Ad Hoc Networks 35 (2015) 127–136

Contents lists available at ScienceDirect

Ad Hoc Networks

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

Intelligent photovoltaic monitoring based on solar irradiance big data and wireless sensor networks

Tao Hu [a], Minghui Zheng [a,*], Jianjun Tan [a], Li Zhu [a], Wang Miao [b]

[a] School of Information Engineering, Hubei University for Nationalities, China
[b] Department of Mathematics and Computer Science, College of Engineering, Mathematics and Physical Sciences, University of Exeter, Exeter EX4 4QF, United Kingdom

ARTICLE INFO

Article history:
Received 9 April 2015
Revised 4 June 2015
Accepted 9 July 2015
Available online 26 July 2015

Keywords:
Multi-data fusion
Semi-supervised SVM
Cloud storage
Residual value
Password-based group key agreement

ABSTRACT

Clean energy technologies, especially photovoltaic, have recently become more and more popular and important due to their substantial benefits for environment, economy, and energy security. How to improve the management and usage efficiency of photovoltaic power stations is a challenging problem that needs to be investigated deeply. In this paper, Wireless sensor networks (WSNs) are utilized to efficiently deliver the monitoring data of the photovoltaic (PV) modules from power stations to the monitoring center located in Cloud datacenter. With the aim of detecting the problems of PV modules from the monitoring big data, a two-class data fusion method is firstly developed to integrate the monitoring data at sensor nodes of WSNs; then an innovative semi-supervised Support vector machine (SVM) classifier is designed and trained by existing solar irradiance big data at the monitor center. With the prediction model provided by the trained classifier, an outlier detection algorithm is devised to classify and locate the problems of PV modules through calculating the average value of the questionable data. In order to evaluate the performance of the proposed methods, a comprehensive experimental platform is set up. The experimental results show that the predicted values match well with the theoretical value of power generation.

1. Introduction

Solar energy, as one of the most important clean energy, has attracted more and more attentions from the society and industry. As a result, a large number of photovoltaic (PV) power stations have been built and are being built in the world. While how to manage these numerous PV power stations and keep the healthy state of these power stations by leveraging of information techniques has been a challenging issue that needs to be solved for the success of the solar energy. With the increasingly rapid combination of PV power stations and power grid, the users and administrators of PV stations are urgently forced to access the power data, which is critical for real-time operation and maintenance. For instance, if there are some failures in PV station operation, power station operators need to quickly access the failure information, accurately locate the position of the failure components and take corresponding actions to recover the faults. The core component of PV station is PV module, which is responsible for converting solar energy into electrical energy based on PV effect. This process is the fundament for a solar cell converting sunlight into electricity and the larger number of PV stations means the higher capacity of power that can be provided. Therefore, PV modules are regarded as the key component in PV power stations.

In this paper, a framework of PV power stations monitor system will be firstly introduced, which can access all key data of each PV power station and locate the questionable PV

* Corresponding author. Tel.: +86 18972401006; fax: +86 07188437881.
E-mail address: mhzheng3@163.com (M. Zheng).

http://dx.doi.org/10.1016/j.adhoc.2015.07.004
1570-8705/

module in real time. Furthermore, this system will be able to predict the power generation/utilization of PV power station and ensure the security of data transmission and authority. In order to achieve this system, this paper will focus on dealing with the following four problems:

Data transmission. In PV power station, a huge amount of monitoring data is captured by various sensors at the same time. Therefore, the efficiency of data transmission is a very important problem in the envisioned system, which requires a unified data format to smoothly deliver data from different PV power stations to a PV monitor center.

Abnormal data detection. Abnormal data at the PV monitor center implies that there are problems in PV modules. So an efficiency algorithm of fault detection is strongly required by this system to not only find the questionable PV module in real time, but more importantly to identify fault alarms of abnormal data.

Power prediction. Power is the core indicator of PV power station. The envisioned system will investigate the problem of predicting the power capacity of a new power station based the station scale and natural conditions.

Data security. Date security involves two types of security issues: one is the security of data transmission from PV power station to monitor center; the other is authority strategy to rule out the un-authorized users to access power station. Furthermore, transmission channel between monitor center and user side application must be secured.

1.1. Related works

In this paper, we propose a solution for distributed PV plant monitor center based on Cloud and WSNs. The monitor center collects the real-time key data of each photovoltaic module including current, voltage, power and light intensity of sunshine. These data are processed by multi-level fusion, transported to monitor center and stored in the private Cloud of monitor center. The monitor center detects the outlier of solar power by leveraging semi-supervised SVM classifier [1]. In this system, users are able to monitor the status of the power stations with their individual privileges. Furthermore, the developed system can innovatively predict the generating capacity of each station by analyzing the existing sunlight data. Finally a password-based group key agreement (PGKA) protocol [2] is employed to ensure the security of data transmission between user's application and the private Cloud.

Using WSNs to monitor the environment and machine is a hot research field [3]. An environmental monitoring framework was developed in [4] based on WSNs, which provided a high quality service, optimal solar energy harvesting, storage and energy awareness. The work in [5] proposed a water quality monitoring system based on WSNs and solar power supply. The current intelligent transportation system relies on WSNs to enhance transportation efficiency and safety [6,7].

Typically, the work in [8] proposed a PV panels monitoring system based on a WSNs, focusing on the design of sensor nodes, paying little attention to the global system design. Some other literatures can be found in [9,10] for designing of sensors and power supply.

Since 2008, the analysis and application of big data have been significant development [11]. It is driven by cloud computing and internet of things especially [12]. And now, it is applied in some professional fields, such as new energy and intelligent transportation. Cloud storage is one of the critical functions of Infrastructure as a Service (IaaS) in Cloud computing architecture [13]. Multiple storage medium were combined to form a pool of storage resources, which shielded the detail of storage hardware configuration, distributed processing, disaster recovery and backup from upper layers [14]. The work [15] is a typical prediction research to extract community activity patterns based on big data. Data security is a key component in Cloud storage [16] and received a lots of research investigations, such as privacy-preserving public auditing [17], toward publicly auditable service [18], two-factor authentication schemes in WSNs [19], etc.

The traditional method in outlier detection is a main trend. The outlier detection based on calculating the distances among objects in the data, such as K-nearest neighbor (K-NN) [20]. The outlier detection based on density estimates the density distribution of the input space and then identifies outliers based on low density [21].

Table 1
Meaning of notations in equations.

Notation	Meaning in equations
$C_i/V_i/P_i$	The current/voltage/power of one PV module, which number is i
sd_i	The fusion data of i PV module
SSD_j	The fusion data of j PV station
d	The outlier detection data set
$\overline{d}$	The mean value of d
α	The controlling parameter of outlier detection
$\partial_{\overline{d}}$	The standard deviation estimation between d and $\overline{d}$
g	The length of d
$f(d)$	The objective equation of outlier detection with data set d
$F(i,j)$	The state of PV module i in PV station j
γ	The detection threshold of monitor data in PV station

1.2. Article structure

The rest of this paper is organized as follows: Section 2 introduces the framework of the whole system and the data flow in this framework. The technique details of data fusion in WSNs, outlier detection, power prediction and data security are investigated in Section 3. Section 4 will build up the test environment in order to validate the performance of the proposed algorithms. Finally, Section 5 concludes this paper. In order to describe the kernel algorithms in our paper better, the Table 1 is employed to explain the meaning of notations.

2. System overview

In this paper, we propose a solution to monitor the PV power stations based on Cloud and WSNs as shown in Fig. 1. In this system, WSNs is employed to deliver the key data of each PV power station to monitor center, and Cloud datacenter is used to store these valuable data. In order to improve the convenience of power station management, several applications for users are developed within this system. This system consists of three components as follows: