

*“Интегрированные модели и мягкие вычисления в искусственном интеллекте”*. – М: Изд-во Физматлит, 2005. – С. 274–280. 7. Курейчик В. М. Применение пчелиного алгоритма для раскраски графов / Курейчик В. М., Кажаров А. А. // *Известия ЮФУ. Технические науки*. – 2010. – № 12 (113). – С. 30–36. 8. Курейчик В.В. Генетический алгоритм определения паросочетаний графа / Курейчик В.В., Курейчик В. М. // *Труды 10-ой Междунар. конф. “Knowledge-dialogue-solution”*, 2003, Варна, Болгария. – Варна, 2003. – С. 246–251. 9. Engelbrecht A. P. *Fundamentals of Computational Swarm Intelligence/ Engelbrecht A. P.* . – , Chichester: John Wiley & Sons, 2005.

УДК 004.652

**О. В. Шулима, В. В. Шендрик, М. О. Шестак**  
Сумський державний університет,  
кафедра комп'ютерних наук

## **ПОБУДОВА СХОВИЩА ДАНИХ СИСТЕМИ ПІДТРИМКИ ПРИЙНЯТТЯ РІШЕНЬ ДЛЯ ПРОЕКТУВАННЯ РОЗПОДІЛЕНИХ ЕНЕРГЕТИЧНИХ СИСТЕМ**

© Шулима О. В., Шендрик В. В., Шестак М. О., 2016

**Розглянуто проблеми, що виникають під час роботи з розрізненими джерелами інформації з використанням баз даних. Наведено модель сховища даних як засобу інтеграції та опрацювання даних з розрізнених джерел при створенні системи підтримки прийняття рішень для проектування розподілених енергетичних систем.**

**Ключові слова:** концептуальна модель, база даних, сховище даних, система підтримки прийняття рішень, альтернативні джерела енергії, розподілена енергетична система.

**This paper discusses the problems that arise when working with disparate data sources using database. The model of data warehouse is presented as way of integrating and processing data from disparate sources while creating a decision support system for the design of distributed energy systems.**

**Key words:** conceptual model, database, data warehouse, decision support system, renewable energy system, distributed energy system.

### **Вступ. Загальна постановка проблеми**

У попередніх роботах було визначено актуальність побудови розподілених енергетичних систем, що використовують альтернативні джерела енергії (вітер та сонце) [1]. Було визначено загальну схему досліджуваної мережі [2]. Це група будівель зі встановленими власними сонячними панелями і батареями для зберігання акумульованої енергії. Існує також загальний парк вітряних генераторів разом з банком зберігання енергії. Крім того, існує зв'язок між системою і зовнішньою мережею для додаткового споживання і продажу надлишків енергії в мережу. В Україні для пересічного користувача дуже важко спланувати роботу такої мережі для власного господарства та оцінити, які саме потужності відновлювальних джерел енергії (ВДЕ) слід використовувати.

Процес прийняття рішень, що стосуються проектування розподіленої мережі, ускладнюється невизначеністю процесу генерації електроенергії: кількість первинної енергії залежить від пори року, доби, погодних умов, а також різних чинників, які важко врахувати. Це збільшує час на прийняття правильного рішення відносно планування вибору місць спорудження систем, визначення її складових з погляду оцінки ефективності роботи.

Методологічною основою процесу проектування розподіленої мережі є системний аналіз, в основу якого покладено процедуру побудови узагальненої моделі системи. На практиці це

пов'язано зі створенням систем підтримки прийняття рішень (СППР) з розвиненими інформаційними зв'язками між її складовими.

Для аналізу процесу проектування розподіленої енергосистеми необхідно забезпечити:

1. Зберігання інформації та керування нею;
2. Опрацювання як структурованої, так і неструктурованої інформації, роботу з картографічними даними;
3. Аналіз різнотипної інформації з використанням консолідаційного та федеративного підходів.

У роботі [3] були визначені підсистеми СППР та інструментарії їх створення, а також загальна архітектура системи. Так, робота системи забезпечується поєднанням для загального використання GIS, Matlab, Web-server та ін., причому кожен компонент програмного забезпечення має власне середовище для зберігання інформації – базу даних (БД).

Важливою науковою проблемою, яка виникає, є розроблення і удосконалення методів опрацювання різнотипних даних з метою підтримки прийняття рішень. Робота в єдиній системі з декількома незалежними інформаційними продуктами, що використовують різні БД, схеми керування, тощо вимагає розроблення уніфікованої моделі предметної області.

### **Аналіз останніх досліджень та публікацій**

Консолідація, або інтеграція даних – це об'єднання даних, які спочатку вводяться в різні системи. Самі ці системи можуть розташовуватися в одній локальній мережі, але мати різні платформи і внутрішню архітектуру [4].

Питання з консолідації розрізної інформації для її подальшого опрацювання з метою прийняття рішень виникли разом із застосуванням поняття “сховищ даних” ще у кінці минулого століття. Значний внесок у вирішення цієї проблеми зробили вчені: С. Lagoze [5], Н. Van de Sompel [5], Н. J. Watson [6], D. Theodoratos [7], J. T. Horng [8], К. В. Антипін [9], Е. Гришенков [10], А. В. Фомичев [9], Н. Б. Шаховська [11] та ін. Це питання найчастіше вирішується при створенні СППР для конкретних прикладних задач, причому питання побудови єдиної моделі консолідації даних до кінця не розкрито. Тому при створенні СППР для проектування розподіленої енергетичної системи стає актуальним розроблення власної моделі предметної області в умовах консолідації розрізнених даних.

Вчені називають такі методи опрацювання даних з джерел з різними структурами даних.

1. Пошук інформаційних джерел на основі метаданих. Це найпоширеніший метод пошуку інформації, необхідної для користувача [12], робота якого ґрунтується на основі посередника. Програма посередника отримує модель вимог користувача до даних, що він шукає. Пошук здійснюється серед джерел даних, інформацію про які розміщено у репозиторії посередника. До цього додаються не функціональні вимоги до інформаційних джерел (наприклад, такі характеристики якості даних, як точність, частота оновлення тощо). Як наслідок, специфікація посередника шукає джерела, що задовольняють вказані вимоги, визначає структури даних та допустимі операції над даними джерела.

2. Web-інтеграція. Web-інтеграція – це методи опрацювання і подання інформаційних ресурсів за допомогою Web-технологій. Зберігання і передавання даних відбувається за допомогою мови XML, спеціально створеної для організації взаємодії з різними застосуваннями [13].

3. Інтеграція на рівні сховищ даних. Метою інтеграції даних на рівні їх сховища є отримання єдиної і цілісної картини корпоративних даних предметної області [14]. Типовим підходом до інтеграції інформації в масштабах єдиної системи, яка складається з різних підсистем, що реалізовані на основі різних програмних засобів з власними базами даних, є побудова сховищ та вітрин даних на основі отримання оперативних даних, їхньої трансформації до єдиної схеми і завантаження даних у сховища. Вони найкраще пристосовані до роботи з великими обсягами інформації, що потрапляє періодично з неоднорідних джерел.

Архітектура розроблюваної СППР, окрім консолідації інформації передбачає ієрархічне впорядкування складових частин та агрегування інформації під час передавання її до верхніх рівнів. Це

вказує на те, що для роботи з інформацією в СППР, яка надходить з різних джерел, необхідно поєднати всі вищеописані типи інтеграції на верхньому рівні ієрархії якого знаходиться сховище даних.

### **Формулювання цілі статті**

СППР для проектування розділеної енергетичної системи, повинна будуватися на підставі збору та опрацювання історичних даних про метеорологічні умови у даних географічних координатах, та даних про типовий графік навантаження споживачів, повинна враховувати можливості використання декількох відновлювальних джерел енергії та підключення до зовнішньої мережі.

**Метою** даної роботи є концептуальне визначення наборів даних, що впливають на прийняття рішення та розроблення концептуальної моделі предметної області.

Розроблення концептуальної моделі сховища даних та забезпечення ефективного управління даними потребують вирішення наступних **завдань**:

- сформулювати цілі процесу збирання інформації;
- розробити інфологічну модель предметної області проектування енергетичних систем;
- розробити логічну модель предметної області з визначенням основних сутностей інформаційної системи та зв'язків між ними.

### **Виклад основного матеріалу**

#### **Структура СППР**

Для супроводу прийняття рішень при проектуванні розподіленої енергетичної системи необхідно використовувати такі інформаційні модулі:

– “Підсистема збирання та попереднього опрацювання даних (SCP)” – збір інформації про погодні умови, технічні та економічні показники обладнання в мережі Інтернет. Дані зберігаються в SQL таблицях та використовуються в інших модулях;

– “Підсистема аналітичного опрацювання даних (SAPD)” підрозділяється на дві частини”:

– “Геоінформаційна система (GIS) регіону” – створення мережевої та погодної карт місцевості. Дані зберігаються у власній БД.

– “Моделювання” – дослідження роботи мережі з використанням Matlab/Simulink. Результати роботи зберігаються у текстових файлах;

– “Підсистема зберігання даних (SDS)” – формування сховища даних на основі даних, що зберігаються в SQL таблицях, GIS та Matlab;

– “Підсистема формування рішень” – генерація альтернативних проектних рішень, їх оцінка та кінцеве формування оптимального рішення. Дані отримуються зі сховища даних.

Підсистема прийняття рішень майже не працює з детальними даними, а в основному використовує агреговані показники, що розраховуються на основі “сирих” даних. Серед способів зберігання даних саме структура сховища даних якнайкраще дозволяє керувати цією ситуацією. В цьому випадку до сховища даних записується агрегована інформація, що розраховується в Matlab. Це дає змогу зручно зберігати і мати швидкий доступ до агрегованої інформації, адже кінцеві результати розрахунків Matlab зберігаються у текстових файлах, що є незручним для подальшої роботи з ними.

– “Представлення результатів” – використовується web-інтерфейс, через який користувачі отримують доступ до системи. Окрім сформованого рішення, що є техніко-економічним обґрунтуванням процесу проектування розподіленої енергетичної системи, користувачі можуть проглядати GIS місцевості, у вигляді кольорової карти. Така карта надає можливість співставити ділянки для побудови розподіленої мережі з потенціалом відновлювального джерела в цій місцевості. Це дає можливість вибрати найкращу локацію для побудови проекту з ВДЕ “з нуля”.

На рис. 1 наведено схему взаємодії компонентів СППР для проектування поділеної енергетичної системи із зазначенням зв'язків між складовими.

#### **Технологія взаємодії баз даних**

З огляду на аналіз попередніх досліджень було визначено, що опрацювання неузгоджених даних потребує застосування різноманітних підходів. Так в СППР проектування енергетичної

системи пропонується інтегрована інформаційна технологія опрацювання інформації, що використовує ієрархічне впорядкування з технологією тиражування:

$$\begin{aligned}
 & \text{Пошук інформаційних джерел } I_{fc} \text{ в SCP} \quad \rightarrow \\
 & \text{Web-інтеграція } I_{web} \text{ в SAPD} \quad \rightarrow \\
 & \text{Інтеграція на рівні сховищ даних } I_{wh} \text{ в SDS} \\
 & I_{int} = \{ I_{wh}, I_{web}, I_{fc} \}
 \end{aligned}$$

Перелічені підсистеми знаходяться під керівництвом різних серверів: Web, GIS, Matlab, Microsoft SQL Server. Як кінцевий метод інтеграції використовується інтеграція на рівні сховища даних. Інтегровані дані використовуються в підсистемах формування рішень та представлення результатів.

Технологія тиражування – копіювання визначеної частини даних з однієї системи в іншу за певним розкладом [15]. Зазвичай для тиражування використовують три методи обміну даними між інформаційними продуктами:

- синхронізація – порівняння даних;
- експорт – копіювання даних таблиці за певний період;
- вибірка – копіювання частин таблиці за певний період за параметрами користувача.

На рис. 2 зображено схему взаємодії основних джерел даних у підсистемах СППР проектування енергетичної системи.

Бази даних призначені для накопичення структурованих даних і метаданих про параметри, що впливають на проектування розподілених енергетичних систем. Вони є центральними частинами системи та її основних підсистем, оскільки надають інформацію іншим модулям.

Оскільки всередині СППР існує інтеграція між складовими підсистемами, то необхідно продумати, як працювати з файлами, що містять дані в різних форматах та інформація з яких необхідна при роботі різним підсистемам. Вирішенням цієї проблеми стало використання Підсистеми збирання та попереднього опрацювання даних. Через цю підсистему наповнюється база даних SQL, що розміщується на web-сервері.

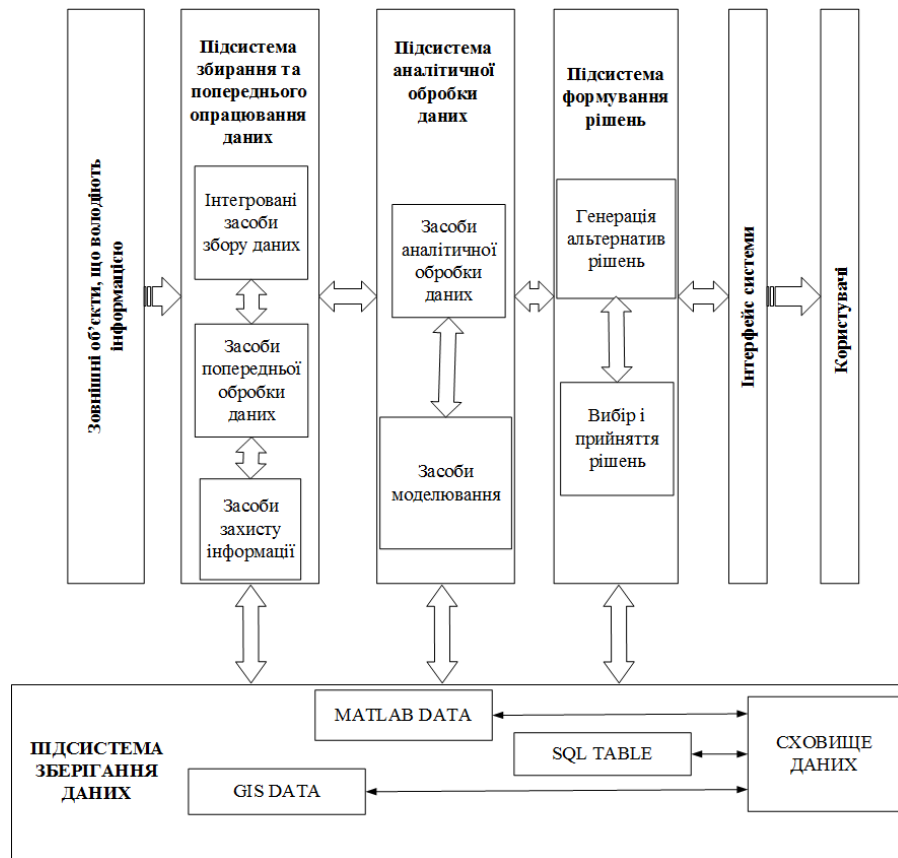


Рис. 1. Структура СППР

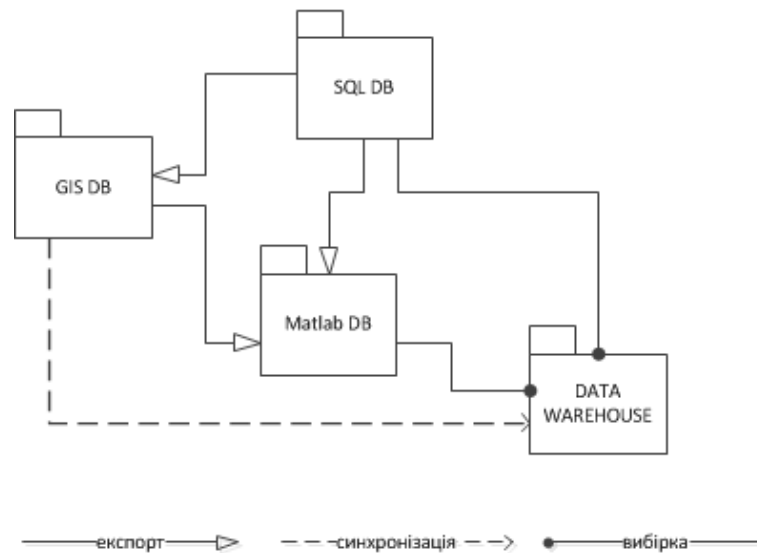


Рис. 2. Схема взаємодії основних джерел даних

Підсистема зберігання даних орієнтована на надійне зберігання великих обсягів даних, що будуть використані в підсистемі прийняття рішень. Використовуються засоби збереження даних, які складаються з двох основних частин: баз даних окремих модулів і СУБД Microsoft SQL Server.

#### Концептуальне проектування сховища даних

Етап концептуального проектування передбачає опис предметної області у термінах формальної мови. Широкого використання під час побудови концептуальної моделі сховища даних набуло ER-моделювання. Першим етапом побудови ER-діаграми є виділення сутностей. Системний аналіз предметної області дав змогу побудувати концептуальну модель сховища даних СППР.

Основними сутностями сховища даних є Вітрові Турбіни, Сонячні Панелі, Батареї, Профіль Сонячної активності, Профіль Вітру, Профіль Географічної Зони, Електричне Навантаження, Зовнішня Мережа, Розрахункові Дані 1, Розрахункові Дані 2, Критерії Оптимізації, Прийняті Рішення. На рис. 2 наведено діаграму “сутність-зв’язок” (ER-діаграму) в нотації IDEF1X, яка використовується для опису сховища даних СППР на концептуальному рівні проектування.

Сутності Вітрові Турбіни, Сонячні Панелі та Батареї містять інформацію про технічні характеристики та ціну обладнання. Ці дані використовуються при розрахунку енергетичної та економічної ефективності системи. Сутності Профіль Сонячної Активності, Профіль Вітру та Профіль Географічної Зони містять дані, необхідні для визначення варіантів конфігурації системи за можливої потужності. Сутність Електричне Навантаження містить дані про типове споживання в будинку, що зіставляються з погодними даними. Сутність Зовнішня Мережа містить дані про пропускну спроможність зовнішньої мережі, що використовуються при визначенні параметрів монтажу системи. Сутність Розрахункові Дані 2 містить дані про величину сонячної інсоляції в регіоні, розраховуються на основі погодинних даних у визначеній місцевості та використовуються при розрахунку параметрів, що входять до сутності Розрахункові Дані 1. Дані цієї сутності визначають величину можливої здобутої електроенергії від ВДЕ. Сутність Критерії Оптимізації містить кількісні значення критеріїв, що впливають на прийняття рішень. Сутність Прийняті Рішення містить дані про оптимальну конфігурацію системи.

Структура сховища даних побудована за схемою “сніжинка”, оскільки планується одержувати інформацію з реляційних баз даних.

Перша сніжинка (таблиця OptimizationCriteria) містить критерії оптимізації, що є агрегованою інформацією з розрахункових даних Matlab, параметрів зовнішньої мережі, технічних/економічних характеристик обладнання, графіка навантаження в мережі. Ці дані надходять до Matlab за методом “експорт” з SQL DB (рис. 2), де на основі їх розраховуються критерії оптимізації, що вже передаються і зберігаються у сховищі даних (перша сніжинка).

Завдання підбору оптимального рішення полягає у розв'язанні задачі оптимізації із застосуванням методу парних порівнянь. Для порівняння даних використовують дані з першої сніжинки, після їх опрацювання кінцеві результати заносяться до другої сніжинки (таблиця OptimalConfigurationDecision).

Діаграму “сутність-зв'язок” системи показано у спрощеному вигляді з таблицями, що належать до різних баз даних, над кожною таблицею вказано, до якої бази її вміщено. Так, всі розрахункові дані містяться у таблицях Matlab, а всі історичні дані – у таблицях SQL, що належать GIS та Microsoft SQL Server. Бази останнього становлять Сховище Даних.

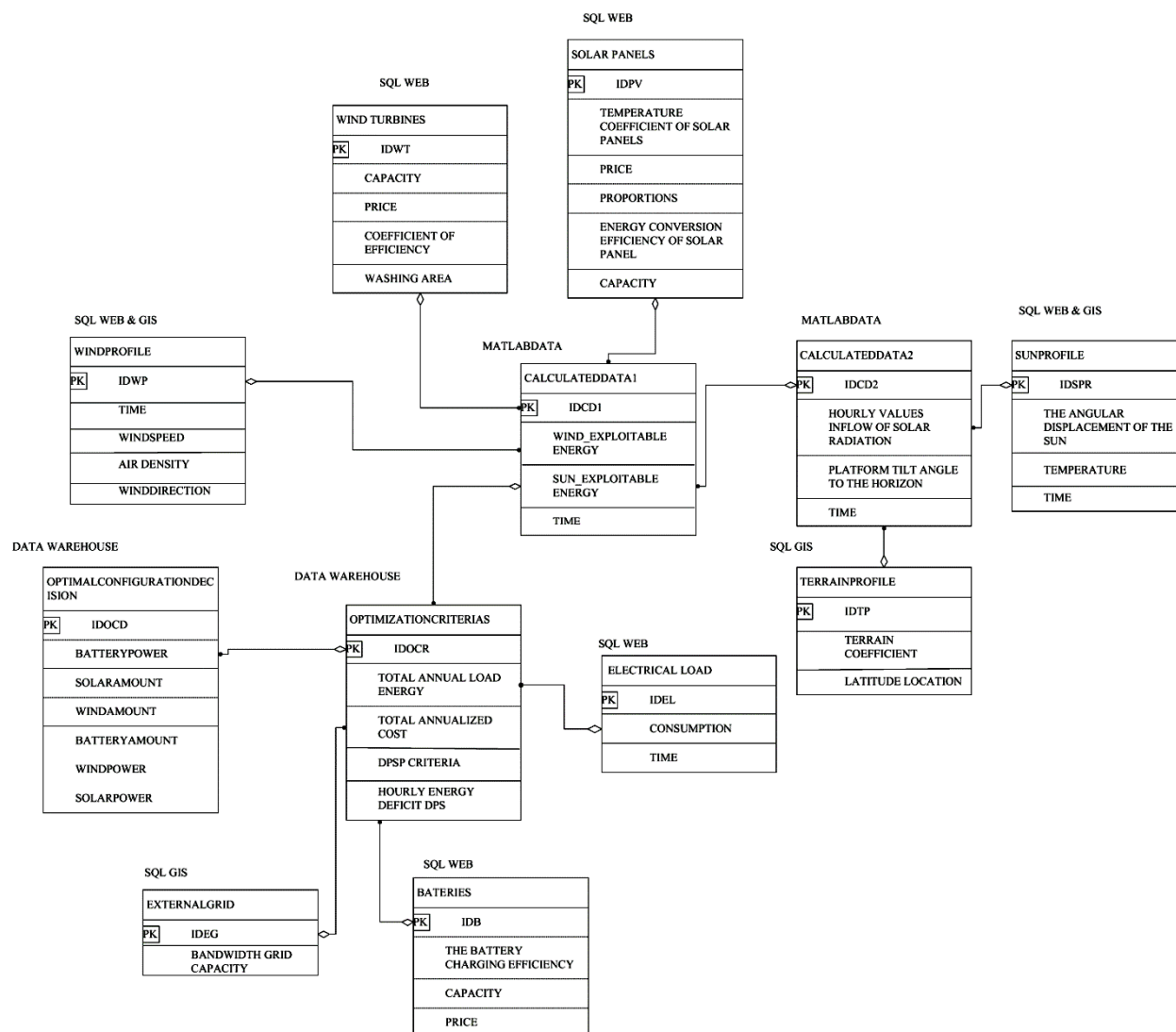


Рис. 2. Діаграма “сутність-зв'язок” системи

### Висновки та перспективи подальших наукових розвідок

У статті описано предметну область сфери проектування розподіленої енергетичної системи з використанням відновлювальних альтернативних джерел енергії. Проаналізовано методи опрацювання даних з джерел з різними структурами даних. Визначено структурну схему СППР та запропоновано використовувати інтегровану інформаційну технологію обробки інформації. Ієрархічне впорядкування використане як засіб зв'язку між частинами системи. Запропоновано схему інтеграції даних між розподіленим базами даних на основі методів тиражування. Описано концептуальну схему сховища даних за типом “сніжинка”.

Наукова новизна полягає у застосуванні технології сховища даних при прийнятті рішень для вирішення задач проектування розподіленої енергетичної системи.

1. Шендрюк В. В., Ващенко С. М., Шульма О. В., Омеляненко К. А. (2013). Актуальность моделирования распределенных энергосистем эффективного использования возобновляемых источников энергии. *Восточно-Европейский журнал передовых технологий*, (5 (8)), 4–8.
2. Shulyma O., Shendryk, V., Baranova, I., & Marchenko, A. (2014). *The Features of the Smart MicroGrid as the Object of Information Modeling*. In *Information and Software Technologies* (pp. 12–23). Springer International Publishing.
3. Olha Shulyma, Paul Davidsson, Vira Shendryk, Anna Marchenko. *The Architecture of an Information System for the Management of Hybrid Energy Grids*. – Accepted on SEN-MAS'15 Workshop.
4. Шаховська, Н. Б. (2011). Методи опрацювання консолідованих даних за допомогою просторів даних. *Проблеми програмування*, (4), 72–84.
5. Lagoze, C., & Van de Sompel, H. (2001, January). *The Open Archives Initiative: Building a low-barrier interoperability framework*. In *Proceedings of the 1st ACM/IEEE-CS joint conference on Digital libraries* (pp. 54-62). ACM.
6. Watson, H. J., & Gray, P. (1997). *Decision support in the data warehouse*. Prentice Hall Professional Technical Reference.
7. Theodoratos, D., Ligoudistianos, S., Sellis, T. (2001). *View selection for designing the global data warehouse*. *Data & Knowledge Engineering*, 39(3), 219–240.
8. Hornig, J. T., & Chen, C. W. (2001). *A mechanism for view consistency in a data warehousing system*. *Journal of Systems and Software*, 56(1), 23–37.
9. Антипин К. В., Фомичев А. В., Гринев М. Н., Кузнецов С. Д., Новак Л. Г., Плешачков П. О., Ширяев, Д. Р. (2004). *Оперативная интеграция данных на основе XML: системная архитектура BizQuery*. Труды Института системного программирования РАН, 5.
10. Грищенко, Е. (2001). *Планирование и консолидация данных многомерной базы*. *Открытые системы*. СУБД, (4), 65–72.
11. Шаховська Н. Б. *Сховища та простори даних : монографія / Н. Б. Шаховська, В. В. Пасічник; Міністерство освіти і науки України, Національний університет “Львівська політехніка”*. – Львів : Видавництво Національного університету “Львівська політехніка”, 2009. – 240 с.
12. Егошина А. А., Вороной, А. С. (2011). *Повышение эффективности извлечения информации из слабо структурированных источников на основе метаданных и базы знаний*. *Научные работы Донецкого национального технического университета*. Сер.: Информатика, кибернетика та обчислювальна техніка, (13), 44–47.
13. Росинский В. В. (2012). *Обеспечение интеграции данных в корпоративных информационных системах на основе прогрессивных web-технологий // Вісник Державного університету інформаційно-комунікаційних технологій*. – (10, № 1), 87–94.
14. Шаховська Н. Б., Тарасов, Д. О. (2009). *Особенности интеграции данных информационных систем национального университета “Львівська Політехніка”*.
15. Третьяк В. Ф., Голубничий Д. Ю., Челенко Ю. В. (2005). *Тиражирование данных в системе управления базами данных*. *Управління розвитком*. – Х.: ХНЕУ, (3), 94–95.