

УДК 658.26

ПРИМЕНЕНИЕ АГЛОМЕРАТИВНО-ИЕРАРХИЧЕСКОГО МЕТОДА КЛАССИФИКАЦИИ В СИСТЕМАХ ОПЕРАТИВНОГО ПРОГНОЗИРОВАНИЯ ЭЛЕКТРИЧЕСКИХ НАГРУЗОК ПРОМЫШЛЕННЫХ ПРЕДПРИЯТИЙ

Васильев Д.А., Сарафанов И.А.

Саратовский государственный технический университет, г. Саратов

E-mail: iwansarafanov@gmail.com

Аннотация. В статье рассматривается применение агломеративно-иерархического метода в системе оперативного прогнозирования электрических нагрузок промышленных предприятий. Данный метод позволяет сформировать классы схожих по статистическим характеристикам выборок электрических нагрузок, которые позволят определить прогнозное значение нагрузки.

Ключевые слова: промышленное предприятие, электрическая нагрузка, оперативное прогнозирование, кластерный анализ, метод прогнозирования, агломеративно-иерархический метод.

Постановка задачи

Существенное место в задаче автоматизации управления нагрузкой промышленного предприятия занимает проблема ее адаптивного прогнозирования, которая является обоюдно важной как для производителя, так и для потребителя электроэнергии.

Проблема прогнозирования нагрузки на промышленных предприятиях возникла в связи с необходимостью выравнивания графиков электрических нагрузок и образованием дефицита по активной мощности в энергетических системах. Для производителей электроэнергии прогноз ожидаемой нагрузки важен с точки зрения оптимизации операционных расходов, резервирования мощностей, удобства проведения профилактических работ и обеспечения безопасности систем. Потребителю такой прогноз необходим для осуществления минимизации своих издержек, в частности, для избежания простоя производства в случае нехватки необходимой мощности, либо уплаты штрафов при превышении лимитов и переплаты за заказанные, но не полностью израсходованные объемы мощности [1,2,3].

В данной статье предлагается на основе математического аппарата кластерного анализа, и в частности, агломеративно-иерархического метода, формировать классы «близких» по статистическим характеристикам выборок электрических нагрузок, являющихся основой для вычисления прогнозных значений нагрузок.

Имеется множество $\{P_1, \dots, P_s, \dots, P_n\}$ реализаций электрических нагрузок, измеренных в дискретные моменты времени $t=r\Delta t$, где $r=0, \dots, R$ – порядковый номер интервала Δt (Δt – интервал сбора информации в системе).

Каждая из реализаций содержит L нагрузок и $(L - k - t^* + 1)$ выборок объемом k , где t^* – интервал прогнозирования, равный количеству интервалов, на которое осуществляется прогноз.

Тогда множество выборок нагрузок по всем реализациям

$$P^p = \bigcup_{s=1}^n P_s = \left\{ P_1^p, \dots, P_j^p, \dots, P_{n(L-k-t^*+1)}^p \right\}.$$

Каждой выборке P_j^p ставится в соответствие фактическое значение нагрузки $P_j^\phi \in \left\{ P_1^\phi, \dots, P_j^\phi, \dots, P_{n(L-k-t^*+1)}^\phi \right\}$, имеющее место в момент времени $k+t^*$, на который выполняется прогноз и вектор признаков $X_j=(x_{1j}, \dots, x_{ij}, \dots, x_{qj})$, характеризующий выборку.

Множество выборок P^p разбивается на классы в пространстве признаков X . Для построения признакового пространства выбираются следующие пять признаков:

1. Последнее значение электрической нагрузки текущей выборки;
2. Среднее значение нагрузки по текущей выборке $P_{cp} = \frac{1}{k} \sum_{i=1}^k P_i$;
3. Дисперсия нагрузки по текущей выборке $D = \frac{1}{k-1} \sum_{i=1}^k (P_i - P_{cp})^2$;
4. Коэффициент неравномерности выборки $K_n = \frac{P_{\max} - P_{\min}}{P_{cp}}$;
5. Коэффициент заполнения $K_z = \frac{P_{cp}}{P_{\max}}$

Построение классов выборок нагрузок позволяет выделить их относительно однородные статистические совокупности, отражающие различные условия функционирования электроэнергетики промышленных предприятий.

Агломеративно-иерархический метод

Построение классов выборок может быть осуществлено с помощью агломеративно-иерархического метода классификации, основой при реализации которого служит матрица расстояний между объектами (выборками нагрузок P^p_j).

Основная идея метода заключается в последовательном объединении группируемых объектов — сначала самых близких, а затем все более удаленных друг от друга. Процедура построения классификации состоит из последовательных шагов, на каждом из которых производится объединение двух ближайших групп объектов (кластеров). Близость двух кластеров определяется как среднее значение расстояния между всеми такими парами объектов, где один объект пары принадлежит к одному кластеру, а другой – к другому

$$D_{pq}^2 = \sum_{i \in X_p} \sum_{j \in X_q} \frac{d_{ij}^2}{n_p n_q},$$

где D_{pq}^2 – мера близости между p -м и q -м кластерами; d_{ij} – евклидово расстояние между i -м и j -м объектами; X_p – p -й кластер; X_q – q -й кластер; n_p, n_q – число объектов в p -м и q -м кластерах, соответственно.

На первом шаге процедуры агломеративно-иерархического метода рассматривается начальная матрица расстояний между объектами, по которой определяется минимальное расстояние; далее, наиболее близкие объекты, находящиеся друг от друга на этом расстоянии, объединяются в один кластер, в матрице вычеркиваются строка и столбец, соответствующие первому из этих объектов, а расстояния от нового кластера до всех остальных кластеров (на первом шаге – объектов) вычисляются по вышеприведенной формуле – как средние из расстояний от объектов первого кластера до всех остальных. Эти вновь вычисленные значения заносятся в строку и столбец матрицы расстояний, соответствующие второму объекту из первого кластера.

На втором шаге по матрице расстояний, уменьшенной на одну строку и один столбец, снова определяют минимальное расстояние и формируют новый кластер. Этот кластер может быть построен в результате объединения либо двух объектов, либо одного объекта с кластером, построенным на первом шаге. Далее, в матрице расстояний вычеркиваются одна строка и один столбец, а одна строка и один столбец пересчитываются и т.д.

Таким образом, агломеративно-иерархический метод кластерного анализа включает $n-1$ аналогичных шагов. При этом после выполнения каждого шага число кластеров уменьшается на единицу, а матрица расстояний уменьшается на одну строку и один столбец.

При прогнозировании электрической нагрузки определяется класс $l \in J_l$ (J_l – количество выборок нагрузок, составляющих l -й класс), к которому принадлежит выборка, используемая для прогнозирования.

Прогнозное значение нагрузки $p(t+t^*)$ определяется как среднее арифметическое значений p_j^ϕ , соответствующих классу, к которому отнесена распознаваемая выборка

$$p(t+t^*) = \frac{1}{J_l} \sum_{j=1}^{J_l} p_j^\phi.$$

Результаты работы

На рис. 1, показаны результаты прогнозирования нагрузки с помощью агломеративно-иерархического метода.



Рис. 1. Результаты прогнозирования нагрузки с помощью агломеративно-иерархического метода.

Анализ результатов работы показал, что прогнозирование электропотребления промышленного предприятия при объемах выборок 5-7 нагрузок на интервалы времени до 3 минут в зависимости от выделенного класса нагрузок обеспечивает точность прогнозирования в пределах от 1,3% до 5,1%.

Литература

1. Васильев Д.А. Концепция формирования структуры АСУ энергетикой промышленных предприятий с непрерывным характером производства / Д.А. Васильев, В.А. Иващенко, А.Ф. Резчиков // Проблемы точной механики и управления: Сб. науч. тр. Саратов: ИПТМУ РАН, 2002. С.108–116.
2. Васильев Д.А. Выбор модели прогнозирования электрической нагрузки промышленных предприятий с непрерывным характером производства / Д.А.Васильев, В.А.Иващенко, А.Ф.Резчиков, Е.С.Курышова // Электротехнические комплексы и силовая электроника. Анализ, синтез и управление: Межвуз. науч. сб. Саратов: ИПТМУ РАН, 2001. С.36–38.
3. Большов Л.А. Прогнозирование электропотребления: Современные подходы и пример исследования / Л.А. Большов, М.Ф. Каневский, Е.А. Савельева и др. // Изв. Академии наук. Энергетика 2004 №6 С.74–93.