Pattern-based monte carlo simulation for AMR electricty load analysis

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*—This paper proposes customer behavior analysis for pattern analysis of AMR electricity customer.**

In this paper univaraite models for short-term load forecasting based on customer's pattern behavior analysis and probabilistic monte carlo simulation are proposed. The proposed method were compared with that of other models based on ARIMA, exponential smoothing and neural networks. Application examples confirm valuable properties of the proposed approaches and their high accuracy.

***Index Terms*—Autometic meter reading, confidence interval**

I. Introduction

Here is introduction. In a revolutionary change in enegy section transform the traditional unidirectional electricty grid replaced by bidirectional or smart grid (SG). As a results of increasing in number of Intelligent Electronic Devices (IEDs) in the power system, especailly metering field. Consequently, there are repidly jump in enormous data volume in power system for storage, mining, sharing and visualization [1]. The advance meter read (AMR) with 15-min read intervals has also been develop to replacve the traditional managtic once a month reading meters. The AMR reads 96 data per day and carries out 2880 data per month, which means that 2880 times customer data are fed to utility. In addition, other states variables also transported.

In previous work, there is observation that the forecasting accuracy highly depend on hourly load patterns incorporate with other variables [2]. In addition, it can also help in long term applications i.e., model customer behavior under various incentive and pricing structures, planning processes [4]. The behavior of applicace in resident customer helps to forecast shorterm load [6].

In this article, we propose to generate behavior pattern for AMR customer consumption using confidence interval and Monte Carlo simulation. In particular, we make the following contributions:

- We show how to extrat a feature of customer consumption behavior by confidence interval with quantile values in order to reduce mumber of data.
- We formulate probabilistic function of individual customer behavior from extracted features.
- We deploy Monte Carlo simulation technique to simulate power consumption using individual probabilistic customer behavior.

PEA, AIT, NEU

II. Literature reviews

Here is Literature reviews.

The AMR data and individual major applicance usage learning are used to predict short-term residential load using Long short-term memory (LSTM) technique [6].

The big data has brought numberous tengible benefits to utilities and electricity uesers, which can be systemically concluded as follows:

- *Increasing System Stability Reliability:* Wide area mornitoring require numberous of measurement units, espécailly phase measurement units (PMUs) to ensure that the operator can manage system stability. In coorporate with AMR or Smart meter could help in this situation[cite ????].
- *Increasing Asset Utilization Efficiency:* With low accuray of GIS input data, the distribution network topology need to be verified, especially the under ground feeder which are difficult to check [8]. The big data proces could help to develop modeling of secondary size of transformer as well as energy theft [7].
- *Better Customer Experience Satisfaction:* Demand response program is an effective way to manage power balance during high congestion period as well as high tariff. The customer who engaging though demand response program could reduce their energy bill or earn incentives [cite ???].

There is several benefits of deploying AMR at homes and office. The mass rollout enables easier billing, fraud detection, forewarning of blackouts, smart real-time pricing schemes, demand response and efficient energy utilization. However, to acheive aboved benefits, there need advanced data analytics, especailly customer behavior analysis, which is the main motivation of this study.

In addition, the customer pattern also was clustered using Markov model with CFSFDP [5] In previous works, electrical customer consumption's pattern is formulated using various approach. Gaussian mixture model (GMM) is proposed to formaulate individual AMR-based electricity comsumption pattern [3].

So far, there is less number of article study on electricity customer behavior. The contribution of this work is ...

III. Problem Formulation

Here is Problem formulation. The overall methodology is shown in Figure1

A. *Pattern formulation using confidence intervals for quantiles calculation*

In this paper 15 minutes based kilowatt data are collected from AMR system. These data are accumulate into 30-minutes based kiloWatt-hour.

$$X = \left\{X^1, X^2, X^3, ..., X^n\right\} \tag{1}$$

$$X^n = \left\{X^n_1, X^n_2, X^n_3, ..., X^n_d, ..., X^n_{366}\right\} \tag{2}$$

$$X^n_d = \left\{X^n_{d,1}, X^n_{d,2}, X^n_{d,3}, ..., X^n_{d,t}, ..., X^n_{d,48}\right\} \tag{3}$$

where X is set of customer, X^n is set of daily consumtion of custome n, X^n_d is set of 30 minutes based power consumption (kWhr) of customer n on day d. $x^n_{d,t}$ is power consumption of customer n on day d at time t. The equation (1)-(3) are cleansing into equation (4). X^{n*} is set of power consumption at individual time step. X^{n*}_t is set of power consumption at time t of customer n.

$$X^{n*} = \left\{X^{n*}_1, X^{n*}_2, X^{n*}_3, ..., X^{n*}_t, ..., X^{n*}_{48}\right\} \tag{4}$$

The X^{n*} is cleansing raw data prepared to feature extraction process. As memntion above, this paper proposed confidential interval at quantile value as extracted feature. The extracted feature processes are shown in equation (5)- (6).

$$Y^n = \left\{Y^n_1, Y^n_2, Y^n_3, ..., Y^n_t, ..., Y^n_{48}\right\} \tag{5}$$

$$Y^n_t = \left\{Y^n_{t,0}, Y^n_{t,0.05}, Y^n_{t,0.1}, ..., Y^n_{t,q}, ..., Y^n_{t,1}\right\} \tag{6}$$

Where Y^n is representing set of extract feature of customer n at individual time period, Y^n_t is set of extracted feature of customer n at time period t which content 20 step of quantile value, q, (0 to 1 at 0.05 step size). $Y^n_{t,q}$ is formulated using equation (7).

$$Y^n_{t,q} = \int_{q-1}^{q} F_{X^{n*}}(q)dq \tag{7}$$

Where $F_{X^{n*}}$ is commulative distribution function of power consumption of customer n at time t. So, $Y^n_{t,q}$ is expected power consumption of customer n at time period t, and quantile q.

Hence, we can extract customer behavior feature as well as reduce number of process data in next step.

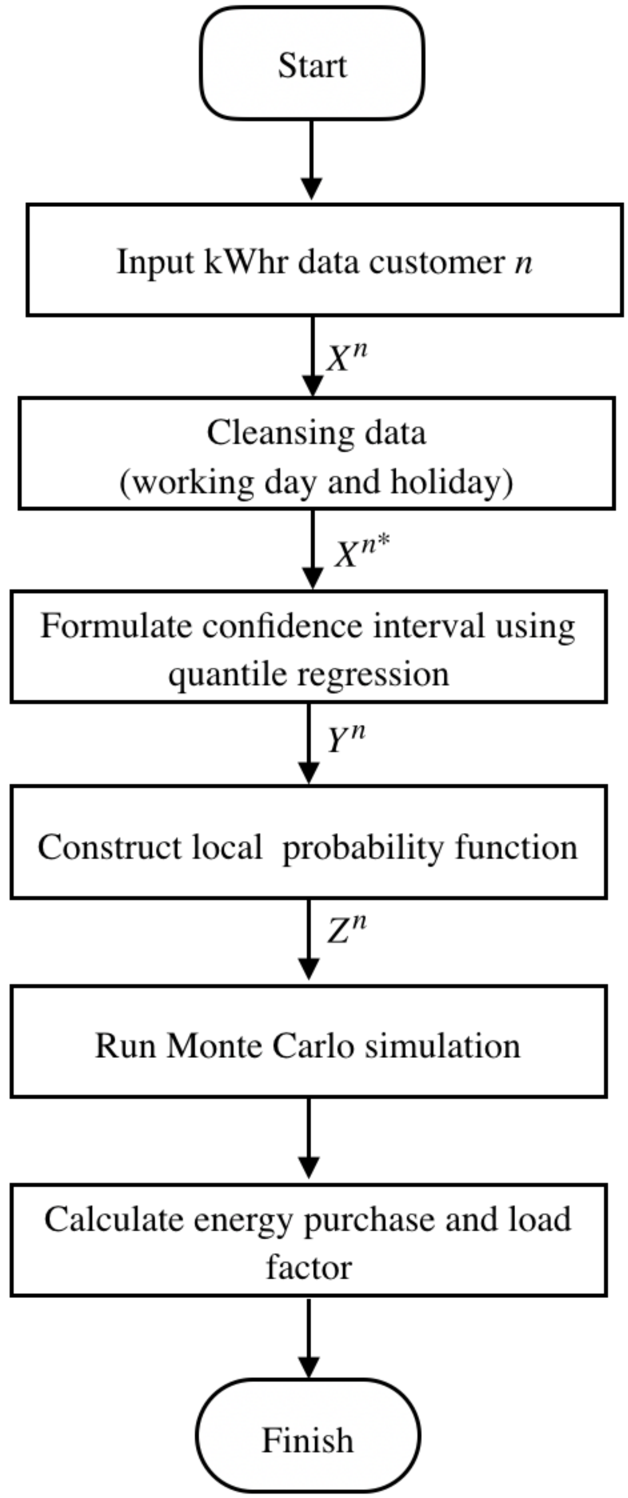

Fig. 1. Conceptual methodology

B. Continuous Probability Distribution constuction

$$Z^n = \{Z_0^n, Z_1^n, Z_3^n, ..., Z_t^n, ..., Z_{48}^n\} \tag{8}$$

$$Z_t^n = \{z_{0,50}^n, z_{50,100}^n, z_{100,150}^n, ..., z_{a,b}^n, ..., z_{19500,2000}^n\} \tag{9}$$

where Z^n is set of continous probability distribution function of power consumption of customer n. Z_t^n is set of continous probability distribution function of power consumption of customer n at time t with difference consumption range (from 0 to 20,000 kiloWatt-hour with 50 kiloWatt-hour step size). $z_{a,b}^n$ is probability of power consumption between lower a and upper b kiloWatt-hour of customer n which is be formulation by equation (10).

$$z_{a,b}^n = \mathrm{P}\big[a \le Y_t^n \le b\big] = \int_a^b Y_t^n dY_t^n \tag{10}$$

where a and b is lower and upper kilowatt-hour in range $\big[a, b\big]$.

C. Monte carlo simulation

From Z^n, monte carlo simulation generate electricity consumption as

$$P^n = \{P_1^n, P_2^n, P_3^n, ..., P_i^n, ..., P_m^n\} \tag{11}$$

$$P_i^n = \{P_{i,1}^n, P_{i,2}^n, P_{i,3}^n, ..., P_{i,t}^n, ..., P_{i,48}^n\} \tag{12}$$

where m is number of samples. $P_{i,t}^n$ is generated electricity consumption of customer n at time t at sample i.

D. Find cost and load factor

$$C_i^n = \sum_{t=1}^{48} P_{i,t}^n \times a_t \tag{13}$$

where C_i^n is cost of energy purchasing of customer n at sample i, a_t is electricity tariff at time t.

$$\mathrm{LF}_i^n = \frac{\mathrm{average}(P_i^n)}{\max(P_i^n)} \tag{14}$$

where LF_i^n is load factor of customer n at sample i.

IV. Test Cases and Results

In this study, AMR data is collected from PEA. This dataset comprehensively records the quarter hourly kilowatt reading of 35 commercial and industrial customers. We accomulate the kilowatt reading into kilowatt hour for every 30 minutes. The AMR customer names are change to alias for information security.

In feature extraction processes, total number of 70,272 raw data for each individual customer (2 years of collections) can be reduce to 1,920 data points (960 point for each working day and holiday).

The result of MC simulation with numbers of samples (20 samples) are consider as witness in this study Here is results. See in I, II

TABLE I
Energy cost per day on workind day

AMR-ID	Raw data		Proposed approach (20 samples)	
	Mean	SD	mean	sd
21652	75,139	14,264	75,502	11,558
136898	153,630	23,736	152,176	12,202
137091	34,053	9,608	37,134	3,980
137138	33,384	16,728	34,907	3,805
42432	236,816	44,446	238,527	11,596
66543	12,456	2,569	13,103	988
21654	6,795	6,970	9,151	1,781
42421	67,473	11,738	65,978	4,034
42423	5,743	3,470	6,245	2,119
43958	70,504	16,975	71,375	6,810
137110	11,973	2,166	13,418	951
21655	5,579	1,591	5,730	397
42431	12,357	3,402	11,882	975
44834	63,665	11,412	64,856	2,979
56452	213,943	35,413	215,219	9,267
56457	36,415	4,492	36,151	1,431
56458	28,128	3,577	27,797	1,401
124642	65,564	9,881	64,834	2,205
124647	56,148	8,723	55,809	1,814
124649	241,325	39,482	238,368	10,597
124656	57,745	7,114	57,947	2,554
124683	15,237	2,342	15,043	750
185767	22,119	5,746	22,564	1,361
56448	51,399	8,024	50,880	2,564
136900	84,763	4,738	84,603	2,969
137094	237,926	33,501	236,738	14,264
164978	11,111	3,533	11,625	1,055
189318	147,381	18,182	148,718	4,839
193781	59,164	37,639	61,509	6,127
44318	31,611	8,814	31,667	2,099
124687	4,702	725	5,637	300
21689	62,114	12,386	62,917	6,148
44831	57,599	8,382	58,446	1,941
56459	11,456	5,235	11,453	1,353
124678	56,828	28,156	57,598	5,405

V. Conclusion

Here is Conclusion.

The major contribution of this work is to propose new simulation univariate monte carlo simulation models based on pattern of customer behavior analysis.

Acknowledgment

I am vary grateful to Mr. Pradya Panyainkeaw, AMR division, PEA, Thailand for supplying data, and AIT, PEA for financial support.

References

References

[1] Depuru SSSR, Wang L, Devabhaktuni V. Smart meters for power grid: challenges, issues, advantages and status. Renew Sustain Energy Rev 2011;15(6):273642.

[2] Srinivasan D. Evolving artificial neural networks for short term load forecasting. Neurocomputing 1998;23:26576.1534.

[3] Chaweewat P., Singh J. G. , Ongsakul W. A Two Stages Pattern Recognition for Time-of-use Customers based on Behavior Analytic by Using Gaussian Mixture Models and K-mean Clustering: a Case Study of PEA, Thailand, 2018 International Conference and Utility Exhibition on Green Energy for Sustainable Development (ICUE)

[4] N. Yu, S. Shah, R. Johnson, R. Sherick, M. Hong, and K. Loparo, Big data analytics in power distribution systems, 2015 IEEE Power Energy Soc. Innov. Smart Grid Technol. Conf., pp. 15, 2015.

TABLE II
LF PER DAY ON WORKIND DAY

AMR-ID	Raw data		Proposed approach (20 samples)	
	Mean	SD	mean	sd
21652	0.458	0.078	0.422	0.076
136898	0.595	0.080	0.422	0.063
137091	0.403	0.098	0.256	0.040
137138	0.459	0.084	0.312	0.042
42432	0.551	0.057	0.428	0.039
66543	0.396	0.051	0.326	0.041
21654	0.332	0.110	0.209	0.036
42421	0.455	0.046	0.369	0.037
42423	0.227	0.111	0.114	0.048
43958	0.753	0.175	0.720	0.070
137110	0.523	0.080	0.401	0.087
21655	0.408	0.091	0.272	0.051
42431	0.482	0.090	0.320	0.044
44834	0.569	0.090	0.524	0.053
56452	0.620	0.089	0.554	0.048
56457	0.567	0.060	0.498	0.031
56458	0.621	0.083	0.570	0.057
124642	0.583	0.059	0.528	0.046
124647	0.543	0.063	0.480	0.044
124649	0.655	0.061	0.523	0.048
124656	0.592	0.082	0.495	0.080
124683	0.520	0.071	0.431	0.053
185767	0.561	0.089	0.412	0.057
56448	0.571	0.051	0.507	0.065
136900	0.710	0.056	0.663	0.060
137094	0.350	0.035	0.310	0.028
164978	0.464	0.111	0.314	0.053
189318	0.682	0.050	0.582	0.050
193781	nan	nan	0.350	0.082
44318	0.613	0.073	0.469	0.057
124687	0.553	0.047	0.522	0.048
21689	0.232	0.030	0.221	0.021
44831	0.533	0.060	0.463	0.059
56459	0.366	0.078	0.257	0.057
124678	0.673	0.094	0.407	0.035

[5] Y. Wang, Q. Chen, C. Kang, and Q. Xia, Clustering of Electricity Consumption Behavior Dynamics Toward Big Data Applications, IEEE Trans. Smart Grid, vol. 7, no. 5, pp. 24372447, 2016

[6] W. Kong, Z. Y. Dong, D. J. Hill, F. Luo and Y. Xu, "Short-Term Residential Load Forecasting Based on Resident Behaviour Learning," in IEEE Transactions on Power Systems, vol. 33, no. 1, pp. 1087-1088, Jan. 2018.

[7] Jokar P, Arianpoo N, Leung VCM. Electricity theft detection in AMI using customers' consumption patterns. IEEE Trans Smart Grid 2016;7(1):21626.

[8] Luan W, Peng J, Maras M, et al. Smart meter data analytics for distribution network connectivity verification. IEEE Trans Smart Grid 2015;6(4), [11].