Estimate Power generation of invisible solar site using State estimation

1st Pornchai Chaweewat
EECC
AIT)
Pathumthani, Thailand
chaweewat.p@gmail.com

2nd Weerakorn Ongsakul
EECC
AIT)
Pathumthani, Thailand
email address

3rd Jai Govind Singh
EECC
AIT)
Pathumthani, Thailand
email address

4th Ali abur
EEC
NEU
Boston, MA, USA
email address

***Abstract*—The number of large scale and rooftop scale solar photovoltaic (PV) systems in electricity grid is growing at exponential rate. Invisible solar PV refers to small-scale and rooftop solar sites in distribution system that are not mornitored by utilities. These invisible solar PV sites cause challenging to utilies and system operators. In this paper, a methodology is proposed to identify location and estimate installation capacity of solar PV sites. With measurement voltage at any buses and power flow in any lines, state estimation (SE) algorithm could find power consumption in any buses. The consumption are fed to find correlation with solar irradaiton data. Thus, the correlative between these data could locate the site of invisible solar PV sites. Futhermore, the change point method and proposed algorithm could estimate installation capacity of invisible solar PV sites. The proposed methodology is test in distribution system i.e, four bus system and CIGRE Medium voltage distribution network with PV and Wind DER with randomly invsisble solar PV sites location. The overall results of the study shows that the correlation between consumption and solar irradiation can perfectly identify invisible solar PV sites with F_1 score aboved 0.96 and Matthews correlation coefficient (MCC) score aboved 0.90. The installation capacity estimation performe good with mean absolute error (MAPE) below 17%.**

***Index Terms*—Solar photovoltaic, invisible solar PV site, Gaussian noise, state estimation, correlation, change point detection, Matthews correlation coefficient**

I. Introduction

Sixty years ago, the price of solar panels was astronomical. The price was over 1,000 US dollars per watt in today's money with 1 percent efficiency [1]. As the cost continues to drop, solar panel systems are becoming much more viable for the average household. Today's price is less than 2.8 US dollars per watt for solar panel system including solar module, inverter, wiring hardware, labor cost, and maintenance cost [2]. As result of exponentially acceralate increasing in number of globally solar PV integration, especailly in residential sector.

Recently, the highest penetration of solar PV, recognized a large number of unauthorized solar PV installation [3]. Unauthorized solar PV installation creates safety risks and lack of visibility may result in incorrect planning, and operation, including over-voltage, back-feeding. Futhermore, in worst case scenario, it may damange system equipment such as transformers, voltage regulators, as well as customer applicants [4], [5]. In previous studies, there are various reasons for unauthorized or incorrectly registered PV system For example, owner decided not to apply for a permit to avoid fees [6]. The difference rules depending on size and type of PV installation can make the owners believe they do not need a permit [7].

Futhermore, many researchers have studied the impacts and risks of PV on distribution systems [8]- [11]. These impacts can cause physical and financial damage to utility and solar PV owners, especially invisible solar PV which utility could not mornitor real-time. However, the detection and estimation of residential PV systems has not been the focus of the studies and related research. This issue is key motivation to this study.

The rest of the paper is structured as follows: Section II reviews related research on state estimation application for power system and previous detection and estimation techniques on invisible solar PV site. Section III formulates the mathenatic model of measurement data as well as the structure of proposed methods as well as evaluation processes. Section IV shows the results of proposed method on test cases. Section V concluedes the paper and dicusses future research oppotunities on invisible PV system detection and estimation.

PEA, AIT, NEU

II. Literature reviews

A. Applications of State estimation in power system

State estimation on distribution power system is reviewed in [12].

State estimation provides the optimum estimate of power system state based on received measurement units and the knowledge of network modeling. The measurement units may include as follows:

- power injection (real/reactive),
- power flow (real/reactive),
- bus voltage magnitude,
- phase angle,
- line current magnitude,
- current injection magnitude.

State Estimation provides various benefits for power system analysis as follows:

1) Real time mornitoring: For power system in transmission and distribution levels [13]. The Advance Meter infrastructure (AMI) data are fed into SE to real-time mornitor distribution power system in [14]. It improved the accuracy comparing to pseudo measurment based on historical load.

2) Load allocation: In distribution power system, SE assigns active and reactive power values to bus downstream [15]. Class curves and individual average demand of clients improved accuracy of load allocation [16]. The data from AMR system, which are used as the pseudo-measurements, was used to estimate load on the distribution transformers. [13].

3) Bad data analysis: In several real application, it happens that some flow measurement has no sign or the sign is wrong. A re-weighting technique was used to reduced the weights association to these measurements [15]. It also could helps to treat of power flow with erroneous sign. Futhurmore, SE was implemented in cyber attack events [17] on power system states. A detection false data injection method is proposed in [18].

4) Topology indentification: the status of branches and switches are indentified in [15].

B. previous solar PV detection techniques

The previous studies has been done on detection of location and estimation of power generation of invisible solar PV sites. Reference [19] proposed a change-point detection algorithm for a time series for residential solar PV detection. These method required smart meter data and historical load profile. The uncertainty of solar PV site and the data generated by a small set of selected representative sites were taken into account to estimates the power gneneration of known solar PV site in [20]. The new hybrid k-means and PCA techniques provided good accuracy of estimating invisible solar PV sites. The big data characterization of smart grids and two-layer dynamuc optimal synchrophasor measurement devices selection algorithm for fault detection, identification, and causal impact analysis was proposed in [21]. The utility data based on random matrix therory (RMT) are used in detection and estimation of invisible solar PV in real case in China in [22].

C. Contributions

Many researchers have studied the detection and estimation location of invisible solar PV installation [19]- [22]. These study depended on smart meter data, data set of solar PV generation, as well as historical demand of customer in each load. This paper proposed method which identify location using state estimation, correlation measured and estimating the installation capacity using change point method of invisible solar PV site. All solar PV's site is assumed to operate at unity power factor. The solar PV is located randomly at any bus excepted slack bus. The measurement units are located at optimum locations. The measured data such as solar irradiation, power flow measurements in lines and voltage measurements at busses are generated with gaussian noise. The proposed method consists of four steps:

- collect solar irradiation, measurement data
- allocate power demend in each buses, nodes.
- indentify invisible solar PV site
- estimate solar PV site's installation capacity

III. Problem Formulation

This section will presents the overall methodology of the study which will be devided into four parts. The conceptual of methodology is illustrated in Fig. 1.

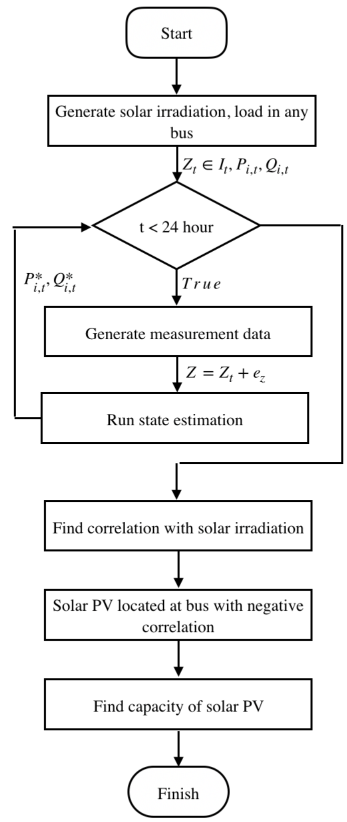

Fig. 1. conceptual methodology

A. Measurement devices, measured data and accuracy

In this study, there are three type of measurements; bus, line, and transformers measurement. Voltage magnitude in per unit is measurement at bus. Active and reactive power flow, and magnitude curreng are measured at line and trasformers. Bus measurement has 1% error and line measurement has 3% error.

To generate measurement data for testing prurposes, measurement error was added to tha actual measurements as shown in Equation (1).

$$Z = Z_a \pm +e_z \tag{1}$$

where Z_a is actual data and e_z is error added base on accuarcy of the measurement.

These error are assumed to be modeled independent Gaussian random variable [23]. where where the error value is expected value from gaussian distribution. noise is guassian distribution,as shown in Equation (2).

$$g(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}((x-\mu)/\sigma)^2} \tag{2}$$

B. Load allocation based state estimation

SE is based on the weighted least square (WLS) approach [24]. The method solves the following WLS problem to obtain an estimate ofthe system operating point defined by the system state x:

$$\min_x J(x) = \sum_{i=1}^{m} w_i(z_i - h_i(x))^2 = \left[z - h(x)\right]^T W \left[z - h(x)\right] \tag{3}$$

where w_i and $h_i(x)$ represent the weight and the measurements function associated with measurement z_i, respectively. For the solution of this problem the conventional iterative method is adape by solving following normal equations at each iteration, to compute the update $x^{k+1} = x^k + \Delta x^k$.

$$\left[G(x^k)\right]\Delta x^k = H^T(x^k)W\left[z - h(x^k)\right] \tag{4}$$

Where

$$G(x) = H^T(x)WH(x) \tag{5}$$

is the gain matrix and H is the jacobian of the measurement function $h(x)$.

C. Solar PV site's location identification

This part will descript location identification of invisible solar PV's site. The correlation-based feature selection (CFS) is key tool here. The basis of the CFS that was introduced in [25]. The CFS will be peformed as a statistical measure of relationship between solar irradiation and load allocation from previous part. The measure is best used in these two variables that demenstrate a linear relationship between each other. The correlation coefficient that indicates the straength of the relationship between these two variables can be found using following formula:

$$\mathrm{r}_{IP} = \frac{\sum(I_i - \overline{I})(P_i - \overline{P})}{\sqrt{\sum(I_i - \overline{P})^2(P_i - \overline{P})^2)}} \tag{6}$$

where r_{IP} is the correlation coefficient of the linear relationship between the solar irradiation (I) and load consumption (P), I_i is the values of the solar irradiation variable in a sample, $\overline{I}$ is the mean of I, P_i is the values of the load consumption variable in a sample, $\overline{P}$ is the mean of P.

The negaive correlation shows that the variables tend to move in opposity directions (i.e., when one variable increases, the other variable decreases). The probability of solar PV site located at bus i is expressed in Equation (7)

$$\mathrm{prob}_i == \begin{cases} 0 & \text{if } r_i >= 0 \\ 1 & \text{if } r_i < 0 \end{cases} \tag{7}$$

D. Estimate solar PV site's installation capacity

Once, we can identify location of invisible solar PV located. Then, we try to estimate size of invisble solar PV units. We deploy change-point detection of change-point analysis. The change-point detection algorithm is a powerfull tools used to detect abrupt changes in time series data. We add solar irradiation and load allocation where invisble solar PV located. Then, we can find chaning point of the dataset.

$$\Delta D_{i,t} = p_e \Delta I_{i,t} \tag{8}$$

where p_e is the size of the unauthorized PV system at bus i, and time t.

Then, we remove noise and find mean, range or other to estimate solar capacity.

E. Evaluation

This section will descripe evaluation methods will being used in this paper. There are two part of evaluate the proposed invisible solar PV size. First, the identification of the invisible solar PV location is evaluated by confusion matrix, $textF_1$ score, and Matthews correlation coefficient (MCC). Second, the estimation of installation of invisible solar PV size is evaluated by mean absolute error (MAE), and mean absolute percentage errot (MAPE).

The results of location of invisible solar PV detection will be represented in confusion matrix as exampled in Table I

TABLE I
CONFUSION MATRIX

		True condition	
	Total population	Condition positive	Condition negative
Prediction condition	Prediction condition positive	True positive (TP)	False positive (FP)
	Prediction condition negative	False negative (FN)	True negative (TN)

where TP, TN, FP, and FN are defined as number of hit, correct rejection, false alarm and miss, respectively.

1) F_1 score: The F_1 score is a measure of a test's accuracy. The F_1 score is the harmonic average of the precision and recall, where an F_1 score reaches its best value at 1 (perfect precision) and worst at 0. The F_1 is formulated in Equation (9).

$$\mathrm{F}_1 = \frac{2 \times \mathrm{TP}}{2 \times \mathrm{TP} + \mathrm{FP} + \mathrm{FN}} \tag{9}$$

2) Matthews correlation coefficient: The Matthews correlation coefficient (MCC) is First introduced by biochememist Brian W. Matthews in 1975 [26]. The MCC is essence an correlation coefficient between theobserved and predicted binary classification; its returns a value between -1 and +1. A coefficient of +1 represetnes a perfection prediction, 0 no better than random prediction and -1 indicates total disagreement between prediction and observation. An Equation (10) shows the formular of MCC for 2 classes. The dervations are express in [27].

$$\mathrm{MCC} = \frac{\mathrm{TP} \times \mathrm{TN} - \mathrm{FP} \times \mathrm{FN}}{\sqrt{(\mathrm{TP}+\mathrm{FP})(\mathrm{TP}+\mathrm{FN})(\mathrm{TN}+\mathrm{FP})(\mathrm{TN}+\mathrm{FN})}} \tag{10}$$

3) Mean absolute error and Mean absolute percentage error: The MAE defines as the difference between the actual and estimated invisible solar PV capacity which is computed by Equation (11). The MAPE expresses the scaled difference between the actual and estimatedd invisible solar PV capacity as a percentage of the actual invisible solar PV capacity. MAPE is scale independent and it can be used to compare estimation performance across different data sets. MAPE can be calculated by Equation (12).

$$\text{MAE} = \frac{1}{N}\sum_{i=1}^{N}|p_a - p_e| \tag{11}$$

$$\text{MAPE} = \frac{100}{N}\sum_{i=1}^{N}\left|\frac{p_a - p_e}{p_a}\right| \tag{12}$$

where p_a and p_e are the actual and estimated solar PV installation capacity. N is the number of sample.

IV. Test cases and Results

Two power system are tested. There two test cases are based on four bus system and CIGRE Task Force C6.04.02 paper [28].

A. four bus system

The one line diagram of the four bus system is shown in Fig. 2.

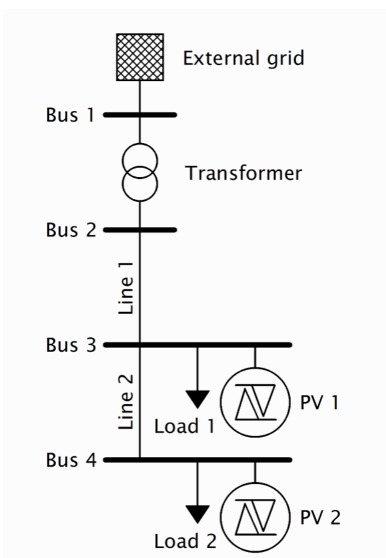

Fig. 2. four bus system

Table II shows the results. From the table, II it can be seen that the proposed method can predict perfectly location of invisible solar PV.

TABLE II
Results of invisible solar PV identification on 4 bus system

		Actual PV located	
		Yes	No
Prediction PV located	Yes	126	0
	No	9	60

The results is tested by F_1 score and MCC using Equation 9, 10. The F_1 score is 0.965. The MCC is 0.901 The false negative (FN) occurs when invisible solar PV's capacity is 1-2 kW where maximum demand is around 30 kW.

TABLE III
Results of invisible solar PV estimation on 4 bus system

Size of invisible solar PV (kW)	MAE(kW)	MAPE (%)
1-10	0.77	13.9
11-20	2.4	16.9
21-30	3.75	14.2
31-40	4.42	12
41-50	3.85	8.4

The overall MAPE is 13.587 % and MAE is 2.919 kW.

B. CIGRE system

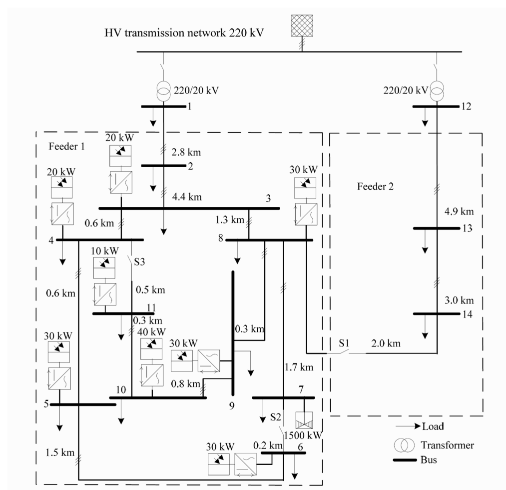

Fig. 3. CIGRE network

Table IV shows the results. From the table, II it can be seen that the proposed method can predict perfectly location of invisible solar PV.

TABLE IV
Results of invisible solar PV identification on CIGREsystem

		Actual PV located	
		Yes	No
Prediction PV located	Yes	154	8
	No	3	159

The results is tested by F_1 score and MCC using Equation 9, 10. The F_1 score is 0.965. The MCC is 0.9325. The false negative (FN) occurs when invisible solar PV's capacity is 10-20 kW.

TABLE V
Results of invisible solar PV estimation on CIGRE system

Size of invisible solar PV (kW)	MAE (kW)	MAPE (%)
1-100	9.01	17.60
101-200	24.3	15.97
201-300	50.00	19.85
301-400	52.55	13.98
401-500	54.92	12.28

The overall MAPE is 16.27 % and MAE is 37.58 kW.

V. Conclusion

Here is Conclusion.

Acknowledgment

References

[1] K. Mathiesen. (Jan. 2016). What is holding back the growth of solar power?. [online]. Available: https://www.theguardian.com/sustainable-business/2016/jan/31/solar-power-what-is-holding-back-growth-clean-energy.

[2] I. Ost I. (May. 2018). The Price of a Solar Panel System over the Years. [online]. Available: https://blog.pickmysolar.com/the-price-of-a-solar-panel-system-over-the-years.

[3] P. Fairley. (Jan. 2015). Hawaii s Solar Push Strains the Grid. [online]. Available: http://www.technologyreview.com/news/534266/hawaiis-solar-push-strains-the-grid/

[4] W. Staff. (Sep. 2014). Heco Customers Asked to Disconnect Unauthorized PV Systems. [Online]. Available: http://khon2.com/2014/09/05/heco-customers-asked-to-disconnect-unauthorized-pv-systems/.

[5] V. Schwarzer and R. Ghorbani, Transient over-voltage mitiga- tion: Explanation and mitigation options for inverter-based dis- tributed generation projects, Elect.Vehicle Transp. Center, Sch. Ocean Earth Sci. Technol., Univ. Hawai i Manoa, Honolulu, HI, US A , Tech. Rep. HNEI-02-15, Feb. 2014.

[6] B. Palma. (Sep. 2017). Is It Illegal in Florida to Power Your Home With Solar After a Storm Because of Lobbying?. [online]. Available: https://www.snopes.com/fact-check/is-it-illegal-florida-power-home-solar-storm/.

[7] B. Sipe. (Jul. 2006). Renewable Energy System Permit Fees in Oregon. [online]. Available: https://www.energytrust.org/wp-content/uploads/2016/11/0609_SolarPermitFees0.pdf.

[8] M. E. Baran, H. Hooshyar, Z. Shen, and A. Huang, Accommodating high PV penetration on distribution feeders, IEEE Trans. Smart Grid, vol. 3, no. 2, pp. 10391046, Jun. 2012.

[9] M. J. Reno, K. Coogan, R. J. Broderick, J. Seuss, and S. Grijalva, Impact of PV variability and ramping events on distribution voltage regulation equipment, in Proc. IEEE Photovolt. Spec. Conf., Tampa, FL, USA, 2014.

[10] A. Samadi, L. Soder, E. Shayesteh, and R. Eriksson, Static equivalent of distribution grids with high penetration of PV systems, IEEE Trans. Smart Grid, vol. 6, no. 4, pp. 17631774, Jul. 2015.

[11] H. Ravindra et al.,Impact of PV on distribution protection system, in Proc. North Amer. Power Symp. (NAPS), Champaign, IL, USA, Sep. 2012, pp. 16.

[12] A. Primadianto and C. Lu, ”A Review on Distribution System State Estimation,” in IEEE Transactions on Power Systems, vol. 32, no. 5, pp. 3875-3883, Sept. 2017.

[13] Haibin Wang and N. N. Schulz, ”A revised branch current-based distribution system state estimation algorithm and meter placement impact,” in IEEE Transactions on Power Systems, vol. 19, no. 1, pp. 207-213, Feb. 2004.

[14] M. Baran and T. E. McDermott, ”Distribution system state estimation using AMI data,” 2009 IEEE/PES Power Systems Conference and Exposition, Seattle, WA, 2009, pp. 1-3.

[15] J. Pereira, J. T. Saraiva and V. Miranda, ”An integrated load allocation/state estimation approach for distribution networks,” 2004 International Conference on Probabilistic Methods Applied to Power Systems, Ames, IA, 2004, pp. 180-185.

[16] R. Sharifian, E. Jafari, M. Rahimi, P. Ghaebi Panah, ”State estimation for load allocation in distribution systems,” International Journal on Technical and Physical Problems of Engineering, vol. 4, no. 1, pp. 86-90, Mar. 2012.

[17] Nathan Wallace, Stanislav Ponomarev, and Travis Atkison, ”Identification of Compromised Power System State Variables”, International Conference on Information and Knowledge Engineering, Nevada, USA, 2014.

[18] G. Chaojun, P. Jirutitijaroen and M. Motani, ”Detecting False Data Injection Attacks in AC State Estimation,” in IEEE Transactions on Smart Grid, vol. 6, no. 5, pp. 2476-2483, Sept. 2015.

[19] X. Zhang and S. Grijalva, A data-driven approach for detection and estimation of residential pv installations, IEEE Transactions on Smart Grid, vol. 7, no. 5, pp. 24772485, Sept 2016.

[20] H. Shaker, H. Zareipour, and D. Wood, Estimating power generation of invisible solar sites using publicly available data, IEEE Transactions on Smart Grid, vol. 7, no. 5, pp. 24562465, Sept 2016.

[21] H. Jiang, X. Dai, D. W. Gao, J. J. Zhang, Y. Zhang, and E. Muljadi, Spatial-temporal synchrophasor data characterization and analytics in smart grid fault detec- tion, identification, and impact causal analysis, IEEE Transactions on Smart Grid, vol. 7, no. 5, pp. 2525 2536, Sept 2016.

[22] X. He, R. Qiu, L. Chu, Q. Ai, Z. Ling, J. Zhan, Detection and estimation of the invisible units using utility data based on random matrix theory, Oct. 2017, [online] Available: https://arxiv.org/abs/1710.10745.

[23] A. Abur and A. G. Exposito, Power System State Estimation: Theory and Implementation. New York: Marcel Dekker, 2004.

[24] Roytelman I. and S.M. Shahidehpour, State estimation for electric power distribution systems in quasi real-time conditions, IEEE Trans. Power Systems, 1993 winter meeting, paper no:090-1-PW RD.

[25] M. A. Hall, Correlation-based feature subset selection for machine learning, Ph.D. dissertation, Dept. Comput. Sci., Univ. Waikato, Hamilton, New Zealand, 1999.

[26] ”Comparison of the predicted and observed secondary structure of T4 phage lysozyme”

[27] Powers, David M W (2011). ”Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness Correlation” (PDF). Journal of Machine Learning Technologies. 2 (1): 3763

[28] CIGRE’ Task Force C6.04.02: Developing benchmark models for integrating distributed energy resourcesS. Barsali, Kai Strunz, Zbigniew StyczynskiPublished 2005