Analysis of Timing Requirements for Data Aggregation and Control in Smart Grids

Kemal, Mohammed Seifu; Olsen, Rasmus Løvenstein

Published in:
Telecommunications Forum (TELFOR), 2014

DOI (link to publication from Publisher):
10.1109/TELFOR.2014.7034380

Publication date:
2014

Document Version
Accepted author manuscript, peer reviewed version

Analysis of Timing Requirements for Data Aggregation and Control in Smart Grids

Mohammed Kemal, *Student Member, IEEE* , Rasmus Olsen, *Member, IEEE,*

***Abstract*—Modern communication mechanisms are at the heart of a smart grid system to ensure that the required information is transmitted within various components of the grid. Throughout this paper, we have studied how communication performance delays and smart grid controller delays effects the overall control system operation. The main goal of this paper is to propose and analyze resource allocation algorithm satisfying delay requirements of the communication infrastructure and the controller. Simulation of a communication network is implemented using OMNeT++ to study and record the corresponding delays. The collected data is used to train resource allocation learning algorithm implemented in Matlab. The method will enable us to have a real number evaluation of time allocation for the communication network and the controller by taking into account constraints present on the system.**

Keywords*—*Smart Grid, Data Collection, Timing Requirement, Control.

I. INTRODUCTION

The smart grid concept has widely been used in recent years in different contexts and with different definitions [1]. In this paper, smart grid is viewed as a network of electric generation units, transmission lines, distribution substations and consumers with sensors and communication devices as part of the communication infrastructure.

Renewable energy sources such as solar energy and wind energy are significantly variable demanding highly sophisticated control mechanisms. This control mechanism requires a faster and efficient communication system between highly distributed energy sources and consumers. Recent improvements seen in communication technologies has opened the door in resolving restrictions on electrical grids caused by communication constraints.

For complex controllers deployed on smart grid, the communication infrastructure must have the capacity to transfer the information within an optimal time frame. The performance of control operation is constrained by capacity of the communication infrastructure. The main motive for this work is the need for a method to optimize the communication and controller delays in-terms of minimizing the total operating time and maximizing the capacity of the system. The communication infrastructure implemented for data collection under the simulation platform is DSL access network.

Mohammed Kemal is a recent graduate of Electronic Systems, Aalborg University, Aalborg, Denmark,e-mail: moseifu@gmail.com.
Rasmus Olsen is an Associate Professor of the Department of Electronic Systems, Aalborg University, Aalborg, Denmark,e-mail: (see http://kom.aau.dk/ rlo).

II. BACKGROUND

A. *Timing Requirement*

For a control system studied, the timing requirement(control loop back time) is constrained by communication delays for data collection, delays for data redistribution and delay for the control operation. This scenario is depicted as high level message sequence diagram shown on figure 1. Shown by the message sequence diagram is one step of the periodic message exchange performed by the control mechanism.

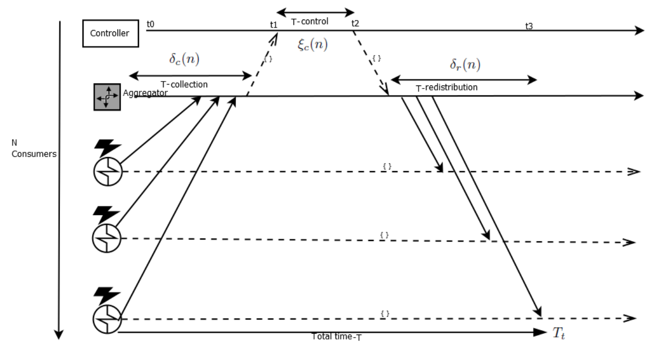

Fig. 1. *High Level Message Sequence Diagram*

The control loop back time is constrained by,

- N_c number of consumers
- N_a number of aggregators;
- d distance;
- B bandwidth;
- β aggregation server capacity;
- ξ Control complexity;

$\delta_c(n)$, $\xi_c(n)$ and $\delta_r(n)$ are T-collection, T-control and T-redistribution respectively(see figure 1). The total control loop back time T_t satisfies a relationship:

$$T_t = \delta_c(n) + \xi_c(n) + \delta_r(n) \tag{1}$$

III. RELATED WORKS

In relation to smart grid timing requirements, [4] states the communication delivery times for different applications in smart grids. It has described a standard defining communication delivery times of information to be exchanged within and external to substation integrated protection, control, and data acquisition systems[4]. The idea of using mixture models is a common practice in studying delay distributions of a communication infrastructure. [5], [6], [7] all use Gaussian mixture distribution with maximum likelihood expectation maximization to model the distributions of communication network delay measurements. [8] Proposes penalized maximum likelihood expectation maximization (PML-EM) algorithm applied on additive Gaussian mixture model. [9] is a paper to simulate both the communication side and the control aspects of a smart grid scenario where OMNeT++ is used to simulate power line communication infrastructure. Our paper here differs from other literatures because delay analysis and estimation methods by Gaussian mixture model is used to implement resource allocation algorithm. To the best of our knowledge, such a method for resource allocation used in smart grid controller is not considered on other literatures.

IV. Resource Allocation using Density Estimation(RADE)

Here we propose a resource allocation algorithm RADE to give us numerical evaluation of allocation of available time resource for the control server and communication mechanism. The first phase is to estimate model of the system by using recorded data sets of the corresponding delays.

1) Collection of data sets : $\{\xi_c(n), \delta_r(n), \delta_c(n)\}$, where $\xi_c(n)$ is collected from the control algorithm, $\delta_r(n)$ and $\delta_c(n)$ are collected from the communication network.
2) Estimate the distribution of each collected(training) data by using parametric density estimation (EM-ML for Gaussian mixture mode).

$$\xi_c(n), \delta_r(n), \delta_c(n) \sim \sum_{i=1}^{K} \mathcal{N}(\mu_i, \Sigma_i) \tag{2}$$

3) Compute the density function $P(x)$

$$P_{\xi_c}(\xi_c), P_{\delta_r}(\delta_r), P_{\delta_c}(\delta_c) = \sum_{i=1}^{K} \phi_i \mathcal{N}(\mu_i, \Sigma_i) \tag{3}$$

4) Estimate the joint PDF

$$\mathbf{P} = \prod_{i=1}^{n} P^i(x) = P_{\xi_c}(\xi_c) * P_{\delta_r}(\delta_r) * P_{\delta_c}(\delta_c) \tag{4}$$

5) Estimate the joint CDF,$F_X(x)$

Now we have the system model, We use it to give us numerical evaluation by setting ε. ε is the accepted threshold or confidence bound our system satisfies according to the trained model. In other words, its the required percentile coverage of the distribution in which the system should satisfy fulfilling the delay requirements to perform in an acceptable manner.

while *test data :* $\{x; {\xi_c}^t, {\delta_d}^t, \delta_c^t\}$ **do**
  Estimate $F_X(x)$;
  if $F_X(x) > \varepsilon$ **then**
    It satisfies the system requirement;
  else
    It doesnt satisfy the system requirement;
  end
end

Algorithm 1: Testing time allocation

V. Experiment and Results

A. Communication Scenario Description

1) Network Topology: Digital subscriber lines(DSL) are known to provide high-speed, low latency and secure infrastructure to be used for smart grid applications. For our implementation as shown with figure 2, It is assumed that there is point to point connection from smart meters to the aggregator and distribution server. An aggregator located at central station will be used for collecting the data and forward it to the controller and redistribution is done in the same manner in which controller signals are sent back to the meters by using the same communication channel.

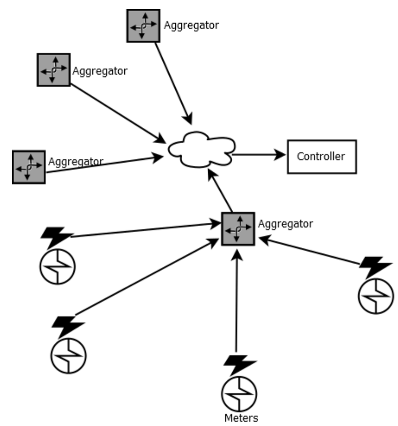

Fig. 2. *Network toplogy used for implementation*

2) Simulation: We have implemented the simulation by setting the parameters for the simulation to create multiple scenarios and see the results by plotting the recorded delay distributions. Since the DSL channel that we are using is a shared medium, We have generated four types of traffics to see how it affects the end to end delay caused by the automatic metering interface(AMI) traffic.

TABLE I. *Scenario parameters*

Parameter	**Value**
TCPapps	3
UDPapps	1
AMI packet size	200B
AMI packet Idle interval	3600s
AMI resource utilization	25%
HTTP resource utilization	25%
FTP resource utilization	25%
Video streaming resource utilization	25%
Number of client	100
QOS	Diff Server
DSL BW	10Mbps
DSL propagation delay	20ms
Back-end BW	100Mbps
Back-end propagation delay	10ms

HTTP and FTP traffics use TCP packets in which they are sent to a specific TCP port on the server where the server listens to those traffics on the specified ports. A normal scenario is tested by using droptail queuing policy in which the last packets are randomly dropped during buffer overflow. To implement video streaming and burst data generation, UDP ports are set for each traffic on the server and intelligent consumer. QOS with queuing and priority scheduling policy is implemented by using differentiated services(diffserv) implementation in Inet. A scenario on table I is tested in which QOS parameter with priority queuing policy is implemented. The message from intelligent consumers is a 200B TCP packet transmitted with an idle interval of every 10 minutes. The full set of parameter is shown in table I. Data is collected for end to end delay($\delta_c(n) + \delta_r(n)$) of the communication network. We can see from the histogram plot that the distributions has a significant degree of randomness because of the queue on the buffer.

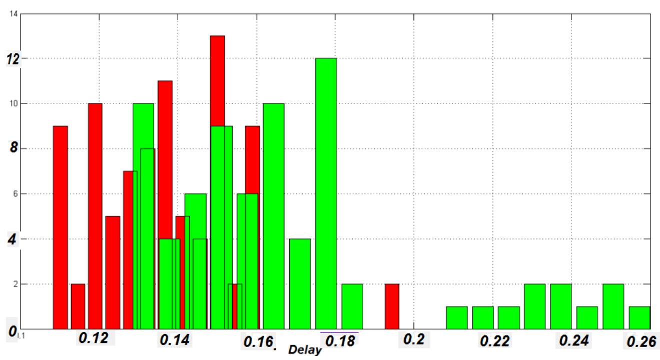

Fig. 3. *Scenario delay distribution*

B. Controller Requirement

Hierarchical predictive control model [3] is used as a use case for the analysis part and the rest of this work. It is chosen because of the prior work done to study computational burdens in relation to traffic caused by the consumers and other loads which presents a tested model to be used in relation to our study. It is proposed for resource distribution to be deployed on smart grids. In smart grid control, consumers or energy consuming units like heat pumps, car batteries, refrigerators and more, require power in a limited value within a specific set of time. This and other constraints add to the complexity of the control system. The high level controller[3] has to deal with the aggregated value of the constraints mentioned. For this model, all distributed loads are seen as one big consumer by aggregating the constraints within.

According to Hierarchical Control Predictive Model(HCPM)[3], The main computational burdens are the vertex generation, the top level distribution and the aggregator level distribution, all of which must be performed at each time step. Vertex generation is generation of a vertex vector representing minimum or maximum power consumption rate at a specific moment for each consumer. Equation 6 taken from [3] is an example we use in this paper to illustrate the complexity of a controller. It is stated that the complexity of the quadratic programming of a distribution task increases approximately with the square of the number of receivers and approximately doubles with an increase in the horizon length.

$$D_c(N_d) = \beta_d 2^{N_h} {N_d}^2 + \beta_o \tag{5}$$

N_d the number of associated consumers or aggregators in the layer directly below where β_o is dependent on the computational power of computer and Horizon length N_h is the number of time steps of the prediction. It is also argued on [3] and shown on Equation 7 that the vertex computation burden is linear with the number of consumers and number of vertices. The total upper computational burden is mainly affected by the vertex computational burden and the distribution computational burden. This can be denoted as the aggregate of the two computational burdens as noted on Equation 8. The relationship is used for the estimation of the controller delay($\xi_c(N_d)$) for the operation. We have used this relationships to estimate the delay distributions for the controller using Gaussian mixture model when testing the models on the later section.

$$V_c(N_d) = \beta_v N_h! 2^{N_h} N_d \tag{6}$$

$$\xi_c(N_d) = V_c(N_d) + D_c(N_d) \tag{7}$$

C. RADE Implementation and Results

The primary goal of RADE is to tell us if our allocation of smart grid time resource for the communication infrastructure and control server are appropriate. For this we have designed a system that will give numerical evaluation of the choice. In this work we focused on time allocation problem where RADE is used to test if our allocation satisfies delay requirements for the communication network and the controller given N number of intelligent consumers connected to the system. On the previous sections we have analyzed how we collect the features that will be used for RADE algorithm. The two features used are the delay distributions collected from the network simulation done on OMNeT++ and the distribution for the controller. Let us discuss the data preparation phase where the collected distributions are trained to model the best fit for the algorithm.

D. Data Preparation and Training

We collected 100 samples from each intelligent consumers(meters) in total 10000 samples by the simulator. The same amount of data is generated from the controller distribution by assuming a mixture of Gaussian distribution. The data is divided in to three namely, training data, validation data, and test data. The training data is used for RADE as input features which is used to train the learning algorithm. The first step is using the method expectation maximization and maximum likelihood method to parametrize the training data with Gaussian mixture parameters. An example fit is shown on the fig 4 which shows the fitting of delay distribution of the tested scenario which was shown fig 3.

The cross validation data is used to make the learning algorithm work efficiently and improve if it doesnt satisfies the confidence level that we want our fit to have. The confidence level is the percentage of coverage of the distribution that our fit is able to cover. For this work we used confidence level of above 90%. We use the test data to confirm if the fit works as desired. The same procedure is taken to divided the delay distribution of the controller. Assumption is taken

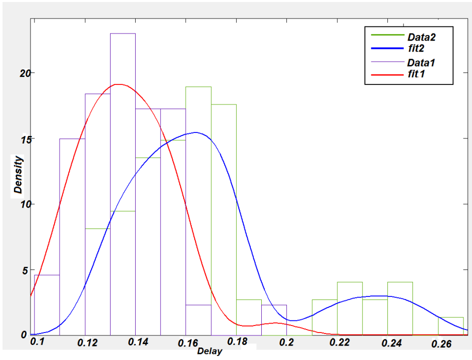

Fig. 4. *Fitting the distribution*

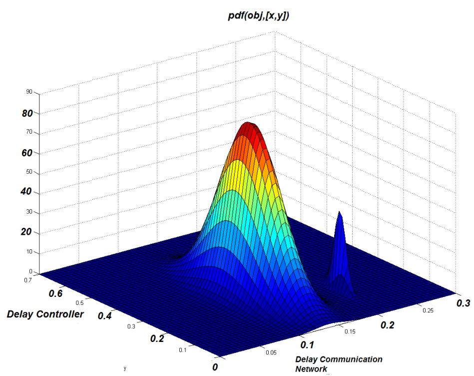

Fig. 5. *PDF plot of the Joint distribution*

that the controller delay distribution can be modeled by using Gaussian mixture model.

E. RADE Test

On previous sections we discussed what input datas we used for the model and how we prepared the data by fitting them to a Gaussian mixture model using EM-ML. By using the method mentioned by RADE algorithm we get the joint distribution of the input features. Fig 5 shows the multivariate distribution function **P(x)** of the data. The delay distribution for the communication network has a bigger peak value around 150ms and a smaller peak value at 220ms. While the controller delay distribution is concentrated at around 300ms.

The multivariate CDF($F_X(x)$) of the distributions is shown on Fig 6. It is found from the test that we can get 100% coverage if we set 220ms for the communication network and 600ms for the controller.

Another test taken to show how the proposed method can be used to evaluate the time allocation problem and assist to make proper allocation choices. It is a test performed to relate communication network mean delay with total processing time given the control distribution is fixed and taking different confidence bounds on the CDF($F_X(x)$). As shown on Table II, The test is taken for three confidence bounds 75% , 85% and 95%. The mean for the Gaussian mixture distribution is varied by sliding the whole Gaussian components by its value.

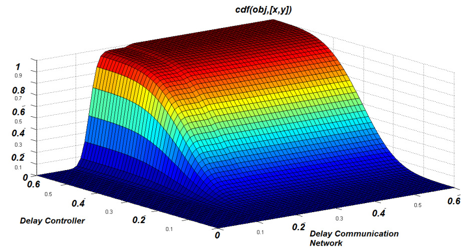

Fig. 6. *CDF plot of the Joint distribution*

TABLE II. *Confidence bound and Mean Communication Network*

Mean Ntk Delay(s)	**0.000**	**0.100**	**0.200**	**0.300**	**0.400**	**0.500**
Total Time(75%)	0.409	0.513	0.609	0.712	0.801	0.891
Total Time(85%)	0.611	0.719	0.803	0.899	1.091	1.183
Total Time(95%)	0.821	0.931	1.010	1.110	1.183	1.300

VI. Conclusion

Throughout the course of this paper, We proposed and implemented Resource allocation algorithm(RADE) specifically for smart grid control timing requirements. RADE provides real number evaluation of allocation of resources which will enable us to check fulfillment of confidence bounds by allocating the proper processing time for smart grid communication network and controller.

The training datas used for this model are communication network delay distribution and controller delay distribution. To get a realistic distribution of the communication delay, we implemented a simulation network for DSL access network using OMNeT++. Since access network is a shared medium, The simulation is tested by generating different network traffics. Controller using HPCM is used for this work for estimating controller delay distributions. RADE Implementation is tested and got results to support our argument that this method can be used for resource allocation problems on smart grids.

References

[1] Energinet and DanskEnergy, Smart Grid in Denmark, May, 2010

[2] Ali Keyhani,Muhammad Marwali, Smart Power Grids , 2011, Springer

[3] Klaus Trangbaek, Mette Pedersen, Jan Bendtsen, Jakob Stoustrup, Predictive smart grid control with exact aggregated power constraints, 2011, From the book Smart Power Grids 2011, -, 655-675

[4] 1646TM IEEE Standard Communication Delivery Time Performance Requirements for Electric Power Substation Automation, 2004, IEEE Power Engineering Society,

[5] Figueiredo M., Jain,, Unsupervised learning of finite mixture models,, Mar. 2002, IEEE Trans. on Pattern Analysis and Machine Intelligence, 51 Issue: 8 , 381396

[6] G. McLachlan and D. Peel, Finite mixture models,, Mar. 2000, New York: John Wiley & Sons,

[7] G. Celeux, S. Chr‘etien, F. Forbes, and A. Mkhadri,, A component-wise EM algorithm for mixtures,, 1999, INRIA Rhone-Alpes,France,

[8] Meng-Fu Shih, Alfred O. Hero, Unicast-Based Inference of Network Link Delay Distributions with Finite Mixture Models, 2010, IEEE Transactions on Signal Processing, 2219 - 2228

[9] J.T.Madsen,T.l.F.Kristensen, R.L.Olsen, T.K.Madsen, K.Balachandran, Simulating Network Adapted Market Controller in a Smart Grid Scenario, 2013,