Flexible Hierarchical System of Automatic Voltage Control

Aleksandr Domyshev¹, Alexey Osak¹, Kirill Zamula²
¹Melentiev Energy Systems Institute SB RAS, Irkutsk, Russia
²“Energy-T” Ltd, Tolyatti, Russia

Abstract. The subsystem for optimal control of voltage and reactive power of EPS is developed. The proposed solution uses state of art methods for state estimation, forecasting and dynamic optimization. A new architecture of an artificial neural network is proposed – a neuro-analytical network. Algorithms are proposed that allow reliable combination of classical automatic control methods and methods using machine learning. The proposed methodology is designed for use in a real power system for automatic voltage control.

1 Introduction

The voltage and reactive power control system is designed to maintain the optimal regime of the Magadan power system. In the Magadan power system, the main generation is concentrated hydro generation, which, due to its concentration, creates difficulties with optimal voltage control in the power system.

In our proposed solution, compensating devices are installed at important substations. The voltage is maintained in the optimal range by coordinated control of compensation devices, in combination with the coordinated control of the voltage on the buses of the power stations.

For reliability the system is designed in two levels of control: the lower level, which provides basic control functions, and the upper one, represented by the optimal control subsystem, which performs intelligent voltage control in the power system. For testing and debugging of the optimal control subsystem, the complex provides a subsystem of quasi-dynamic modeling of system behavior. This subsystem is not directly involved in the process of optimal voltage control.

Fig. 1. Power system under consideration.

Fig. 2. Structure of the voltage and reactive power control system.

Functions of the centralized low-level control system:
- Local controllers monitoring
- Transfer of information about the state of local controllers to the optimal control subsystem.
- Interconnection with optimal control subsystem via standard protocol (IEC 60870-5-104 or IEC 61850).
- Processing of the control task from the optimal control subsystem.
- Transfer of control actions to local controllers. Conversion of control actions to the type required for a particular type of regulator is provided.
- Control of technological limits. Technological limits protect the system against errors in the optimal control subsystem, communication errors, or digital attacks.
- Controlling consistency of constraints.
- Logging of all incoming commands from the optimal control subsystem.
- Archiving of all measurements and states of local controllers.

* Corresponding author: domyshev@isem.irk.ru
Functions of the optimal control subsystem:
- Collecting information on the current state of the system.
- Creating bus/branch model.
- State estimation.
- State prediction.
- Shaping constraints.
- Dynamic optimization with constraints.
- Analysis of optimization results.
- Forming of control actions.

2 State estimation

The state estimation problem is to obtain such a steady state that would be closest to the measurements:

$$\bar{v} - v(x) \rightarrow 0$$

where $\bar{v}$ – measured parameters, $v(x)$ – calculated mode parameters according to the values of the system state vector $x$.

An original state estimation method was developed using the classical state estimation algorithm together with the dynamic state estimation algorithm based on the Kalman filter. These two algorithms mutually rescue each other, ensuring the reliability of the result. The flowchart of the combined algorithm is shown in Figure 3.

The state assessment works on the bus/branch model formed by the ANARES topological processor. The source data for the topological processor is a graphical single-line diagram of the electrical network and the status of switching devices. Binding to telemetry is carried out interactively from a graphic diagram.

2.1 Static estimation

To solve the problem of static estimation, the method of weighted least squares is used [1, 2], which minimizes the following objective function:

$$J = [\bar{v} - v(x)]^T R_\bar{v}^{-1} [\bar{v} - v(x)]$$

where $\bar{v}$ – are the measured parameters of the state; $v(x)$ – are the calculated parameters according to the values of the state vector of the system $x$. The matrix $R_\bar{v}$ is the covariance matrix, in the absence of correlation between different measurements, it is a diagonal matrix of measurement variances $R_\bar{v} = \text{diag}(\sigma_{v1}^2, \sigma_{v2}^2, ..., \sigma_{vm}^2)$, $m$ is the number of measurements.

The results of the state estimation are the state vector

$$\hat{x} = [\hat{0}_1, \hat{0}_2, ..., \hat{0}_n, \hat{\delta}_1, \hat{\delta}_2, ..., \hat{\delta}_n]^T,$$

containing the estimated values of voltage angles for each node, $n$ is the total number of buses.

It should be noted that due to the insufficiency of measurements, so-called pseudo measurements are also used, that is, expert values of the load and generation obtained, for example, based on load schedule.

2.2 Dynamic state estimation

The dynamic state estimation problem is solved using an Unscented Kalman filter. The forecast horizon can be from several seconds to several minutes, which corresponds to the interval between the receiving of measurement data at time $k$ and at time $k+1$.

This method eliminates the need to explicitly calculate the Jacobian in contrast to the Extended Kalman filter. Unscented transformation implies the choice of a certain minimum set of sigma points for the data expectation and covariance of the desired vector with random noise superimposed on it. Using sigma-points, nonlinear forecast functions are constructed, which are then used to calculate the cross-covariance matrix and Kalman filter coefficients. Sigma-points are calculated using the following formulas:

$$\chi^{(i)} = \hat{x} + \left(\sqrt{N_x Z_x}\right)^i$$

$$W^{(i)} = W^{(i+N_x)} = \frac{1}{2N_x}$$

$$\chi^{(i+N_x)} = \hat{x} - \left(\sqrt{N_x Z_x}\right)^i$$

where $N_x$ is the dimension of the desired state vector; $i$ is the index of sigma point ($i \in \{1, ..., N_x\}$; $W^{(i)}$ is the weight of $i$-th sigma point; $\sqrt{N_x Z_x}$ is the $i$-th column of the matrix square root of the covariance matrix of the initial nonlinear transformation models multiplied by the dimension of the phase vector.

3 State prediction

A combined algorithm using forecasting based on load schedules and machine learning using artificial neural
networks performs state prediction in the optimal control subsystem.

3.1 State prediction based on load schedule

As a hierarchical load schedule, a combination of daily schedule and “winter coefficient” is used:

$$P_n = P_{n0}(P_{cr,s}K_z + P_{cr,c}(1 - K_z)),$$

where $P_{n0}$ is the initial value of load power or generation in basic mode; $P_{cr,s}$ is the value of the coefficient of the winter daily schedule of active power in relative units for the time of prediction; $P_{cr,c}$ is the value of the coefficient of the summer daily schedule of active power in relative units for the time of prediction; $K_z$ is the “winter coefficient”, set for each month. A value of 1 corresponds to the winter schedule, 0 for the summer schedule. The value for a specific day of each month is calculated by linear approximation.

$$\sigma = \sigma_0 e^{\alpha r},$$

where $i$ is the current moment; $\tau$ is the forecasting depth. Variance $\sigma(\tau)$ exponentially depends on the forecasting depth: $\sigma(\tau) = \sigma_0 e^{\alpha \tau}$.

3.2 State prediction using artificial neural network

As a basic ANN architecture, a recurrent neural network LSTM was used for state prediction. LSTM cell is shown in figure 6.

![Fig. 6. LSTM cell of ANN.](image)

The use of recurrent LSTM networks for prediction allows obtaining results not inferior to FCRBM networks [3]. Moreover, based on LSTM networks, it is possible to build deep networks, which in turn allows us to identify more complex correlations in the source data. As regards the choice of a particular type of cells of a recurrence network, the work [4] can be cited in favor of the choice of LSTM, in which various RNN variants (in particular, LSTM and GRU) are analyzed and it is revealed that LSTM using dropout technology gives the best result in all tests.

One of the important features of the source data for solving the problem of state prediction is their probabilistic nature. Input measurements are specified not just by a value, but also by variances. This provides additional information for training the neural network. Besides, at the output of the neural network we will have not only the values of the parameters of the steady state, but also the probability distributions for these parameters. This information will allow us to make more adequate control decisions based on the results of forecasting. Therefore, to solve the state prediction problem, the Bayesian ANN was applied. The Bayes by Backprop (BBB) method [5-9] provides a posterior distribution of neural network weights. Using it for recurrent ANN is described in [10].

The architecture of the ANN on which studies of the proposed approaches were carried out is presented in Figure 7. Features of this architecture: several fully connected input layers for revealing correlations between parameters, several B-LSTM layers for revealing time dependencies, output fully connected layer. The number of layers may vary depending on the complexity of the problem being solved, but in practice it did not exceed 3 for each type of layer.

![Fig. 7. Architecture of proposed deep ANN.](image)

3.3 Case study of the proposed prediction algorithm

Studies of the presented architecture were carried out on the scheme of the Magadan energy system (Figure 1). If a network is trained to predict a small number of parameters, then training is effective (Figure 8) and the results of the forecast using a trained neural network are close to real (Figure 9). In the presented figures, the number of predicted parameters is 6.

In the case when we used this architecture to predict all the required parameters of the considered electric grid, the results became insufficient (Figures 10, 11).
electrical grid into clusters in each of which only a small number of parameters are predicted.

For each of the clusters, a reliably working architecture is used, shown in Figure 7. The results of the forecast are the initial data for calculating the power flow: injection of active and reactive power at buses and voltage at the PU buses.

Based on the output parameters of each cluster, one step of the power flow calculating algorithm is performed. This calculation is performed at every step of the iterative process of training and forecasting. Thus, the analytical method for calculating the power flow is embedded in a single model (Figure 12) participating in the process of backpropagation of gradients.

### 3.4 Neuro-analytical network

To solve the above problem, an architecture of the neuro-analytical network was developed. The proposed approach involves clustering the entire investigated

Fig. 8. Learning error. Number of predicted parameters is 6.

Fig. 9. The result of prediction is close to real data. Number of predicted parameters is 6.

Fig. 10. Learning error on real sized problem.

Fig. 11. The result of prediction is insufficient on real sized problem.

Fig. 12. Neuro-analytical network.

Clustering on subgrids is performed using a hierarchical clustering algorithm by principal component analysis. A measure of distance is the electrical resistance between grid nodes. The criterion for ending the partition into clusters is the acceptable maximum cluster size.

The components of the gradient vector of the analytical part of the grid are calculated based on the implicit differentiation of the error of output vector of predicted variables $y$ relative to measured ones:

$$\varepsilon = (y(u) - \bar{y})$$

For this, a system of linear equations is solved:

$$J \left[ \frac{dx}{du} \right] = \left[ \frac{dx}{dy} \right]$$

where $J = \left[ \frac{dx}{du} \right]$ is the transposed Jacobi matrix; $x_{ck}$ is the vector of independent variables for cluster $k$; $u = [u_{r1}, u_{i1}, ..., u_{rN}, u_{in}]$ is the vector of real and imaginary parts of bus voltages obtained as a result of load flow calculation.

#### 3.5 State prediction results

The forecasting unit using the ANN as the forecast gives the vector $\hat{x}(i+\tau)$ of the predicted measurements, as well as the standard deviation of each predicted value from the actual measurement in the test sample:
\[ \sigma = \sqrt{\frac{\sum_{i=1}^{n_t} (\bar{x}_i - \bar{x})^2}{n_t}} \]

where \( n_t \) is the number of slices in the test sample, \( \bar{x}_i \) is the test value, \( \bar{x} \) is the measured value.

The final indicator of the accuracy of the forecast is the total standard deviation for all predicted measurements. The standard deviation is used to calculate the variance \( \sigma^2 \), which in turn is used to obtain a generalized prediction result by the state estimation method.

## 4 Shaping constraints

The next stage of optimal control will be the formation of constraints, which will be further taken into account in optimization. Constraints are voltage constraints; restrictions on the injection of reactive power at the buses of the electric grid, which are generators and compensating devices; current and power flow constraints.

Constraints are set in advance manually for each variant of the topological state of the electric grid, and are also calculated by heavy load algorithm. To select the current set of constraints, it is necessary to classify the current or predicted state of EPS (normal, emergency, etc.). By taking into account different topological states of the electric grid, there can be several dozen classes of states and, accordingly, sets of constraints.

The classification of the EES operation mode is carried out using the random forest machine learning method [11-13], which is a development of the decision tree method. An important advantage of decision trees is that they not only relate the regime to a particular class, but also allow numerically substantiating the classification of the regime. The decision tree is a tree whose leaves correspond to the values of the objective function, and in the other nodes, the transition conditions that determine which of the edges of the node to pass the algorithm.

The algorithm for creating a random forest consisting of \( N \) trees is as follows:

- For every \( i \)-th tree (\( i \in [1..N] \)):
  - Sample \( X_i \) is generated using the bootstrap algorithm.
  - A decision tree for sample \( X_i \) is formed
    - Based on a given criterion, we select the best attribute and do a partition in the tree by it until the selection is exhausted.
    - A tree is formed until each sheet contains no more than a specified number of objects or until a certain tree height is reached.
    - For each partition, \( m \) random attributes are first selected from \( n \) initial ones, and then the optimal division is sought among them.

- Among the trees, the final decision is selected by voting using the final classifier:

\[ a(x) = \frac{1}{N} \sum_{i=1}^{N} b_i(x) \]

Classification algorithm training is performed during the system setup process, however, this does not require explicit specification of the attributes of the EPS state to one or another class. Attributes are generated automatically in the learning process.

Depending on which class the current mode is assigned to, in the optimization problem the constraints corresponding to this class are applied in the form of logarithmic barriers.

## 5 Dynamic OPF

In the subsystem of optimal control of voltage and reactive power, an algorithm of dynamic optimization is used to obtain optimal control actions at a given depth. The initial data for dynamic optimization are:

- Current estimated state.
- Forecast in the form of a set of predicted states.
- List of possible control actions.
- Current constraints.

The proposed solution takes into account the cost of control actions, which depends not only on the state vector of the system, but also on time. The full objective function of the optimization problem can be written as:

\[ f_d = \sum_j^{p} \left( \sum_{i=b_p}^{e_p} f_i(X_i) + C(\xi_j - \xi_{j-1,1}) \right) \]

where \( \xi_j = \text{argmin} \sum_{i=b_p}^{e_p} f_i(X_i) \) is the vector of optimal control parameters on the subrange: \( b_p \) and \( e_p \) are indices of the beginning and end of a subrange in a time range; \( C \) is the function of the cost of the control actions depending on the time specified by the index \( t \); \( \xi_0 \) is taken equal to the initial value of the control parameters \( X_0 \); \( f_i \) is the objective function of each optimization subrange, including active power loss, outages from permissible voltage ranges and other constraints.

The cost of control action depends on such factors as residual resource of equipment, priority of using certain control action, minimum permissible time between operating by the same device.

Generally, the dynamic optimization algorithm is presented in Fig. 13.

![Fig. 13. Structure of dynamic optimization algorithm.](image)

As a method of local optimization used for each mode in the process of dynamic optimization, the L-BFGS-B method is used, which is quasi-Newtonian and
takes into account restrictions on control parameters [14 - 16].

Results of the optimal control:
- Constraints in the form of a voltage boundary (maximum and minimum) for each local controller.
- The control task for each local controller with the required time for its processing.
- Results of prediction.
- Optimization Results.
- Calculated values of residual resource of control devices.

7 Software implementation

The subsystem for optimal control of voltage and reactive power of EPS is developed based on the ANARES software [17]. This subsystem can be used both as part of an automatic voltage control system, and for other similar tasks. The structure of the embedded version of the optimal control subsystem is shown in Figure 14.

![Figure 14. Embedding version of subsystem for optimal control of EPS.](image)

The acknowledgments. This research is fulfilled within the contract with «Energy-T» Ltd for "Magadanenergo PJSC.

References

1. A. Monticelli, State estimation in electric power systems: a generalized approach. – Springer Science & Business Media, 2012.
2. A.Z. Gamm, Statistical methods for state estimation of electric power systems. – Nauka, 1976.
3. D. L. Marino, K. Amarasinghe, M. Manic, Building energy load forecasting using deep neural networks //Industrial Electronics Society, IECN 2016-42nd Annual Conference of the IEEE. – IEEE, 2016. – C. 7046-7051.
4. R. Jozefowicz, W. Zaremba, I. Sutskever. An empirical exploration of recurrent network architectures //International Conference on Machine Learning. – 2015. – C. 2342-2350.
5. A. Graves, Practical variational inference for neural networks //Advances in neural information processing systems. – 2011. – C. 2348-2356.
6. C. Blundell et al., Weight uncertainty in neural networks //arXiv preprint arXiv:1505.05424. – 2015.
7. M.J. Wainwright et al., Graphical models, exponential families, and variational inference //Foundations and Trends® in Machine Learning. – 2008. – T. 1. – №. 1–2. – C. 1-305.
8. A. Graves, Practical variational inference for neural networks //Advances in neural information processing systems. – 2011. – C. 2348-2356.
9. C. Blundell et al., Weight uncertainty in neural networks //arXiv preprint arXiv:1505.05424. – 2015.
10. M. Fortunato, C. Blundell, O. Vinyals, Bayesian recurrent neural networks //arXiv preprint arXiv:1704.02798. – 2017.
11. T.K. Ho, Random decision forests //Proceedings of 3rd international conference on document analysis and recognition. - IEEE, 1995. - T. 1. - C. 278-282.
12. L. Breiman, Random forests //Machine learning. - 2001. - T. 45. - №. 1. - C. 5-32.
13. E.M. Kleinberg, On the algorithmic implementation of stochastic discrimination //IEEE Transactions on Pattern Analysis and Machine Intelligence. - 2000. - T. 22. - №. 5. - C. 473-490.
14. A. Domyshev et al., Optimal Power Flow Calculation Using BFGS-Based Optimisation Scheme //2018 2nd IEEE Conference on Energy Internet and Energy System Integration (EI2). - IEEE, 2018. - C. 1-6.
15. R.H. Byrd et al., A limited memory algorithm for bound constrained optimization //SIAM Journal on Scientific Computing. - 1995. - T. 16. - №. 5. - C. 1190-1208.
16. P. Armand, P. Segalat, A limited memory algorithm for inequality constrained minimization. - Technical Report 2003-08, University of Limoges (France) 2003, 2003.
17. Software package for power system modeling, [Online]. Available in RUSSIAN: http://anares.ru/software