Expert Systems With Applications 58 (2016) 36–56

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

Efficient task assignment for spatial crowdsourcing: A combinatorial fractional optimization approach with semi-bandit learning

Umair ul Hassan*, Edward Curry

The Insight Centre for Data Analytics, National University of Ireland Galway, The DERI Building, IDA Business Park, Lower Dangan, Galway, Ireland

ARTICLE INFO

Article history:
Received 29 June 2015
Revised 10 March 2016
Accepted 10 March 2016
Available online 1 April 2016

Keywords:
Spatial crowdsourcing
Task assignment
Combinatorial fractional programming
Multi-armed bandit

ABSTRACT

Spatial crowdsourcing has emerged as a new paradigm for solving problems in the physical world with the help of human workers. A major challenge in spatial crowdsourcing is to assign reliable workers to nearby tasks. The goal of such task assignment process is to maximize the task completion in the face of uncertainty. This process is further complicated when tasks arrivals are dynamic and worker reliability is unknown. Recent research proposals have tried to address the challenge of dynamic task assignment. Yet the majority of the proposals do not consider the dynamism of tasks and workers. They also make the unrealistic assumptions of known deterministic or probabilistic workers' reliabilities. In this paper, we propose a novel approach for dynamic task assignment in spatial crowdsourcing. The proposed approach combines bi-objective optimization with combinatorial multi-armed bandits. We formulate an online optimization problem to maximize task reliability and minimize travel costs in spatial crowdsourcing. We propose the *distance-reliability ratio* (DRR) algorithm based on a combinatorial fractional programming approach. The DRR algorithm reduces travel costs by 80% while maximizing reliability when compared to existing algorithms. We extend the DRR algorithm for the scenario when worker reliabilities are unknown. We propose a novel algorithm (DRR-UCB) that uses an *interval estimation* heuristic to approximate worker reliabilities. Experimental results demonstrate that the DRR-UCB achieves high reliability in the face of uncertainty. The proposed approach is particularly suited for real-life dynamic spatial crowdsourcing scenarios. This approach is generalizable to the similar problems in other areas in expert systems. First, it encompasses online assignment problems when the objective function is a ratio of two linear functions. Second, it considers situations when intelligent and repeated assignment decisions are needed under uncertainty.

1. Introduction

Crowdsourcing systems have emerged as a new form of multi-agent systems where three types of agents interact with each other, namely requesters, workers, and platform. The requesters submit tasks on platform that are performed by workers. The platform serves as a mediator that provides appropriate assignment of tasks to workers to maximize the utility of crowd work (Deng, Shahabi, & Demiryurek, 2013; Hassan & Curry, 2014). Crowdsourcing has been applied in variety domains, such as machine learning (Chen, Lin, & Zhou, 2013b), natural language processing (Tarasov, Delany, & Mac Namee, 2014), and mobile-based sensing (Ganti, Ye, & Lei, 2011). Spatial crowdsourcing is a form of crowdsourcing that employs crowd workers for performing tasks in the physical world at various locations (Cheng et al., 2015; Kazemi and Shahabi, 2012; To, Ghinita, and Shahabi, 2014; To, Shahabi, & Kazemi, 2015). The use of spatial crowdsourcing is further illustrated with the help of the following scenario:

Consider the situation, where a requester is interested in collecting high quality and representative photos of disaster hit areas in a country. The locations of interest are spread across the country. The requester designs a task for each location i.e. spatial crowdsourcing task. The requester is interested in the coverage of all locations with high quality results. Fig. 1 illustrates such a scenario on a map.

Most of the existing crowdsourcing platforms serve as a marketplace with rudimentary task assignment functionality, such as Amazon Mechanical Turk, TaskRabbit, and ClickWorker (Horton & Chilton, 2010; Musthag & Ganesan, 2013). In fact, these platforms largely rely on workers to assign tasks to themselves when visiting the platform. Tasks may not be assigned to appropriate workers in this manual approach, also known as *worker selected tasks*

* Corresponding author. Tel.: +353 85 729 8904.
E-mail addresses: umair.ulhassan@insight-centre.org (U. ul Hassan), edward.curry@insight-centre.org (E. Curry).

http://dx.doi.org/10.1016/j.eswa.2016.03.022
0957-4174/

Fig. 1. Example of spatial crowdsourcing on the map of Haiti after the 2010 earth quake. A new spatial task (in blue) requests recent photos of a building at the indicated location.

(WST), due to search friction issues (Kittur et al., 2013; Kulkarni et al., 2012). Recent research has focused on developing algorithmic approaches to task assignment with the aim of addressing the limitations of the WST approach. The algorithmic approach, also known as the *server assigned tasks* (SAT) approach, formulates the dynamic task assignment as a sequential decision making problem. The sequential decision making involves matching dynamically arriving tasks with dynamically arriving workers over time (Abraham, Alonso, Kandylas, and Slivkins, 2013; Hassan and Curry, 2014; Ho, Jabbari, & Vaughan, 2013).

In this paper, we focus our attention to the SAT-based dynamic task assignment problem in spatial crowdsourcing. Specifically, we introduce and formulate the *minimum-cost maximum reliability assignment* (MC-MRA) problem in spatial crowdsourcing. The MC-MRA problem aims to maximize reliability and minimize the travel costs of spatial tasks. Existing literature on spatial crowdsourcing generally assumes deterministic settings for task assignment i.e. each assignment is assumed to result in task completion with high quality (Deng et al., 2013; Kazemi and Shahabi, 2012; To et al., 2014, 2015). By comparison, the MC-MRA problem considers stochastic settings where the reliability of an assignment is defined in terms of the probability that spatial task will be completed with high quality by the assigned worker (Cheng et al., 2015). Assuming that the reliabilities of different assignment choices are known, we reduce the MC-MRA problem to *minimum-cost maximum weight bipartite matching* problem and adapt two existing approaches to address the reduced problem (To et al., 2015). Our experimental evaluation shows that the adapted approaches do not jointly optimize reliability and travel costs; therefore, resulting higher travel costs under various conditions. To address the limitations of existing approaches, we propose a novel approach based on *combinatorial fractional programming*. The proposed approach aims to dynamically assign tasks such that both the reliability of assignments is maximized and the travel costs are minimized, for all spatial tasks.

Further, we relax the assumption of known reliabilities and consider the MC-MRA problem with online learning. Online learning necessitates the dynamic estimation of worker reliabilities based on the observed outcomes of task assignments. A fundamental challenge of dynamic assignment with estimated reliabilities is to address the dilemma of learning versus optimization, also known as the *exploration-exploitation* trade-off (Barto, 1998) in literature. *Exploration* involves choosing workers for the purpose of learning their reliability. *Exploitation* entails using the gained knowledge to optimize the assignment objective. Existing research works have employed primal-dual (Ho et al., 2013) and multi-armed bandit (Abraham et al., 2013; Chen et al., 2013b; Tran-Thanh, Stein, Rogers, & Jennings, 2014) techniques to address the trade-off. These techniques have primarily focused on the non-spatial crowdsourcing scenarios and use simple cost constraints (Slivkins & Vaughan, 2013). Instead, we address the exploration-exploitation trade-off for the MC-MRA problem that optimizes both the reliability and the travel costs in spatial crowdsourcing. The specific research contributions of this article are summarized below:

- We introduce and formalize the MC-MRA problem based on the SAT-based spatial crowdsourcing. We reduce the MC-MRA problem to the *minimum-cost maximum weight bipartite matching* problem and adapt two existing approaches to address the reduced problem.
- We propose a novel *distance-reliability ratio* (DRR) approach for the MC-MRA problem, that is based on combinatorial fractional programming. The DRR approach employs Newton's method to transform the fractional assignment problem to an equivalent parameterized linear assignment problem.
- We extend the DRR approach, for adaptive assignment, to enable the estimation of worker reliabilities from observed outcomes of previous task assignments. We propose two adaptive DRR algorithms based on combinatorial multi-armed bandit model with semi-bandit learning. The DRR-GRD is inspired by the *greedy exploration* approach and the and DRR-UCB algorithm follow an *interval estimation* approach.
- We extensively evaluate the performance of proposed algorithms against adapted algorithms on synthetic and real-world datasets. The performance results establish the effectiveness of the DRR algorithm and its variants in terms of reliability and travel costs. The results also establish the effectiveness of adaptive DRR algorithms in terms of estimating worker reliabilities.

This paper extends the research on experts systems in two key ways. First, it considers the online algorithms that aim to optimize a bi-objective objective function. One set of optimization variables are probabilistic and the second are deterministic. Second, it combines semi-bandit learning with the bi-objective optimization. Semi-bandit learning assumes that the probabilistic variables are unknown and must be approximated by observing the outcomes of optimization decisions.

The rest of this article is organized as follows. In Section 2, we summarize the related research on the assignment problem in spatial crowdsourcing. We highlight the research gaps in existing literature on spatial crowdsourcing and its related topics. Section 3 provides the necessary definitions of concepts in SAT-based spatial crowdsourcing and Section 3.2 describes the basic *maximum reliability assignment* (MRA) problem. In Section 4, we introduce the *minimum-cost maximum reliability assignment* (MC-MRA) problem and present our proposed DRR approach for efficient solution to the MC-MRA problem. Section 5 extends the proposed DRR approach with online learning based algorithms. We evaluate our proposed algorithms using real-world and synthetic datasets in Section 6. Section 7 discusses the implications of the proposed approaches and their performance results. We conclude the paper in Section 8 and layout plans for future work.

2. Related work

Spatial crowdsourcing is distinguished from other forms of crowdsourcing by the fact that workers are required to visit locations in the physical world to perform tasks. One primary challenge of spatial crowdsourcing is matching tasks with appropriate workers on the ground (Kazemi and Shahabi, 2012; To et al., 2014, 2015). Kazemi and Shahabi proposed a taxonomy of spatial crowdsourcing that highlights two modes of task assignment: