce-related reputation management problems. The second research type mainly focuses on reputation management
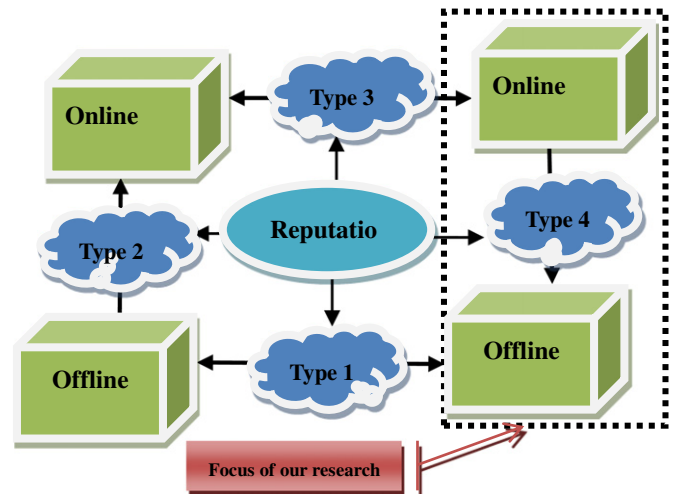


**Fig. 1.** Four types of reputation management research.

problems from offline to online channels. Studies of this type can be summarized by the following questions: Can offline merchants' trust and reputation be successfully transferred from offline to online channels? What factors will affect this transformation? What should be done to ensure this transformation [9–14]? Different from the first two research types, the third research type only focuses on online reputation-related problems; for example, What are the antecedents of online reputation management [12,13]? What are the constructs of online reputation [14,15]? What are the consequences of online reputation management [16]? In addition, it also discusses some extant reputation management systems used by *eBay* [17,18], *Amazon* [19,20] and *Alibaba* [21]. More information about this research type can be found in [3,4,6,22,23]. Research type 4 in Fig. 1 mainly concentrates on new reputation management problems that stem from online to offline channels. To the best of our knowledge, the majority of extant studies of this type focus on the following issues: Can successful online companies transfer their reputation from online to offline settings? How do successful online companies transfer their trust/reputation? Studies on how to develop a system to improve reputation management in the emerging O2O e-commerce setting are limited. As a result, one of the aims of this paper is to fill this gap.

### 2.2. Probabilistic models used in reputation management systems

Although researchers have introduced a variety of trust/reputation management models [5,6,24,25], we only focus on probabilistic reputation management models in this study.

#### 2.2.1. Belief-based reputation management model

Based on the belief and probability theory, this model assumes that the sum of probabilities over all possible outcomes is not necessarily equal to one [6]. In this model, an agent $A$'s personal perceived value on believing in agent $B$ is usually expressed by a quadruple belief metric $(b, d, u, a)$ and can be calculated by $(b + a \cdot u)$. (Here, parameters $b$, $d$ and $u$ reflect agent $A$'s belief, disbelief, and uncertainty value toward agent $B$, respectively. Parameter $a$ is the relative atomicity that represents the base rate probability in the absence of evidence [26,27]). If the belief values of all the participants are aggregated, the reputation score of each participant in this model can be obtained [28]. This model has been extended and widely used in the literature since it was systematically introduced by Falcone et al. [29–32].

The participants in this model are always assumed to be homogeneous. They have similar interactions with each other and use the same metric to evaluate their belief values. However, in many real e-commerce markets (including the O2O market), there exist many