

УДК 004.032.26 + 338.27
ББК 65.23.22.18.22.172.31.29

ПРИМЕНЕНИЕ КОННЕКТИВИСТСКИХ СИСТЕМ ДЛЯ ПРОГНОЗИРОВАНИЯ ПОТРЕБЛЕНИЯ ЭЛЕКТРОЭНЕРГИИ¹

**Камаев В.А.², Щербаков М.В.³, Панченко Д.П.⁴,
Щербакова Н.Л.⁵**

*(Волгоградский государственный технический
университет, Волгоград, Россия)*

Бребельс А.⁶

(Высшая школа Кемпен, Гиль, Бельгия)

В статье рассматривается решение проблемы краткосрочно-го прогнозирования потребления электроэнергии в торговых центрах Голландии и Бельгии. Предлагается методика прогнозирования, основанная на использовании коннективистских систем. Дается общее описание методики прогнозирования и специфика ее применения для рассматриваемой задачи прогнозирования. Приводится обобщение, описание и результаты применения различных коннективистских моделей: модели

¹ Работа поддержана грантом РФФИ № 10-07-97008-р_поволжье_a

² Камаев Валерий Анатольевич, доктор технических наук, профессор (Волгоград, пр. Ленина, д. 65, тел. (8442) 24-81-00, kamaev@unix.cad.vstu.ru).

³ Щербаков Максим Владимирович, кандидат технических наук (Волгоград, пр. Ленина, д. 65, тел. (8442) 24-81-00, vstu@gebeus.ru).

⁴ Панченко Дмитрий Петрович, кандидат технических наук (panchenkodr@gmail.com).

⁵ Щербакова Наталия Львовна, кандидат технических наук (snl@gebeus.ru).

⁶ Бребельс Адриаан, руководитель проектов высшей школы Кемпен (Гиль, Бельгия, adriaan.brebels@portacapena.com).

скользящего среднего, модели линейной регрессии, модели нейронных сетей. Кроме этого показаны изменения в качестве прогноза в зависимости от различных наборов входных переменных. Обсуждаются результаты и приводятся выводы о специфике применения той или иной модели.

Ключевые слова: прогнозирование потребления электроэнергии, коннективизм, коннективистские системы, скользящее среднее, линейная регрессия, нейронные сети.

1. Введение в проблему потребления электроэнергии

В работе рассматривается актуальная проблема прогнозирования суточного потребления электроэнергии в торговых центрах, размещенных в Бельгии и Голландии. Оплата за электроэнергию, как правило, является одной из основных расходных статей бюджета и управляющим необходимо как можно точнее определить возможное потребление электроэнергии для формирования заявок компаниям, поставляющим электроэнергию. Адекватный прогноз позволяет оценить затраты на электроэнергию, оценить временные интервалы в которых ожидается максимальная нагрузка на сеть. Кроме этого, повышение точности прогнозирования, может рассматриваться как один из источников экономии энергоресурсов [1].

Рассмотрим специфики данной предметной области. Первая заключается в зависимости потребления электроэнергии от графика работы (часы открытия, закрытия, выходные дни и праздники). Так потребление в ночные часы или в праздники может отличаться от дневного в 10-12 раз. Вторая связана с тем, что данные о потреблении электроэнергии собираются и передаются на сервер сбора данных (система EcoSCADA, разработанная Бельгийско-Польской компанией Порта Капена) каждые 15 минут [2]. Таким образом, в течение дня число временных отсчетов со значениями потребления равно 96. Третья специфи-

ка связана с пропусками в данных. В EcoSCADA также учитывается информация о параметрах, влияющих на потребление электроэнергии, таких как, погодные условия (температура воздуха, влажность, интенсивность солнечного освещения) и условия использования здания (время открытия, закрытия, календарь).

Будем считать, что имеющиеся временные ряды являются интервальными, полными и случайными.

2. Формализация коннективистских систем

Коннективистская система представляет собой множество связанных между собой элементов обработки информации e , объединенных в слой l [3]. Каждый элемент выполняет преобразование F (часто функцию взвешенного суммирования и нелинейного преобразования) над входными сигналами, поступающих от элементов, связанных с рассматриваемым элементом через веса r . Кроме этого, выделен функционал оптимизации Φ на основе функций потерь.

$$(1) S = \langle e, r, l, F, \Phi, P \rangle.$$

Схематично такая структура представлена на рисунке 1.

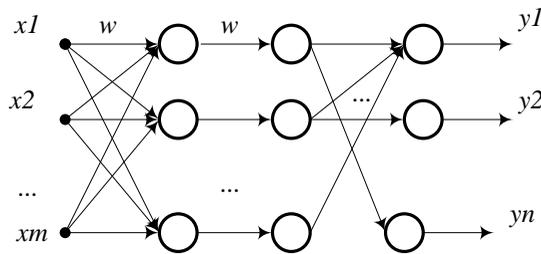


Рис. 1. Структура коннективистской системы

Значения весов связей r могут настраиваться в процессе P минимизации функционала Φ .

Таким образом, формально коннективистская система при данной формализации обобщает в себе различные модели: например скользящего среднего (рис 2 а), линейной регрессии (рис 2 б), нейросетевые модели (рис 1).

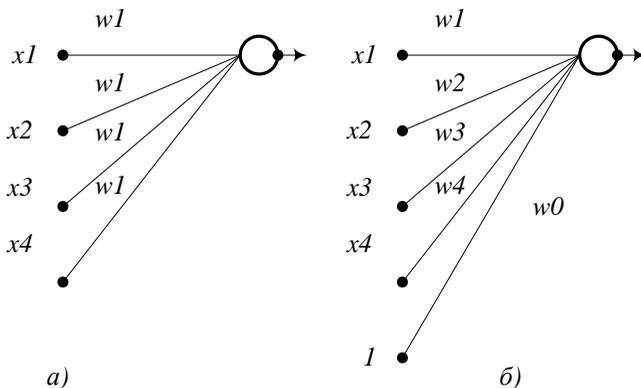


Рис. 1. Структура коннективистских систем: а) скользящего среднего: элемент осуществляет взвешенное суммирование входов и $w_1 = 0.25$; б) линейной регрессии: элемент осуществляет взвешенное суммирование входов и весов, где w_0 – свободный член

Отметим, что коннективистская система, в общем случае, может представлять собой совокупность «слотов» или фреймов, в которых располагаются различные модели, а решение формируется на основе консолидации модели (рис. 3).

Такое обобщение позволяет автоматизировать процесс подбора прогнозных моделей, в процессе которого уменьшается значение функции потерь.

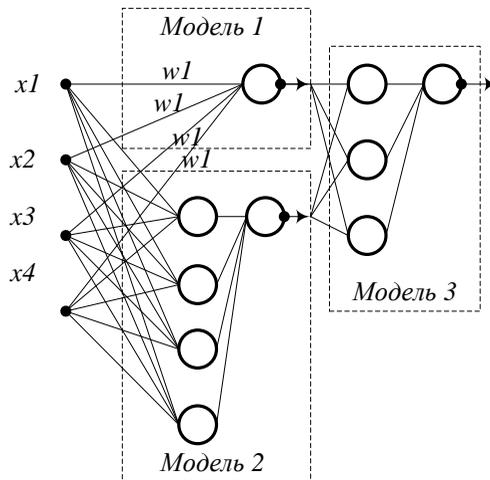


Рис. 3. Пример интеграции 2-х коннективистских моделей в единую сеть. Консолидированное решение принимается нейронной сетью (модель №3)

3. Методика формирования прогнозной модели на основе коннективистской системы

3.1. ОПИСАНИЕ МЕТОДИКИ

Для применения коннективистских моделей предлагается методика формирования прогнозной модели, основанная как на классических подходах к прогнозированию электроэнергии, так и на подходах извлечения знаний из данных. Ниже дано общее описание методики.

1. Формирование цели создания модели прогнозирования или определение функции потерь. Для удовлетворения потребности пользователя в анализе ситуации о потреблении электроэнергии в режиме реального времени важнейшим моментом является отображение отклонения прогнозируемого сигнала от реального (измеренного) значения. Таким образом, первой целью создания модели является минимизация невязки в каждой точке горизонта прогнозирования, т.е. в качестве функции

потерь используется среднее квадратичное отклонение. Другая оценка связана с дневной стоимостью электроэнергии, рассчитанной как произведение потребленной электроэнергии (кВт·ч) и тарифа. Таким образом, качество модели определяется разностью прогнозируемых и реальных затрат на потребление электроэнергии.

2. Анализ потребителей электроэнергии и выявление прямых и косвенных признаков, влияющих на потребление. Потребление электроэнергии связано с деятельностью людей (производства) и прямо пропорционально числу электроприборов. Как правило, информация о числе приборов, их технических характеристиках, программах функционирования не учитывается в SCADA системах. В связи с этим возникает задача определения наблюдаемых параметров, влияющих на работу электроприборов. Целью данного этапа является принятие решения о формировании множества независимых переменных для включения в модель.

3. Предварительный анализ данных. Данный шаг содержит 4 основных процедуры: исключение аномалий, восстановление пропусков в данных, кодирование категориальных переменных и нормализация. Здесь следует ввести понятие выброса (аномалии), под которым будем понимать отклонение потребления электроэнергии в текущий момент времени от среднего на N кВт·ч в течении T дискретных временных отсчетов. Значения N и T устанавливаются управляющим в процессе наблюдения за потреблением электроэнергии. Самая простая процедура - исключение аномальных значений. В этом случае проблема аномалий сводится к решению проблемы заполнения пропусков в данных. Восстановление пропусков осуществляется посредством подстановки среднего значения из значений потребленной электроэнергии в предыдущие моменты времени.

4. Построение модели и ее настройка. На данном шаге осуществляется выбор типа модели, составление плана экспериментов испытания модели, параметрическая оптимизация модели.

5. Определение качества модели. Последний шаг связан с определением качества модели и принятие решения о качестве модели в соответствии с целями прогнозирования (функции потерь). Выделим следующие классические оценки определения точности модели: среднеквадратичное отклонение RMSE:

$$(2) \quad RMSE = \sqrt{\frac{\sum_{i=1}^n (y_m(i) - y_p(i))^2}{n}},$$

Где y_m – измеренное значение энергопотребления, y_p - прогнозное значение, n – горизонт прогнозирования ($n=96$), средняя абсолютная процентная ошибка (MAPE):

$$(3) \quad MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_m(i) - y_p(i)}{y_m(i)} \right|,$$

индекс корреляции R , значение дисперсии модуля разницы прогнозного и реального значений EV .

3.2. СПЕЦИФИКА МЕТОДИКИ ПРИ ПРОГНОЗИРОВАНИИ ПОТРЕБЛЕНИЯ ЭЛЕКТРОЭНЕРГИИ В ТОРГОВЫХ ЦЕНТРАХ

Отметим специфику применения модели для прогнозирования потребления электроэнергии в торговых центрах.

Данная методика формально относится к методике краткосрочного прогнозирования, однако, горизонт прогнозирования равен 96 точкам.

В качестве функции потерь выбрано среднеквадратичное отклонение прогнозного и измеренного значения потребления электроэнергии.

Потребители электроэнергии выделяются в 2 группы: системы климат контроля (HVAC системы) и другие. Все электроприборы связаны с графиком работы торгового центра. HVAC системы зависят от состояния внешней среды: температуры воздуха, относительной влажности, программы работы HVAC и других параметров.

Каждое здание имеет график работы: время открытия, время закрытия и календарь выходных и праздничных дней.

Построение коннективистских систем носит итерационный характер. Если модель не удовлетворяет установленному критерию качества, то осуществляется возврат к шагу 2 методики и/или изменение структуры модели.

Для тестирования моделей были использованы разработки авторов [3, 4, 5].

4. Применение коннективистских систем для прогнозирования

3.1. МОДЕЛЬ СКОЛЬЗЯЩЕГО СРЕДНЕГО

Предполагается, что потребление энергии на текущий момент может быть рассчитано как среднее значение энергопотребления в тот же момент времени одну, две, три и четыре недели назад.

Входные переменные (независимые): 1) значение потребления энергии ($t-1$ неделя); 2) значение потребления энергии ($t-2$ недели); 3) значение потребления энергии ($t-3$ недели); 4) значение потребления энергии ($t-4$ недели).

Выходные переменные (зависимые): 1) потребление энергии.

Таблица 1. Сводные оценки для модели скользящего среднего

Критерий	Значение критерия
RMSE, kWh	37,4
MAPE	25,4%
R	0,9258
EV	0,5897

3.2.МОДЕЛЬ ЛИНЕЙНОЙ РЕГРЕССИИ (ЛР)

3.2.1 МОДЕЛЬ ЛР, ВКЛЮЧАЮЩАЯ ТОЛЬКО ПОТРЕБЛЕНИЕ ЭНЕРГИИ

Предполагается зависимость потребления энергии от текущей температуры, энергопотребления и температуры одну, две, три и четыре недели назад.

Входные переменные (независимые): 1) значение внешней температуры T ($t-1$ неделя); 2) значение внешней температуры T ($t-2$ недели); 3) значение внешней температуры T ($t-3$ недели); 4) значение внешней температуры T ($t-4$ недели); 5) текущее значение внешней температуры T , 6) значение потребления энергии ($t-1$ неделя); 7) значение потребления энергии ($t-2$ недели); 8) значение потребления энергии ($t-3$ недели); 9) значение потребления энергии ($t-4$ недели).

Выходные переменные (зависимые): потребление энергии.

Таблица 2. Сводные оценки для модели ЛР

Критерий	Значение критерия обучающей выборки	Значение критерия тестовой выборки
RMSE, kWh	35,28	24,89
MAPE	38%	21,65%
R	0,9342	42,77%
EV	0,4657	0,1137

3.2.2. ЛР МОДЕЛЬ, ВКЛЮЧАЮЩАЯ ПОТРЕБЛЕНИЕ ЭНЕРГИИ, ВНЕШНИЕ ВОЗМУЩЕНИЯ И ПАРАМЕТРЫ ИСПОЛЬЗОВАНИЯ ЗДАНИЙ

Предполагается зависимость потребления энергии от текущей температуры, энергопотребления и температуры в тот же момент времени одну и две недели назад. Кроме того, в параметры линейной регрессии добавляется информация о профиле использования здания на текущий момент и предыдущие мо-

менты времени (неделю назад, и две недели назад) и статус (открытый / закрытый).

Входные переменные (независимые): 1) значение внешней температуры T ($t-1$ неделя); 2) значение внешней температуры T ($t-2$ недели); 3) текущее значение внешней температуры; 4) статус; 5) профиль использования здания; 6) профиль использования здания ($t-1$ неделя); 7) профиль использования здания ($t-2$ недели); 8) значение потребления энергии ($t-1$ неделя); 9) значение потребления энергии ($t-2$ недели).

Выходные переменные (зависимые): потребление энергии.

Таблица 4. Сводные оценки для модели ЛР с учетом внешних возмущений и параметра использования зданий

Критерий	Значение критерия обучающей выборки	Значение критерия тестовой выборки
RMSE, kWh	22,89	18,49
MAPE	19,27%	17,96%
R	0,9728	0,9832
EV	0,1021	0,0192

3.2.3 МОДЕЛЬ КУСОЧНО НЕПРЕРЫВНОЙ ЛИНЕЙНОЙ РЕГРЕССИИ

Предполагается зависимость потребления энергии от текущей температуры, энергопотребления и температуры одну, две, три и четыре недели назад. Каждый день разбивается на несколько фрагментов. Фрагментам соответствуют разные динамики изменения энергопотребления.

Входные переменные (независимые): 1) значение внешней температуры T ($t-1$ неделя); 2) значение внешней температуры T ($t-2$ недели); 3) значение внешней температуры T ($t-3$ недели); 4) значение внешней температуры T ($t-4$ недели); 5) текущее значение внешней температуры; 6) значение потребления энергии ($t-1$ неделя); 7) значение потребления энергии ($t-2$ недели);

8) значение потребления энергии ($t-3$ недели); 9) значение потребления энергии ($t-4$ недели).

Выходные переменные (зависимые): потребление энергии.

Таблица 3. Сводные оценки для модели КНЛР

Критерий	Значение критерия обучающей выборки	Значение критерия тестовой выборки
RMSE, kWh	16,32	46,46
MAPE	14,88%	21,65%
R	0,9327	0,8131
EV	0,3996	0,7214

3.2.4. ЛР МОДЕЛЬ, ВКЛЮЧАЮЩАЯ КОМБИНАЦИЮ ПАРАМЕТРОВ

Предполагается зависимость потребления энергии от нелинейной комбинации следующих параметров: текущей температуры, энергопотребления и профиля использования здания в тот же момент времени одну и две, три и четыре недели назад, статуса здания (открытый / закрытый).

Таблица 5. Сводные оценки для модели ЛР, включающей комбинацию параметров

Критерий	Значение критерия обучающей выборки	Значение критерия тестовой выборки
RMSE, kWh	20,93	20,2
MAPE	18,97%	27,59%
R	0,9773	0,98
EV	0,0863	0,1336

3.3. МОДЕЛЬ НА ОСНОВЕ ИСКУССТВЕННЫХ НЕЙРОННЫХ СЕТЕЙ (НС)

3.3.1. НС МОДЕЛЬ, ВКЛЮЧАЮЩАЯ ПОТРЕБЛЕНИЕ ЭНЕРГИИ И ПАРАМЕТР ИСПОЛЬЗОВАНИЯ ЗДАНИЙ

Предполагается зависимость потребления энергии от текущей температуры, энергопотребления и профиля использования здания в тот же момент времени одну и две, три и четыре недели назад. Также предполагается использование статуса (открытый / закрытый).

Таблица 6. Сводные оценки для модели НС

Критерий	Значение критерия обучающей выборки	Значение критерия тестовой выборки
RMSE, kWh	17,97	14,78
MAPE	14,01%	11,06%
R	0,9833	0,9893
EV	0,05	0,0099

3.3.2. НС МОДЕЛЬ, ВКЛЮЧАЮЩАЯ КОМБИНАЦИЮ ПАРАМЕТРОВ

Предполагается зависимость потребления энергии аналогично п. 3.2.4.

Таблица 7. Сводные оценки для модели НС, включающей комбинацию параметров

Критерий	Значение критерия обучающей выборки	Значение критерия тестовой выборки
RMSE, kWh	12,36	24,51
MAPE	11,15%	23,17%
R	0,9921	0,9704
EV	0,0335	0,1118

5. Результаты и выводы

5.1. РЕЗУЛЬТАТЫ

Приведем основные результаты применения моделей для решения задачи прогнозирования.

1. Модель скользящего среднего является наименее трудоемкой в вычислениях, однако обладает самой высокой из рассмотренных методов ошибкой (RMSE 37,4 kWh, MAPE 25%), за исключением кусочно-непрерывной линейной регрессии на тестовой выборке (RMSE на 24,2%). Существенным недостатком данного метода является высокое влияние исключений в календаре, которые необходимо учитывать при формировании четырех предыдущих значений энергопотребления. Таким образом, если, к примеру, один из предыдущих четырех понедельников был нерабочим, а значения энергопотребления этого дня будут задействованы для получения прогноза на рабочий понедельник, то прогнозируемое энергопотребление будет существенно занижено. Более того, данный способ прогнозирования не имеет целевых показателей, и поэтому не может быть использован в модели «Что-Если» для выявления потенциала энергосбережения.

2. Модель линейной регрессии обладает меньшей ошибкой (RMSE на 5,6%, EV на 12,4%) по сравнению с предыдущей моделью, однако обладает теми же недостатками. Модель не может быть использована для построения модели «Что-Если» в силу отсутствия управляющего (целевого) регрессора.

3. Кусочно-непрерывная линейная регрессия. Благодаря выделению фрагментов со схожей динамикой изменения энергопотребления этот метод позволяет существенно снизить ошибку прогнозирования (RMSE в 2,2 раза, MAPE на 23,12%). Однако данный метод является чувствительным к расписанию. Так, если процессы, характерные для одного фрагмента, будут протекать в отрезок времени, характерный для другого фрагмента профиля энергопотребления, это вызовет ошибку в про-

гнозировании. Более того, если через неделю, эти данные будут использованы для прогнозирования, это может вызвать еще большую ошибку, что прослеживается по показателям качества модели на тестовой выборке. Этот недостаток привел к невозможности построения уравнения регрессии для определенных участков профиля энергопотребления в силу нестабильности процессов, проходящих в здании в данные промежутки времени.

4. Линейная регрессия с профилем использования здания. Внесение более подробных знаний об уровне использования здания позволило использовать одно уравнение линейной регрессии для всех участков профиля энергопотребления в течение дня, а также дало потенциал для использования одного уравнения для разных дней недели. Более того, это позволило избавиться от недостатков предыдущих моделей, таких, как зависимость от календаря и расписания: в данной модели календарь и расписание являются управляющими факторами. Показатели качества модели свидетельствуют об удовлетворительных обобщающих свойствах модели (RMSE на тестовой выборке на 19% ниже чем на обучающей выборке, MAPE на тестовой выборке на 1,31% ниже чем на обучающей выборке). Модель может быть использована не только для прогнозирования, но и для объяснения прогноза, поскольку линейный характер уравнения позволяет проследить степень влияния факторов на прогнозируемое значение.

5. Линейная регрессия с нелинейной комбинацией параметров. Применение подхода нелинейной комбинации к параметрам, которые были использованы для предыдущего метода, позволило сократить ошибку прогнозирования на обучающем множестве (RMSE снизился на 8,6%, MAPE снизился на 0,3%), однако, повышение количества регрессоров отрицательно сказалось на обобщающих способностях модели (RMSE на тестовой выборке выросла на 9,2%). Данная модель требует более тщательной проработки: сокращение количества регрессоров,

отбор регрессоров по физическому смыслу комбинаций входных параметров.

6. Искусственные нейронные сети. Применение аппарата искусственных нейронных сетей позволило существенно сократить ошибку прогнозирования (RMSE понизился на 21,5% по сравнению с линейной регрессией на базе профиля использования здания и в 2 раза ниже чем в методе прогнозирования скользящего среднего). Данный метод по всем, используемым в данном исследовании оценкам качества моделей, превосходит остальные рассмотренные модели. Использование календаря и уровня использования здания дает возможность использовать данную модель для нахождения потенциала энергосбережения.

7. Искусственные нейронные сети на основе нелинейной комбинации параметров. Использование в качестве входных параметров факторов, полученных комбинацией исходных параметров, привело к снижению ошибки прогнозирования на обучающем множестве (RMSE снизился на 31,2%), однако, увеличение количества входных параметров снизило обобщающие способности модели (RMSE на тестовой выборке повысился на 65,8%).

5.1. ВЫВОДЫ

1. Внесение детерминированной составляющей в виде знаний о процессах, происходящих в здании, таких, как время открытия, закрытия и т.д., с одной стороны дает возможность управлять прогнозным значением в зависимости от знаний о расписании в следующий день, и расширяет использование данного метода для нахождения потенциала энергосбережения. С другой стороны, возникают проблемы, связанные с необходимостью получения данных о подобном изменении энергопотребления в тот же момент времени в прошлом.

2. Наиболее точной, из рассмотренных методов прогнозирования, является модель искусственных нейронных сетей. Данная модель может быть использована для построения контура управления энергопотреблением. Недостатком данной моде-

ли является невозможность извлечения знаний с целью объяснения полученных прогнозных значений. Наиболее точным из рассмотренных методов прогнозирования, дающим возможность объяснения полученного прогнозного значения является модель линейной регрессии с применением профиля использования здания в качестве входного параметра.

3. При использовании подхода, основанного на консолидированном решении, результаты прогнозирования улучшились в отдельных случаях на 2 - 8 % (RMSE), и на 1 - 15% (MAPE). Однако в некоторых случаях консолидированные решения оказались хуже решения самой лучшей модели.

4. Данная методика может быть использована для решения задачи краткосрочного суточного прогнозирования потребления электроэнергии различных объектов.

Литература

1. МЕЛАМЕД М.А. *Современные методы анализа и прогнозирования режимов электропотребления в энергетических системах* // Итоги науки и техники. – Сер. Энергетические системы и их автоматизация. – 1988. – Т. 4. – С. 4–11.
2. *Система сбора информации о потреблении электроэнергии ЭкоСкада* [Электронный ресурс]. – Режим доступа: <http://www.ecoscada.com>
3. KASABOV N. *Evolving connectionists systems. The Knowledge Engineering Approach* // Originally published in the Series: Perspectives in Neural Computing 2nd ed., XXII, 451 p. 185 illus., Softcover ISBN: 978-1-84628-345-1
4. ЩЕРБАКОВ М.В., ЩЕРБАКОВА Н.Л. *Сравнение моделей аппроксимации безразмерной радиальной скорости течения жидкостей во вращающихся криволинейных каналах* // Известия Волгоградского государственного технического университета. – 2007. – №. 2. – С. 32-33.

5. ЩЕРБАКОВ М.В. *ICDMS – программное обеспечение как сервис для решения задач идентификации на основе неконнективистских систем* // Известия Волгоградского государственного технического университета. – 2009. – Т.12, №. 7. – С. 88-91.
6. BREBELS, A. *Mathematical and statistical framework for comparison of neural network models with other algorithms for prediction of Energy consumption in shopping centres* / Shcherbakov M.V., Kamaev V.A., et all // Proceedings of the 37 Int. Conf. Information Technology in Science Education Telecommunication and Business, suppl. to Journal Open Education. Yalta – Gurzuf. – 2010. – P. 96-97.

IMPLEMENTATION OF CONNECTIONISTS SYSTEM IN ENERGY CONSUMPTION DOMAIN

Valery Kamaev, Volgograd State Technical University, Volgograd, CAD Dept., professor (kamaev@unix.cad.vstu.ru).

Maxim Shcherbakov, Volgograd State Technical University, Volgograd, CAD Dept., Senior lecturer (vstu@gebeus.ru).

Dmitry Panchenko, Volgograd State Technical University, Volgograd, CAD Dept., Senior lecturer (panchenkodp@gmail.com).

Natalia Shcherbakova, Volgograd State Technical University, Volgograd, Physics Dept., Senior lecturer (snl@gebeus.ru).

Adriaan Brebels, Katholieke Hogeschool Kempen, Geel, Belgium, Project Manager (adriaan.brebels@portacapena.com)

Abstract: The problem of energy consumption forecasting in trade centers in Holland and Belgium is considered in this paper. The forecast frameworks based on connectionists systems is suggested here. Paper presents the general description of the framework and the list of specific features connected with observed domain. There are several connectionists systems are implemented to solve forecast problem. They are based on average model, linear regression,

neural networks. Models were tested with different set of input variables. In conclusion results and summary were presented.

Keywords: energy consumption forecasting, connectionism connectionists systems, average model, linear regression, neural networks.