Neurocomputing 000 (2018) 1–15

Contents lists available at ScienceDirect

Neurocomputing

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

Bayesian filter based on the wisdom of crowds

Behzad Bakhtiari, Hadi Sadoghi Yazdi*

Department of Computer Engineering, Ferdowsi University of Mashhad, Mashhad, Iran

A R T I C L E I N F O

Article history:
Received 18 September 2016
Revised 30 November 2017
Accepted 21 December 2017
Available online xxx

Communicated by Prof. Zidong Wang

Keyword:
Bayesian filter
Crowdsourcing
Dynamic model
Time series data
Truth discovery
Wisdom of crowds

A B S T R A C T

Nowadays, the wisdom of crowds aids in data labeling via crowd sensors. One of the successful tools worth mentioning is Amazon Mechanical Turk. Uncertainty in crowd labels, however, deteriorates the result of the learning algorithm. In some applications, such as weather and stock forecasting and object tracking, the temporal dependency of data (dynamics) is effective in decreasing label uncertainty. In order to benefit from the existing knowledge in label dynamics, the current study first employs a traditional state-space model and it shows that these models have serious drawbacks, for instance, sensors with a low coverage rate and the existence of a random labeler are the main challenges posed in this process. Then, an appropriate dynamic model for crowd sensors is presented and the Bayesian filter is applied so that true label inference and system parameter learning are performed jointly. The present work will show that the proposed method is robust enough to meet these challenges and performs better in comparison to the existing methods. The results of experiments on synthetic and real data confirm this issue.

1. Introduction

In recent years, using the wisdom of crowds in machine learning for the purposes of data labeling has been a great success in comparison to traditional methods. In this way, it is recognized that not all the people in the crowd are experts, but that they can have knowledge and skill pertinent to some different parts of the problem. In general, each human in a crowd can be a sensor that observes a special system (such as temperature in weather forecasting system), that is here called a crowd sensor. These sensors can be human or information resources related to humans, such as cellphone sensors, sites, etc. In this situation, the assumptions of crowd sensors are different than to traditional sensors. Some of the assumptions of crowd sensors are as follows:

- Each sensor's operation is unknown, such as the observation function or its precision.
- Sensor knowledge is not perfect and so their data may be unreliable.
- There are malicious and adversarial sensors that give the wrong information.
- There are random labelers (RL) that do not provide useful information.
- Sensors can present observation for limited data (or time), which is called a coverage rate.

Due to the differences that exist between traditional and crowd sensors, however, a lot of effort has been made to understand sensor operation and true label inference. While the majority of these efforts are in classification and regression fields, there has unfortunately been no serious work on Time Series Data (TSD) [1]. Despite, the existence of myriad of studies on time dependent data, the temporal information of labels has not yet been taken into account by these works, even though it is known that there is knowledge in the temporal dependency (dynamics) of labels that could reduce crowd label uncertainty.

On the other hand, there are numerous methods for processing TSD. The State-Space Model (SSM) and Bayesian filter are the most important of these methods, but the direct use of traditional TSD processing methods for crowdsourced TSD (CTSD) employing crowd sensors is not appropriate, because of the differences between the crowd and traditional sensor assumptions. Therefore, the present paper will present a dynamic model for crowd sensors and introduce a useful tool for CTSD processing. The contributions made by this work can be summarized as follows:

- Consideration of dynamics: Some previous works in the crowdsourcing field jointly estimated precision and the true label (Fig. 1(a)). In another group of works, furthermore, the knowledge of the dependency of features and labels is used to estimate precision and the true label in a supervised manner (Fig. 1(b)). In the current study, the temporal dependency of the unknown true label will be taken into consideration in the inference and parameter learning procedure (Fig. 2).

* Corresponding author.
E-mail addresses: bakhtiari@mail.um.ac.ir (B. Bakhtiari), h-sadoghi@um.ac.ir (H.S. Yazdi).

https://doi.org/10.1016/j.neucom.2017.12.037
0925-2312/

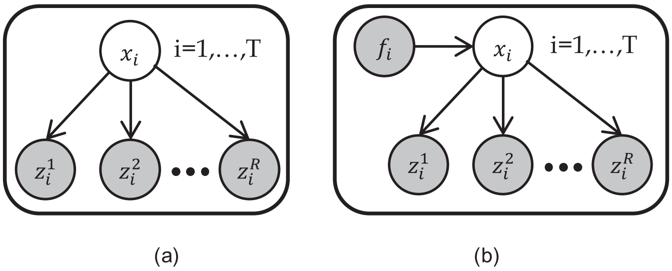

Fig. 1. Comprehensive view for GT inference methods from crowd, (shaded nodes are observations) (a) DS model (b) DS model plus dependency between features and GT, where x_i is the GT (true label), z_i^r is observation of the rth sensor and T is the total number of instances. In these models, labels are independent and each observation can be missed (in these figures, missing observations are not shown).

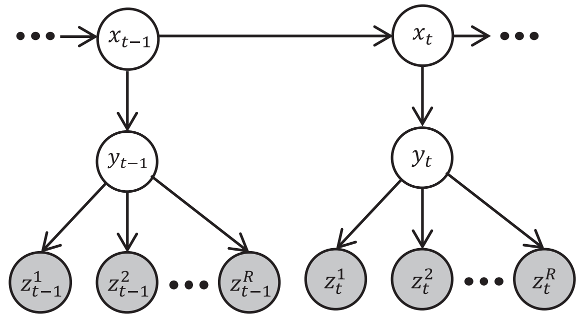

Fig. 3. Dynamic model based on the wisdom of crowds (Shaded nodes are observations. Note that in this figure present/absent of sensors are not shown).

- Dynamic model for crowd sensors: It will be indicated that, in the field of crowdsourcing, the newly introduced problem (Fig. 2) is similar to 'joint inference and learning' in SSM into which much effort has been put into solving. Despite this similarity, these works have not been used in the crowdsourcing field.
- Dynamic model based on the wisdom of crowds: Many problems are posed when using the existing dynamic model for crowd sensors in real crowdsourcing scenarios. To resolve these problems, a dynamic model based on the wisdom of crowds is presented (Fig. 3) along with a suitable Bayesian filter designed for it. This new filter is consistent with crowd sensor assumptions and can be employed for CTSD processing.

For unification purposes, the term 'crowd sensor' (or sensor) is utilized for a set of resources that provide information and the term 'observation' for the information given. In addition, the term 'Ground Truth' (GT) is used instead of true label. In Section 2, related work is investigated. Section 3 presents the proposed method. Section 4 discusses experimental results and Section 5 will be the conclusion.

2. Related work

As the current paper deals with crowdsourcing and TSD processing tools, the related work of each of these is reviewed. First, the previous works on GT inference from crowd observations will be given a comprehensive look. Here, GT modeling and its inference are reviewed without their incidental aspects. For this purpose, all works have been divided into three categories:

Category 1: Methods of this category are baseline methods. Each GT is only estimated by the direct crowd observations of itself without any consideration of other crowd issues. Majority Voting (MV) and Mean are instances of this category for discrete and continuous data, respectively.

Category 2: These methods attempt to estimate jointly the parameters of crowd sensors (such as, sensor precision) and the GT. Numerous methods with different details fall into this category. One of the most well-known (and the oldest) is the work by Dawid and Skene (DS) [2]. In the DS model, precision is considered for each sensor. By employing the Expectation–Maximization (EM) algorithm, the precision of sensors and the GT are estimated jointly. Fig. 1(a) illustrates the DS model where x is the GT and z^r is the observation of the rth sensor. Since this algorithm is sensitive to the initial value, Zhang et al. [3] presents a method to determine the initial value for the EM algorithm in the DS model. Also, because the GT in the DS model is multivalued, Zhao and Han [4] presents a similar method for a continuous GT.

There are many other works like [5–15]. Although different in distinct aspects, from a comprehensive view and focus on the GT model, all works are direct extensions of the DS model that can be interpreted as a weighted MV.

Category 3: In addition to the previous category's assumptions, Category 3 adds some dependency of data to the problem. The knowledge of this dependency helps to more efficiently infer the GT. Jim and Ghahramani [16] present method for determining the dependency between data and the unknown label in the multiple label problem so that the sensor's ID is unknown and it is assumed that just one of the labels is true. For the first time, in the classification and regression problem, Raykar et al. [17] could estimate the precision of sensors, the GT, and the dependency between the GT and features jointly.

In general, the nature of knowledge inference from crowds in the methods that fall into this category is similar to the model in Fig. 1(b). In this model, f is the features. For example, in the classification problem, f is the feature of instance and x is its label. There are many methods that consider distinct aspect, such as [18–21].

In none of the papers mentioned in these three categories is the temporal dependency of GT considered. While in the TSD problem, the dynamics of the GT can be added into GT inference and the parameter learning procedure, based on our knowledge and [1] there are a few works on this issue. For instance, in [22], with the given GT correlations, the GT and precision of the sensors are jointly estimated in consideration of GT correlations. In [23], a tool is presented for temporal alignment for sequence labeling by multiple annotation settings. In this work, GT, the precision of sensors and the temporal dependency between observation frames are jointly estimated. Also, Wang and Yeung [24] presents

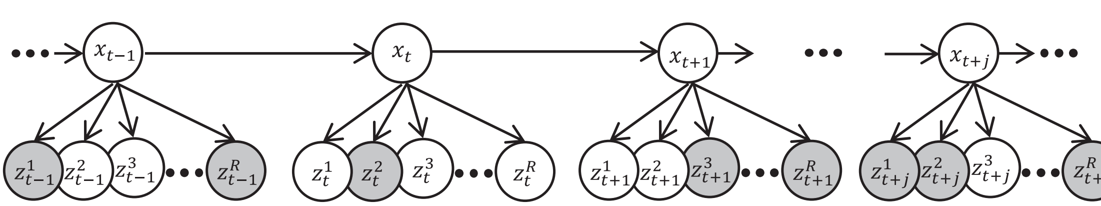

Fig. 2. Dynamic model with crowd sensors, where x_t is the GT (true label), z_t^r is observation of the rth sensor and T is the total number of time indexes. Here labels are dependent and this model incorporates temporal dependence of labels. (Shaded nodes are observations.)

Please cite this article as: B. Bakhtiari, H.S. Yazdi, Bayesian filter based on the wisdom of crowds, Neurocomputing (2018), https://doi.org/10.1016/j.neucom.2017.12.037