

ОЦЕНИВАНИЕ СОСТОЯНИЯ ЭЭС С ИСПОЛЬЗОВАНИЕМ ГЛУБОКОГО ОБУЧЕНИЯ

А.В. Домышев

Институт систем энергетики им. Мелентьева СО РАН
664033, Иркутск, ул. Лермонтова, д. 130
E-mail: domyshev@isem.irk.ru

Ключевые слова: машинное обучение, искусственные нейронные сети, оценивание состояния, рекуррентные нейронные сети, Байесовские нейронные сети.

Аннотация: В данной работе рассматриваются современные методы глубокого обучения нейронных сетей, применяемые для решения задач управления электроэнергетическими системами и, в частности, для задачи оценивания состояния. Предложена методика оценивания состояния в темпе процесса с использованием глубокого обучения искусственных нейронных сетей (ИНС). Данная методика основана на глубоких рекуррентных сетях LSTM. Предложено применение байесовских рекуррентных сетей для решения задачи динамического оценивания состояния. Предложенная методика реализована в программном комплексе, позволяющим проводить оценивание состояние и дообучение ИНС в темпе процесса.

1. Введение

Многослойные нейронные сети, обучаемые с помощью методик глубокого обучения, применяются для решения различных задач, таких как распознавание образов, обработка аудиоинформации, автоматическое управление, прогнозирование и др.

Модели глубокого обучения – это искусственные нейронные сети (ИНС) с несколькими слоями скрытых нейронов, которые имеют связи только между нейронами, принадлежащими к последовательным слоям, но не имеют связей внутри одного слоя.

В энергетике использование ИНС рассматривается прежде всего для решения задач прогнозирования и автоматического оптимального управления в тех случаях, когда входные данные задачи являются стохастическими или сильно зашумленными. Это особенно проявляется в развитии так называемых умных сетей, когда функции автоматического управления спускаются на уровень распределительных электрических сетей. Для автоматического управления распределительными сетями требуется система управления распределением с соответствующими функциями телекоммуникации, мониторинга, вычислений и управления. Однако, пока не ясно, как будет развиваться соответствующая коммуникационная инфраструктура, с тем чтобы передать предполагаемый объем измерений в систему мониторинга. Таким образом, несмотря на внедрение цифровых устройств будет наблюдаться нерегулярность или отсутствие измерений, требуемых для управления распределительной сетью с использованием детерминированных моделей. Кроме того, актуальность самих моделей (схем замещения) тоже может быть под вопросом. С другой стороны, для выполнения автоматического управления требуется задача оценивания состояния в реальном времени, которая, со своей стороны, нуждается в актуальных взаимно согласованных измерениях. В связи с этим, не-

обходимо внедрить вычислительные методы, компенсирующие отсутствие мониторинга в системах распределения и повышающие точность оценивания состояния. Такими методами могут являться искусственные нейронные сети.

Большинство современных подходов по оцениванию состояния и краткосрочному прогнозированию нагрузки с использованием ИНС рассматривают полносвязанные ИНС с достаточно малым количеством скрытых слоев [1 – 6]. Данный подход зачастую не позволяет реализовать универсальное решение, работающее для различных топологий электрической сети.

Проблемы оценки с полносвязанным слоем:

- Изменение топологии сети, например, в результате коммутаций может приводить к необходимости повторного обучения ИНС.
- Пропадание части входных данных при перебоих в телеизмерениях также значительно меняет картину вектора состояний и соответственно влияет на результаты оценивания состояния, получаемые с помощью нейронной сети.
- Плановое развитие электрической сети тем более приводит к необходимости переобучения ИНС в случае традиционно использовавшихся алгоритмов и архитектур построения ИНС.

Таким образом, задача заключается в том, чтобы получить надежно работающий алгоритм оценивания состояния, устойчивый к значительным изменениям входных данных (как топологии, так и измерений).

2. Модель оценивания состояния

Задача оценивания состояния заключается в получении такого установившегося режима, который был бы наиболее близок к имеющимся измерениям.

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

где \bar{v} – измеренные параметры режима, $v(x)$ – вычисленные параметры режима по значениям вектора состояния системы x .

В классической постановке наиболее распространенным методом решения задачи оценивания состояния является метод взвешенных наименьших квадратов [9, 10], основанный на минимизации следующей целевой функции:

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

Матрица R_v – ковариационная матрица, при отсутствии корреляционной зависимости между различными измерениями представляющая собой диагональную матрицу дисперсий измерений $R_v = \text{diag} \{ \sigma_{v_1}^2, \sigma_{v_2}^2, \dots, \sigma_{v_m}^2 \}$, m – количество измерений.

Результатом оценивания состояния является вектор состояния $x = [\hat{U}_1, \hat{U}_2, \dots, \hat{U}_n, \hat{\delta}_1, \hat{\delta}_2, \dots, \hat{\delta}_n]^T$, содержащий оцененные значения напряжений и углов напряжений для каждого узла, n – общее количество узлов электрической сети.

Входными значениями алгоритма оценивания состояния для каждого рассматриваемого момента времени являются замеренные значения напряжений в узлах электрической сети, перетоки активной и реактивной мощности в ветвях и измерения инъекций активной и реактивной мощности в узлах (нагрузка и генерация). Также при наличии векторных измерений в исходных данных могут присутствовать замеренные углы напряжений. Вектор входных измерений может быть отсортирован в соответствии с оптимальным с точки зрения сходимости порядком узлов [12] или быть неупорядоченным. В любом случае, порядок входных данных не отражает топологию сети. Поэтому при использовании ИНС для решения задачи оценивания состояния, изменения в со-

ставе измерений, в случае пропадания части из них, достаточно сильно влияют на картину получаемую на входном слое ИНС.

Задача динамического оценивания состояния, когда мы рассматриваем не только единичный режим, но и предыдущую историю изменения параметров режима, становится очень близка к задаче прогнозирования параметров режима. Поэтому при анализе возможных вариантов построения ИНС для динамического оценивания состояния анализировались архитектуры ИНС, применяемые, в том числе, и для прогнозирования параметров режима.

3. Рекуррентные искусственные нейронные сети LSTM

В качестве базовой архитектуры ИНС для динамического оценивания состояния была принята рекуррентная нейронная сеть LSTM.

Классические рекуррентные нейронные сети (RNN) [7] плохо справляются с ситуациями, когда необходимо выявить и запомнить долговременные тренды. Кроме того, они достаточно склонны к «переобучению». Так, в работе [Mocanu и др.] показано, что для задачи предсказания данных (при обучении с учителем) ограниченные машины Больцмана, и в частности факторизованные условно ограниченные машины Больцмана (FCRBM) превосходят по точности рекуррентные нейронные сети. Однако, в данной работе рассматривались только классические RNN. Кроме того, в приведенных примерах не указана конкретная архитектура RNN и количество скрытых слоев. При этом, в большинстве случаев результаты сравнения точности на тестовой выборке были близкими. Только на некоторых сериях данных (причем, для среднего временного диапазона) ошибка по RNN была выше. Таким образом, можно сделать вывод, что применение RNN допустимо, но необходимо учитывать иерархичность по времени, т.е. проявление различных трендов в зависимости от рассматриваемого масштаба времени и их наложение друг на друга.

Одна из самых широко известных и часто применяющихся конструкций ячеек рекуррентной сети – это ячейки LSTM (Long Short-Term Memory), что можно перевести, как «долгая краткосрочная память». Архитектура LSTM была предложена еще в середине 1990-х годов [8]. Наиболее часто используемая архитектура LSTM описана в работе Герса Шмидхубера 2000 года [11]. Вместе с тем, подходы, основанные на глубоком обучении, применительно для задач прогнозирования, получили популярность достаточно недавно.

Использование рекуррентных LSTM сетей для прогнозирования нагрузки позволяет получить результаты не уступающие сетям FCRBM [13]. При этом на основе сетей LSTM возможно построение глубоких сетей, что в свою очередь позволяет выявить более сложные зависимости в исходных данных. Что касается выбора конкретного типа ячеек рекуррентной сети, то в пользу выбора LSTM можно привести работу [14], в которой проанализированы различные варианты RNN (в частности LSTM и GRU) и выявлено, что LSTM с применением технологии dropout дает лучший результат на всех тестах.

Структурная схема ячейки LSTM показана на рисунке 1. В LSTM есть три основных вида элементов, называемых «гейтами»: входной (i_t), забывающий (f_t) и выходной (o_t), а также рекуррентная ячейка со скрытым состоянием c_t . Если обозначить через x_t – вектор входной информации во время t , через h_t – вектор скрытого состояния, W – матрицы весов для соответствующего уровня, определяющегося индексом, b –

векторы свободных членов, то можно сформулировать математические преобразования, выполняемые ячейкой LSTM следующим образом.

$$c'_t = \tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_{c'}),$$

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + b_i),$$

$$f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + b_f),$$

$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + b_o),$$

$$c_t = f_t \circ c_{t-1} + i_t \circ c'_t.$$

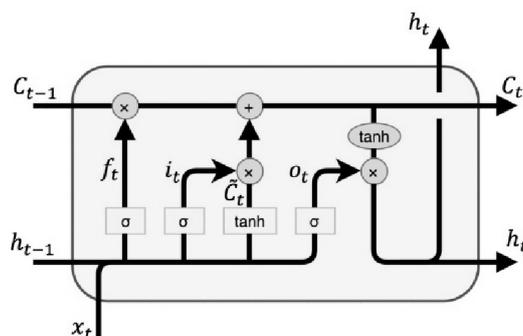


Рис. 1. Ячейка LSTM.

Такая архитектура ячейки LSTM решает проблему исчезающих градиентов, которая мешала рекуррентным сетям обучать долгосрочные зависимости.

4. Применение Байесовских нейронных сетей

Одной из важных особенностей исходных данных для оценивания состояния является их вероятностный характер. Входные измерения, как правило, задаются не просто значением, но также сопровождаются дисперсиями. Это дает дополнительную информацию для обучения нейронной сети. Кроме того, если на выходе нейронной сети мы будем иметь не только значения параметров режима, но и распределения вероятностей для этих параметров, то такая информация позволит принимать более обоснованные решения по управлению режимами на основе результатов оценивания состояния. Таким образом, обоснованным будет применение Байесовских ИНС.

Построить Байесовскую нейронную сеть, учитывающую вероятности параметров ИНС, можно применив метод Байесовского обратного распространения (Bayes By Backprop или сокращенно BBB). Метод Bayes by Backprop (BBB) [15, 16] позволяет получить апостериорное распределение весов нейронной сети $\theta \in \mathbf{R}^d$, где d – размерность пространства параметров нейронной сети [17]. Это распределение, как правило, является распределением Гаусса.

Пусть $\log p(y|\theta, x)$ – логарифм правдоподобия модели, тогда обучение модели можно выполнять минимизируя свободную вариационную энергию:

$$(1) \quad \ell(\theta) = E_{q(\theta)} \left[\log \frac{q(\theta)}{p(y|\theta, x)p(\theta)} \right],$$

где $p(\theta)$ – априорно заданная дисперсия параметров.

Минимизация (1) эквивалентна максимизации логарифма правдоподобия $p(y | \theta, x)$ с учетом расстояния Кульбака-Лейблера между вычисленной плотностью вероятности параметров ИНС и действительной апостериорной вероятностью [18, 19]:

$$\ell(\theta) = -E_{q(\theta)}[\log p(y | \theta, x)] + \text{KL}[q(\theta) \| p(\theta)].$$

Данную минимизацию можно рассматривать, как затухание отклонения параметров к математическому ожиданию, регулируемое стандартным отклонением параметров ИНС. Применение метода Bayes By Backprop для рекуррентных нейронных сетей описано в [20].

5. Архитектура ИНС

Архитектура нейронной сети для оценивания состояния с применением глубокого обучения представлена на рис. 2

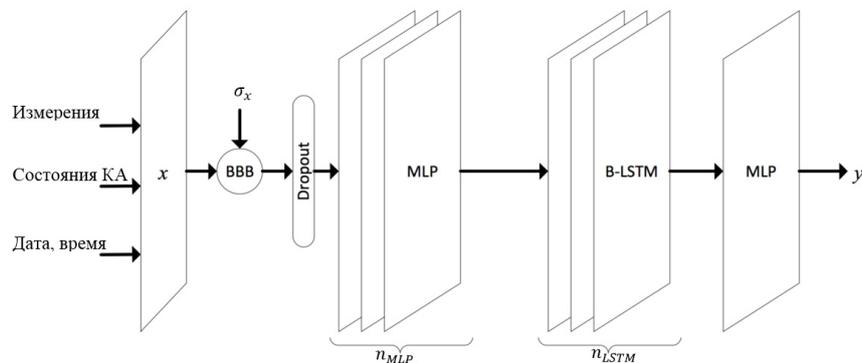


Рис. 2. Архитектура нейронной сети для глубокого оценивания состояния.

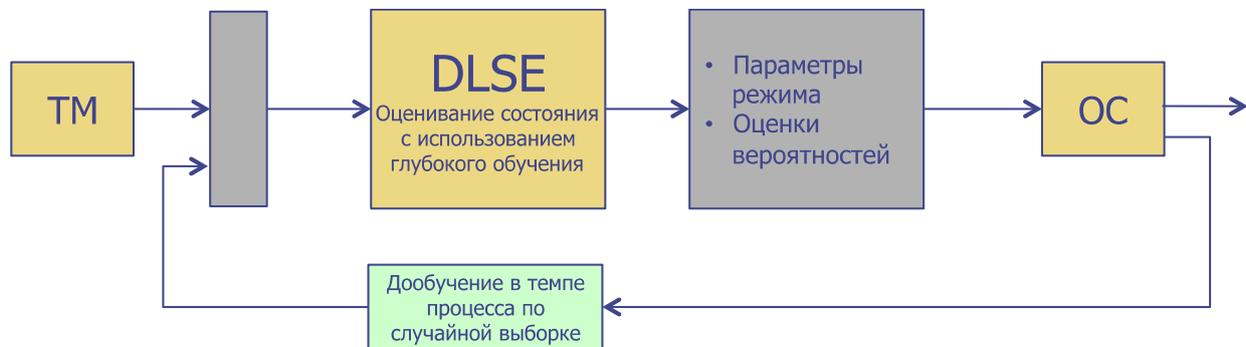


Рис. 3. Архитектура программного обеспечения для динамической оценки состояния.

Особенности применяемой архитектуры:

1) Входной полносвязанный слой. Возможно несколько входных слоев, для того, чтобы сформировать устойчивую взаимосвязь между параметрами одного режима.

Весы Байесовской нейронной сети инициализируются с использованием случайного значения r . Если для скрытых слоев случайное значение задается нормальным распределением $r \sim N(0,1)$, то для входного слоя нормальное распределение ограничено дисперсией измерений, являющихся входными данными для нейронной сети $r \sim N(0, \sigma^2)$.

2) В модели применяется несколько скрытых слоев, сети LSTM, что позволяет выявить более сложные взаимосвязи параметров в разрезе времени. При этом на вход

ИНС подается также информация о текущем моменте времени. Подобный подход применялся в [21] только для сетей прямого распространения типа MLP.

3) Полносвязанный выходной слой, обеспечивает масштабирование выходного вектора, не зависящее от времени.

Проблему учета коммутаций с одной стороны частично решает Dropout на входном слое. Вместе с тем наиболее продуктивным оказался подход имитационного моделирования с перебором коммутационного состояния от текущего режима методом Монте-Карло с последующим моделированием с использованием классического оценивания состояния. Архитектура динамической оценки состояния представлена на рис. 3.

6. Заключение

Предложенная методика опробована на данных реальной энергосистемы. В результате разработан программный комплекс, позволяющий проводить оценивание состояния в темпе процесса для сети с наличием зашумленных или несогласованных во времени данных. При этом ИНС, используемая в данном комплексе, постоянно дообучается при работе системы, используя результаты детерминированного оценивания состояния на основе архивных выверенных данных.

Список литературы

1. Manitsas E. et al. Distribution system state estimation using an artificial neural network approach for pseudo measurement modeling // *IEEE Transactions on Power Systems*. 2012. Vol. 27, No. 4. P. 1888-1896.
2. Ashraf S. M. et al. Voltage stability monitoring of power systems using reduced network and artificial neural network // *International Journal of Electrical Power & Energy Systems*. 2017. Vol. 87. P. 43-51.
3. Glavic M., Fonteneau R., Ernst D. Reinforcement learning for electric power system decision and control: Past considerations and perspectives // *IFAC-PapersOnLine*. 2017. Vol. 50, No. 1. P. 6918-6927.
4. Mocanu E. et al. Deep Learning for Power System Data Analysis // Chapter in book Arghandeh R., Zhou Y. (ed.). *Big Data Application in Power Systems*. Elsevier, 2017. P. 125-158.
5. Hernández L. et al. Artificial neural networks for short-term load forecasting in microgrids environment // *Energy*. 2014. Vol. 75. P. 252-264.
6. Mosbah H., El-Hawary M.E. Optimization of neural network parameters by Stochastic Fractal Search for dynamic state estimation under communication failure // *Electric Power Systems Research*. 2017. Vol. 147. P. 288-301.
7. Николенко С., Кадурин А., Архангельская Е. Глубокое обучение. Погружение в мир нейронных сетей // СПб.: Питер, 2018. С. 232-242.
8. Hochreiter S., Schmidhuber J. Long short-term memory // *Neural computation*. 1997. Vol. 9, No. 8. P. 1735-1780.
9. Monticelli A. State estimation in electric power systems: a generalized approach. Springer Science & Business Media, 2012.
10. Гамм А.З. Статистические методы оценивания состояния электроэнергетических систем. Наука, 1976.
11. Gers F. A., Schmidhuber J. Recurrent nets that time and count // *Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks. IJCNN 2000. Neural Computing: New Challenges and Perspectives for the New Millennium*. IEEE, 2000. Vol. 3. P. 189-194.
12. Ersavas B. F. Sparse Matrix Ordering and Gaussian Elimination // *ECE*. 2002. Vol. 3652. P. 1-12.
13. Marino D. L., Amarasinghe K., Manic M. Building energy load forecasting using deep neural networks // *Industrial Electronics Society, IECON 2016 – 42nd Annual Conference of the IEEE*. IEEE, 2016. P. 7046-7051.
14. Jozefowicz R., Zaremba W., Sutskever I. An empirical exploration of recurrent network architectures // *International Conference on Machine Learning*. 2015. P. 2342-2350.
15. Graves A. Practical variational inference for neural networks // *Advances in neural information processing systems*. 2011. P. 2348-2356.

16. Blundell C. et al. Weight uncertainty in neural networks // arXiv preprint arXiv:1505.05424. 2015.
17. Wainwright M. J. et al. Graphical models, exponential families, and variational inference // Foundations and Trends® in Machine Learning. 2008. Vol. 1, No. 1-2. P. 1-305.
18. Graves A. Practical variational inference for neural networks // Advances in neural information processing systems. 2011. P. 2348-2356.
19. Blundell C. et al. Weight uncertainty in neural networks // arXiv preprint arXiv:1505.05424. 2015.
20. Fortunato M., Blundell C., Vinyals O. Bayesian recurrent neural networks // arXiv preprint arXiv:1704.02798. 2017.
21. Ryu S., Noh J., Kim H. Deep neural network based demand side short term load forecasting // Energies. 2017. Vol. 10, No. 1. <https://doi.org/10.3390/en10010003>