

Power System State Estimation with Dynamic Optimal Measurement Selection

Jinghe Zhang, Greg Welch and Gary Bishop
Department of Computer Science
University of North Carolina at Chapel Hill
Chapel Hill, NC, USA
Email: {jing2009, welch, gb}@cs.unc.edu

Abstract—Power system measurement devices continue to evolve towards higher accuracy and update rate. On the other hand, the computation required for processing the enormous amounts of measurement data associated with large complex power systems makes real-time estimation a major challenge. In this paper we present the Lower Dimensional Measurement-space (LoDiM) state estimation method for large-scale and wide-area interconnected power systems. We present the method in the context of the Kalman filter and Extended Kalman filter, however our *measurement selection procedure* is not filter-specific, *i.e.* it can also be applied on other state estimation methods such as particle filters and unscented filters. Our method can also take advantage of large-scale parallel computation techniques for further improvement. Moreover, the concept of LoDiM should be applicable to other large-scale, real-time and computationally-intensive state tracking systems beyond the power systems, such as weather forecasting systems, gas-pipeline systems, and other critical infrastructure.

I. INTRODUCTION

State estimation plays a basic yet very important role in modern industries. Particularly for power systems, state estimation provides critical data for driving other operation functions including real-time security monitoring, load-forecasting, economic dispatch, and load-frequency control.

Because of the increasing size and complexity of interconnected power networks, state estimation computation remains one of the primary obstacles to overcome. Even with modern supercomputers, the massive data processing is still time consuming and memory challenging. Much previous work has also been dedicated to the acceleration of calculations, *e.g.*, via parallel mathematical systems such as OpenMP [1] and Global Arrays [2].

However in this paper, we introduce a different approach to reduce the computational burden. In our previous research [3], a tracking method called single-constraint-at-a-time (SCAAT) was proposed. The idea behind the SCAAT approach, which is typically implemented with a Kalman filter, is to estimate a globally observable system using only measurements from locally unobservable systems. As a result, in each filtering cycle SCAAT deals with much lower dimensional measurement data. Although SCAAT does not inherently dictate any particular measurement selection, it inspired our design of LoDiM, a SCAAT-based algorithm featuring dynamic measurement selection for state tracking of large-scale systems (power systems, weather forecasting systems, etc.). Furthermore, our

method is expected to significantly facilitate power system operations, supporting *hierarchical* and *distributed* estimation, as well as parallel computation techniques for more improvement.

The remainder of this paper is organized as follows. Section II introduces the background. In Section III, we present our new approach to perform Kalman filtering upon dynamically select measurements, in order to reduce the computational requirement dramatically, without sacrificing the tracking quality. Our method is tested on a multi-machine system model in Section IV to demonstrate how it could benefit the power system state estimation process. Finally, conclusions and acknowledgement are stated in Section V and Section VI respectively.

II. BACKGROUND AND RELATED WORK

As described in [4], Kalman filtering techniques are extensively used for power system state estimation. Recent work [5] also employed them for the detection of bad data (outliers) in power systems.

However in applications such as power systems, computational load is always a concern. Significant previous work has attempted to reduce the computational complexity. For example, in [8] the authors proposed a method for systems that have more measurements than states. They described how an equivalent state-space system can be used where the number of measurements equals the number of states. However, they assumed the system is static and linear, and did not take the preparation overhead into account.

Another alternative to solving the computational challenge was presented in [6], where the authors used Petri net (PN) theory to achieve the optimum utilization of processors, in a state estimator based on the Kalman filter.

[9] illustrated a measurement selection procedure for Extended Kalman filters. Nevertheless, as stated by the authors, an inherent limitation of the proposed method is that measurement selection is based entirely on the steady-state sensitivity matrix. The shortcomings of this approach include: (1) the actual information content of the candidate measurements under typical operating conditions are not considered; (2) the measurement rankings obtained are local and dependent on the steady state chosen as the base case; and (3) dynamic and nonlinear effects are neglected.

Previously in [3], we presented SCAAT, a Kalman-filter-based incremental tracking algorithm using incomplete information. It estimates a globally observable systems using only measurements from locally unobservable systems. The underlying principle is that the single observations provide *some* information about the user's state, and thus can be used to incrementally improve a previous estimate. Here we present work on extending SCAAT with a principled measurement selection procedure, so that the estimation is more stable and reliable, with higher report rates with lower latency than a batch-based Kalman filter.

III. LOWER DIMENSIONAL MEASUREMENT-SPACE STATE ESTIMATION

The LoDiM method employs a Kalman filter that incorporates lower dimensional (sub-space) measurements in each cycle. It employs a special measurement selection procedure to strategically reduce the measurement-space dimension. As a result, it yields a much smaller computational load, lower latency, and most importantly: reliable performance. In this section, we will first give a brief introduction to the Kalman filter, then discuss our dynamic measurement selection method and the structure of LoDiM.

A. The Kalman Filter

The Kalman filter [7] has been used in a wide range of applications from radar tracking to weather forecasting. It is an efficient recursive filter that estimates the state of a process, in a way that minimizes the mean of the squared error.

An assumed *linear* system can be modeled as a pair of linear stochastic process and measurement equations

$$x_k = Ax_{k-1} + w_{k-1} \quad (1)$$

$$z_k = Hx_k + v_k \quad (2)$$

where $x \in \mathcal{R}^n$ is the state vector, $z \in \mathcal{R}^m$ is the measurement vector, A is a $n \times n$ matrix that relates the state at the previous time step $k-1$ to the state at the current step k in the absence of either a driving function or process noise¹, and H is a $m \times n$ matrix that relates the state to the measurement z_k . The process noise w_k and measurement noise v_k are assumed to be mutually independent random variables, spectrally white, and with normal probability distributions

$$p(w) \sim N(0, Q) \quad (3)$$

$$p(v) \sim N(0, R), \quad (4)$$

where the process noise covariance Q and measurement noise covariance R matrices are often assumed to be constant.

In reality, the process to be estimated and (or) the measurement relationship to the process is usually *nonlinear*, especially when the objective is to estimate the dynamic states

¹In practice, the matrix A may change with each time step, but it is assumed to be constant here.

of a power system. A nonlinear system can be modeled using nonlinear stochastic processes and measurement equations

$$x_k = a(x_{k-1}, w_{k-1}) \quad (5)$$

$$z_k = h(x_k, v_k). \quad (6)$$

One can approximate the states and measurements by

$$x_k = a(x_{k-1}) \quad (7)$$

$$z_k = h(x_k). \quad (8)$$

These nonlinear functions can then be linearized about the point of interest x in the state space. To do so one need to compute either or both of the Jacobian matrices

$$A = \left. \frac{\partial a(x)}{\partial x} \right|_x \quad (9)$$

$$H = \left. \frac{\partial h(x)}{\partial x} \right|_x \quad (10)$$

where A and H are the partial derivatives of a and h (respectively) with respect to x .

We define $\hat{x}_k^- \in \mathcal{R}^n$ to be the *a priori* state estimate at time step k given the knowledge of the process prior to k , so $e_k^- \equiv x_k - \hat{x}_k^-$ is called the *a priori* estimate error and $P_k^- \equiv E[e_k^- e_k^{-T}]$ is called the *a priori* estimate error covariance.

Similarly, We define $\hat{x}_k \in \mathcal{R}^n$ to be the *a posteriori* state estimate at time step k given measurement z_k , so $e_k \equiv x_k - \hat{x}_k$ and $P_k \equiv E[e_k e_k^T]$ are called the *a posteriori* estimate error and the *a posteriori* estimate error covariance respectively.

The Kalman filter estimates the state by minimizing the *a posteriori* estimate error covariance, in a recursive prediction-correction manner. The prediction step is realized by a set of time update equations:

Prediction :

$$\begin{cases} \hat{x}_k^- = A\hat{x}_{k-1} \\ P_k^- = AP_{k-1}A^T + Q \end{cases} \quad (11)$$

The time update equations are responsible for projecting forward (in time) the previous state x_{k-1} and error covariance estimates P_{k-1} to obtain the *a priori* estimates for the next time step k .

The correction step is carried out by a set of measurement update equations:

Correction :

$$\begin{cases} K_k = P_k^- H^T (HP_k^- H^T + R)^{-1} \\ \hat{x}_k = \hat{x}_k^- + K_k(z_k - H\hat{x}_k^-) \\ P_k = (I - K_k H)P_k^- \end{cases} \quad (12)$$

where K is an $n \times m$ matrix called the Kalman gain matrix, z_k is the actual measurement at time step k , $H\hat{x}_k^-$ is the predicted measurement at time step k , and $(z_k - H\hat{x}_k^-)$ is called the residual.

B. Principle of Design

As one would expect, performing dynamic state estimation with a KF/EKF (and in fact, any filter) is a rather computationally intensive process. For small systems, the computation could be fast enough for real time control applications. However there are three factors that increase the computational effort: the size of the system, the complexity of model components, and the number of measurements to be processed. In the ‘‘Correction’’ phase, corresponding to equation (12), we have noticed the expensive cost of calculating the Kalman gain K_k , primarily because it involves the inversion of a $m \times m$ matrix ($HP_k^-H^T + R$), with complexity of $O(m^3)$. This makes the computation intractable when the number of measurements m is too large, which unfortunately, is true for the modern power systems.

On the other hand, if we could reduce the measurement-space dimension, i.e. only use a subset of measurements in each Kalman filter cycle to update (perhaps a subset of) the states, the computation can be sped up dramatically, according to SCAAT. Now the question boils down to deciding which measurement subset to use at each cycle. Previous studies suggested pre-determined measurement subsets; however the dynamic nature of power systems compelled us to seek an improved method. We set our sights on the state space. Through principal component analysis (PCA) of the dynamic error covariance we can ‘‘pin down’’ the the subspace with the largest estimation uncertainty, i.e. the subset of the state space that needs to be updated more urgently.

Due to the fact that covariance matrices are always symmetric and positive semidefinite, there are several important properties about them. Before the measurement update (‘‘Correction’’ phase (12)) begins, let us consider the PCA of the *a priori* error covariance matrix $P^- = U \cdot D \cdot U^T$:

- 1) There exists an orthonormal basis U ($UU^T = U^TU = I$, where I is the identity matrix), whose columns are the eigenvectors of P^- , such that the error covariance matrix expressed in this basis is diagonal. The axes of this new basis are called the *Principal Components* of P^- .
- 2) As the off-diagonal elements of this new diagonal covariance matrix D are 0, the new variables defined by this new basis (the projections of the *a priori* estimate error $e^- = x - \hat{x}^-$ on the Principal Components) are uncorrelated.
- 3) The diagonal elements of this new matrix D are the eigenvalues of P^- . So the variances of the projections of error e^- on the Principal Components are equal to the corresponding eigenvalues of P^- .
- 4) The eigenvalues in D are ordered decreasingly. The m^{th} eigenvalue corresponds to the m^{th} eigenvector.

The principal components that correspond to the largest elements of D , indicate the axes in the state space that have the largest estimation uncertainties. We prefer to target these uncertainties first, i.e. we want to find the set of measurements that can be used to reduce those uncertainties most efficiently.

This makes perfect sense in any type of state estimation process: in a 3D tracking application example, if we use several cameras to estimate the location of a certain object, and we noticed the uncertainty is growing rapidly in one direction, then in the next cycle of filtering we would ideally use a camera which is looking in an orthogonal direction.

C. Measurement Selection Procedure

Similar to SCAAT, LoDiM also constrains the unknowns *over time* and refines the estimation continually, rather than waiting for a complete collection of observations to form. Nonetheless, it is the measurement selection that distinguishes LoDiM from SCAAT. In LoDiM, we select the measurements that benefit our estimate the most (i.e. reduce estimation uncertainty most effectively) during each iteration.

After the ‘‘Prediction’’ phase (11) of each Kalman filter iteration, we have the $n \times n$ *a priori* error covariance matrix

$$P^- = U \cdot D \cdot U^T \quad (13)$$

where D is the diagonal matrix consisting of the eigenvalues of P^- in decreasing order, and U is the orthonormal basis whose columns are the corresponding eigenvectors. Notice that computing the full PCA can be a time consuming process, especially if the state space is large. For this reason, in this paper we only investigate the first eigenvector u_1 in U , which represents the directions that we are most uncertain about in the state space. The largest eigenvalue and the corresponding eigenvector u_1 can be conveniently obtained by existing algorithms such as the *power method* [13].

Now consider the measurements. We can rewrite the measurement equation as

$$\begin{aligned} z &= Hx + v \\ &= HUU^Tx + v \\ &= (HU)x' + v \\ &= H'x' + v \end{aligned}$$

where $x' = U^Tx$ is the new state vector defined by the new basis U , H' is the corresponding $m \times n$ new measurement Jacobian matrix, and v is the unchanged measurement noise vector with $p(v) \sim N(0, R)$. Here we assume R to be a $m \times m$ diagonal covariance matrix, i.e. the measurement noise sources are uncorrelated.

Because the basis U is composed of unit vectors, H'_{ij} can also be considered the magnitude of the projection (i.e. the scalar projection) of the i^{th} measurement direction vector in the direction of the j^{th} basis in U . Notice that we now have

$$Hu_1 = [H'_{11} \quad H'_{21} \quad H'_{31} \dots H'_{m1}]^T \quad (14)$$

Intuitively, for the same basis, say u_1 , the larger H'_{i1} is, the better the corresponding i^{th} measurement could reduce the uncertainty in this basis direction. However we should also keep in mind that different measurement has different amount of noise. Thus for u_1 we create a ‘‘ranking’’ vector r_1 from

H' and R using the following adjustment:

$$r_1 = \begin{bmatrix} \frac{H_{11}^2}{R_{11}} & \frac{H_{21}^2}{R_{22}} & \frac{H_{31}^2}{R_{33}} & \dots & \frac{H_{m1}^2}{R_{mm}} \end{bmatrix}^T \quad (15)$$

where R is the measurement noise covariance matrix. The $m \times 1$ vector r_1 evaluates the uncertainty calibration abilities of each measurement regarding the most significant uncertainty component, scaled by the corresponding measurement noise level.

We are able to easily locate the m_σ measurements with the largest values among these m elements in vector r_1 (15). Thus when our budget of computation time and memory space is tight, we could use only m_σ measurements in the next step but still achieve stable estimation results. Using this approach to reduce state estimation uncertainty is similar to fighting the *Hydra* in Greek mythology: if we are not able to destroy all the “heads” at once, at least we can aim at and cut off the most threatening “head” during each round!

D. LoDiM Architecture

In this section, we explore the idea of employing modern shared-memory multi-core high performance computers, to improve the computational efficiency in large-scale Kalman-filter-based state estimations. We propose a better architecture for our new method, LoDiM, consisting of algorithm parallelization and parallel implementation.

To generalize our approach, let us consider a nonlinear system described by equations (5) and (6), with a presumably large measurement space. LoDiM has its main state estimation process, which is similar to SCAAT algorithm, running in the foreground:

- 1) Compute the time Δ_t since the previous estimate.
- 2) Predict the state and error covariance. Share the predicted error covariance P^- with the background process.

$$\begin{cases} \hat{x}^- = a_{\Delta_t}(\hat{x}_{t-\Delta_t}, 0) \\ P^- = A_{\Delta_t} P_{t-\Delta_t} A_{\Delta_t}^T + Q_{\Delta_t} \end{cases} \quad (16)$$

- 3) If this is the first cycle, choose m_σ measurements (which is a much smaller measurement subset) randomly; otherwise, choose the m_σ best measurements nominated by the background process. Predict the measurement and compute the corresponding Jacobian.

$$\begin{cases} \hat{z} = h_\sigma(\hat{x}_t^-, 0) \\ H = H_\sigma(\hat{x}_t^-, 0) \end{cases} \quad (17)$$

- 4) Compute the *Kalman gain*.

$$K = P^- H^T (H P_k^- H^T + R_{\sigma,t})^{-1} \quad (18)$$

- 5) Correct the predicted state estimate and error covariance from (16) using the actual sensor measurement $z_{\sigma,t}$.

$$\begin{cases} \hat{x}_t = \hat{x}^- + K(z_{\sigma,t} - \hat{z}) \\ P_t = (I - KH)P^- \end{cases} \quad (19)$$

Concurrently, LoDiM uses an auxiliary measurement selection process as described in subsection III-C, running in the background:

- 1) Compute the principal component u_1 of the error covariance P^- predicted in the foreground.
- 2) Compute Hu_1 according to (14) and the *ranking* vector r_1 according to (15).

$$r_1 = (Hu_1) .* (Hu_1) ./ \text{diag}(R) \quad (20)$$

where $.*$ and $./$ denote element-by-element multiplication and division (respectively).

- 3) Select the m_σ measurements that correspond to the m_σ largest elements in r_1 , to be used by the foreground process.

IV. SIMULATION RESULTS

In this section, we investigate the performance of the proposed LoDiM state estimation using a 16-generator 68-bus system, which represents the New England/New York interconnected system [11]. We simulate an emergency event: a three-phase fault at bus 29 that happens at time $t = 1.1$, and is then cleared in 0.05 seconds. This event represents a relatively large disturbance emergency.

For this 68-bus system, we simulate the 135 state variables and record them as the “true” system states. There are 272 simulated Phasor Measurement Unit (PMU) measurements, combined with random noise. PMUs, as their name implies, provide both voltage phasor and current phasor measurements, at a higher frequency and accuracy level [12].

Without a loss of generality, we visualize the state estimate results for bus 60, a generator bus. Fig. 1 depicts the voltage magnitude tracking result of bus 60 during the 10 seconds from $t = 0$ to $t = 10$, using the conventional Kalman filter. The black solid line plots the true state, while the red dot-dot line plots the estimated state. The entire measurement set is used as input in the regular EKF dynamic state estimation, resulting in lower estimation rate.

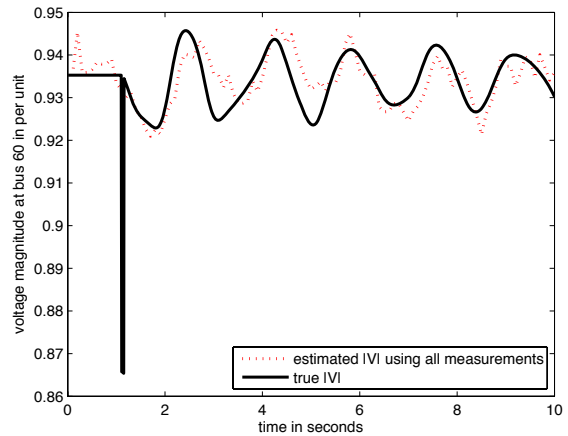


Fig. 1. Bus 60 Voltage magnitude estimation using conventional Kalman filter

Next, we exam the performance of the reduced measurement-space state estimation with a naive approach:

a small subset of the measurements (70 measurements in our experiment) is chosen randomly during each shorter estimation cycle. The tracking result is demonstrated in Fig. 2: the green dash-dot line records the estimation of the true state, which is plotted by black solid line.

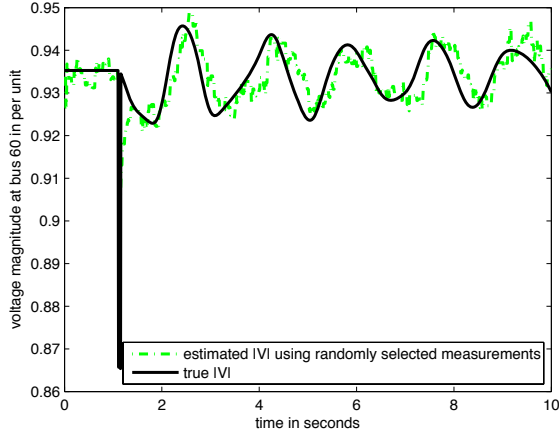


Fig. 2. Bus 60 Voltage magnitude estimation using randomly chosen measurements

Our LoDiM state estimation method use the same number of dynamically selected measurements (70 measurements also) as input during each short estimation cycle. In contrast, the small measurement subset is selected by the *measurement selection procedure* described in section III-C. Fig. 3 shows the voltage magnitude tracking result of the same bus during the same period, using LoDiM method. The black solid line still represents the true state, while the blue dash-dash line plots the estimated state.

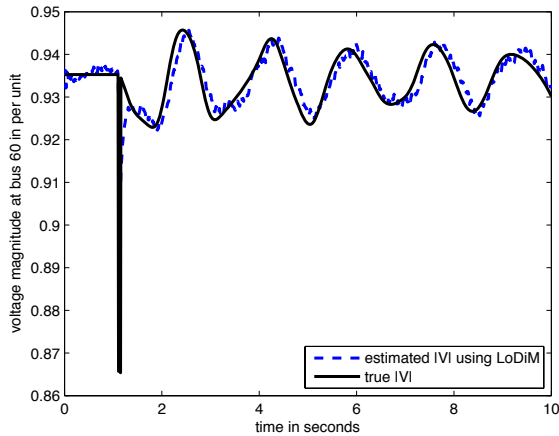


Fig. 3. Bus 60 Voltage magnitude estimation using our LoDiM method for measurement selection.

Finally, let us take a closer look at the performance of the different state estimation methods described above. Within a

short time period from $t = 4$ to $t = 6$ seconds at bus 60, the true state and the estimated states using these three approaches respectively, are plotted in Fig. 4 for a better comparison.

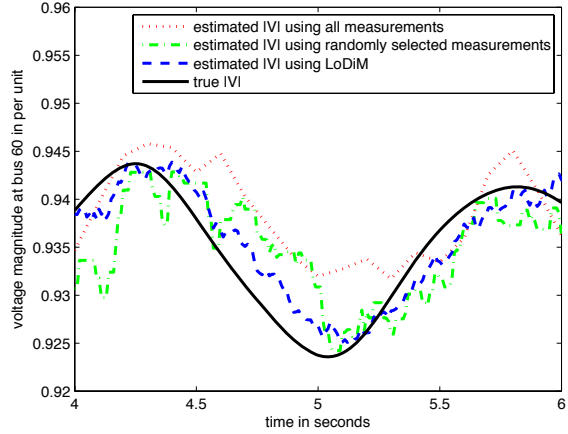


Fig. 4. The performance comparison of generator bus 60 state estimation with full measurement-space (in red), with randomly selected measurements (in green) and with LoDiM (in blue), from $t = 4$ to $t = 6$

Another comparison is performed similarly at bus 26, a load bus. The zoom-in figure (Fig. 5) also demonstrates how our proposed LoDiM method, appearing both smooth and accurate, outperforms the other two methods.

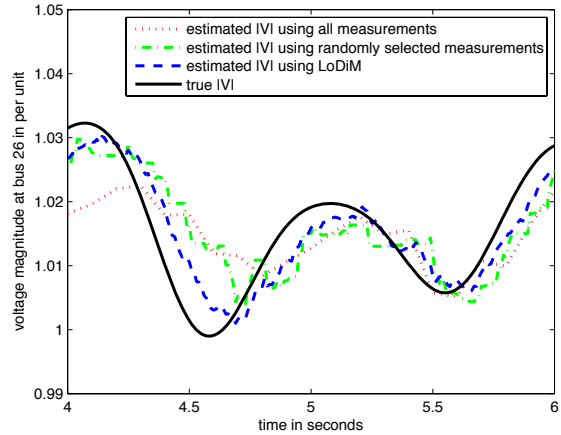


Fig. 5. The performance comparison of load bus 26 state estimation from $t = 4$ to $t = 6$

V. CONCLUSIONS

We have presented LoDiM, a new method for measurement selection in state estimation. During each estimation cycle, unlike the traditional batch methods of KF/EKF state estimation methods, which handle the entire measurement-space, it only needs to deal with a lower dimensional measurement-space thus reducing the computation time and increasing the

estimate rate. And unlike previous methods that process lower dimensional (or sequential) measurements, LoDiM seeks to choose the most valuable measurement at each cycle. The smaller measurement-space results in higher reporting rates, and the associated measurement selection results in more accurate estimates. LoDiM can be further parallelized and optimized, to benefit large-scale real-time state estimations, for the power systems and beyond.

VI. ACKNOWLEDGEMENTS

At the Pacific Northwest National Laboratory we acknowledge Zhenyu Huang, Ning Zhou, Pengwei Du, and Ruisheng Diao for helpful background discussions. This work was supported by a U.S. Department of Energy (grant DE-SC0002271) "Advanced Kalman Filter for Real-Time Responsiveness in Complex Systems," PIs Zhenyu Huang at PNNL and Greg Welch at UNC. At DOE we acknowledge Sandy Landsberg, Program Manager for Applied Mathematics Research; Office of Advanced Scientific Computing Research; DOE Office of Science. This work was also supported in part by the Office of Naval Research (award N00014-09-1-0813, "3D Display and Capture of Humans for Live-Virtual Training." At ONR we acknowledge Dr. Roy Stripling, Program Manager.

REFERENCES

- [1] Barbara Chapman, Gabriele Jost, and Ruud van der Pas, "*Using OpenMP: Portable Shared Memory Parallel Programming.*" the MIT Press, 2007
- [2] Jarek Nieplocha, Bruce Palmer, Manojkumar Krishnan, and P. Sadayappan, " *Overview of the global arrays parallel software development toolkit*". 2006 ACM/IEEE conference on Supercomputing, Tampa, Florida, 2006.
- [3] Greg Welch and Gary Bishop, "*SCAAT: Incremental Tracking with Incomplete Information*". Computer Graphics. T. Whitted, Los Angeles, CA, USA (August 3 - 8), ACM Press, Addison-Wesley: 333-344. 1997
- [4] Amit Jain, Shivakumar N. R. "*Power System Tracking and Dynamic State Estimation*" IEEE PES Power Systems Conference Exposition (PSCE) 2009
- [5] Chien-Hung Huang, Kuang-Kong Shih, Chien-Hsing Lee and Yaw-Juen Wang "*Application of Kalman Filter to Bad-Data Detection in Power System*". InTech 2010
- [6] Anant Oonsivilai and Kenedy A. Greyson "*Optimization of Distributed Processors for Power System: Kalman Filters using Petri Net*". World Academy of Science, Engineering and Technology 53 2009
- [7] R. E. Kalman "A new approach to linear filtering and prediction problems". Transaction of the ASME Journal of Basic Engineering, 82(Series D):35-45, 1960.
- [8] M.J. Goris, D.A. Gray and I.M.Y. Mareels "*Reducing the computational load of a Kalman filter*". Electronic Letters Vol. 33 No. 18 1539 -1541, 28th August 1997
- [9] Shoujun Bian, Michael A. Henson "*Measurement selection for on-line estimation of nonlinearwave models for high purity distillation columns*". Chemical Engineering Science 61 (2006) 3210-3222
- [10] Frank L. Lewis "*Optimal Estimation with an Introduction to Stochastic control theory*". John Wiley and Sons, Inc. 1986
- [11] R. Graham "*Power System Oscillation*". Kluwer Academic Publishers, Boston, 2000
- [12] A. Phadke, J. Thorp, and M. Adamiak, "A new measurement technique for tracking voltage phasors, local system frequency, and rate of change of frequency" Power Apparatus and Systems, IEEE Transactions on, vol. PAS-102, no. 5, pp. 1025-1038, may 1983.
- [13] Golub, G. H. and Van Loan, C. F "*Matrix Computations*". Johns Hopkins Univ. Press, Baltimore, MD, 1989.