

Міністерство освіти і науки України  
Харківський національний університет імені В.Н. Каразіна  
Факультет комп'ютерних наук  
Кафедра теоретичної та прикладної системотехніки

---

# Сучасні інформаційні технології

Матеріали засідань школи-семінару  
Випуск 4

Харків – 2019

УДК 004/681.3

У збірнику подано матеріали засідань школи-семінару «Сучасні інформаційні технології», які проводилися у осінньому семестрі 2020 – 2019 навчального року.

Затверджено до друку Вченою радою факультету комп'ютерних наук Харківського національного університету імені В.Н. Каразіна, протокол №3 від 19 листопада 2019 року.

C23 Сучасні інформаційні технології: матеріали засідань школи-семінару. Випуск 4. – Харків : ХНУ імені В.Н. Каразіна, 2019. – 53 с.

## ОРГАНІЗАЦІЙНИЙ КОМІТЕТ ШКОЛИ-СЕМІНАРУ

*Голова оргкомітету*

Шматков Сергій Ігорович (доктор технічних наук, професор).

*Члени оргкомітету*

Доля Григорій Миколайович (доктор технічних наук, професор);

Лосєв Юрій Іванович (доктор технічних наук, професор);

Угрюмов Михайло Леонідович (доктор технічних наук, професор);

Толстолузька Олена Геннадіївна (доктор технічних наук, доцент);

Бердніков Анатолій Георгійович (кандидат технічних наук, доцент).

*Секретар оргкомітету*

Стрілець Вікторія Євгенівна (кандидат технічних наук).

## Зміст

<i>Оразмурадова А. Р., Бердніков А. Г.</i> Методика розробки коригувальних рішень при управлінні ІТ-проектом .....	3
<i>Романская А.С.</i> Методика управління ризиками проекту .....	10
<i>Донець В.В.</i> Методи класифікації стану пацієнтів в системах медичного моніторингу .....	15
<i>Кузьменко М.В.</i> Система розпізнавання аудіоінформації за допомогою штучних нейронних мереж .....	17
<i>Маришев П.Д.</i> Обробка та збереження великих обсягів даних у хмарних сховищах .....	19
<i>Нікіфорова С.Р.</i> Комп'ютерна модель стратифікації пацієнтів в системах медичного моніторингу на основі методу нечітких с-середніх .....	22
<i>Передерин А.В.</i> Модели и методы информационного обеспечения процесса управления .....	23
<i>Петухов М.В., Шейко О.Г.</i> Аналіз інформативності змінних стану пацієнта в системі медичного моніторингу.....	26
<i>Семенченко О.В.</i> Система розпізнавання фрагментів обличчя за допомогою методів штучного інтелекту .....	28
<i>Семенюг Е.А.</i> Оценка качества волоконно-оптической СПД с оптическими усилителями .....	30
<i>Харченко В.О.</i> Метод управління ризиком для підвищення надійності складних систем.....	33
<i>Артюх С.О.</i> Багатоканальна система збору даних і управління технічним об'єктом з дистанційним доступом .....	39
<i>Садретдінов Я.С.</i> Програмна модель вибору оптимального маршруту служби таксі .....	41
<i>Зарицький О. В., Костенко О. Б.</i> Побудова моніторингу регіональних ресурсів на основі ранжування інформаційних систем управління територіями .....	44

## Доповіді школи-семінару «Сучасні інформаційні технології»

### МЕТОДИКА РОЗРОБКИ КОРИГУВАЛЬНИХ РІШЕНЬ ПРИ УПРАВЛІННІ ІТ-ПРОЕКТОМ

*Оразмурадова А.Р., Бердніков А.Г.*

*Харківський національний університет імені В.Н. Каразіна*

Управління проектом по ресурсним параметрам включає у себе планування календарних термінів, бюджетних показників, інформаційних та інших матеріальних і/або не матеріальних ресурсів, а також дотримання їх у запланованих межах. У разі відхилення будь-якого з показників, мають бути надані коригуючі впливи, що в даному випадку здійснюється за допомогою відомих методів мережевого планування.

З огляду на те, що ми можемо виділити ресурси як основні обмежуючі фактори реалізації проекту, робимо логічний висновок, що ресурсний контроль дає змогу завершити всі проектні роботи в термін та без втрат. Оскільки, витрати і час є взаємодоповнюючими, з одного боку, та альтернативними складовими загальних ресурсних показників проекту, з іншого, це обумовлює або на орієнтацію на скорочення витрат проекту, або на скорочення життєвого циклу проекту. В той же час, ми маємо розуміти, що під час ресурсного планування, потрібно брати до уваги усі ресурсні складові проекту. Тож, необхідно сформулювати конкретну стратегію реалізації ІТ-проекту, що дозволяє тим, чи іншим, шляхом отримати переваги у часі, витратах, інформаційних, людських чи інших ресурсах. Данна стратегія має бути сформована у методиці, що обумовлює альтернативу вибору ресурсного фактору, дає змогу наглядно продемонструвати етапи та хід реалізації проекту та дозволяє вчасно виробити коригуючі впливи на основі відповідних конкурентних переваг.

Таким чином, правильний розподіл ресурсів на етапі планування проекту та моніторинг його розвитку при реалізації є запорукою успішного і своєчасного завершення робіт. Можемо зробити логічний висновок, що завдання управління часом є надзвичайно актуальною і стає перед кожним керівником будь-якого ІТ-проекту.

#### **Постановка задачі**

Кожен керівник ІТ-проекту має розуміти, який саме критерій є порушеним та вчасно внести певні необхідні корективи. У випадку, коли швидкість реалізації і витрати низькі, існує перевага у виконанні бюджету. При високих витратах і швидкості реалізації, очевидно стає абсолютна часова перевага. Якщо розглянути випадок з великою, чи достатньою, кількістю інформаційних ресурсів, але з великою нестачею людських, можемо діагностувати проблему з людино/годинами, та навпаки, їх надлишок у зворотній ситуації. Таким чином, при формуванні стратегії реалізації проекту, слід взяти до уваги інтегральну перевагу ресурсного планування і розглядати весь план реалізації проекту як єдину складну синтезовану модель.

Перехід до однієї з переваг при балансі інших критеріїв дає велику цінність як для керівника ІТ – проекту, так і для споживача. Цінність інтегральної переваги стає очевидною при розгляді її на основі переваги в часі, коли за рахунок підвищення швидкості потоку реалізації, призводить до скорочення запасів і підвищенню їх оборотності, вивільненню оборотних коштів і зменшення обсягу нереалізованих товарів, тобто до зниження витрат як системного ефекту. Тож, проектуючи дану модель на різні ресурсні критерії, ми можемо розуміти теоретичну і практичну цінність такого дослідження. Але слід пам'ятати, що на стратегічному рівні саме скорочення життєвого циклу проекту (за інших рівних умов) є ключовим фактором економії загальних витрат проекту.

Саме тому, необхідно вибрати метод, за допомогою якого можна вдало відслідкувати хід реалізації проекту за будь-яким із ресурсних критеріїв. Аналізуючи існуючі методи, зроблений

висновок, що найбільш повним та ефективним є метод освоєного обсягу. Оскільки початкова версія даного методу включає у себе календарний план, показники різних видів обсягу, розрахування індексів, показників відхилення, прогнозних показників, а також графічне відображення ходу реалізації проекту, цей метод стає оптимальним для вирішення проблеми ресурсного планування. Але, оскільки у основі методу освоєного обсягу лежить модель розрахунку показників за часовими та бюджетними параметрами, робимо логічний висновок, що даний метод потребує доопрацювання.

Якщо зробити метод освоєного обсягу придатним не тільки для демонстрації відхилень, а і для їх коригування, слід додати до розроблюваної методики метод оцінки і аналізу проекту PERT, а декомпозицію IT-проекту провести не за роботами, а за етапами реалізації, що ще на рівні планування допоможе нам скоротити життєвий цикл проекту. Для даної задачі добре підійде алгоритм Демукрона.

Роблячи висновки, слід зазначити, що управління проектом за ресурсними параметрами включає в себе планування календарних термінів проекту, контроль дотримання всіх параметрів і надання коригувальних впливів у разі будь-яких відхилень під час реалізації проекту.

### **Розробка методики виробітку коригувальних рішень при реалізації IT-проекту**

Контроль проекту за ресурсними параметрами потребує мати у арсеналі кожного керівника IT-проекту оптимальний метод для аналізу ходу реалізації проекту та коригування відхилень показників. В даний час до складу інструментарію управління ресурсами входить метод освоєного обсягу. Він складає основну частину управління вартістю і тривалістю проекту, але за допомогою певного доопрацювання і доповнення існуючої версією методу, ми зможемо розглядати його не тільки як засіб для моніторингу та обліку поточного стану робіт, а й для прогнозування і оперативного планування.

Метод освоєного обсягу дозволяю чітко показати поточні тенденції розвитку проекту. Він базується на фундаментальному принципі, який наголошує, що стан і тенденції минулого є хорошими провісниками майбутнього.

За допомогою методу, будь-який керівник IT-проекту отримує впорядковану систему контролю якості, швидкості, бюджету, інформаційної зіставної проекту, тощо.

Метод дозволяє отримати обґрунтовані і своєчасні відповіді на наступні, важливі для успіху всього проекту питання:

- Чи відстає проект від графіку або випереджає його?
- Наскільки ефективно використовується час?
- Яка ймовірна тривалість проекту?
- Чи знаходиться проект в рамках або ж виходить за рамки бюджету?
- Наскільки ефективно використовуються ресурси?
- Наскільки ефективно повинні використовуватися ресурси для успішного завершення проекту?
- Які очікувані загальні витрати по проекту?
- Чи буде проект завершений в рамках або ж за рамками бюджету?

Відповідь на кожне з цих питань керівник проекту може отримати, розраховуючи основні показники методу освоєного об'єму. Кожний з показників несе у собі ту чи іншу інформацію, жоден з них не є "проміжним".

Якщо в ході використання методу освоєного об'єму буде виявлено відставання від календарних планів або ж перевитрата бюджету, то керівник проекту, використовуючи цей метод, зможе вирішити такі задачі:

- Де (на якому етапі реалізації) виникли проблеми?
- Чи є проблеми критичною?

– Що необхідно зробити для вирішення виявлених проблем?

Але, повертаючись до питання ресурсного планування, ми маємо розуміти тенденції виконання щодо кожного з ресурсних показників, не акцентуючи увагу тільки на бюджетних чи часових показниках. Тож, необхідно додати до існуючої версії методу освоєного обсягу інструментарій для більш повної оцінки стану проекту за ресурсними показниками.

Після проведених досліджень, було вирішено, що для цієї мети добре підійде техніка оцінки та аналізу проекту PERT. Більш детально метод PERT та його застосування в рамках методу освоєного обсягу буде розглянуто у наступних розділах.

Базовими показниками методу освоєного обсягу є:

- планові обсяги (Planned Value – PV);
- освоєння обсяги (Earned Value – EV);
- фактичні витрати (Actual Cost – AC).

Показник планових або базових обсягів PV показує, скільки робіт повинно бути вироблено в конкретний момент часу відповідно до плану проекту, тобто який прогрес у виконанні проекту повинен бути на певну дату. Тобто, планові обсяги – це чисельне вираження запланованих до виконання об'ємів робіт на поточну дату. Базові обсяги мають бути незмінними впродовж усього проекту. Показник планових обсягів (так само як і інші базові показники методу освоєного об'єму) графічно подаються у вигляді кумулятивних графіків, що показують планові об'єми виконання робіт за проектом наростаючим підсумком. Такі кумулятивні графіки зветься S-криві. І стосовно до планових об'ємів говорять про S-криву планових обсягів, тобто про кумулятивний графіку планових обсягів.

Показник освоєних обсягів EV являє собою знімок проекту в поточний момент часу. Освоєний обсяг є чисельним виразом фактично виконаних об'ємів робіт, виражених в показниках планової вартості цих робіт. Цей показник часто називають планової (кошторисної) вартістю фактично виконаних робіт, бюджетною вартістю виконаних робіт. Таким чином, EV – це бюджетна, планова, кошторисна вартість робіт, фактично виконаних на поточну дату.

Показник фактичних AC витрат показує фактичну (реальну) вартість виконаних обсягів робіт, тобто сукупність всіх витрат, що виникли в ході виконання робіт на поточну дату. Іноді про фактичні витрати говорять як про кількість ресурсів, які необхідно було використати в ході робіт, фактично виконаних на поточну дату або протягом визначеного періоду часу.

Якщо говорити про діагностування поточного стану проекту, то такий аналіз, а також прогнозування його майбутніх тенденцій проводиться на основі похідних показників методу освоєного обсягу. Наведені в табл. 1 показники використовуються для відповіді на основні питання, що виникають в ході контролю та оперативного управління проектом на стадії реалізації. Взаємозв'язок між цими питаннями і показниками методу освоєного обсягу, а також їх математичне значення показані в табл. 1.

Показник ВАС показує повну суму всіх планових обсягів робіт проекту, підсумкове значення базового плану виконання проекту.

Оскільки у класичній версії методу освоєного обсягу основні розрахунки проводяться на базі бюджетних та часових показників, необхідно адаптувати їх та зробити відповідними до усіх ресурсних критеріїв. Враховуючи всі особливості методу освоєного обсягу можемо зробити логічний висновок, що за основу ми можемо брати будь-яку пару взаємопов'язаних показники. Позначимо їх як X та Y відповідно. У загальному вигляді значення основних показників буде виглядати таким чином, як показано у табл. 2.

Таблиця 1 – Основні показники методу освоєного обсягу

Тип відхилення	Показник освоєного обсягу	Питання оперативного управління проектом	Математичне значення
Variances Відхилення	<b>Schedule Variance – SV</b> Відхилення за розкладом (термінами)	Чи відстає проект від графіка або випереджає його?	$SV = EV - PV$
	<b>Cost Variance – CV</b> Відхилення за витратами (за вартістю)	Чи знаходиться проект в рамках або ж за рамками бюджету?	$CV = EV - AC$
Indices Індекси	<b>Schedule Performance Index – SPI</b> Індекс виконання розкладу	Наскільки ефективно використовується час?	$SPI = EV / PV$
	<b>Cost Performance Index – CPI</b> Індекс виконання бюджету	Наскільки ефективно використовуються ресурси?	$CPI = EV / AC$
	<b>To-Complete Performance Index – TCPI</b> Індекс необхідної ефективності	Наскільки ефективно повинні використовуватися ресурси для успішного завершення проекту?	$TCPI = (BAC - EV) / (BAC - AC)$
Forecasts Прогнози	<b>Time Estimate at Completion – EACt</b> Прогнозна тривалість проекту	Яка ймовірна тривалість проекту?	$EACt = (BAC / SPI) / (BAC / N)$ ,
	<b>Estimate at Completion – EAC</b> Прогнозна вартість проекту	Які очікувані загальні витрати по проекту?	$EAC = BAC / CPI$
	<b>Variance at Completion – VAC</b> Відхилення при завершенні	Чи буде проект завершений в рамках або ж за рамками бюджету?	$VAC = BAC - EAC$
<b>Budget At Completion – BAC</b> - показник повного бюджету проекту			

Таблиця 2 – Адаптація основних показників методу освоєного обсягу за ресурсним критерієм

Тип відхилення	Показник освоєного обсягу
Variances Відхилення	<b>SV(X)</b> - Відхилення за першим видом ресурсів X
	<b>CV(Y)</b> - Відхилення за другим видом ресурсів Y
Indices Індекси	<b>SPI(X)</b> - Індекс виконання X
	<b>CPI (Y)</b> - Індекс виконання Y
	<b>TCPI</b> - Індекс необхідної ефективності
Forecasts Прогнози	<b>EAC (X)</b> - прогнозна необхідність ресурсу X
	<b>VAC (Y)</b> - відхилення Y при завершенні

Розглянемо математичне значення основних показників методу освоєного обсягу.

*Відхилення за розкладом* (за термінами) (Schedule Variance – SV) визначається як різниця між освоєним об'ємом (Earned Value – EV) і плановим об'ємом (Planned Value – PV):

Відхилення за розкладом може бути розраховане у відсотках:

$$SV\% = (EV - PV) / PV = SV / PV$$

Це значення можна інтерпретувати як свідчення того, на скільки % проект відстає від графік, тобто який % запланованих робіт не виконано в термін.

*Індекс виконання розкладу* (Schedule Performance Index – SPI) розраховується шляхом ділення освоєного об'єму (Earned Value – EV) на плановий об'єм (Planned Value – PV).

Індекс виконання розкладу SPI показує, яка кількість планових годин було в середньому виконано за проектом на 8 робочих годин, тобто з якою ефективністю виконувалися роботи.

Слід зазначити, що використання показників відхилення за розкладом і індексу виконання розкладу ґрунтується не на тимчасових показниках, а на показниках об'ємів робіт.

Використовуючи індекс виконання розкладу SPI і середні планові об'єми в розрахунку на одиницю часу, можна визначити *оцінку тривалості проекту* при завершенні (за умови, що поточні тенденції зберігатимуться і в майбутньому).

Таким чином, прогнозна тривалість проекту (TimeEstimate at Completion – EACt) визначається за формулою:

$$EACt = (BAC / SPI) / (BAC / N),$$

де N – це кількість періодів, на якій розподілені роботи.

Така оцінка є приблизною і грубою, але тим не менше вона відображає об'єктивну ситуацію.

*Відхилення за витратами* (за вартістю) (Cost Variance – CV) визначається як різниця освоєного об'єму (Earned Value – EV) і фактичних витрат (ActualCost – AC).

Показник відхилення за витратами може бути розрахований в відсотковому відношенні:

$$CV\% = (EV - AC) / EV = CV / EV$$

Це означає, що витрати проекту на поточну дату перевищують плановий бюджет на ту ж дату.

*Показники освоєного обсягу* (EV) і *фактичних витрат* (AC) використовуються для визначення індексу виконання бюджету (Cost Performance Index – CPI), одного з найважливіших показників ефективності проекту. Індекс виконання бюджету визначається як частка освоєного обсягу (EV) і фактичних витрат (AC):

Відхилення при завершенні може бути виражено і в відсотковому відношенні:

$$VAC\% = (BAC - EAC) / BAC = VAC / BAC$$

Таким чином, можемо зробити висновок, що кожна за приведених формул може бути адаптована для конкретного проекту.

Взаємозв'язок між основними показниками методу основного об'єму у класичному вигляді графічно показаний на рис. 1.

Оскільки, ми маємо брати до уваги не тільки планові об'єми, а і можливість виробітку коригуючих впливів, потрібно на етапі планування робіт ввести поняття раннього та пізнього закінчення робіт.

Раннє закінчення роботи – визначається сумою раннього початку і тривалості розглянутої роботи.

Пізнє закінчення роботи – визначається різницею тривалості критичного шляху і максимальної тривалості шляху від завершального події всієї моделі до кінцевого події розглянутої роботи.

Внесення до планових робіт такого діапазону дозволить керівнику проекту ще на етапі планування робіт скорегувати у кращу сторону хід реалізації проекту.

Графічно такий вид планування буде мати вигляд, показаний на рис. 2.

Таким чином, під час планування, менеджеру слід застосувати метод, що допоможе правильно встановити ранні та пізні строки завершення робіт у умовах невизначеності. Саме таким методом є метод PERT.

PERT – це абревіатура від англійських Project Evaluation and Review Technique, тобто технологія оцінки, яка базується на ідеї мережевого планування. Для того щоб розрахувати для кожної роботи дату початку і дату закінчення, потрібно оцінити тривалості всіх робіт з оцінкою того, скільки вони займуть часу. Тільки після цього менеджер зможе цю модель розрахувати і зрозуміти терміни закінчення робіт.

У даній техніці за основу береться не один показник терміну робіт, а 3:

- оптимістичну;
- песимістичну;
- найбільш вірогідну.

Виходячи з початкових даних, із заздалегідь заданою вірогідністю проводиться розрахунок відповідних термінів.

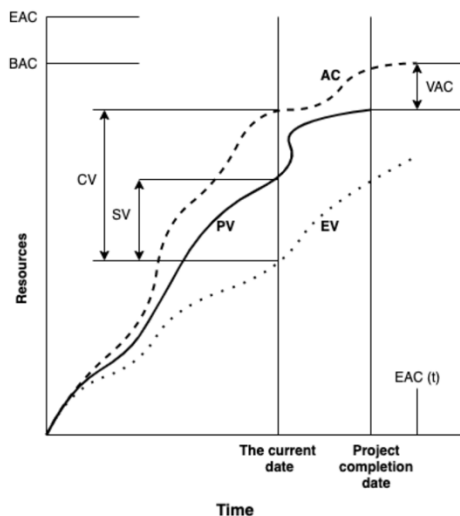


Рисунок 1 – Взаємозв'язок показників методу освоєного обсягу

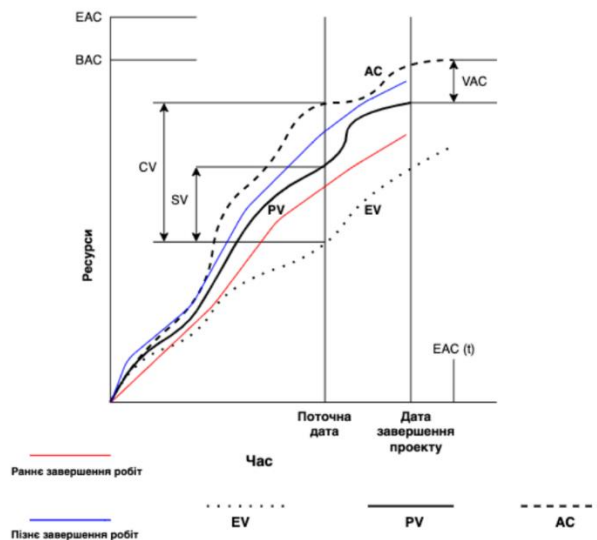


Рисунок 2 – Графічне відображення показників методу освоєного об'єму в умовах невизначеності

Тривалість кожної роботи в проекті має обмеження від найкращого до найгіршого значення, а середній показник підлягає розрахунку. Відставання від графіка, викликані помилкою в момент встановлення тривалості або суб'єктивними причинами, мають властивість тривати і далі. Причому час операцій може відхилитися або в сторону нижнього, або в сторону верхньої межі.

Таким чином, метод PERT призначений для оптимізації тривалості проектів за рахунок логіко-управлінських рішень і, більшою мірою, завдяки застосуванню статистичних методів. Техніка дозволяє при розрахунку тривалості робіт застосувати імовірнісний підхід з використанням середнього значення розподілу.

Ряд розподілених показників тривалості доступний до аналізу як сума середньозважених значень для операцій, включених до складу критичного шляху. Отримання такого середньозваженого показника разом з відхиленням дає менеджеру можливість розрахуватися ймовірності можливої тривалості як операцій, так і проекту. У методі застосовується наступна формула розрахунку середньозваженого часу операції:

$$tE = (tO + 4tM + tP) / 6,$$

де  $(tM)$  – найбільш ймовірний час операції;  $(tO)$  – оптимістичне час операції;  $(tP)$  – песимістичний час операції.

Будучи залежною від передбачуваного розподілу значень в діапазоні трьох оцінок, очікувана тривалість  $tE$  розраховується за представленою формулою.

Оскільки розрахунки методу PERT базуються у тому числі на значенні критичного шляху, ми маємо взяти до уваги алгоритм для впорядкування початкового графу залежності робіт. Надійним доповненням до методу PERT є алгоритм Демукрона. Таке поєднання надасть менеджеру надійний інструментарій для реалізації проекту за ресурсними параметрами.



Цей метод полягає в обчисленні порядкової функції та її упорядкуванні. Передбачається, що вершини мережі пронумеровані від 1 до  $n$ . У нашому випадку кожна за вершин символізує окрему роботу.

Наочно процес визначення рівнів вершин можна представити таким чином. Нульовий рівень утворюють входи. Видаливши з мережі всі вершини нульового рівня і виходять з них дуги, знову отримаємо мережу, входами якої будуть вершини першого рівня вихідної мережі. Зазначений процес "пошарового" видалення вершин слід продовжувати до тих пір, поки всі вершини вихідної мережі не будуть розподілені за рівнями. Таким чином ми зможемо провести поетапну декомпозицію робіт.

Алгоритм Демукрона використовує матрицю суміжності вершин як засіб уявлення мережі і заснований безпосередньо на визначенні рівня вершини та їх сортування. Графічно результат роботи алгоритму відтворений на рис. 3.

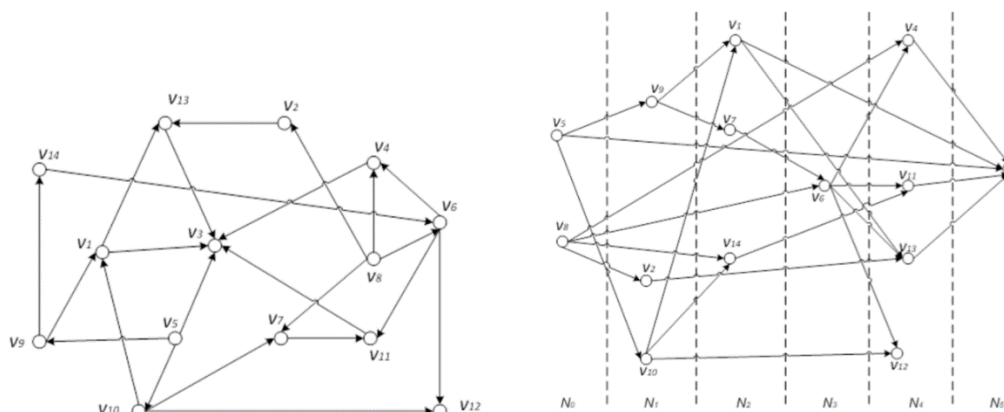


Рисунок 3 – Упорядкування графа алгоритмом Демукрона

### Висновки

Таким чином, у даній роботі мною була запропонована методика розробки коригувальних рішень при реалізації ІТ-проекту.

Тож, тепер ми маємо повний інструментарій для успішного моніторингу та аналізу будь-якого ІТ-проекту за ресурсними показниками. Можемо підбити висновки та сформулювати деякий алгоритм моніторингу реалізації ІТ-проекту за розробленою методикою.

На початку проекту ми маємо невпорядкований граф залежності робіт. За допомогою алгоритму Демукрона реалізуємо поетапну декомпозицію та отримуємо основу для формування календарного плану реалізації проекту.

Наступним кроком є встановлення термінів реалізації робіт та розрахування песимістичного, оптимістичного та найбільш вірогідного сценарію реалізації проекту за допомогою методу PERT.

Останнім етапом є обчислення основних показників освоєного обсягу за вибраними критеріями та виробка коригуючих рішень для успішної реалізації ІТ-проекту.

Розроблена методика є ефективною та може бути використана керівниками будь-яких проектів, а також застосована у навчальному процесі.

### Список літератури

1. Devaux, Stephen A. Managing Projects as Investments: Earned Value to Business Value. CRC Press. pp. 160–171, 2014.
2. Marshall, Robert. The Contribution of Earned Value Management to Project Success of Contracted Efforts. Journal of Contract Management, pp. 21-331, 2007.
3. Либерзон В. Основные понятия и процессы управления проектами. Директор ИС, 2000. № 3.

4. Harold Kerzner, Project Management: A Systems Approach to Planning, Scheduling, and Controlling (8th ed.), 2003.

5. Белоусов А. И., Ткачев С. Б. Дискретная математика. — М.: МГТУ, 2006.

## МЕТОДИКА УПРАВЛЕНИЯ РИСКАМИ ПРОЕКТА

*Романская А.С.*

*Харьковский национальный университет имени В.Н. Каразина*

Процессы принятия решений в управлении проектами происходят, как правило, в условиях неопределенности, которая зависит от следующих факторов:

- неполное знание всех параметров процесса управления или обстоятельств при ситуации выбора решения;
- наличие вероятностного характера поведения окружающей среды и, соответственно, необходимости учета факторов случайности;
- наличие субъективных факторов противодействия, когда принятие решений идет в ситуации взаимодействия партнеров с противоположными или не совпадающими интересами [1].

Под риском понимается потенциальная, численно измеримая возможность неблагоприятных ситуаций и связанных с ними последствий в виде потерь, ущерба, убытков.

Проблема исследования рисков в процессах разработки программного обеспечения усложняется вследствие возрастания разнообразия и сложности разрабатываемых программных продуктов. Современные проекты создания программного обеспечения характеризует невозможность четко описать продукт проекта на начальных стадиях его реализации.

Неспособность правильно управлять рисками может привести к ощутимым экономическим потерям, которые могли бы быть значительно уменьшены своевременным анализом, прогнозированием и сокращением рисков, что делает это направление одним из самых актуальных, особенно для крупных производителей программного обеспечения, разрабатывающих сотни программных продуктов. Использование риск-менеджмента на практике позволит повысить обоснованность управленческих решений и экономическую эффективность деятельности предприятий-производителей программного обеспечения [2, 3].

Создание экономико-математических методик управления рисками в сфере программной инженерии является на сегодняшний день важной и недостаточно изученной задачей, что указывает на актуальность проведения исследования по данной тематике.

Повышение достоверности количественной оценки риска проекта и ускорение процесса принятия решений за счет создания методики управления рисками. Для достижения поставленной цели необходимо решить следующие задачи:

- сформировать концепцию риска инновационного проекта;
- провести анализ существующих методик оценки риска проекта;
- сформировать и классифицировать существенные факторы риска проекта на всех этапах его жизненного цикла;
- разработать модель процесса анализа риска, адаптированную для проектов;
- развить методику принятия решений в условиях риска.

Методика управления рисками проекта на основе метода PERT и дерева решений позволит ответить на два вопроса при разработке проекта: разрабатывать ли проект и за какое время его разработка будет эффективнее в итоге.

Предположим, что перед руководителем проекта стоит вопрос вводить ли в эксплуатацию новую технологию? Если новая технология будет работать безотказно, компания получит прибыль 200 млн. у.е. Если же она откажет, компания может потерять 150 млн. у.е. По оценкам главного инженера, существует 60% шансов, что новая технология не заработает. Можно создать экспериментальный прототип, а затем уже решать, вводить технологию в реальный проект или нет. Эксперимент обойдется в 10 млн. у.е.

Главный инженер считает, что существует 50% шансов, что экспериментальный прототип будет работать. Если экспериментальный прототип будет работать, то 90% шансов за то, что реальный проект с новой технологией также будет работать. Если же прототип не будет работать, то только 20% шансов за то, что реальный проект с новой технологией заработает.

Следует ли разрабатывать прототип? Следует ли разрабатывать реальный проект с новой технологией? Какова ожидаемая стоимостная оценка наилучшего решения? Стоит ли уменьшать количество времени на разработку?

Определим длительность разработки прототипа и в дальнейшем реального проекта по экспертным оценкам из таблицы расчета сетевой модели проекта (табл. 1).

Таблица 1 – Расчет сетевой модели проекта

Работа	Кодировка работы по номерам событий i-j	Оптимистическая оценка (a)	Наиболее вероятная оценка (t)	Пессимистическая оценка (b)	Ожидаемая средняя продолжительность (t <sub>ср</sub> )	Дисперсия оценки продолжительности работы	Раннее начало i-ой работы	Раннее окончание i-ой работы	Позднее начало i-ой работы	Позднее окончание i-ой работы	Полный резерв i-ой работы
1	2	3	4	5	6	7	8	9	10	11	12
1	1-2	2	2,25	2,3	2,22	0,00	0,00	2,22	0,00	2,22	0,00
2	2-3	1,6	1,75	1,95	1,76	0,00	2,22	3,98	2,22	3,98	0,00
3	1-3	3	4	4,5	3,92	0,06	0,00	3,92	0,06	3,98	0,06
4	3-4	0,9	1,7	1,9	1,60	0,03	3,98	5,58	3,98	5,57	0,00
5	4-5	4,8	5,2	5,6	5,20	0,02	5,58	10,78	5,57	10,78	0,00
6	5-6	2,4	2,8	3,2	2,80	0,02	10,78	13,58	10,78	13,58	0,00
7	6-7	1,9	2,1	2,2	2,08	0,00	13,58	15,66	13,58	15,66	0,00
8	7-8	2	2,4	3,6	2,53	0,07	15,66	18,19	15,66	18,19	0,00
9	8-9	5,6	6	8	6,27	0,16	18,19	24,46	18,19	24,46	0,00
10	9-10	0,3	0,4	0,7	0,43	0,00	24,46	24,89	24,46	24,89	0,00
11	1-12	4	4,8	5,6	4,80	0,07	0,00	4,80	1,43	6,23	1,43
12	12-13	4,4	5,2	6	5,20	0,07	4,80	10,00	6,23	11,43	1,43
13	13-14	1,2	2,4	3,6	2,40	0,16	10,00	12,40	11,43	13,83	1,43
14	14-15	12,8	14	15,6	14,07	0,22	12,40	26,47	13,83	27,89	1,43
15	1-10	1,6	2	2,8	2,07	0,04	0,00	2,07	22,83	24,89	22,83

Вместо одной детерминированной величины продолжительности для работ проекта задаются три оценки длительности: **оптимистическая** (работа не может быть выполнена быстрее, чем за  $t_{\text{опт}}$ ); **пессимистическая** (работа не может быть выполнена медленнее, чем за

$t_{\text{песс}}$ ) и **наиболее вероятная**  $t_{\text{вер}}$ . Затем вероятностная сетевая модель превращается в детерминированную путем замены трех оценок продолжительностей каждой из работ одной величиной, называемой **ожидаемой продолжительностью**  $t_{ie}$  и рассчитываемой как средневзвешенное арифметическое трех экспертных оценок длительностей данной работы:

$$t_{ie} = \frac{t_{i\text{опт}} + 4t_{i\text{вер}} + t_{i\text{песс}}}{6}.$$

Для характеристики степени неопределенности оценки продолжительности отдельной работы служит **дисперсия**, вычисляемая по формуле:

$$\sigma_i^2 = \left( \frac{t_{i\text{песс}} - t_{i\text{опт}}}{6} \right)^2.$$

Надо отметить, что вероятности завершения проекта в сроки, определяемые оптимистической или пессимистическими величинами, стремятся к нулю. Если подставить в детерминированную модель ожидаемые продолжительности каждой работы, то можно получить общую продолжительность проекта, вероятность окончания разработки к сроку, определяемому которой, равна 50 %, т.е. так называемую **ожидаемую продолжительность проекта** [4].

Однако разработчиков больше интересует вероятность завершения проекта к определенному моменту времени, например, к договорному сроку  $T_{\text{план}}$ . Она также может быть определена с помощью соотношения:

$$Z = \frac{T_{\text{ПЛАН}} - T_E}{\sigma T_e}.$$

На основании полученной величины  $Z$  по специальным таблицам можно найти соответствующую вероятность, выраженную в процентах или в долях единицы.

Критический путь:

(1,2)(2,3)(3,4)(4,5)(5,6)(6,7)(7,8)(8,9)(9,10)(10,11)(11,15)(15,16)(16,17)(17,18)(18,19)(19,20)(20,21)(21,22)(22,23)(23,24)(24,25)(25,26)(26,27)(27,29)(29,30)

Продолжительность критического пути: 67.29.

Для определения максимально возможного срока выполнения всего комплекса работ с надежностью 95 % будем использовать [5] формулу:  $T = T_{\text{кр}} + Z \cdot S_{\text{кр}}$ .

Для решения поставленной задачи найдем значение аргумента  $Z$ , которое соответствует заданной вероятности 95% (значению графы  $\Phi(Z)$   $0.95 \cdot 100\%$  в таблице соответствует  $Z=1.96$ ).

$$T = 67.29 + 1.96 \cdot 1.24 = 69.72$$

Максимальный срок выполнения всего комплекса работ при заданном уровне вероятности 95% составляет всего 69.72 дня.

Известно, что при сроке выполнения работ в 65 дней проект принесет прибыль 200 млн. у.е. с вероятностью 40 % и станет убыточен с вероятностью 60% и убытком в 150 млн. у.е.

При сроке выполнения работ за 50 дней с вероятностью 60% проект принесет прибыль 150 млн. у.е. и убыток в 100 млн. у.е. с вероятностью 40%.

Используем метод дерева решений для выбора эффективного развития проекта. Построим дерево решений проекта (рис. 1).

После создания дерева решений проведем обратный анализ: идя по дереву справа налево, для каждого узла принятия решения рассчитываем ожидаемую величину ожидаемой денежной стоимости. Расчет ожидаемой денежной стоимости EMV [6] сделано по следующей формуле:

$$EMV_j = \sum_{i=1}^n u_i p_i.$$

В узле F возможны исходы «не работает» с вероятностью 0,4 (что приносит прибыль 200) и «не работает» с вероятностью 0,6 (что приносит убыток — 150) => оценка узла F.  $EMV(F) = 0,4 \times 200 + 0,6 \times (-150) = -10$ . Это число мы пишем над узлом F.

$$EMV(G) = 0.$$

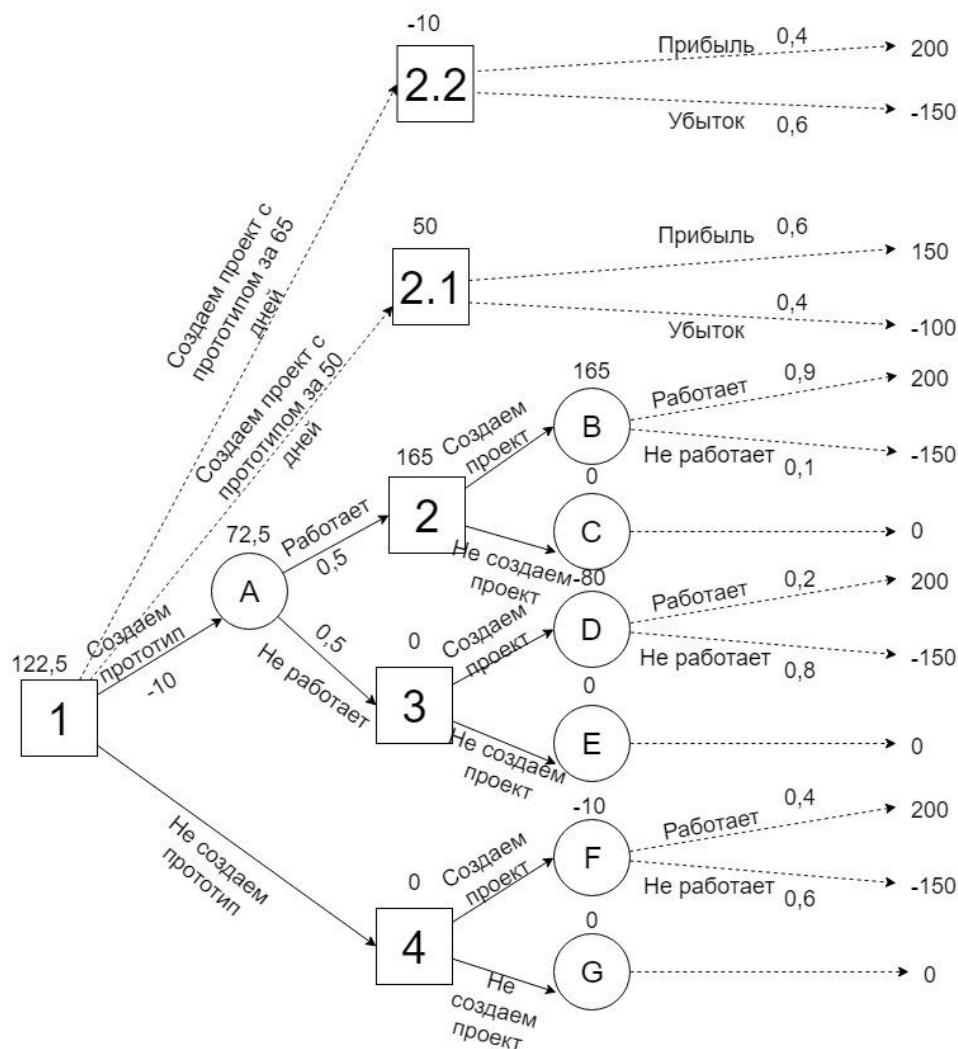


Рисунок 1 – Дерево решений проекта

В узле 4 мы выбираем между решением «создаем проект» (оценка этого решения  $EMV(F) = -10$ ) и решением «не создаем проект» (оценка этого решения  $EMV(G) = 0$ ):  $EMV(4) = \max \{EMV(F), EMV(G)\} = \max \{-10, 0\} = 0 = EMV(G)$ . Эту оценку мы пишем над узлом 4, а решение «создаем проект» отбрасываем и зачеркиваем.

Аналогично:

$$EMV(B) = 0,9 \times 200 + 0,1 \times (-150) = 180 - 15 = 165.$$

$$EMV(C) = 0.$$

$EMV(2) = \max \{EMV(B), EMV(C)\} = \max \{165, 0\} = 165 = EMV(2)$ . Поэтому в узле 2 отбрасываем возможное решение «не создаем проект».

$$EMV(D) = 0,2 \times 200 + 0,8 \times (-150) = 40 - 120 = -80.$$

$$EMV(E) = 0.$$

$EMV(3) = \max \{EMV(D), EMV(E)\} = \max \{-80, 0\} = 0 = EMV(E)$ . Поэтому в узле 3 отбрасываем возможное решение «создаем проект».

$$EMV(A) = 0,5 \times 165 + 0,5 \times 0 - 10 = 72,5.$$

Отдельно рассчитаем  $EMV$  для узлов 2.1 и 2.2:

$$EMV(2.1) = 0,6 \times 150 + 0,4 \times (-100) = 50. \text{ Пишем над узлом 2.1}$$

$$EMV(2.2) = 0,4 \times 200 + 0,6 \times (-150) = -10. \text{ Пишем над узлом 2.2}$$

$EMV(1) = \max \{EMV(A), EMV(4)\} + \max \{EMV(2.1), EMV(2.2)\} + \max \{72,5; 0\} + \max \{50; -10\} = 72,5 + 50 = 122,5 = EMV(A)$ . Поэтому в узле 1 отбрасываем возможное решение «не создаем прототип».

Ожидаемая стоимостная оценка наилучшего решения равна 122,5 млн. у.е. Строим прототип. Если прототип работает, то создаем проект. Если прототип не работает, то проект создавать не надо. Для увеличения прибыли строим прототип и проект за 50 дней.

Таким образом, мы понимаем, что необходимо ускорить разработку проекта.

Для оценки вероятности выполнения всего комплекса работ за 50 дней нам необходима формула:  $P(t_{кр})$ , где  $Z = (T - T_{кр}) / S_{кр}$ ,  $Z$  – нормативное отклонение случайной величины,  $S_{кр}$  – среднее квадратическое отклонение, вычисляемое как корень квадратный из дисперсии продолжительности критического пути.  $\Phi(Z)$  – значение дифференциальной функции нормального распределения вероятностей, называемой функцией Лапласа, определяют в зависимости от ее аргумента  $Z$  по таблице [7].

Критический путь проходит по работам.

Дисперсия критического пути:  $S^2(L_{кр}) = S^2(1,2) + S^2(2,3) + S^2(3,4) + S^2(4,5) + S^2(5,6) + S^2(6,7) + S^2(7,8) + S^2(8,9) + S^2(9,10) + S^2(10,11) + S^2(11,15) + S^2(15,16) + S^2(16,17) + S^2(17,18) + S^2(18,19) + S^2(19,20) + S^2(20,21) + S^2(21,22) + S^2(22,23) + S^2(23,24) + S^2(24,25) + S^2(25,26) + S^2(26,27) + S^2(27,29) + S^2(29,30)$

$S^2(L_{кр}) = 0 + 0,07 + 0,16 + 0 + 0,07 + 0,07 + 0,16 + 0,22 + 0,07 + 0,07 + 0 + 0,04 + 0,16 + 0,04 + 0,04 + 0,22 + 0 + 0,07 + 0 + 0,01 + 0,04 + 0 + 0 + 0 + 0,04 + 0,16$

$S(L_{кр}) = 1,24$

$P(t_{кр} < 50) = 0,5 + 0,5\Phi((50 - 67,29)/1,24) = 0,5 + 0,5\Phi(-13,94) = 0,5 + 0,5 * (-0,49999) = 0,25$

Вероятность того, что весь комплекс работ будет выполнен не более чем за 50 дней, составляет 25%.

Соответственно, после применения мер по ускорению разработки проекта прибыль от разработки проекта окажется максимальной.

### Выводы

Разработанная методика показывает, что метод PERT в сочетании с методом дерева решений является удобным способом работы с событиями риска, которые считаются независимыми друг от друга (которые могут не соответствовать действительности), не только в малых, но и в больших проекты, формально или неформально.

Методика интегрирована в общепринятую процедуру управления рисками и применяется на этапе оценки и анализа рисков. Эта методика эффективна благодаря тому, что она связывает наиболее опасные и возможные риски со способностью управлять ими [8].

Таким образом, использование рассмотренной методики управления проектными рисками на предприятии позволит структурировать саму систему управления рисками, учитывать процедурный аспект управления рисками путем разработки алгоритма действий на каждом этапе проекта. Внедрение централизованной системы управления рисками также может решить проблему непредвиденных рисков для предприятия и его проектов, упростить принятие решений о рисках и снизить затраты на управление рисками по сравнению с существующим уровнем.

### Список литературы

1. Качалов Р.М. Управление экономическим риском: теоретические основы и приложения: монография. М., СПб: Нестор-История, 2012. 248 с.
2. Терюхов В.Е. Риск-анализ инвестиционного проекта // Управление риском. 2002. № 3. С. 18–20.
3. Чернов В.Б. Анализ риска комплексного инвестиционного проекта // Управление риском. 2003. № 3. С. 56–64.

4. Смоляк С.А. Оценка эффективности инвестиционных проектов в условиях риска и неопределенности. Теория ожидаемого эффекта. М.: Наука, 2002. 182 с.
5. Лепешкина М.Н. Методологические аспекты оценки рисков // Менеджмент в России и за рубежом. 2001. № 6. С. 88–98.
6. Сорокина О.В. Имитационное моделирование как метод оптимизации инвестиционных рисков // Традиции, инновации и инвестиции современной рыночной экономики: матер. Междунар. конф. Казань, 2004 г. Ч. 2. Казань: Академия управления «ТИСБИ», 2004. С. 351–355.
7. Секерин А.Б., Селютин В.Д., Строев С.П. Нечетко-множественная модель управления риском экономической несостоятельности производственного предприятия // Управление риском. 2008. № 2. С. 28–35.
8. Студеникина С.А., Заславская С.Е. Использование математических моделей в управлении рисками // Современные проблемы текстильной и легкой промышленности: тезисы докладов Межвуз. науч.-тех. конф. Ч. 2. М.: РосЗИТЛП, 2004. 160 с.

## МЕТОДИ КЛАСИФІКАЦІЇ СТАНУ ПАЦІЄНТІВ В СИСТЕМАХ МЕДИЧНОГО МОНІТОРИНГУ

*Донець Володимир Віталійович*

*Харківський національний університет ім. В. Н. Каразіна*

Поява та розвиток помилок у складних системах є завжди складним динамічним процесом, тому зазвичай експерти не завжди можуть прогнозувати їх виникнення чи тип. Але це дуже важливо – правильно визначити, або спрогнозувати помилку для повернення системи в нормальний режим роботи. Контроль та прогнозування динаміки процесу допомагає спеціалістам приймати ефективні рішення. Одним з прикладів складної системи – є медико-біологічна система, що складається з пацієнта – динамічної системи, в якій виникають помилки, лікар – в ролі експерта та система діагностики стану пацієнтів – система, що допомагає експерту проводити прогнозування динаміки процесу. На кожному етапі лікування пацієнта може виникати множина кінцевих станів різної розмірності, що визначається лікарем. Проблема розпізнавання поточного стану пацієнта зводиться до проблеми класифікації значень змінних стану пацієнта.

Велика увага приділяється проблемам класифікації стану складних систем. На сьогоднішній день опубліковано багато робіт, що описують методи вирішення класифікаційної проблеми як технічних так й біомедичних систем. На основі цих робіт можливо розробити власний метод класифікації стану пацієнтів в системах медичного моніторингу.

Були розглянуті такі задачі:

- 1) порівняння ефективності роботи методів класифікації стану складних динамічних систем;
- 2) розробка методу класифікації стану пацієнтів в системах медичного моніторингу на основі штучних нейронних мереж із використанням логістичної регресії;
- 3) порівняння ефективності роботи існуючих методів із розробленим методом.

Розглянемо алгоритми, які застосовуються при розв'язанні задачі класифікації.

*Naive Bayes Classifier* – це простий імовірнісний класифікатор, заснований на теоремі Байеса з «наївним» припущенням незалежності. Байєсівські класифікатори можуть бути дуже ефективними [1].

*K-Nearest Neighbors Classifier* – це метричний алгоритм автоматичної класифікації об'єктів або регресії, він також є одним з найбільш часто використовуваних і відомих методів виконання завдань розпізнавання та вирішення задач виведення даних [2]. Серед переваг цього

методу можна виділити простоту впровадження та високу продуктивність та не високу ймовірність помилки [3].

*Random Forest Classifier* – алгоритм машинного навчання, запропонований Лео Брайманом [4], суть якого полягає у використанні певного набору дерев рішень. Цей алгоритм поєднує метод брейманської упаковки та метод випадкового простору, запропонований Тінг Кам Хо. Завдяки тому, що дерево рішень є основою алгоритму, алгоритм забезпечує порівняння важливості параметрів і допомагає виявити надзвичайно важливі властивості для визначення кінцевого класу або стану [5, 6]. Наведено роботи, в яких цей алгоритм був використаний для визначення стану пацієнтів, розпізнавання мови та почерку [5]. Випадковий лісовий класифікатор також допомагає визначити тип раку, виділяючи важливі параметри, на які лікарі можуть звернути увагу за власною класифікацією, що може збільшити ймовірність діагностування раку навіть за відсутності доступу до цієї системи [6].

*Artificial Neuronal Network* – атематична модель, а також її програмне або апаратне втілення, побудована за принципом організації та функціонування біологічних нейронних мереж – мереж нервових клітин живого організму. Це поняття виникло при вивченні процесів, що протікають в мозку, і при спробі змодельовати ці процеси. Першою такою спробою були нейронні мережі У. Маккалок і У. Питтса. Після розробки алгоритмів навчання одержувані моделі стали використовувати в практичних цілях: в задачах прогнозування, для розпізнавання образів, в задачах управління та ін.

*Logistic Regression* – це статистична модель, яка в своїй базовій формі використовує логістичну функцію для моделювання бінарної залежної змінної, хоча існує ще багато складних розширень [4].

Для розв'язання задачі класифікації стану пацієнтів в системах медичного моніторингу була обрана радіально-базисна штучна нейронна мережа.

Структура радіально-базисної штучної нейромережі з одним прихованим рівнем ( $K=1$ ) зображена на рис. 1. Тут  $\{Y_{ph}^{(0)}\}$  – це набір входів,  $\{Y_i^{(k)}\}$  – набір виходів  $k$ -ого рівня,  $k=1..(K+1)$ ;  $K$  – кількість скритих рівнів;  $\{w_{ij}^{(k)}\}$  – набір вагових коефіцієнтів на  $k$ -ому рівні;  $N_0$  – кількість нейронів на вході мережі;  $N_1$  – кількість скритих нейронів;  $N_2$  – кількість нейронів на виході.

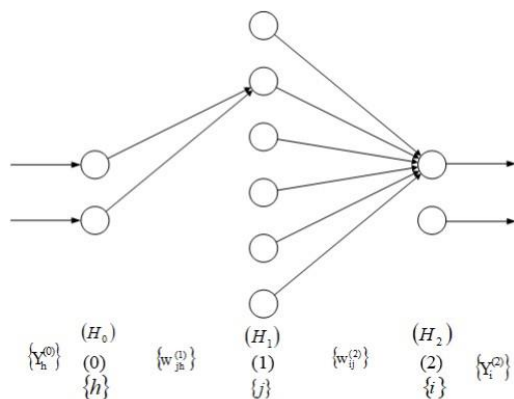


Рисунок 1 – Структура нейронної мережі

Алгоритм навчання радіально-базисної штучної нейронної мережі детально описаний у [7].

Для оцінки роботи нейромережі використаємо зразок урологічних даних, що містить 47 змінних стану, значення яких були трьох типів: дійсні числа, булеві і порядкові. Навчальний зразок складався з даних про 30 пацієнтів, тестовий зразок містив дані про 10 пацієнтів. Результатом класифікації є визначення стану пацієнта (здорового або хворого). Для класифікації використовували ANN з логістичною функцією. В результаті навчання помилка склала 0,122 і стандартне відхилення 0,0024. Результат розпізнавання на контрольній

вибірці становить 100%.

В результаті навчання та мережевого тестування зразка урологічних даних була побудована матриця плутанини та крива ROC для аналізу якості класифікації (рис. 2).



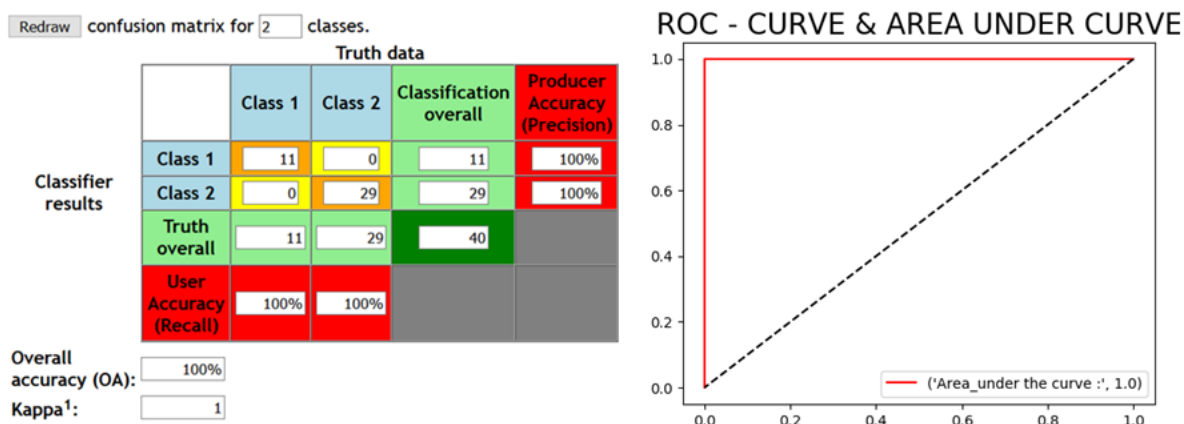


Рисунок 2 – Матриця плутанини та ROC-крива

### Висновки

В роботі було розглянуто деякі методи класифікації стану складних систем, наведено приклад оцінки якості класифікації медичних даних з використанням розломленої для цього штучної нейронної мережі. На майбутнє в роботі будуть більш детально розглянуті та порівняні існуючі методи класифікації стану складних систем, а також порівняна їх ефективність для класифікації стану пацієнтів. На основі цих даних буде розроблений власний метод класифікації стану, котрий буде порівняно з існуючими методами

### Список літератури

1. Zhang Harry. The Optimality of Naive Bayes. American Association for Artificial Intelligence, 2004.
2. Rennie J, Shih L, Teevan J, Karger D. Tackling the poor assumptions of Naive Bayes classifiers. Proceedings of the Twentieth International Conference on Machine Learning, 2003, Washington DC.
3. Denisko D, Hoffman MM. Classification and interaction in random forests. Proceedings of the National Academy of Sciences of the United States of America, 115(8), pp. 1690–1692, February 2018.
4. Leo Breiman, Random Forest, Machine Learning: Proceedings of the Thirteenth International conference, 45, 5–32, 2001.
5. Denisko D, Hoffman MM. Classification and interaction in random forests. Proceedings of the National Academy of Sciences of the United States of America, 115(8), pp. 1690–1692, February 2018.
6. Wouter G Touw, Jumamurat R Bayjanov, Lex Overmars, Lennart Backus, Jos Boekhorst, Michiel Wels, Sacha A F T van Hijum. Data mining in the Life Sciences with Random Forest: a walk in the park or lost in the jungle? Brief Bioinform, 14(3), pp. 315–326, May 2013.
7. Системное совершенствование элементов сложных технических систем на основе концепции обратных задач [Текст] : монография/ В. Е. Стрелец, А. А. Трончук, Е. М. Угрюмова и др.; под общ. ред. М. Л. Угрюмова. – Х. : Нац. аэрокосм. ун-т им. Н. Е. Жуковского «Харьк. авиац. ин-т», 2013. – 148 с.

## СИСТЕМА РОЗПІЗНАВАННЯ АУДІОІНФОРМАЦІЇ ЗА ДОПОМОГОЮ ШТУЧНИХ НЕЙРОМЕРЕЖ

Кузьменко М.В.

Харківський національний університет ім. В. Н. Каразіна

Кожен день пошуковими системами користуються мільйони людей. Однак всі пошукові системи працюють в розрахунок на пошук тексту і іноді зображень, і до сих пір існує дуже мало програмних засобів для пошуку музичних композицій за фрагментами їх мелодії або гармонії, багато які з них або малоефективні, або їх розробка перебуває в зародковому стані. Проблема

пошуку музики стає актуальною в сучасному світі, коли її велика кількість існує в електронному вигляді.

Поряд з цим існує й інша проблема, яка існує набагато довше, що йде корінням в історію музики і з розвитком цифрового звукозапису знаходить своє рішення. Це – розпізнавання музичного запису і перетворення його в початкове, нотне подання.

Метою роботи є розробка методики та програмної реалізації розпізнавання аудіоінформації за допомогою чисельних методів та нейронних мереж.

На даний момент існує кілька програмних рішень, призначених для розпізнавання аудіо. Amazing Midi – програма для розпізнавання аудіо та перетворення в midi;

Digital Ear – близька до першого варіанту система, за кількома додатковими інструментами;

D'Accord iChords 2.0 – програмне рішення для розпізнавання акордів;

IntelliScore Ensemble WAV to MIDI Converter – найбільш ефективне програмне рішення, що дозволяє отримати корисний результат.

Коли ми говоримо про розпізнавання, ми маємо на увазі здатність машини, тобто нашого з Вами комп'ютера, "розуміти", що представлено йому на розгляд. Типова задача автоматичного розпізнавання формулюється приблизно так: вся множина аудіоінформації, яка підлягає обробці, деяким способом розбивається на кінцеве число класів, званих образами [1]. Автоматичному пристрою, забезпеченому механізмом сприйняття образів, пред'являється аудіофрагмент і пристрій повинен вирішити, до якого класу він належить. Таким чином, завдання розпізнавання зводиться до самостійної класифікації машиною аудіоінформації, без допомоги людини.

Кожен звук, що входить в співзвуччя, має в своєму спектрі, крім основного тону, набір коливань більш високої частоти – обертони. Співвідношення частот обертонів для тону визначають тембр звуку.

Основною проблемою розпізнавання є те, що у кожного тембру – своє співвідношення частот, причому обертони різних нот можуть накладатися. Тому відокремити обертони від основних тонів стає серйозною проблемою.

Після аналізу штучних нейромереж та чисельних методів було обрано застосування:

1. Перетворення Фур'є і отримання вхідних частот.
2. Виділення найбільш сильних гармонік (в рамках «роздільної здатності» програми) і знаходження їх обертонів.
3. Передача обертонів нейронній мережі, яка виконує розпізнавання нот.

Штучні нейронні мережі (ШНМ) – математичні моделі, а також їх програмні або апаратні реалізації, побудовані за принципом організації та функціонування біологічних нейронних мереж. Перевага нейромережевого підходу перед іншими методами розпізнавання образів полягає в тому, що нейромережі здатні відновлювати гіперповерхні, які розділяють, як завгодно складної форми, не спираючись на гіпотези про компактності або лінійної роздільності класів.

Для вирішення наведеної задачі пропонується використовувати класичний штучний багатошаровий перцептрон без зворотного зв'язку.

Нейронна мережа складається з шару рецепторів, які представляють собою входи матриці обертонів, прихованого шару нейронів і вихідного шару нейронів. Кожен вихід є індикатором присутності певної ноти в спектрі, який активується в разі присутності цієї ноти в співзвуччі. Топологія мережі – повнозв'язний граф.

Задача прихованого шару – складання вирішальної гіперповерхні, тобто, виділення нот. Задача вихідного шару – перетворення виходу з прихованого шару в зрозумілий людині вид. Нейрони, що перемогли, будуть відповідати нотам. У якості функції активації використовується класична сигмоїда.

### Висновки

- Застосування перетворення Фур'є дозволяє отримати спектральне подання аудіо.
- Застосування нейронної мережі дозволяє створити автоматизоване (а в теорії – і автоматичне, в динаміці) програмне рішення для розпізнавання музичного аудіо.
- Існує ряд обмежень, що накладаються на розпізнавання, пов'язаних як з апаратно-програмними обмеженнями, так і з особливостями музики.
- У процесі розробки створено програмне рішення на основі чисельних методів і нейронної мережі.

### Список літератури

1. Бодянский Е.В., Руденко О.Г. Искусственные нейронные сети. – Харьков: ТЕЛТЕХ, 2004. – 369 с.: ил.
2. Уоссермен Ф. Нейрокомпьютерная техника – 1992. – с. 144-162.
3. Белобородов А.Ю. Распознавание аудиообразов с применением обертонового ряда. – 2010.
4. Хайкин С. Нейронные сети: Полный курс. — М.: Вильямс, 2008, 1103 с.
5. Шерман Н. Формирование равномерно-темперированного строя. — М., 1964.
6. FFT анализ [Электронный ресурс] / WebSound.ru – режим доступа: <http://websound.ru/articles/theory/fft.htm>
7. Процедура обратного распространения ошибки [Электронный ресурс] / Интернет-университет INTUIT – режим доступа: <http://www.intuit.ru/department/ds/neuronnets/4/1.html>
8. Алгоритм обратного распространения ошибки [Электронный ресурс] / AIPortal – режим доступа: <http://www.aiportal.ru/articles/neural-networks/back-propagation.html>

## ОБРОБКА ТА ЗБЕРЕЖЕННЯ ВЕЛИКИХ ОБСЯГІВ ДАНИХ У ХМАРНИХ СХОВИЩАХ

*Маришев П.Д.*

*Харківський національний університет імені В. Н. Каразіна*

Веб-сервіси Amazon (AWS) надають широкий спектр послуг, щоб допомогти партнерам швидко та легко створювати та розгортати великі програми та аналітичні програми. AWS надає швидкий доступ до гнучких і недорогих ІТ-ресурсів, щоб користувач міг швидко масштабувати практично будь-які великі програми передачі даних, це може включати:

- 1) зберігання даних;
- 2) аналітику;
- 3) виявлення шахрайства;
- 4) рекомендаційні двигуни та керовані події Extract, Transform, Load (ETL) та багато іншого.

Завдяки AWS користувачеві не потрібно вкладати великих інвестицій у час або гроші для побудови та підтримки інфраструктури, натомість можна забезпечити правильний тип та розмір ресурсів, необхідних для живлення великих програм та даних для аналізу. Можна отримати доступ до такої кількості ресурсів, яка потрібна, майже миттєво, і платити лише за те, що використовується. Якщо подивитись на те, як традиційно розробляються та розгортаються аналітичні бази, то користувач може мати декілька баз даних та систем запису, і ці реляційні бази даних можуть зберігати великі обсяги даних для аналізу. Користувач може співпрацювати з сервісами для створення сховища даних, яке містить таблиці з фактами та їх розміри, і цей склад даних буде використовуватися для управління операційними та спеціальними звітами та інформаційними панелями.

Проекти зберігання великих обсягів даних працюють так, що для створення сховища даних потрібно витратити багато грошей наперед, а потім на оплату за кожен гігабайт пам'яті.

Але зараз існують нові сучасні архітектури даних, і однією з таких архітектур є Data Lake (DL). DL може містити реляційні та нереляційні дані та масштабувати від терабайт до екзабайтів розмірів даних. DL робить унікальним те, що схема визначається під час аналізу. Завдяки DL даних багато різних службових інструментів та аналітичних двигунів можуть бути використані для отримання розуміння даних, які вони містять, DL також розроблена для того, щоб скористатися недорогим сховищем, щоб сприяти більш аналітичним та швидким інноваціям. Для реалізації DL на AWS використовується Amazon Simple Storage Service (S3) як основу DL. S3 – довговічна послуга зберігання об'єктів, яка має декілька різних способів введення даних:

- 1) з терміналу .
- 2) через API.
- 3) за допомогою потокової передачі даних.

S3 забезпечує безперебійну довговічність і доступність в надзвичайних масштабах, а можливості відповідності, безпеки та аудиту вбудовані. Використання S3 як основи DL означає, що можна запускати будь-яку аналітику на одних і тих же даних, фактично не переміщуючи дані навколо. В той же час можна масштабувати зберігання та обчислювати самостійно, це означає, що ця нова архітектура DL є дуже економічно вигідною. Основна проблема використання S3 в тому, що для спрощення передачі даних потрібен API на основі фреймворку та сервісу на Amazon, який би підтримував цей API, ось тут Apache Hadoop Framework і Amazon Elastic MapReduce (Amazon EMR).

Apache Hadoop – це програмне забезпечення з відкритим кодом для зберігання та широкомасштабної обробки наборів даних на кластерах. Hadoop – це проект вищого рівня Apache, який будується та використовується глобальною спільнотою учасників та користувачів. Apache Hadoop складається з наступних модулів:

- 1) загальний Hadoop: містить бібліотеки та утиліти, необхідні для інших модулів Hadoop.
- 2) розподілена файлова система Hadoop (HDFS): розподілена файлова система, яка зберігає дані на товарних машинах, забезпечуючи дуже високу сукупну пропускну здатність у кластері.

3) Hadoop YARN: платформа управління ресурсами, відповідальна за управління обчислювальними ресурсами в кластерах і використання їх для планування програм користувачів.

- 4) Hadoop MapReduce: модель програмування для широкомасштабної обробки даних.

Усі модулі в Hadoop розроблені з принциповим припущенням, що збої в апаратному забезпеченні (окремих машин або стелажів машин) є загальними і тому повинні автоматично оброблятися програмним забезпеченням в рамках.

В основі Apache Hadoop є два основні компоненти: розподілена файлова система Hadoop (HDFS) і паралельна обробка даних MapReduce.

HDFS – це розподілена файлова система Hadoop для зберігання великих файлів з можливістю потокового доступу до інформації по блоку, розподіленої по вузлах обчислювального кластеру, яка може складатися з довільного обладнання. HDFS, як і будь-яка файлова система, є ієрархією каталогів з вбудованими в них підкаталогами та файлами.

Кластер HDFS включає наступні компоненти

- 1) Вузол управління, вузол імен або сервер імен (NameNode) - це окремий, унікальний на сервері кластерів програмний код для управління простором імен файлової системи, зберіганням дерева файлів, а також метаданими файлів і довідники. NameNode - це необхідний компонент кластера HDFS, який відповідає за відкриття та закриття файлів, створення та видалення каталогів, контроль доступу з боку зовнішніх клієнтів та відповідність між файлами та блоками, що дублюються (копіюються) на вузлах даних. Сервер імен розкриває для всіх розташування блоків даних на кластерних машинах.

2) Secondary NameNode - вторинний вузол імені, окремих сервер, єдиний в кластері, який копіює зображення HDFS та журнал транзакцій операцій блоку файлів у тимчасову папку, застосовує зміни, накопичені в журналі транзакцій, до зображення HDFS, і записує його у вузол NameNode та очищає журнальні транзакції. Вторинний NameNode необхідний для швидкого відновлення NameNode вручну в разі відмови.

3) Вузол або сервер даних (DataNode, Node) - один із багатьох серверів кластерів з програмним кодом, відповідальним за файлові операції та роботу з блоками даних. DataNode - необхідний компонент кластера HDFS, який відповідає за запис та зчитування даних, виконання команд з NameNode для створення, видалення та копіювання блоків, а також періодично надсилає повідомлення про стан (серцебиття) та обробляє запити для читання та запису з файлові клієнти HDFS-системи. Варто зазначити, що дані переходять від решти вузлів кластера до клієнта через вузол NameNode.

4) Клієнт - користувач або додаток, що взаємодіють через спеціальний інтерфейс (API - інтерфейс програмування прикладних програм) з розподіленою файловою системою. Якщо є достатньо прав, клієнтові дозволено наступні операції з файлами та каталогами: створювати, видаляти, читати, писати, перейменовувати та переміщувати. Створюючи файл, клієнт може чітко вказати розмір блоку файлів (за замовчуванням 64 Мб) та кількість створених реплік (значення за замовчуванням – 3).

Завдяки блоковій реплікації через вузли даних, розподілена файлова система Hadoop забезпечує високу надійність зберігання даних та швидкість обчислення. Крім того, HDFS має ряд відмітних властивостей:

- великий розмір блоку в порівнянні з іншими файловими системами (> 64 Мб), оскільки HDFS призначений для зберігання великої кількості величезних (> 10 ГБ) файлів;
- зосередження уваги на недорогих і, отже, не найнадійніших серверах – толерантність відмов всього кластера забезпечується реплікацією даних;
- реплікація відбувається в асинхронному режимі – інформація поширюється на декілька серверів прямо під час завантаження, тому вихід з ладу окремих вузлів даних не тягне за собою повну втрату даних;
- HDFS оптимізовано для потокового зчитування файлів, тому використовувати його для нерегулярного та довільного читання не представляється практичним;
- клієнти можуть читати та записувати файли HDFS безпосередньо через інтерфейс програмування Java;
- файли записуються один раз, що обмежує вільні зміни до них;
- принцип роботи WORM (пишіть один раз і читайте багато, пишіть один раз – читайте багато разів) повністю звільняє систему від блокування запису-читання. Запис у файл одночасно доступний лише одному процесу, що виключає численні конфлікти запису;
- HDFS оптимізовано для потокової передачі даних;
- стиснення даних та раціональне використання простору диска знизили навантаження на канали передачі даних, які найчастіше є вузьким місцем у розподілених середовищах;
- самодіагностика – кожен вузол даних через певні проміжки часу надсилає діагностичні повідомлення у вузол імені, який записує журнали операцій над файлами у спеціальний журнал;
- всі метадані серверу імен зберігаються в оперативній пам'яті.

#### **Висновки**

Виходячи з вищесказаного можна стверджувати, що на даний момент Amazon є найоптимальнішим рішенням для зберігання і обробки великих даних в хмарі. Завдяки доступності і інтуїтивному інтерфейсу розібратися як працювати зі сховищем може навіть рядовий користувач, не кажучи вже про програмістів. Також з огляду на їхню цінову політику

користувач ніколи не буде переплачувати за інструменти, якими не користується, і за гігабайти, які використовуються. У Amazon є звичайно конкуренти, ближче з усіх підібрався Фаєрбейс, але він набирає популярність за рахунок того, що є можливість безкоштовно користуватися сховищем до певного обсягу даних. Якщо ж дивитися на надійність, то Amazon на даний момент немає рівних.

### Список літератури

1. Сбарські П., Дебойс П., Фергюсон Д. Безсерверна Архітектура AWS, 2015.
2. Сенко О. Робота з великими даними у хмарах, 2016 .
3. Гюттен Д. AWS: Навчальний посібник з веб-служб Amazon Підручник для початківців, 2017.
4. Шонбергер В., Кукір К. Великі дані: революція, яка перетворить, як ми живемо, працюємо та думаємо, 2014.
5. Мінер Д., Шука А. Шаблони дизайну MapReduce, 2015.
6. Лублінський Б., Сміт К., Якубович А. Професійні Рішення Hadoop, 2015.

## КОМП'ЮТЕРНА МОДЕЛЬ СТРАТИФІКАЦІЇ ПАЦІЄНТІВ В СИСТЕМАХ МЕДИЧНОГО МОНІТОРИНГУ НА ОСНОВІ МЕТОДУ НЕЧІТКИХ С-СЕРЕДНІХ.

*Нікіфорова С.Р.*

*Харківський національний університет імені В.Н. Каразіна*

Організація даних у осмислені групування – це один із найбільш фундаментальних способів розуміння та навчання. Кластерний аналіз – це формальне вивчення методів та алгоритмів згрупування або кластеризації об'єктів відповідно до вимірних або сприйнятих внутрішніх характеристик або подібності.

У багатьох сучасних медичних установах є інформаційні системи для зберігання різних медичних даних про стан здоров'я пацієнтів, які лікарі використовують для розпізнавання (діагностики) патологічних процесів. Однак, аналізуючи медичні дані, виявляючи закономірності цих даних та витягаючи їх, доводиться стикатися з проблемою виміру. Розмір збережених даних, що визначається кількістю різних ознак, що описують стан здоров'я пацієнта, дуже великий і іноді досягає кількох десятків і сотень показників. У такій інформації важко орієнтуватися, що може призвести до помилок або неточностей діагностування. Медична сфера є сферою наукових досліджень з дуже високою вартістю помилки. Неточний діагноз або зволікання з його визначенням можуть вартувати пацієнту життя.

Основна мета цієї роботи – вивчення та удосконалення існуючих та нових методів кластеризації у галузі великих даних, та їх практичне застосування для діагностування пацієнтів медико-біологічних закладів.

Проблема кластеризації об'єктів полягає у розбитті всієї сукупності, що аналізується, на невелике число однорідних, схожих між собою, кластерів. Кластерний аналіз не використовує мітки категорій, які позначають об'єкти за допомогою попередніх ідентифікаторів, тобто мітки класів. При чому, поняття кластеру, у даному випадку, використовується як опис певної категорії цих об'єктів, наприклад «здорові пацієнти», «пацієнти з високою ймовірністю захворювання», «хворі», тощо. Для цього можуть використовуватися різні методи.

Загалом, методи машинного навчання можна розділити на дві великі підгрупи. Це навчання з учителем (Supervised Learning) та без учителя (Unsupervised Learning). Ключова різниця між ними полягає у тому, що у першому випадку користувач повинен сам визначати приналежність об'єкта до певного кластеру, і вже на основі дій користувача і будується неймережа. Користувач виступає як вчитель, він повинен навчати мережу. Нажаль, у реальному житті при роботі з пацієнтами дуже важко сходу сказати, до якої групи пацієнт належить, до того ж у якості невідомої виступає кінцева кількість пацієнтів. Тому, для

вирішення цієї задачі доцільно розглядати алгоритми навчання без учителя, яскравим представником яких є метод нечіткої кластеризації *C*-середніх.

Кластеризація має довгу і багату історію в різних наукових галузях. Один з найпопулярніших і простих алгоритмів кластеризації, *C*-середніх, вперше був опублікований у 1955 році. Незважаючи на те, що *C*-середніх був запропонований понад 50 років тому, і з тих пір були опубліковані тисячі алгоритмів кластеризації, *C*-середніх – це широко використовуваний метод. Це говорить про труднощі в розробці алгоритму кластеризації загального призначення та неправомірну проблему кластеризації.

Метод нечіткої кластеризації *C*-середніх (Fuzzy clustering) дозволяє розбити наявну множину елементів потужністю  $N$  на задане число нечітких множин  $k$ . Метод нечіткої кластеризації *C*-середніх можна розглядати як вдосконалений метод  $k$ -середніх, при якому для кожного елемента з розглянутої множини розраховується ступінь його приналежності кожному з кластерів. Мета методу — розділити  $n$  спостережень на  $k$  кластерів, так щоб кожне спостереження належало до кластера з найближчим до нього середнім значенням. Метод базується на мінімізації суми квадратів відстаней між кожним спостереженням та центром його кластера.

Насправді і досі ведуться дослідження можливостей використання цього алгоритму. Регулярно виходять публікації з новими доповненнями та модифікаціями до нього. У даній роботі було вирішено удосконалити цей метод кластеризації за допомогою використання метрики Махаланобіса для визначення відстані між об'єктом та кластерами.

Відстань Махаланобіса — відстань у евклідовому просторі, що узагальнює поняття евклідової відстані. Вона дозволяє взяти до уваги значущість певних метрик для кінцевого результату, що веде до підвищення точності діагностики.

### Список літератури

1. Горячая В. А. Solutions synthesis method to the condition prediction problem of patients in the medical monitoring systems / В. А. Горячая, М. Л. Угрюмов, О. А. Гайденко, Н. С. Бакуменко, И. А. Трофимова // Radioelectronic and computer systems. – 2017. – № 3. – С. 68–76.
2. Obermeyer Z., Emanuel E.J. Predicting the future—big data, machine learning, and clinical medicine // The New England journal of medicine, 2016.
3. Oskoueï R.J., Kor N.M., Maleki S.A. Data mining and medical world: breast cancers' diagnosis, treatment, prognosis and challenges // American journal of cancer research, 2017.
4. Deo R.C. Machine learning in medicine // Circulation, 2015.
5. Hamet P., Tremblay J. Artificial intelligence in medicine // Metabolism, 2017.
6. Kheir N. Systems modeling and computer simulation // Marcel Dekker, NewYork, 2018.

## МОДЕЛИ И МЕТОДЫ ИНФОРМАЦИОННОГО ОБЕСПЕЧЕНИЯ ПРОЦЕССА УПРАВЛЕНИЯ

*Передерин А. В.*

*Харьковский национальный университет имени В. Н. Каразина*

Первым этапом управления является его информационное обеспечение. Для уменьшения времени сбора информации о состоянии элементов сети этот процесс желательно выполняться параллельно по всем управляемым объектам. Общее время сбора информации будет определяться интервалом времени получения данных от наиболее удаленного объекта.

Сбор информации о состоянии сети может осуществляться как по инициативе центра коммутации, на котором решается задача управления сетью, так и периодически по

установленному графику и по инициативе всех узлов, изменение состояния которых может повлиять на эффективность функционирования сети.

### Управление параметрами канала передачи данных

Процесс управления потоками информации в КС является достаточно сложным. Нужно учитывать особенности передачи информации. Ее можно передавать по одному каналу или по нескольким. Чтобы определить параметры, которые влияют на эффективность системы управления, подробно рассмотрим модель управления каналом передачи данных.

Для повышения качества обслуживания в применяемых протоколах применяется обратная связь. Поэтому рассмотрим процессы приема и передачи сообщений в системах с обратной связью. Устройство-отправитель передает в КС пакет данных. Время, которое тратится на передачу данных выражается через его длину ( $n$ ), ширину окна ( $W_0$ ) и скорость модуляции ( $B$ )

Принятое сообщение проверяется на ошибки на устройстве-получателе. Производится проверка контрольной суммы. После ее вычисления в пакетах данных может не содержаться ошибок (с вероятностью  $P_{пр}$ ), находятся обнаруженные –  $P_{оо}$  или необнаруженные ошибки  $P_{но}$ .

Время, затраченное на передачу одного пакета в соответствии с формулой (1)

$$T_{cp} = W_0 \cdot T_{п} + T_{ож} + \frac{(T_{та})(P_{пот} + P_{оо})}{1 - (P_{пот} + P_{оо})} \quad (1)$$

Время доставки одного пакета возрастает при увеличении времени ожидания ( $T_{ож}$ ). Минимальное время доставки будет при продолжительности тайм-аута, равной  $T_{ож}$ . Из вышесказанного можно сказать, что управляющими параметрами протокола канала передачи данных, которые влияют на эффективность управления, является ширина окна и продолжительность тайм-аута.

Самым простым способом управления скоростью передачи с помощью перехода на другой вид модуляции. Например, при переходе с ВФМ на АВФМ или ВФМ на ДВФМ скорость передачи увеличивается вдвое. Но при этом увеличивается вероятность ошибки.

Согласно (1) увеличение скорости в два раза уменьшает в два раза  $T_{п}$ , но при этом увеличивается  $P_{оо}$ . Зависимости построены при  $P_{пот} = 0$ ,  $T_{ож} = 0$  и  $W_0 = 1$ . По полученным графикам можно сказать, что управляя скоростью передачи посредством изменения вида модуляции, можно получить уменьшение времени доставки только в случае определенного соотношения сигнал / шум ( $q$ ). Величина данного отношения зависит от продолжительности переданного пакета и вида модуляции. Несмотря на полученные графики можно сказать, что повышение скорости передачи можно обеспечить, если переходить от ВФМ к ДВФМ, при условии, что отношение сигнал / шум удовлетворяет неравенству  $q_1 \geq 3.1$

### Управление параметрами буферных ресурсов

Одни из главных показателей качества обслуживания компьютерной сети являются время доставки сообщения и вероятность его потери. Эти два показателя зависят от:

- выделенной емкости БЗУ (буферного запоминающего устройства);
- количества узлов коммутации, которые входят в маршрут доставки сообщения.

Чтобы эффективно распределить ресурсы компьютерной сети, нужно правильно оценить влияние каждого указанного фактора на вероятность потери сообщения.

Благодаря управлению буферными ресурсами, можно повлиять на одни из важных характеристик сети, такие как вероятность потери сообщения и время задержки на узле. Следовательно, можно повлиять на скорость доставки пакета.



Нужно определить требуемую емкость БЗУ на этих узлах и время их задержки. Известно, что вероятность потери пакета на узле из-за переполнения БЗУ находится по формуле:

$$P_{\text{пот}} = \frac{1-\rho}{1-\rho^{w+1}} \cdot \rho^w \quad (2)$$

Коэффициент загрузки узла  $\rho$ , емкость БЗУ  $W$ . Изменяя данные параметры можно повлиять на вероятность потери пакета. Следовательно, можно повлиять на эффективность управления. Если маршрут состоит из  $n$  промежуточных пунктов, то вероятность потери возрастет и найти ее можно по формуле

$$P_{\text{пот,тр}} = 1 - (1 - P_{\text{пот1}})^n \quad (3)$$

где  $P_{\text{пот1}}$  – вероятность потери на одном промежуточном пункте.

С помощью полученных выражений можно увидеть, как влияют параметры на вероятность потери. Это позволяет влиять на эффективность управления буферными ресурсами. Например, задавая значения  $n$  и  $P_{\text{пот,тр}}$  можно предъявить требование к вероятности  $P_{\text{пот1}}$ , а по формуле (2), при известной величине  $\rho$  найти необходимую емкость БЗУ.

На рис. 1 изображены зависимости вероятности потерь от емкости БЗУ при разных коэффициентах загрузки. Из графиков видно, что при уменьшении емкости БЗУ начиная с 6 можно видеть увеличение вероятности потери сообщений в широком диапазоне изменения коэффициента загрузки  $\rho$ . Увеличение емкости больше 6 практически никак не влияет на вероятности потери.

На рис. 2 изображены зависимости вероятности потерь в тракте ( $P_{\text{пот,тр}}$ ) от числа промежуточных пунктов ( $n$ ), при разной емкости БЗУ и коэффициентах загрузки. Если задана вероятность потери сообщения в тракте и коэффициент загрузки, то с помощью данной графической зависимости можно предъявить требования к емкости БЗУ и допустимому количеству промежуточных пунктов.

На рис. 2 изображено применения графика для нахождения емкости БЗУ зная вероятность потери на одном узле. Мы проводим две перпендикулярные линии как показано стрелками и определяем, какая вероятность БЗУ необходима.

Таким же способом определим количество промежуточных узлов, зная только вероятность потери в тракте. С помощью двух перпендикулярных стрелок как показано на рис. 1 можно точно определить, сколько узлов используется.

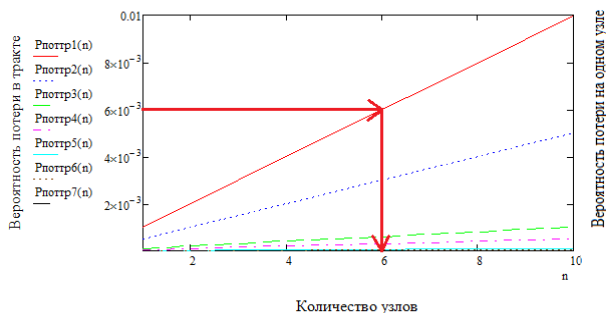


Рисунок 1 – Зависимость  $P_{\text{пот,тр}} = f(n)$

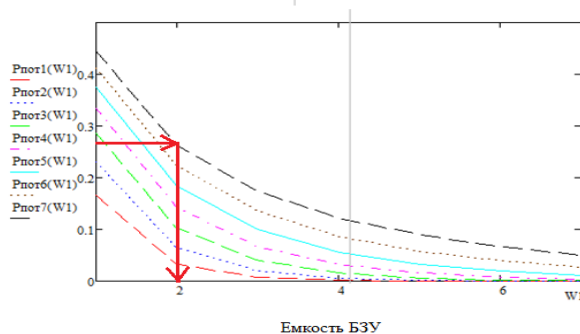


Рисунок 2 – Зависимость  $P_{\text{пот}} = f(W)$

### Управление информационными ресурсами

Предположим, что сообщение, которое передается, содержит  $M$  информационных разрядов. Обозначим длину пакета  $m_n$ . К каждому пакету прилагается  $k_{\text{сл}}$  служебных разрядов.

Из сообщения, которое передается, формируется  $\frac{M}{m_n}$  пакетов. Если время распространения сигнала по сравнению с продолжительностью пакета мало, то относительное время задержки одного пакета (по продолжительности одной посылки) определяется по формуле:

$$T_{з1} = (m_n + k_{сл}). \quad (4)$$

Время передачи всех  $\frac{M}{m_n}$  пакетов равно:

$$T_3 = \frac{M}{m_n} T_{з1} \quad (5)$$

Таким образом, относительное время сообщение будет равно:

$$T_d = (m_n + k_{26сл}) + (M + \frac{Mk_{сл}}{m_n}) \frac{1}{(1-P_{00}-P_{пот})P_{КВ}}. \quad (6)$$

Благодаря формуле (6) найдем оптимальную длину пакета. Начальные данные будут следующие:  $m_n$  будет меняться от 5 до 40;  $k_{26сл} = 10$ ;  $M$  будет равняться 100 и 300;  $P_{00} = 0.01$ ;  $P_{пот} = 0.1$ ;  $P_{КВ} = 0.1$ . используя эти параметры и формулу (6) построим графики для нахождения оптимальной длины пакета (рис.3).

Анализируя полученный график на рис.3 можно сказать, что наиболее оптимальная длина пакета для передачи сообщения при заданных значениях равна 20. Увеличение длины пакета не дает достаточного уменьшения времени передачи сообщения.

В работе были предложены разные методы процесса управления. Так же было практически доказано, что рассмотренные методы увеличивают эффективность управления, что и являлось целью данного исследования.

### Список литературы

1. Гребешков А. Ю. Стандарты и технологии управления сетями связи. – М.: Эко-Трендз, 2003. – 288 с.
2. Хамед С. М. Способы и средства оптимизации нагрузки в корпоративных компьютерных сетях: дис. канд. техн. наук. – К.: НТУ України "КПІ", 1996. – 152 с.
3. Стеклов В. К., Кільчицький Є. В. Основи управління мережами та послугами телекомунікацій. – К.: Техніка, 2002. – 438 с.
4. Олоту О. Д. Методы и модели управления ресурсами цифровых информационных сетей с интеграцией телеинформационных служб: дис. кандидата тех. наук. – К.: 2016. – 175 с.

## АНАЛІЗ ІНФОРМАТИВНОСТІ ЗМІННИХ СТАНУ ПАЦІЄНТА В СИСТЕМІ МЕДИЧНОГО МОНІТОРИНГУ

*Петухов М. В., Шейко О.Г.*

*Харківський національний університет імені В. Н. Каразіна*

Машинне навчання швидко проникає в багато галузей серед яких і охорона здоров'я – від діагностування та прогнозування до розробки ліків та епідеміології. Тому машинне

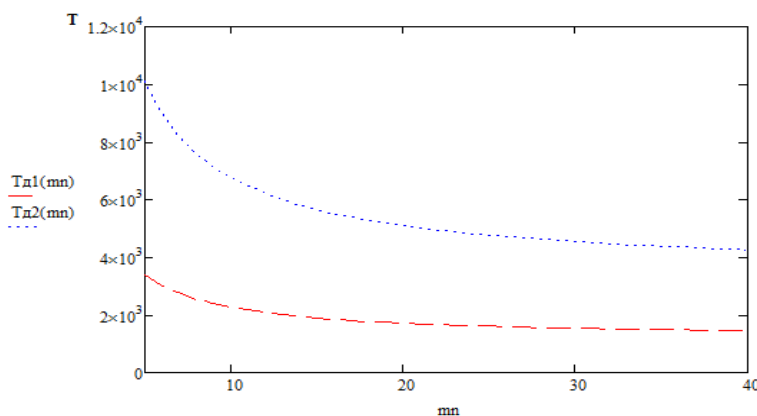


Рисунок 3 – Залежність часу передачі повідомлення від довжини пакета

навчання має значний потенціал для вирішення проблем пов'язаних з охороною здоров'я.

На сьогоднішній день галузь медицини значною мірою спирається на евристичні підходи, завдяки чому знання набуваються завдяки досвіду та самонавчанню, що є обов'язковим у дуже мінливому середовищі охорони здоров'я. Збільшення знань та розуміння захворювань пов'язане із зростанням інформації та даних, частково завдяки прогресу інструментів, які генерують кількісні та якісні вимірювання фізіологічних показників. Таке велике поле даних дозріло для застосування машинного навчання (ML). Дійсно, зростає усвідомлення потенціалу ML як платформи, яка може отримати інформацію з численних джерел в інтегровану систему, яка може суттєво допомогти процесам прийняття рішень висококваліфікованим працівникам. Коли ML все ще знаходилася в стані народження, було припущено, що успіх інтелектуальної системи, яка може вчитися та вдосконалюватись, залежатиме від здатності збирати та зберігати велику кількість даних у базі знань. Поліпшення, що були внесені в обчислювальні ресурси, а також зберігання та обмін даними за останнє десятиліття, стали вагомим фактором у використанні потенціалу систем машинного навчання в медицині.

Зараз створюється багато інформаційних систем для зберігання різних медичних даних про здоров'я пацієнтів, які використовуються лікарями для встановлення (діагностики) стану здоров'я. Розмірність даних, визначають кількістю різних показників, що описують стан здоров'я пацієнта, ця величина досить велика і деколи досягає декількох десятків або сотень. Проблема зменшення розмірності простору показників і визначення найбільш інформативних показників є актуальною для медичних інформаційних систем.

Інформативність (значимість) змінних стану – поняття відносне. Деяка підмножина змінних може бути інформативною для вирішення однієї задачі аналізу станів або класифікації даних і не інформативною для іншої. Вибір критеріїв оцінювання інформативності (значущості) змінних стану залежить від того, що від чого потрібно відрізнити, тобто від типу і розмірності безлічі змінних стану  $S$  розглянутих елементів системи або процесу, а також від типу вирішальних функцій  $P$ . Для кожного завдання слід знаходити свою інформативну підмножину змінних стану.

Проблемами в процесі прийняття рішень в системах медичного моніторингу на етапі діагностування, є:

- висока ймовірність помилки при розпізнаванні стану пацієнтів на основі даних моніторингу, бо в залежності від стану пацієнта призначається індивідуальна програма лікування;
- вибір контрольованих змінних стану індивідуально для кожного пацієнта з метою подальшого аналізу ефективності призначеної програми лікування.

Регресійний аналіз використовується у випадку, якщо відношення між змінними можуть бути виражені кількісно у вигляді деякої комбінації цих змінних. Отримана комбінація використовується для передбачення значення, що може приймати залежна змінна, яка обчислюється на заданому наборі значень незалежних змінних. У найпростішому випадку для цього використовується лінійна регресія.

Необхідно зазначити, що серед регресійних моделей виділяють:

- однопараметричні моделі (залежність від однієї змінної);
- багатопараметричні моделі (залежність від декількох змінних);

З точки зору практики оптимального проектування, або діагностування систем і процесів логічним є твердження про доцільність витрат на вимір інформативних змінних, для яких значимо ставлення сигнал / шум. У той же час, згідно з загальним тенденціям розвитку науково-технічного прогресу, такі вимірювання стають все більш дорогими.

Таким чином, необхідно виконати редукцію простору контрольованих змінних стану, тобто, виконати пошук підмножини інформативних (значущих) змінних стану  $S_\beta$  мінімальної розмірності  $S_\beta \subset S_0$ .

Уявімо  $Y_i(S), S = \{s_l\}, l = 1 \dots L$  у вигляді ряду Тейлора, зберігши в розкладанні тільки доданки першого порядку малості. Отримана функція є лінійною. Для дисперсії довільної лінійної функції декількох випадкових змінних має місце оцінка.

Пол	4.25351166683838E-05
АД сист	0.0107246560790197
АД диаст	0.0849701707839148
Ин Кетле	0.0327672366239706
ЧССHR удмин	0.0643442846337191
ДисперсияD мс2	0.0150946920004092
CKOSDNN мс	0.0371935644656698
KBCV %	0.0360711232616579
RMSSD мс	0.0729712069170252
PNN50 %	0.0904733071714854
Амо %	0.119041508019545
Мо мс	0.0413798483147553
МхDMn мс	0.110819884327535
ИНСI у е	0.131603330817463
TP мс2	0.013642558503735
HF мс2	0.0771833221137461
LF мс2	0.0278987327187455
VLF %	0.0337780381309364

Рисунок 1 – Приклад розрахунку інформативності показників

Харківського національного університету. Збірник наукових праць. Серія: «Математичне моделювання. Інформаційні технології. Автоматизовані системи управління». – 2015. – Випуск 26 (№1156). – С. 5-16.

2. Горячая В.А. Информационная технология диагностирования состояния пациентов в системах медицинского мониторинга / В.А. Горячая, М.Л. Угрюмов, С.В. Черныш // Бионика интеллекта. – 2017. – №2 (89). – С. 84– 91.

3. Terry, N. P. Regulatory disruption and arbitrage in health-care data protection. *Yale J. Health Pol'y L. & Ethics* 17, 143–208 (2017).

4. Widmann, A., Schröger, E., and Maess, B. (2015). Digital filter design for electrophysiological data—a practical approach. *J. Neurosci. Methods* 250, 34–46.

Визначимо енергію сигналу:

$$E_i = \sum_{l=1}^L |D_{Y_i|s_l}|,$$

$$\text{де } D_{Y_i|s_l} = \left( \sum_{n=1}^L r_{ln} \frac{\partial Y_i}{\partial s_n} \sigma_{s_n} \right) \frac{\partial Y_i}{\partial s_1} \sigma_{s_1}.$$

Визначимо коефіцієнт інформативності (значимості вкладу)  $s_l$  в  $Y_i$ :

$$\beta_{il} = \frac{D_{Y_i|s_l}}{E_i}, \sum_{l=1}^L \beta_{il} = 1$$

Приклад розрахунку інформативності контрольованих змінних стану показаний на рис. 1.

В даний час рішення про вибір програми лікування для кожного пацієнта приймаються суб'єктивно, виходячи з особистого досвіду лікаря.

Застосування на практиці автоматизованих комп'ютерних систем контролю і діагностування стану може допомогти лікарям і/або пацієнтам приймати рішення, які забезпечать краще лікування та більш точне діагностування.

### Список літератури

1. Антонян И. М. Метод оценивания информативности переменных нейросетевых моделей систем и процессов при неопределенности данных / И. М. Антонян, В. А. Горячая, А. И. Зеленский, Е. М. Угрюмова // Вісник

## СИСТЕМА РОЗПІЗНАВАННЯ ФРАГМЕНТІВ ОБЛИЧЧЯ ЗА ДОПОМОГОЮ МЕТОДІВ ШТУЧНОГО ІНТЕЛЕКТУ

Семенченко О.В.

Харківський національний університет імені В. Н. Каразіна

На сьогоднішній день більшість об'єктів, розважальних комплексів або підприємств потребують застосування систем ідентифікації особистості. Ідентифікація людини за зображенням обличчя є одним з найбільш пріоритетних і поширених напрямів розвитку

біометричних систем.

Об'єкт дослідження – інтелектуальні системи відеоспостереження.

Предмет дослідження – методи та засоби комп'ютерної ідентифікації фрагментів обличчя людини.

Метою роботи є дослідження існуючих методів обробки, фільтрації та сегментації зображень, а також методів виявлення та ідентифікації обличчя.

Практична цінність – програмне забезпечення, що реалізує розроблений в даній роботі метод виділення об'єктів інтересу і модифікований метод виявлення осіб, дозволяє:

- підвищити ефективність і якість виявлення осіб;
- скоротити час виявлення осіб на зображенні;
- скоротити кількість помилкових виявлень обличчя.

Ідентифікація особистості по обличчю може бути проведена різними способами, наприклад, фіксуючи зображення в зоні видимості, використовуючи звичайну відеокамеру, або за допомогою використання теплового малюнка особи. У загальному випадку, завдання автоматичної ідентифікації людини по фрагментами його обличчя складається з двох етапів. Перший - це визначення місця розташування обличчя на зображенні. Для цього вихідне зображення сканують вікном меншого розміру, і кожен раз визначають деяку ступінь схожості зображення у вікні з людським обличчям. Цей етап найбільш обчислювально трудомісткий, так як потрібно проводити повне сканування для різних розмірів вікна, а так само кожен раз знаходити ступінь схожості зображення у вікні з обличчям.

Процес ідентифікації є встановлення тотожності невідомого об'єкта відомому на підставі збігу ряду ознак.

*Недоліки існуючих рішень ідентифікації:*

- великі витрати часу на обробку та отримання кінцевих результатів;
- трудомісткість і складність застосованих методів;
- неточність ідентифікації.

Для розв'язання задачі комп'ютерної ідентифікації людини за фрагментами його обличчя, що використовуються в ІСВ для обліку кількості людей або відвідувачів, в цілому і на окремих її етапах застосовуються різні методи нормалізації, фільтрації і сегментації зображень. Перелік основних етапів обробки зображення для рішення задачі комп'ютерної ідентифікації людини за фрагментами його особи наведено на слайді 5.

На наступних слайдах наведені методи, що застосовуються при сегментації, фільтрації, виявлення та ідентифікації обличчя на зображеннях.

У якості методу сегментації зображення обрано – метод низькочастотного фільтру рекурсивного згладжування. А у якості виявлення та ідентифікації обличчя на зображеннях – модифікований метод Віоли-Джонса.

Метод складається з трьох ключових особливостей:

- представлення зображення у вигляді «інтегрального зображення»;
- використання класифікатора засноване на алгоритмі AdaBoost;
- каскад класифікаторів дозволяє швидко відкидати НЕ-обличчя.

Таким чином, розроблений метод виділення об'єктів інтересу складається з наступних етапів:

- отримання початкового зображення;
- сегментація та фільтрація;
- отримання областей об'єкта;
- об'єднання областей і отримання об'єкта інтересу.

Була проведена модифікація методу Віоли-Джонса. Для підвищення ефективності і якості виявлення похилих рис обличчя пропонується використання нового набору ознак.

Додатковий набір ознак розширює базовий набір прямокутних ознак, підвищуючи при цьому ефективність виявлення об'єктів, розташованих під кутом.

Результатом роботи став програмний комплекс «Hawk» розроблений на сучасному об'єктно-орієнтованій мові Java. Для опису поетапної роботи приведено блок-схему роботи програмного комплексу

Розроблений програмний комплекс «Hawk» працює з кольоровим відеозображенням.

У режимі виявлення ОІ (об'єкти інтересу) проводиться обробка відеозображення в пошуках ОІ. Даний режим містить в собі виконання всіх етапів програмного комплексу, крім процедур виявлення та ідентифікації обличчя. Результатом роботи програми в режимі виявлення ОІ є виділення об'єктів інтересу в області обмежені чорними рамками, що позначають ОІ.

У режимі виявлення та ідентифікації обличчя відбувається повне сканування зображення в пошуках обличчя для ідентифікації в обхід виявлення ОІ. Режим виконує тільки останні етапи програмного комплексу і в разі успішного виявлення та ідентифікації записує статистичну інформацію.

Таким чином, результатами дослідницької діяльності є:

1) метод виділення об'єктів: координати об'єкта інтересу на зображенні; кількість пікселів, з яких складається об'єкта інтересу на зображенні

2) модифікація методу виявлення обличчя Віюлі-Джонса: підвищення ефективності і якості виявлення похилих рис обличчя.

## **ОЦЕНКА КАЧЕСТВА ВОЛОКОННО-ОПТИЧЕСКОЙ СПД С ОПТИЧЕСКИМИ УСИЛИТЕЛЯМИ**

*Семенюг Е.А.*

*Харьковский национальный университет им. В. Н. Каразина*

На данный момент информация представляет самый ценный ресурс для любой сферы жизни в самом широком понимании этого слова. Именно потому главным является то, с какой максимальной скоростью будет передаваться эта информация, а также какого качества она будет, и насколько полной ее будет видеть получатель.

Использование оптических усилителей позволят передавать информацию на более дальнее расстояние, без использования ретрансляторов. Это же в свою очередь влияет и на стоимость всей линии передачи данных. Именно поэтому данные устройства активно используются в большом количестве современных сетей – городские и региональные, междугородные и трансконтинентальные, и пр.

При изучении данной темы была выявлена потребность в разработке программного продукта, который позволил бы оценить качество малых и средних СПД в процессе проектирования. Это в свою очередь позволяет снизить риски финансовых потерь во время реализации плана в жизнь.

Особенностью программы можно считать довольно широкий диапазон варьирования параметров системы и условий ее работы, что в свою очередь делает возможным скорректировать необходимые показатели для наиболее качественной передачи информации в канале. Прикладной значимостью разрабатываемого программного продукта является использование ее в ходе изучения дисциплины «Оптоинформатика».

### **Постановка задачи**

Разработать программный продукт, помогающий оценить качество системы передачи данных с использованием оптических усилителей.

### Анализ факторов, влияющих на процесс передачи данных

Основным преимуществом именно оптических усилителей можно считать то, что в отличие от регенераторов, усиление сигнала происходит путем увеличения амплитуды входных импульсов при использовании оптических технологий.

Среди довольно большого числа преимуществ присутствуют и недостаток – возникновение шума в канале. При этом стоит понимать, что оптические усилители также увеличивают и его амплитуду. Происходит накопление шума в канале и в какой-то момент использование усилителя будет бесполезным.

Набор основных параметров оптических усилителей.

Для того, чтобы исследовать поведение лазерного луча в оптоволокне и определить эффективность использования усилителей, стоит обратить внимание на несколько основных факторов, каждый из которых влияет на определение общей картины передачи данных.

**Коэффициент усиления** – параметр, который зависит от длины волны входного сигнала. На нее же влияет концентрация, распределение и форма энергетических уровней ионов в усилителе, а также длина волны лазера накачки. Данное значение можно определить по формуле:

$$G = \frac{P_{S\ out}}{P_{S\ in}}, \quad (1)$$

где  $P_{S\ out}$  и  $P_{S\ in}$  – показатели мощности сигналов на входе и выходе усилителя.

**Мощность насыщения в усилителе** – показатель мощности сигнала на выходе из усилителя. Но при этом стоит учитывать, что значение  $G$  должно быть ниже, как минимум в 2 раза,  $G_{\max}$ .

На качество сигнала, который попадает в усилитель также значительное воздействие оказывает такой параметр, как отношение сигнал/шум (OSNR). Современные стандарты передачи данных требуют достаточно больших значений этого параметра, так как именно он является основным для определения коэффициента появления ошибок BER.

**Шум-фактор** – показатель возрастания шума в усилителе, в сравнении с полезным сигналом. Данное значение рассчитывается по следующей формуле:

$$NF = \frac{OSNR_{in}}{OSNR_{out}}, \quad (2)$$

где  $OSNR_{in}$  и  $OSNR_{out}$  – значение сигнал-шум на входе и выходе усилителя.

Еще одним из базовых показателей передачи информации можно считать **коэффициент битовых ошибок BER**:

$$BER = \frac{1}{2} \operatorname{erfc} \left( \frac{Q}{\sqrt{2}} \right) \approx \frac{e^{-\frac{Q^2}{2}}}{Q\sqrt{2\pi}} \quad (4)$$

Необходимо обратить внимание на тот фактор, что значение данного коэффициента не должно превышать допустимый уровень, иначе работа всей системы будет недопустима. Более того, для определения данной величины используется Q-фактор. Эта величина является характеристической величиной качества сигнала, поступающего на вход усилителя.

$$Q = \frac{|\mu_1 - \mu_0|}{\sigma_1 + \sigma_0}, \quad (5)$$

где  $\mu_1$  и  $\mu_0$  – уровень сигнала и шума,  $\sigma_1$  и  $\sigma_0$  – среднеквадратичное отклонение данных сигналов и шумов.

### Особенности моделирования СПД

Для создания имитационной модели было рассмотрено несколько вариантов методологий расчетов. Самым оптимальным и грамотным является метод разделения ветки сигнала и шума. Благодаря этому можно отдельно смоделировать усиление мощности сигнала и накопление мощности шума.

Для ветки сигнала характерны следующие параметры -  $I_0$ - значение мощности в начале канала,  $R$  – длина участка канала,  $\alpha$  – показатель затухания .

1. Значение мощности сигнала на входе первого усилителя  $I_1 = I_0 * e^{-\alpha R}$ .
2. Значение мощности сигнала на выходе каждого сегмента усилителя  $I_1^{n'} = I_1 * e^{x_0^{n-1}d}$ .
3. Ненасыщенный показатель усиления  $x'_0 = \frac{x_0}{1 + \frac{I_1}{I_{\text{нас}}}}$ , где  $I_{\text{нас}}$  – мощность насыщения.

Для ветки шума используются следующие параметры –  $I_{\text{ш}}$ - мощность шума в канале,  $d$  – размер одного сегмента усилителя,  $M$  – количество сегментов.

1. Значение мощности шума на входе 2 усилителя  $I_{\text{ш}3} = I_{\text{ш}2} * e^{-\alpha R}$ .
2. Значение мощности шума на выходе 2 усилитель --  $I_{\text{ш}4} = I_{\text{ш}3} * e^{x_0 d M} + I_{\text{ш}2}$

### Оценка качества СПД

Для того, чтобы оценить качество передачи данных в моделируемой системе, был разработан программный продукт. Благодаря широкому диапазону варьирования параметров, пользователи могут задать характеристики канала связи и отследить, насколько будет эффективно использование того или иного количества усилителей.

Таблица 1 – Параметры для расчета оценки качества СПД

Кол-во усилителей	1	2	3	4	5	6
OSNR	28	25	23	22	21	20
Реальное OSNR	631	316	200	158	126	100
Q-фактор	18	17.8	17.5	17.2	17	16
BER	$10^{-15}$	$3.5 * 10^{-15}$	$2.2 * 10^{-14}$	$10^{-14}$	$7.4 * 10^{-13}$	$1.43 * 10^{-10}$

Опираясь на материал, изложенный выше, в табл. 1 представлены результаты моделирования, полученные из разработанного программного продукта.

Из табл. 1 следует, что при возрастании числа усилителей падает отношение сигнал шум и возрастает вероятность появления ошибки при передаче информации. Поэтому разработчики обычно используют 3-4 усилителя на участке СПД, затем требуется использование ретрансляторов.

Физическими причинами ухудшения отношения сигнал шум являются:

- накопление шумов от каждого усилителя;
- тот факт, что, шум усиливается сильнее, чем сигнал, так как имеет меньшую мощность, чем сигнал по сравнению с мощностью насыщения.

При заданном количестве усилителей и определенных значения отношений сигнал-шум, можно получить значения Q-фактор, которые приведены в таблице, в соответствующей строке. Также учитывая имеющиеся значения  $\mu_0 = 631$  и значение показателя шума в диапазоне от 1 до 6.3, рассчитывается значение показателя коэффициента появления ошибок BER.

### Выводы

Для того, чтобы получить максимально качественный канал передачи данных, необходимо учитывать несколько важных факторов. Среди них – протяженность ВОЛС, ее характеристические особенности, уровень шума в канале и количество усилителей, расположенных на линии.

При проектировании сети большее внимание стоит уделять и такому показателю, как BER. Именно исходя из его значений можно понять, насколько эффективна данная система и до какого момента в ней можно использовать оптические усилители.

### Список литературы

1. Убайдуллаев Р.Р. Протяженные ВОЛС на основе EDFA / Р.Р. Убайдуллаев // Lightwave, 2003. – №1. – С. 22-29.
2. Андрійв В.А. Будівництво та технічна експлуатація волоконно-оптичних ліній зв'язку: підручник для ВУЗів / В.А. Андрійв, В.А. Бурдін, Б.В. Попов. – М.: Радіо та зв'язок, 1995. – 200 с.



3. Андрушко Л.М. Посібник по ВОЛЗ / Л.М. Андрушко, В.А. Вознесенський. – К. : Техника, 1988. – 239 с.
4. Бейлі Д. Волоконна оптика: теорія і практика / Д. Бейлі, Е. Райт. – М. : Кудіц-прес, 2008. – 320 с.
5. Волоконно оптичні датчики. Вступний курс для інженерів і наукових працівників / під. ред. Е. Удда. – М. Техносфера, 2008. – 520 с.
6. Склярів О.К. Сучасні волоконо-оптичні системи передачі, апаратура та елементи / О.К. Склярів. – М. : Солон-Р, 2001. – 240 с.
7. Фріман Р. Волоконно-оптичні системи зв'язи / Р. Фріман. – М. : Видавництво Техносфера, 2003.

## **МЕТОД УПРАВЛІННЯ РИЗИКОМ ДЛЯ ПІДВИЩЕННЯ НАДІЙНОСТІ СКЛАДНИХ СИСТЕМ**

*Харченко В.О.*

*Харківський національний університет імені В.Н. Каразіна*

Управління ризиком – невід'ємна частина ефективного управління будь-якою системою. Особа, яка планує, управляє або здійснює постійний контроль за всіма розділами проекту, повинен чітко і ефективно управляти ризиками системи. В даний час, як правило, на етапі проектування при розрахунках показників працездатності враховують тільки безумовні ймовірності відмов елементів системи. Такий підхід, дозволяє оцінити тільки ймовірність відмови системи від взаємодії між її елементами, не є цілісним аналізом. Запропонований метод враховує як взаємодії між елементами системи, так і вплив зовнішнього середовища на ймовірність відмови елементів (підсистем) складної системи.

Вимоги до безпеки виробничих процесів є одним з найважливіших факторів, що стимулюють розвиток науково-методичного апарату аналізу та управління ризиком при експлуатації потенційно небезпечних об'єктів (ПНО). При цьому, коли витрати на забезпечення безпеки ПНО становлять значну частку матеріальних ресурсів суспільства, важливого значення набуває проблема оптимізації показників безпеки складних технічних систем (СТС) з урахуванням витрат і прогнозованих значень показників прийняттого ризику аварій і катастроф.

Прийнятний ризик поєднує в собі технічні, економічні, соціальні та політичні аспекти і являє собою певний компроміс між рівнем безпеки та можливостями її досягнення. Тому розробка методики управління дозволяє з використанням своєчасних обчислювальних засобів і методів оптимізації ефективно управляти організаційними системами, виробництвом і потенційно небезпечними об'єктами.

Під складною технічною системою розуміється безліч взаємопов'язаних елементів, що взаємодіють між собою і утворюють певну цілісність, яка має певні властивості, властивими тільки цій системі і відсутніми у кожного елемента окремо (властивість емерджентності).

Складні системи характеризуються, перш за все, великим числом складових елементів, безліччю різноманітних зв'язків, різноманітністю структурних елементів і різноманітністю їх фізичної природи. Крім того, складні системи мають властивість оптимальності. Прикладами СТС можуть служити енергетичні комплекси, гідротехнічні вузли, автоматичні і автоматизовані системи, що управляють технологічними процесами в хімічній, металургійній промисловості та машинобудуванні, атомні електростанції, ракетно-космічні комплекси, системи військового призначення і так далі.

У загальному випадку обґрунтування вимог до безпеки і надійності СТС пов'язано з порівнянням виробничих, експлуатаційних витрат та існуючим ризиком аварій і катастроф.

Тому в якості критерію оптимізації можна розглядати мінімум сумарних економічних витрат (витрат), пов'язаних як з попередженням можливих подій (аварій і катастроф), так і з ліквідацією наслідків (збитків) від їх виникнення.

Безпека, як і інші властивості технічних систем, забезпечується властивостями окремих компонентів (елементів), що вимагає проведення великої кількості контрольних-профілактичних і ремонтно-відновлювальних заходів, виконуваних на всіх елементах і на всіх етапах життєвого циклу систем. Заходи забезпечення необхідного рівня безпеки характеризуються показниками вартості, ефективності, ресурсоемності тощо.

Оцінка надійності складних систем виконується за допомогою методів, пов'язаних з використанням даних про відмови та відновлення елементів, отриманих в процесі експлуатації систем.

Для розрахунку показників надійності технічних систем застосовуються аналітичні методи. До них відносяться методи теорії випадкових процесів, теорії експертних оцінок (евристичного прогнозування), декомпозиції (еквівалентування), логіко-імовірнісні, асимптотичні, аналітико-статистичні методи. На практиці використовують методи імітаційного і статистичного моделювання (метод Монте-Карло).

– Метод побудови марківських моделей. На основі інформації про структуру та принципи функціонування досліджуваної системи визначається безліч її можливих станів, яке розділяється на два підмножини: працездатних станів і станів відмови. Будується граф переходів, вершинами якого є стани системи, а ребрами – можливі переходи між станами. Інтенсивності переходів визначаються характеристиками безвідмовності і ремонтпридатності елементів системи. З графу переходів складається система рівнянь, рішення якої дозволяє отримати необхідні показники надійності.

– Метод теорії випадкових процесів. Моделі, засновані на математичному апараті випадкових подій і марківських, напівмарковських випадкових процесів, дозволяють врахувати вплив на надійність електропостачання структурного і тимчасового резерву, обмежень і ступеня незалежності джерел живлення, узгодженості роботи захистів і автоматики, рівня контролю за станом основних елементів систем електропостачання.

– Логіко-імовірнісний метод розрахунку надійності з використанням дерева відмов застосовується, коли число різних відмов системи відносно невелике (наприклад, для аналізу надійності автоматизованої системи диспетчерського управління електропостачанням). Вони використовують математичний апарат бінарної алгебри логіки і теорію ймовірності. Методи теорії масового обслуговування, до яких відносяться диференційний метод розкладання на фази, метод Кендалла, дозволяють зводити немарківську модель до марківської.

– Метод евристичного прогнозування оцінки надійності відновлюваних систем полягає в об'єднанні груп елементів цієї системи в один еквівалентний елемент, який характеризується альтернуючим процесом відновлення. Тим самим відбувається зменшення числа елементів в системі. Метод не дозволяє встановити похибку обчислень і застосовується виключно для випадку високонадійних елементів і систем.

– Метод декомпозиції (еквівалентування) складних технічних систем заснований на побудові математичних моделей, що дозволяють отримувати досить точні верхню і нижню межі оцінюваного показника надійності. Метод еквівалентування набув широкого поширення для розрахунку надійності систем з великим числом елементів при паралельному і послідовному їх з'єднанні.

– Метод Монте-Карло застосовується для дослідження поведінки імовірнісних систем в умовах, коли невідомі в повній мірі внутрішні взаємодії в цих системах. Цей метод полягає у відтворенні досліджуваного фізичного процесу за допомогою імовірнісної математичної моделі і обчисленні характеристик цього процесу.

– Методи імітаційного моделювання в цілому є універсальними і допускають розгляд систем з великою кількістю елементів. Однак їх використання в якості методу дослідження задач надійності доцільно лише тоді, коли важко або неможливо отримати аналітичне рішення. Основними етапами такого дослідження є: побудова формальної моделі, розробка програм імітації траєкторій моделі, проведення імітаційних експериментів.

В основі оцінки надійності лежить розрахунок показників, які враховують інтенсивність відмов елементів на основі безумовних ймовірностей. Запропонований метод може включати в себе не тільки безумовні, а й умовні ймовірності. Під умовними ймовірностями, будемо розуміти взаємозв'язок елементів з зовнішнім середовищем, та вплив одного елемента на інший. Будь-яку складну систему, будь-якої природи можливо представити, як технічну систему.

Показник ефективності залежить від структури системи, значень її параметрів, характеру взаємодії із зовнішнім середовищем, тобто показник ефективності визначається процесами функціонування системи. Показник ефективності можна вважати функціоналом від процесу функціонування системи. Найбільш широко використовуються такі показники ефективності: ефективність (вартісна), ймовірність виконання системою поставленого завдання, живучість, швидкодія, пропускна здатність, здатність до перебудови. Дані показники взаємопов'язані і вибір того чи іншого показника ефективності визначається специфікою системи.

Як приклад розглянемо виробництво, в якому окремі підприємства при кооперації передають один одному продукти, що відповідають окремим стадіям процесу, службовців типовим прикладом системи послідовно з'єднаних елементів. Кожному елементу відповідає свій показник надійності, в результаті надійність всього виробничого процесу зменшується з подовженням ланцюга кооперації. Однак її можна підняти введенням відповідних резервних елементів (запасів сировини, резервних потужностей обладнання, запасів готової продукції і т.п.). Резерви можна розглядати як паралельне з'єднання елементів, які вводяться в дію за наявності відмови основного елемента. При цьому очевидно, що накопичення і зберігання резервів зажадає додаткових витрат.

Найбільш часто вживається вартісний критерій ефективності функціонування системи – максимум прибутку в одиницю часу:

$$\max_{n \in N} F(n) = \max_{n \in N} [W_1(n) + W_2(n)], \quad (1)$$

де  $W_1(n)$  – економічний ефект в одиницю часу від функціонування системи;  $W_2(n)$  – приведена до відповідної одиниці часу сума експлуатаційних витрат і капітальних витрат з урахуванням нормативного терміну окупності;  $n$  – вектор, що характеризує варіанти побудови системи і їх характеристики;  $N$  – область визначення вектора  $n$ .

У деяких випадках зручніше інший запис критерію (1) (критерію повних витрат):

$$\min_{n \in N} F(n) = \min_{n \in N} [W_2(n) + W_3(n)], \quad (2)$$

де  $W_3(n)$  – втрати зумовлені недосконалістю системи в порівнянні з деякою ідеальною;

З огляду на те, що зазвичай важко скласти математичну модель системи і діючих на неї збурень введемо характеристику помилки. Величина помилки буде характеризувати відхилення системи від ідеальної. Тоді критерій може бути представлений у вигляді:

$$\min_{n \in N} F(n) = \min_{n \in N} [W_3(p(n)) + W_2(n)], \quad (3)$$

де  $p(n)$  – помилка, яка характеризує складову вектора і веде до втрат.

Як показник ефективності складної системи виберемо деякий показник  $R$ , значення якого залежить не тільки від структури і параметрів системи, але і від значень характеристик надійності її елементів. Нехай  $R_1$  – значення показника ефективності, обчислене в припущенні,

що відмови елементів мають інтенсивність, відповідну заданим характеристикам, а  $R_2$  – в припущенні, що всі елементи абсолютно надійні. Тоді як показник надійності складної системи може бути величина:

$$\Delta R = |R_2 - R_1|, \quad (4)$$

яка показує, наскільки знижується ефективність системи внаслідок можливих відмов її елементів ( $R_1$ ) в порівнянні з ефективністю ідеальної системи, елементи якої абсолютно надійні ( $R_2$ ). Таким чином, завдання зводиться до відшукування мінімуму  $\Delta R$ .

Отже, завдання обґрунтування параметрів безпеки складних систем є оптимізаційною завданням, а її рішення припускає знаходження таких значень параметрів  $P_1, P_2, \dots, P_i, \dots, P_n$ , які забезпечать мінімум функціоналу

$$C_{\Sigma}(P) = \sum_{i=1}^n [C_i(P_i) + R_i(P_i)] \rightarrow \min(P_i) \quad (5)$$

де,  $P_i$  – параметри оптимізації, що представляють собою показники безпеки  $i$ -х систем (компонентів, елементів);  $C_i(P_i)$  – затрати на реалізацію комплексу організаційно-технічних заходів, зв'язаних з забезпеченням обраного показника надійності  $P_i$ ;  $R_i(P_i)$  – затрати, зв'язані з ризиком аварій і катастроф в процесі експлуатації (функціонування)  $i$ -ої системи;  $P$  – вектор з компонентами  $P_1, P_2, \dots, P_i, \dots, P_n$ .

Розв'язання оптимізаційної задачі (1) пов'язано з деякими труднощами через істотну невизначеності вихідної інформації, обумовленої різними режимами і умовами експлуатації СТС; складності структурних схем і безлічі елементів різної фізичної природи; невизначеності процесів старіння і, як наслідок, технічного стану окремих компонентів і системи в цілому. Відомо, що, з одного боку, витрати (вартість) на забезпечення заданого рівня надійності і безпеки СТС нелінійно збільшуються в міру зростання вимог до цих показників.

Відповідно, для кожної  $i$ -ї системи (елемента) в загальному випадку функція вартості  $C(P)$  (для системи в цілому або для її  $i$ -ї компоненти) може бути апроксимована залежністю виду

$$C(P) = C_{\Delta} + C_0 \frac{\ln(1-P)}{\ln(1-P_0)}, \quad (6)$$

де  $C_{\Delta}$  – постійна величина затрат, яка не залежить від надійності та безпеки СТС;  $C_0$  – затрати (вартість), зв'язані з забезпеченням технічних можливостей максимального значення показників безпеки  $P_0$ .

З іншого боку, експлуатація системи з низькими показниками надійності і безпеки пов'язана з високими експлуатаційними витратами, зумовленими підвищеною інтенсивністю відмов, збільшенням числа несправностей, ремонтів і технічних обслуговувань, а також з витратами по усуненню наслідків (збитків) аварій і катастроф.

Інтегральним показником небезпеки систем, що володіють параметрами  $P$ , є показники ризику, які являють собою добуток імовірності  $Q = 1 - P$  деякої несприятливої події на величину збитку  $C_y$  від цієї події:

$$R(Q) = QC_y = (1 - P)C_y. \quad (7)$$

З урахуванням випадкового характеру експлуатаційних процесів сумарна вартість всіх витрат, зв'язаних з досягненням прийнятного рівня ризику експлуатації СТС, дорівнює

$$C_{\Sigma}(P) = C(P) + R(P) = C_{\Delta} + R_{\Delta} + C_0 \frac{\ln(1-P)}{\ln(1-P_0)} + R_0 \frac{\ln P}{\ln P_0} \quad (8)$$

Для визначення  $P^*$  необхідно взяти похідну від виразу (8)

$$P^* = \frac{\frac{R_0}{\ln P_0}}{\left[ \frac{C_0}{\ln(1-P_0)} + \frac{R_0}{R_0 + C_0} \frac{\ln P_0}{\ln(1-P_0)} \right]}, \quad (9)$$

де  $P^*$  – оптимальний рівень безпеки системи за критерієм мінімуму сумарних витрат.

Для ілюстрації можливостей розробленого математичного апарату наведемо приклад

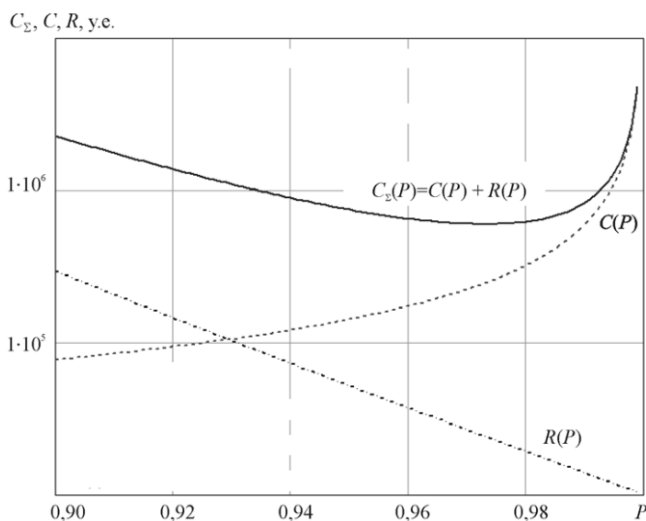


Рисунок 1 – Результати оптимізації показників безпеки складних систем за критерієм мінімуму сумарних затрат

розв'язання задачі оптимізації параметрів безпеки деякої ідеалізованої СТС за критерієм мінімуму сумарних витрат. В якості даних при моделюванні функцій експлуатаційного ризику сумарних витрат, прийняті значення параметрів, що не суперечать сучасній виробничій та експлуатаційній практиці в різних галузях економіки:  $C_{\Delta} = 970$  у.е.;  $R_{\Delta} = 2300$  у.е.;  $C_0 = 1\,773\,950$  у.е.;  $C_y = 6\,957\,575$  у.е.;  $P_0 = 0,9999$ . Отримані результати рішення, представлені на рис. 1 у вигляді графіку, отримали наявність оптимального значення параметра безпеки СТС  $P^* = P_{\text{опт}} = 0,973062$ , забезпечує мінімум сумарних затрат  $C_{\Sigma}(P) = 8,894 \cdot 10^5$  у.е.

У загальному випадку вимоги до безпеки СТС в цілому трансформуються у

вимоги до показників надійності і безпеки окремих елементів системи. Відповідно розподіл вимог до показників надійності і безпеки функціонування елементів  $P_i(t)$  системи протягом часу  $t$  також передбачає вирішення оптимізаційної задачі.

Завдання полягає в розподілі вимог між окремими підсистемами (елементами) таким чином, щоб були задоволені вимоги безпеки до всієї системи в цілому. У тих випадках, коли система складається з  $n$  підсистем приблизно еквівалентного обсягу (підсистеми близькі за складністю, наприклад за кількістю елементів), а їх відмови є небезпечними і незалежними подіями (тобто відмова будь-якої з підсистем може привести до події), необхідне значення показника  $P_i(t)$  може бути знайдено методом рівномірного розподілу квот безпеки.

Наприклад, якщо в найпростішому випадку надійність характеристики елементів приблизно рівні, а відмова будь-якого елемента  $n$  системи є небезпечним і незалежним, то

$$P_c(t) \geq P_{\text{тр}}(t) = \prod_{i=1}^n P_i(t) \approx P_i^n(t), \quad (10)$$

$$P_i(t) = \sqrt[n]{P_{\text{тр}}(t)},$$

де  $P_i(t)$  – ймовірність безпечного функціонування елемента систем протягом часу  $t$ ;  $P_c$  – показник безпеки системи;  $P_{\text{тр}}(t)$  – необхідні значення показнику безпеки СТС.

В інших випадках між показниками надійності і безпеки окремих елементів і систем в цілому існує більш складна залежність  $P_c(t) = f[P_i(t)]$ . Для її встановлення та формалізації використовуються методи системного аналізу і моделювання небезпечних процесів, наприклад апарат логіко-ймовірнісної теорії безпеки складних систем.

У тих випадках, коли між показниками безпеки системи є складна функціональна залежність  $P_c(t) = f\{P_1(t), P_2(t), \dots, P_n(t)\}$ , тобто задача є багатовимірною (багатопараметричною) завданням оптимізації, на першому етапі її рішення виникає необхідність в розробці математичної моделі управління зв'язку.

Модель будується шляхом системного аналізу небезпеки СТС і логічного представлення сценаріїв розвитку небезпечних станів, що відбивають причинно-наслідковий механізм виникнення події. В якості вихідних даних подій розглядаються окремі передумови (відмови, порушення, відхилення, впливу і т.п.), що призводять до події даного виду. Для

формалізації розроблених сценаріїв небезпечних станів застосовуються математичні методи обчислення істинності функцій алгебри логіки, що дозволяють безпосередньо отримати аналітичні залежності для функції виникнення події. При цьому отримана математична модель безпеки СТС дозволяє проводити оптимізацію параметрів безпеки для кожного елемента з урахуванням їх структурної значимості, з точки зору їх впливу на показники ризику аварій і катастроф.

Як приклад розглянемо деякий спрощений сценарій переходу СТС небезпечний стан (рис. 2) В результаті виникнення безлічі різних передумов до події з вірогідністю  $\{Q_1, Q_2, Q_3, Q_4, Q_5\}$ .

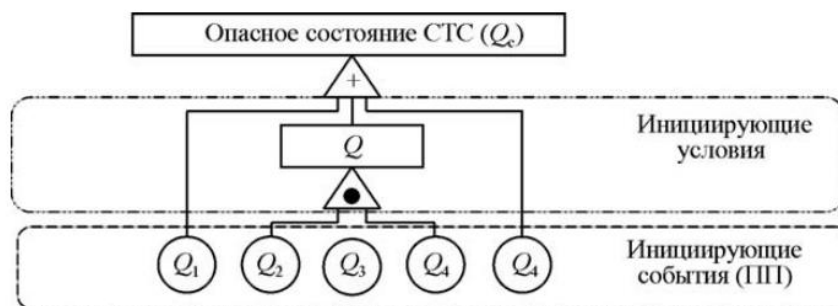


Рисунок 2 – Логічна модель сценарію небезпечного стану складної системи

Логістична функція небезпечного стану (ФНС) системи має вигляд:

$$Q_c = Q_1 \vee Q_5 \vee (Q_2 \wedge Q_3 \wedge Q_4), \quad (11)$$

де  $Q_1, Q_2, Q_3, Q_4, Q_5$  - відповідно ймовірності виникнення передумов до події.

Логістична ФНС (11) є неповторною і монотонною, тому ймовірнісна функція стану СТС матиме вигляд

$$Q_c = 1 - (1 - Q_1)(1 - Q_5)(1 - Q_2 Q_3 Q_4),$$

або

$$Q_c = 1 - P_1 P_5 [1 - (1 - P_2)(1 - P_3)(1 - P_4)] \forall P_i = 1 - Q \quad (12)$$

Для наглядності результатів моделювання вихідну багато параметричну задачу оптимізації  $C_2(P_1, P_2, P_3, P_4, P_5) \rightarrow \min$  зведемо до двопараметричної задачі шляхом прийняття наступного припущення:

$$P_1 = P_5; P_2 = P_3 = P_4$$

Аналіз результатів моделювання показав, що запропонований підхід враховує як структурну значимість окремих елементів системи з точки зору їх впливу на надійність системи в цілому, так і вартісні витрати на забезпечення необхідних значень параметрів надійності. Так, з урахуванням структурної значимості елементів і прийнятих вихідних даних мінімум сумарних затрат  $C_2(P_1, P_2) = 9,075 \cdot 10^5$  у.е. буде забезпечений при оптимальних значеннях показників безпеки  $P_1 = P_5 = 0,9667853$  та  $P_2 = P_3 = P_4 = 0,66620427$ .

Розроблені математичні моделі надають змогу враховувати ситуаційні особливості потенційно небезпечних об'єктів та проводити оптимізацію показників безпеки складних систем шляхом порівняння експлуатаційних затрат з прогнозуванням ризиків аварій і катастроф. Представлений спосіб рішення задачі оптимізації показників надійності СТС за критерієм мінімальних сумарних затрат. Розроблені моделі дозволяють автоматизувати процес рішення задач оптимізації вимог до надійності складних систем.

### Список літератури

1. Майструк А.В., Майструк А.А., Резчиков Е.А. Моделирование задачи оптимизации показателей безопасности сложных систем с учётом эксплуатационных затрат и показателей риска // Машиностроение и инженерное образование, 2013. – №3.

2. Юрков Н. К. Риски отказов сложных технических систем / Н. К. Юрков // Надежность и качество сложных систем. – 2014. – № 1(5). – С. 18–24.
3. Aven T. How to define and interpret a probability in a risk and safety setting // Discussion paper, with general introduction by Associate Editor, Genserik Reniers Safety Science, 51 (2013), P. 223-231
4. Spiegelhalter D.J., Riesch H., Don't know, can't know: Embracing deeper uncertainties when analysing risks, Philosophical Transactions of the Royal Society A, 369 (2014), pp. 4730-4750
5. Rodrigues M.A., Arezes P., Leão S.P., Risk, Criteria in occupational environments: Critical overview and discussion // Procedia - Social and Behavioral Sciences, 109 (2014), P. 257-262

## **БАГАТОКАНАЛЬНА СИСТЕМА ЗБОРУ ДАНИХ І УПРАВЛІННЯ ТЕХНІЧНИМ ОБ'ЄКТОМ З ДИСТАНЦІЙНИМ ДОСТУПОМ**

*Артюх С.О.*

*Харківський національний університет імені В.Н. Каразіна*

Прискорювачі заряджених часток (ПЗЧ) – це клас пристроїв для отримання заряджених часток (іонів, електронів, позитронів тощо) великих енергій за допомоги електричного та магнітного полів. Перші найпростіші ПЗЧ були розроблені в 20-30-х роках ХХ століття для проведення досліджень в області атомної та ядерної фізики. Швидкий розвиток технологій та енергії прискорення дозволив перенести фундаментальні дослідження в область фізики елементарних часток, але із цим зростали розміри самих прискорювачів та витрати на їх побудову та експлуатацію. Логічним розвитком став Великий адронний колайдер, котрий є найбільшим у світі прискорювачем елементарних часток. Паралельно із прискорювачами орієнтованих на великі енергії, розроблялись різноманітні типи прискорювачів малих та середніх енергій для вирішення задач прикладних областей. Зараз можна спостерігати швидкий ріст чисельності подібних пристроїв у сферах медицини, промисловості, сільського господарства.

Із ростом потужностей прискорювачів, росли вимоги до безпеки, надійності, технічного обладнання, професіоналізму обслуговуючого персоналу. Окремої уваги вимагає технічне обладнання прискорювача, яке включає в себе багато систем, які забезпечують функціонування самого прискорювача, а саме системи високовольтного та високочастотного живлення, охолодження, датчиків температури, тиску, струму пучка, вакууму; механічних переміщень, управління магнітами тощо. Контроль роботи прискорювача заряджених часток здійснюється системою керування, складність якої залежить від масштабів та типу прискорювача. Важливою частиною системи керування є програмна частина, яка повинна забезпечувати взаємодію оператора із об'єктом управління, а саме виконання різних сценаріїв, пов'язаних із увімкненням та вимкненням установки, вимірюванням параметрів пучка, обробкою аварійних ситуацій, ведення журналів подій у системи, а також виконання інших процедур, пов'язаних із роботою прискорювача.

Як і будь-які складні експериментальні установки, прискорювачі розробляються тривалий час, термін їх експлуатації достатньо великий, а вимірювальна й управляюча електроніка та програмне забезпечення зі с часом морально й фізично застарівають. Проста заміна інформаційно – вимірювальної й управляючої системи (ІУВС) попереднього покоління на розроблене із застосуванням сучасної елементної та програмної бази дозволяє не тільки продовжувати термін експлуатації, але і отримати нові результати за рахунок більш прецизійних відмірювань. Вимогою часу є і більш широке коло користувачів високовартісного часу унікальної установки.

Тому розробка сучасних багатоканальних систем збору даних і управління складних технічних об'єктів, у тому числі, для прискорювачів заряджених часток є актуальною задачею.

Метою даної статті є розробка структурної схеми, комп'ютерне моделювання й створення апаратно-програмних модулів для побудови підсистем багатоканального вимірювання та керування у складі інформаційно – вимірювальної й управляючої системи складних технічних об'єктів, окремо, прискорювача заряджених часток Ван де Граафа.

### Структурна схема багатоканальної системи збору даних і управління з дистанційним доступом

Структурна схема багатоканальної системи збору даних і управління представлена на рис. 1. Така схема визначає основні функціональні частини системи і призначена для відображення загальної структури приладу, який розробляється, його основних блоків, вузлів, частин і зв'язків між ними [1].

Схема зібрана на основі 32 – розрядного мікроконтролера STM32F4, модулях зв'язку по радіоканалам ESP32-SX1278-LoRa й SIM800L та DP83848 Ethernet Board для роботи в дротовій локальній мережі.

Мікроконтролер STM32F4 запрограмований на прийом інформації як від аналогових, так і від цифрових датчиків, їх аналого-цифрове перетворення, фільтрацію, передачу даних та прийом команд управління від інтерфейсних модулів, а також на управління цифровими і аналоговими регуляторами. Як видно зі схеми робота контролера може здійснюватися в складі бездротової сенсорної мережі – протокол каналного рівня LoRaWAN, Wi-Fi технологія бездротової локальної мережі з пристроями на основі стандартів IEEE 802.11 (ESP32-SX1278-LoRa), GSM мережі (SIM800L) і найбільш поширеною технологією організації дротових локальних мереж Ethernet (DP83848 Ethernet Board).

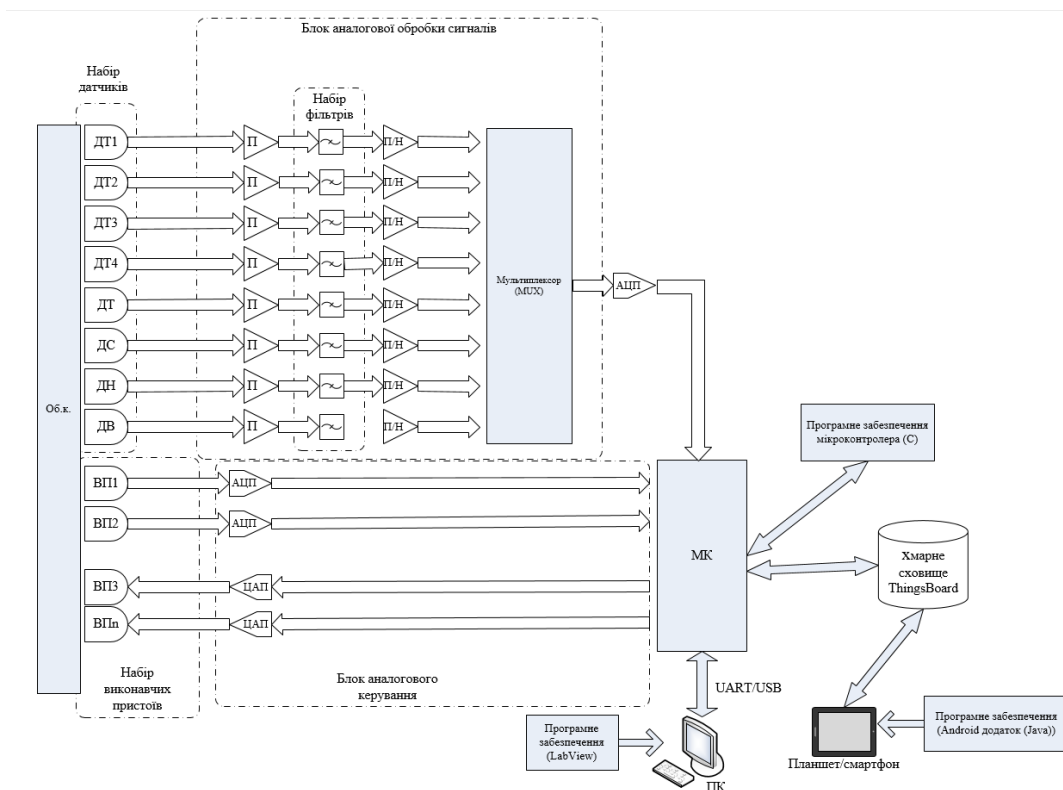


Рисунок 1 – Структурна схема багатоканальної системи збору даних і управління:  
 Об.к. – об'єкт керування, МК – мікроконтролер, ВП - виконуючий пристрій, ДТ – датчик температури, ДН – датчик напруги, ДС – датчик струму, ДВ – датчик вакууму, ЦАП - цифро-аналоговий перетворювач, АЦП - аналого-цифровий перетворювач, ПК - персональний комп'ютер, П – підсилювач, П/Н – пристрій посилення/нормування сигналу.



### Програмне забезпечення

Для розробки інтерфейсу керування, використано середовище розробки LabVIEW. LabVIEW (LV) (Laboratory Virtual Instrumentation Engineering Workbench) – це середовище та платформа розробки програм, на графічній мові програмування G фірми National Instruments. На даний час є версії для UNIX, Linux, Mac OS, але найпоширенішими та найпопулярнішими є версії під ОС Windows. LV використовується в системах збору й обробки даних, а також для управління технічними об'єктами і технологічними процесами. Програма LV є віртуальним приладом (ВП) і складається з двох частин:

- 1) блочної діаграми, що описує логіку роботи віртуального приладу;
- 2) лицьовій панелі, що описує зовнішній інтерфейс віртуального приладу.

Лицьова панель ВП містить засоби введення-виведення: кнопки, перемикачі, світлодіоди, верньєри, шкали. Вони використовуються людиною для управління ВП, а також іншими ВП для обміну даними.

Блокова діаграма містить функціональні вузли, які є джерелами, приймачами і засобами обробки даних. Також компонентами блокової діаграми є термінали («задні контакти» об'єктів лицьовій панелі) і керуючі структури (є аналогами таких елементів текстових мов програмування, як умовний оператор «IF», оператори циклу «FOR» та «WHILE» тощо). Функціональні вузли та термінали об'єднані в єдину схему лініями зв'язків. LV підтримує широкий спектр обладнання і має в своєму складі численні бібліотеки таких компонентів, як: для підключення зовнішнього обладнання з найбільш поширених інтерфейсів і протоколів (RS-232, TCP/IP USB тощо); для візуалізації даних та результатів їх обробки; для моделювання складних систем; для генерації звітів; для генерації та цифрової обробки сигналів і т.д.

На рисунку додано, що для обміну даними між ПК та МК використано USB інтерфейс. Для роботи з віддаленими користувачами запропоновано хмарний сервіс ThingsBoard і розроблено Android-додаток.

### Список літератури

1. Моделирование информационной системы мониторинга и контроля параметров технически сложного объекта / Гудков К.В., Гудкова Е.А., Володина М.А. // Современные информационные технологии. – 2014. – № 19. – С. 191-195.
2. Забаев В.Н. Применение ускорителей в науке и промышленности: учебное пособие / В.Н. Забаев. – Томск: Изд-во ТПУ, 2008. – 5 с.
3. Шведунов И.В. Программное обеспечение системы контроля и управления ускорителями электронов НИИЯФ МГУ: автореф. дис. на соиск. уч. степ. канд. физ.-мат. наук: 01.04.20 / НИИ ядерной физики им. Д.В. Скобельцына – 3 с.
4. Никифоров А.С. Опыт использования FlexCtrl SCADA для автоматизации ускорителей / В.В. Алейников, К.П. Сычев, И.В. Борина, А.А. Рукавишников // Письма в ЭЧАЯ. – 2012. – Т.9. – №4-5 (174–175). – 684 с.

## ПРОГРАМНА МОДЕЛЬ ВИБОРУ ОПТИМАЛЬНОГО МАРШРУТУ СЛУЖБИ ТАКСІ

*Садретдінов Я.С.*

*Харківський національний університет імені В. Н. Каразіна*

Дорожня галузь є однією з найважливіших галузей економіки будь-якої промислово розвиненої країни. Недарма автомобільні дороги називаються «кровоносною системою» будь-якої держави. Вони грають величезну соціально-економічну роль в житті сучасного суспільства. Диспетчерське управління переслідує мету підвищення ефективності використання рухомого складу і підтримки якості транспортного обслуговування пасажирів на нормативному рівні.

Пасажирські перевезення, здійснювані службами таксі, в сучасних соціально-

економічних умовах, потребують розробки комплексу заходів щодо вдосконалення процесів управління перевізного процесу. Сьогодення вимагає якісно нового підходу до вирішення питань управління роботою маршрутних таксі.

*Метою роботи є створення автоматизованої системи пошуку і оптимізації маршрутів служби таксі.*

Для досягнення поставленої мети необхідно вирішити наступні задачі:

- провести аналіз руху на дорогах міста;
- виконати огляд існуючих програмних реалізацій;
- проаналізувати алгоритми пошуку на графах і способи представлення графів в інформаційних системах;
- спроектувати базу даних для зберігання інформації про основні елементи вулично-дорожньої мережі міста;
- розробити програмний комплекс, що дозволяє розраховувати найкоротший шлях і виробляти перерахунки виходячи з завантаженості доріг.

Основним завданням служби таксі є забезпечення клієнтам комфортної та швидкої поїздки з мінімальними витратами. Не завжди час, витрачений в дорозі, обумовлений пересуванням автомобіля по найкоротшому шляху. У таких випадках паливо, витрачене водієм на об'їзд заторів, компенсується швидкістю поїздки.

Зараз використовуються такі системи таксі.

City Guide (СітіГІД) — автомобільна система GPS / ГЛОНАСС-навігації. Програма поширюється як на платній основі, так і існують безкоштовні варіанти з урізаною функціональністю. Визначає оптимальний маршрут руху транспортного засобу.

Програмний комплекс "Таксі-Координатор" — розроблений на основі унікальної растрово-векторної технології надшвидкої візуалізації великих обсягів графічної інформації.

Проаналізувавши їх, була розроблена схема системи управління службою таксі, яка побудована на основі технології клієнт-сервер.

Система управління служби таксі складається з:

- підсистеми звітності по водіях, телефоністів і диспетчерам. У бухгалтера є можливість створити звіт по здачі комісійних для водіїв, і розрахувати заробітну плату за прийнятими замовленнями телефоністів і по обробленим замовленнями диспетчерів;
- підсистеми вибору маршрутів дозволяє вибрати оптимальний маршрут, виходячи з завантаженості доріг;
- підсистеми ведення бази клієнтів, яка дозволяє вести облік замовлень по кожному клієнту.

У даній роботі розглядається підсистема вибору маршрутів ІУС служби таксі. Далі більш детально розглянемо алгоритми які використовувалися.

Алгоритм Дейкстри будує найкоротші шляхи, що ведуть з вихідної вершини графа до решти вершин цього графа (якщо такі є). Алгоритм працює тільки для графів без ребер з негативними вагами, застосовується в програмуванні і технологіях.

Спочатку позначається вихідна вершина, далі наступна, очевидно, що буде позначена вершина, найближча до вихідної, й суміжна з нею.

Нехай на якомусь етапі вже позначено кілька вершин. Відомі найкоротші шляхи, що ведуть з вихідної вершини до поміченим. Для кожної з непомічених вершин виконаємо наступне:

- 1) розглянемо всі дуги, що ведуть з помічених вершин в одну непомічену. Кожна така дуга є останньою дугою на шляху з вихідної вершини в цю непомічену;
- 2) виберемо серед них самий короткий до всіх непомічених вершин, і позначимо вершину, до якої він веде.

Алгоритм завершиться, коли будуть помічені всі досяжні вершини. В результаті роботи алгоритму Дейкстри будується Дерево найкоротших шляхів.

Пошук в ширину (BFS – від англ. Breadth first search) – один з найлегших алгоритмів пошуку в графі. Пошук в ширину – це класичний метод розв'язання задачі пошуку найкоротшого шляху між двома конкретними вершинами графа.

Алгоритм пошуку в ширину із заданою стартовою вершиною  $a$ . Спочатку всі вершини позначаються як нові. Першою відвідується вершина  $a$ , вона стає єдиною відкритою вершиною. Надалі кожен черговий крок починається з вибору деякої відкритої вершини  $x$ . Ця вершина стає активною. Далі досліджуються ребра, інцидентні активній вершині. Якщо таке ребро з'єднує вершину  $x$  з новою вершиною  $y$ , то вершина  $y$  відвідується і перетворюється у відкриту. Коли всі ребра, інцидентні активній вершині, досліджені, вона перестає бути активною і стає закритою. Після цього вибирається нова активна вершина, і описані дії повторюються. Процес закінчується, коли безліч відкритих вершин стає порожнім.

Пошуком в глибину (DFS – depth first search) називається один з методів обходу графа, суть якого полягає в тому, щоб йти "вглиб" поки це можливо. У процесі пошуку в глибину вершинам графа присвоюються номери, а ребра позначаються. Обхід вершин графа відбувається згідно з принципом: якщо з поточної вершини є ребра, що ведуть у непройдені вершини, то йдемо туди, інакше повертаємося назад.

Обхід починається з відвідин заданої стартовою вершини  $a$ , яка стає активною і єдиною відкритою вершиною. Потім вибирається інцидентне вершині  $a$  ребро  $(a, y)$  і відвідується вершина  $y$ . Вона стає відкритою і активною. Надалі, як і при пошуку в ширину, кожен черговий крок починається з вибору активної вершини з безлічі відкритих вершин. Якщо всі ребра, інцидентні активній вершині  $x$ , вже досліджені, вона перетворюється в закрити. В іншому випадку вибирається одне з недосліджених ребер  $(x, y)$ , це ребро досліджується. Якщо вершина  $y$  нова, то вона відвідується і перетворюється у відкриту.

Спроектвана база даних складається з 5 таблиць.

Через особливості зображення даних в системі як графічної інформації, практично вся робота з картою пов'язана з використанням миші. При роботі задіяні різні дії: натискання кнопки миші, клацання, подвійне клацання, прокрутка колеса миші і його натискання, переміщення миші та інші. На малюнку приведено головне вікно програми.

Для початку пошуку найкоротшого шляху необхідно встановити початкову та кінцеву точки. Після введення точок автоматично визначається оптимальний шлях за обраним алгоритмом. При цьому знайдений шлях відображається на карті, і на панелі праворуч виводяться його характеристики: відстань і час на проїзд. Для зміни початкового або кінцевого адреси потрібно просто заново його ввести.

### **Висновки**

В роботі проаналізовані програмні аналоги системи, яка розробляється. Детально розглянуто та застосовано методи пошуку оптимального маршруту: алгоритм Дейкстри, методи пошуку в глибину та ширину. Реалізована програмна система пошуку оптимального маршруту служби таксі. Проведено порівняльний аналіз запропонованого алгоритму з існуючими аналогами.

### **Список літератури**

1. Хаханов В.И., Чумаченко С.В. Дискретная математика (теоретическое и практическое содержание курса)
2. Глушаков, С.В. Базы данных [Текст] / С.В. Глушаков, Д.В. Ломотько Худож.-оформитель А. С. Юхтман. – М.: ООО «Издательство АСТ», 2002. – 504с.
3. Методическое пособие по дисциплине "Проектирование информационных систем в экономике". 071900: Курс лекций/ В. А. Щербанов; Министерство общего и профессионального

- образования Российской Федерации, Томский государственный университет систем управления и радиоэлектроники, Кафедра автоматизированных систем управления. – Томск: ТМЦДО, 1999. – 153 с.
4. Бойков В. Н., Федотов Г. А., Пуркин В. И. Автоматизированное проектирование автомобильных дорог (на примере IndorCAD/Road). – М.: Изд-во МАДИ (ГТУ), 2005. – 223 с.
  5. Проект ДубльГИС [Электронный ресурс]. – Режим доступа: <http://www.2gis.ru>, свободный.
  6. Лопатин А. А. Комплекс программ поиска оптимального маршрута движения в дорожно-транспортной сети города // Инновационные недра Кузбасса. ИТ-технологии: сборник научных трудов. – Кемерово: ИНТ, 2007. – С. 392–393.
  7. Рапопорт Э.Я. Структурное моделирование объектов и систем управления [Текст] / Э.Я.Рапопорт. — М: финансы и статистика, 2003.
  8. Чекалов, А.П. Базы данных: от проектирования до разработки приложений [Текст]. – Санкт-Петербург: БХВ-Петербург, 2003. 384 с.
  9. Мартин, Дж. Организация баз данных в вычислительных системах [Текст] /пер. с англ. под ред. А.Л.Щерса. – М.: Мир, 1980 г. – 662 с.
  10. Бойко, В.В. Проектирование баз данных информационных систем [Текст] / В.В. Бойко, В.М. Савинков. М.: Мир, 1989. – 462 с.
  11. ДСанПиН 3.3.2-007-98. Государственные санитарные правила и нормы работы с визуальными дисплейными терминалами электронно-вычислительных машин [Текст].
  12. ДСН 3.3.6.037-99. Санітарні норми виробничого шуму, ультразвуку та інфразвуку [Текст].
  13. ДСН 3.3.6.042-99. Санітарні норми мікроклімату виробничих приміщень [Текст].
  14. НПАОП 40.1-1.32-01. Правила будови електроустановок. Електрообладнання спеціальних установок [Текст].

## **ПОБУДОВА МОНІТОРИНГУ РЕГІОНАЛЬНИХ РЕСУРСІВ НА ОСНОВІ РАНЖУВАННЯ ІНФОРМАЦІЙНИХ СИСТЕМ УПРАВЛІННЯ ТЕРИТОРІЯМИ**

*Зарицький О. В., Костенко О. Б.*

*Харківський національний університет міського господарства імені О. М. Бекетова*

В інформаційну епоху гостро постали питання належного управління регіональними ресурсами.

Сталий та стрімкий розвиток сучасного суспільства не можливий без належного моніторингу та адміністрування будь-якої сфери господарювання. Геопросторова зорієнтованість кожної такої галузі господарської діяльності є запорукою її ефективності. Саме тому широкого застосування набувають геоінформаційні системи (ГІС).

Як відомо [1], [2], [3], метою проведення моніторингу регіональних ресурсів є створення відкритої динамічної інформаційної системи, яка відображає стан розвитку регіонів України та забезпечує їх прозорість.

Міжвідомчої однорідності даних не досягнуто в діючих інформаційних системах України. Існує значне дублювання та конфліктність інформації в реєстрах та системах як у паперовому так і в електронному вигляді. Окремі відомості досі є відсутніми в інформаційних системах.

Серед професіоналів земельного ринку існують складнощі використання значних масивів архівних матеріалів при переході до простору державної системи координат УСК-2000 [4], [5].

Перехід на європейські стандарти ведення банку географічної інформації потребує однозначних інструкцій до наповнення Картографо-геодезичного фонду України. Також необхідні єдині вимоги до електронних документів про окремі геопросторові об'єкти.

Для здійснення ефективного управління регіональними ресурсами в світі розвиваються багато напрямів. Деяким підходам приділяється увага і в Україні.

З 1916 року у місті Нью-Йорк впроваджується Зонінг [6] і набуває широкого застосування у всьому світі. Зонінг – це динамічна ГІС з комплексом геометрично-правових правил, яка сприяє сумісності між різними видами землекористування [7].

Законодавчий аспект (Стаття 24 [8]) також актуалізує питання Зонінгу, оскільки на сьогодні без плану зонування передача (надання) земельних ділянок у власність чи користування для містобудівних потреб забороняється у більшості випадків земельних відносин. З січня 2013 року до квітня 2018 року ця заборона стосувалася всіх суб'єктів земельних відносин.

Через зростаючу складність об'єктів інфраструктури і щільну забудову територій з 1980-х років починають впроваджуватися 3D кадастри для належної реєстрації правового статусу земельних ділянок та інших об'єктів нерухомості [9].

На прикладі досвіду Польщі, з 2018 року в Україні відбувається імплементація понять про консолідацію земель. Законом [10] пропонуються механізми обміну орендованими земельними ділянками без згоди орендодавця з метою консолідації земель кожного з орендарів. Межі земельних ділянок, розташованих у масиві земель сільськогосподарського призначення, можуть бути змінені у процесі консолідації земель на умовах та в порядку, визначених законом (Пункт 8 статті 37-1 [11]).

Даний підхід регулювання земельних відносин може бути використаний у процесі планування та впорядкування об'єднаних територіальних громад. Цей інструмент є актуальним з огляду на плани запуску ринку землі до кінця 2020 року.

З 2007 року здійснюється розробка стандартів для розвитку інфраструктури геопросторових даних в Європейському Союзі (INSPIRE) [12] з 2007 року. До 2019 року відбувалася імплементація цих стандартів в Україні, що стало початком розвитку Національної інфраструктури геопросторових даних [13].

У 2018 році за підтримки Нідерландів запущено геопортал відкритого просторового планування (PMAP) [14].

Виникає необхідність у розробленні нових підходів побудови ГІС регіональних ресурсів. Нова технологія має врахувати ті ризики, що існують в інформаційній взаємодії, та сформулювати запобіжники в місцях їх прояву.

#### **Постановка задачі**

В дослідженні були поставлені такі завдання:

- дослідити інструменти управління регіональними ресурсами;
- реалізувати технологію уніфікації з ранжуванням для сприяння узгодженої взаємодії інформаційних системах України;
- застосувати інструмент валідації до контенту баз даних інформаційних систем;
- розробити технологію конвертування архівних баз даних до оновлених стандартів;
- створити продукційну модель відтворення пропущених відомостей у предметному середовищі;
- розробити методику, що уможливило б візуалізацію геопросторових даних;
- розробити методику автоматизації документообігу стосовно геопросторових об'єктів.

#### **Аналітичний огляд управління регіональними ресурсами**

Через стале й тривале застосування найбільша увага була приділена інформаційній системі Зонінг. Всі нормативи, регламенти та обмеження геометрично можна найкращим чином представити як зони. Саме така сукупність діючих зон і дали назву цій системі.

Згідно світової практики, інформаційна система Зонінг відображає регламентований рух розвитку адміністративно-територіальної одиниці. Ця система є інтерактивною.

Досліджуючи розвиток та досвід застосування понять зонування, зонування територій, зонінгу (містобудівна документація) запропоновано нове тлумачення та концепцію Зонінгу. Зонінг сьогодні – це динамічна геоінформаційна система (ДГІС) з комплексом геометрично-правових правил, що охоплює: землекористування, планування та зонування. Зонінг – це інструмент сприяння веденню узгодженої господарської діяльності.

Ідея Зонінгу – користувач виділяє територію, яка його цікавить, та переглядає весь пласт пов'язаних регіональних ресурсів, всі пов'язані нормативні, регламентні, обмежувальні зони. Наприклад, переглядаючи відомості про земельну ділянку окремої житлової зони населеного пункту, можна отримати інформацію про:

- межу земельної ділянки;
- мінімальну довжину фасаду;
- глибину вільної зони фасаду;
- розміри вільних бічних зон біля будинку;
- максимальну висоту фасаду будівлі;
- максимальну висоту всієї будівлі;
- варіанти розміщення всіх будівель та споруд в межах допустимих співвідношень площ та відстаней;
- весь ряд показників оцінки майна тощо.

Враховуючи актуальні проблеми управління територією, ДГІС Зонінг запропоновано використовувати як для населених пунктів, так і для об'єднаних територіальних громад.

Для впровадження динамічної геоінформаційної системи в галузі містобудування, землеустрою, геодезії та картографії необхідно вирішити ряд проблем, що ускладнюють її реалізацію, оскільки:

- більшість проектних рішень в тематичних інформаційних системах досі реалізуються без урахування майнової структуризації території;
- існує потреба у відкритості та актуальності відомостей про майнові об'єкти;
- впроваджується реформа адміністративно-територіального устрою України;
- не досягнуто міжвідомчої однорідності даних в діючих інформаційних системах України;
- спостерігається дублювання та конфліктність інформації в реєстрах та системах;
- існують великі архіви диференційованої інформації, з якою необхідно працювати, але яка не відповідає сучасним вимогам її організації.

#### **Відбір та уніфікація інформаційних систем в галузі господарської діяльності**

Оскільки досліджувана господарська діяльність майже повністю регламентовані міжнародними стандартами, нормативно-правовими актами українського законодавства, то детальний аналіз варто розпочинати саме з цих документів. Такий підхід зумовить реалізацію універсальної інфонологічної моделі даних ДГІС.

Наведені вище принципи ДГІС Зонінг лежать в основі побудови БГД для реалізації моніторингу регіональних ресурсів на будь-якому адміністративно-територіальному рівні України. Відповідно геодезичну, картографічну, землевпорядну та містобудівну діяльності запропоновано ототожнювати з ДГІС Зонінг.

ДГІС Зонінг формує свої відомості з множини інших тематичних інформаційних систем (ІС)  $DGISZ = \{IS_1, IS_2, IS_3, \dots, IS_n\}$ , де  $n$  – кількість інформаційних підсистем. Кожна така ІС заснована на множині нормативних документів  $Z = \{Z_1, Z_2, Z_3, \dots, Z_d\}$ , де  $d$  – кількість нормативних документів. В кожному нормативному документі декларується множина відомостей  $I = \{i_1, i_2, i_3, \dots, i_a\}$ , де  $a$  – кількість відомостей в ІС, та множина класифікаторів  $C = \{c_1, c_2, c_3, \dots, c_b\}$ , де  $b$  – кількість класифікаторів в ІС.

Відомість – це певні факти, дані про кого-, що-небудь; знання (тільки множина) в якій-небудь галузі. Класифікатор – будь-який систематизований перелік відомостей.

Стосовно однієї й тієї ж відомості визначається певна інформаційна взаємодія між ІС. Досліджуючи законодавство, було виявлено заплутаність зв'язків, що ускладнює функціонування ДГІС.

Такі обставини призводять до появи наступних чинників:

- дублювання, коли одна й та сама відомість формується різними відомствами або підсистемами, може мати різний зміст за часом, повнотою, якістю створення;
- нетотожне дублювання, коли, наприклад, сутність в назві представлена різними комбінаціями словосполучень;
- неузгодженість між відомствами;
- неоднозначність тлумачень, коли, наприклад, одній і тій же сутності даються в різних підсистемах наводяться відмінні визначення;
- неактуальність.

Така структура «інформаційного» скелету не є оптимальною. На практиці це суттєво затримує обмін відомостями між державними міністерствами та службами. Це також ускладнює погоджувальні процедури та збільшує час реалізації проектних рішень. А також недостовірно відображає дійсність.

Одним із засобів структуризації є процедура декомпозиції первинних, законодавчо визначених ІС, які відносяться до галузей господарської діяльності. Декомпозиція — це процес поділу систем на зручні для обробки елементи, які приймаються за неподільні об'єкти.

Результат декомпозиції – це виявлення основних відомостей, на базі яких буде побудовано несуперечливу модель даних ІС.

Декомпозицію доцільно здійснити через уніфікацію з ранжуванням. Уніфікація інформаційних систем [15] — це формування єдиного методологічного підходу до створення, функціонування ГІС, зберігання геопросторових даних, визначення ступеня їх точності, забезпечення можливості автоматизованого обміну даними між ГІС в режимі реального часу, усунення розбіжностей та забезпечення достовірності ГІС.

У дослідженні взято ряд ІС із законодавства та уніфіковано як компоненти ДГІС Зонінг.

Наприклад, запропоновано об'єднати державне еколого-геологічне картування України (ІС ДЕГКУ), державний інформаційно-геологічний фонд (ІС ДІГФ), державний кадастр родовищ і корисних копалин (ІС ДКРКК) до уніфікованої ІС ДІГФ<sup>U</sup>.

За результатом нормативно-правового обґрунтування вдалося уніфікувати виявлену кількість ІС з 26 до 19 одиниць (рис. 1).

За проведеним дослідженням у ДГІС Зонінг сформовано підсистеми, що впливають на структуру концепції  $N = [1: 19]$ . Опрацьовано нормативно-правові акти законодавства України  $Z = [1: 38]$ . Виявлено множину всіх відомостей  $I = [1: 8682]$  та класифікаторів  $C = [1: 347]$ . Також законодавчо визначено взаємозв'язок через обмін відомостями між обраними підсистемами.

Через щільну інформаційну взаємодію між ІС та через велику кількість відомчих установ важко встановити однозначні судження. Таким чином виникає багато суперечностей. Частково це може бути зумовлене і результатом уніфікації складу ІС.

Проведений аналіз та дослідження [16], [17], [18] «підказали» можливість реорганізації процесу впровадження ДГІС Зонінг: найбільш оптимальною на етапі інфологічного моделювання є необхідність виключення дублюючих елементів на всіх рівнях системи.

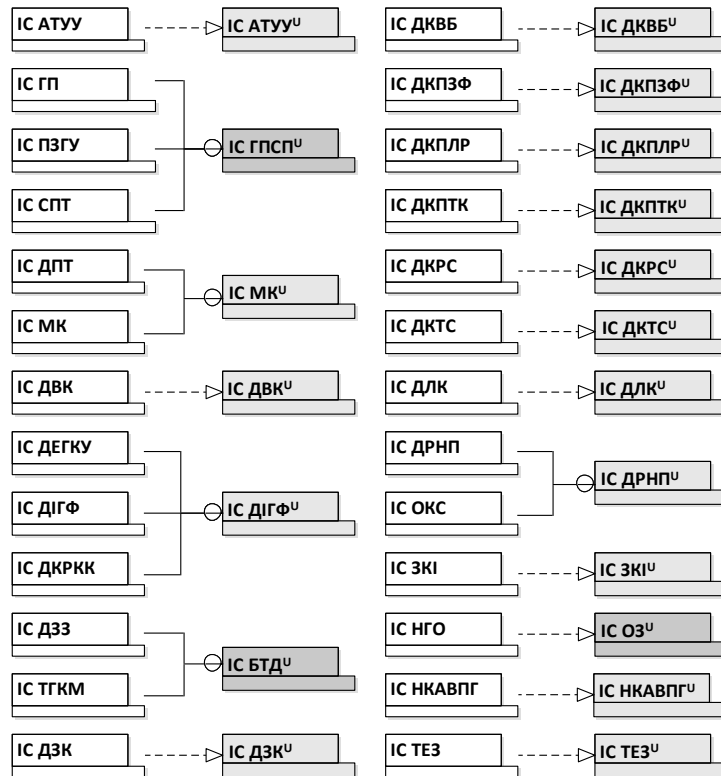


Рисунок 1 – Уніфікація підсистем в складі динамічної геоінформаційної системи Зонінг

**Уніфікація з ранжуванням інформаційних систем галузі господарської діяльності**

В галузі може бути не визначено прямих домінант між ІС. Тому таку неоднозначність запропоновано вирішити методом ранжування для ймовірного визначення рангів між складовими ДГІС Зонінг.

Для характеристики рангу  $IC^U$  є доцільним ввести критерії асоціації, зв'язку, домінування. Асоціація визначається, як відношення суми вагових коефіцієнтів асоційованих законів до суми всієї множини нормативних документів  $K_l^A = \frac{\sum_{k=1}^m a_{k,l}}{\sum_{k=1}^m Z_k}$ . Зв'язність визначається, як відношення суми виявлених попарних зв'язків між системами до суми всі інших систем  $K_k^J = \frac{\sum_{l=1}^n j_{k,l}}{(\sum_{l=1}^n IS_l) - 1}$ . Домінування визначається, як відношення суми виявлених домінант попарно порівняних систем до суми всі інших систем  $K_k^D = \frac{\sum_{l=1}^n d_{k,l}}{(\sum_{l=1}^n IS_l) - 1}$ . Добуток цих критеріїв формує коефіцієнт значущості  $K_k^R = K_l^A \cdot K_k^J \cdot K_k^D$ . Найбільший рейтинговий коефіцієнт відповідає першому рангу, а найнижчий коефіцієнт — останньому рангу за шкалою ранжування  $IC^U$  (табл. 1).

В результаті ранжування для множини  $IC^U$ , що складають ДГІС Зонінг, можна визначити інформаційний потік відомостей (табл. 1). Допускається також еквівалентність  $IC^U$ , якщо вони природньо не перетинаються між собою.

Визначені ранги  $IC^U$  дозволяють провести уніфікацію з ранжуванням їх складових.

У дослідженні розроблено алгоритм уніфікації (дедублювання) відомостей ДГІС Зонінг, який відповідає такій послідовності:

1. Формуються масиви відомостей, попередньо визначені в законодавстві (згідно рангів).
2. Послідовно обираються відомості в  $IC^U$  нижчого рангу (які ще не розглядалася).



3. Серед множини відомостей  $IS^U$  вищих рангів визначаються відомості із аналогічними поняттями та спільнокореневими словами відносно активної відомості.

4. Використовуючи метод парних порівнянь, зіставляються слова або словосполучення відомостей в різнорангових  $IS^U$ . При побудові БГД ДГІС Зонінг співпадіння замінюються в  $IS^U$  вищого рівня лінком/ключем першоджерела (першовідомості). Кроки 2-4 повторюються з поступовим відсіюванням дублів та переходом до  $IS^U$  вищого рангу.

5. Впорядковуються масиви уніфікованих відомостей. В результаті утворюється уніфікований склад відомостей у ранжованих  $IS^U$ .

Таблиця 1

Результат ранжування  $IS^U$  для динамічної геоінформаційної системи Зонінг

Ранги IS	Код IS	Назва IS (аббревіатура)	Рейтинговий коефіцієнт	Нормалізований рейтинговий коефіцієнт
Ранг 1	$IS_{16}$	IS МК <sup>U</sup>	0,3523	1,0000
Ранг 2	$IS_5$	IS ДЗК <sup>U</sup>	0,3420	0,9709
Ранг 3	$IS_{18}$	IS ОЗ <sup>U</sup>	0,1462	0,4150
Ранг 4	$IS_3$	IS ГПСР <sup>U</sup>	0,1201	0,3410
Ранг 5	$IS_6$	IS ДГФ <sup>U</sup>	0,1008	0,2860
Ранг 6	$IS_{10}$	IS ДКПТК <sup>U</sup>	0,0997	0,2830
Ранг 7	$IS_8$	IS ДКПЗФ <sup>U</sup>	0,0922	0,2616
Ранг 8	$IS_{19}$	IS ТЕЗ <sup>U</sup>	0,0776	0,2202
Ранг 9	$IS_{13}$	IS ДЛК <sup>U</sup>	0,0723	0,2054
Ранг 10	$IS_4$	IS ДВК <sup>U</sup>	0,0679	0,1927
Ранг 11	$IS_{15}$	IS ЗКІ <sup>U</sup>	0,0604	0,1716
Ранг 12	$IS_9$	IS ДКПЛР <sup>U</sup>	0,0254	0,0721
Ранг 13	$IS_{12}$	IS ДКТС <sup>U</sup>	0,0197	0,0558
Ранг 14	$IS_{17}$	IS НКВІП <sup>U</sup>	0,0125	0,0356
Ранг 15	$IS_{11}$	IS ДКРС <sup>U</sup>	0,0063	0,0180
Ранг 16	$IS_{14}$	IS ДРНП <sup>U</sup>	0,0057	0,0161
Ранг 17	$IS_7$	IS ДКВБ <sup>U</sup>	0,0048	0,0137
Ранг 18	$IS_1$	IS АТУУ <sup>U</sup>	0,0037	0,0104
Ранг 19	$IS_2$	IS БТД <sup>U</sup>	0,0000	0,0000

Для характеристики результату уніфікації складових IS введено показник ефекту ранжування для відомостей  $E_I$  та для класифікаторів  $E_C$ . Показник ефекту ранжування розраховується як відношення множин складових одного типу відповідно до та після уніфікації з ранжуванням. Як результат, ефект ранжування в підсистемах:

– для відомостей коливається в інтервалі  $1,00 \leq E_I \leq 1,76$  (загальна множина уніфікованого складу відомостей  $I = [1: 7157]$ );

– для класифікаторів коливається в інтервалі  $1,00 \leq E_C \leq 1,58$  (загальна множина уніфікованого складу класифікаторів  $C = [1: 294]$ ).

Порівняльні результати уніфікації з ранжуванням складових ДГІС Зонінг представлені в на рис. 2.

У ході дослідження виявлено, що близько 1/5 відомостей та класифікаторів не є унікальними, що потенційно призводить до створення неоднозначності та дублювання екземплярів класів БГД ДГІС Зонінг.

У підсумку, уніфікований склад відомостей дозволяє побудувати нормативно прийнятні складові підсистем (класи, домени).

#### **Побудова моделей перспективного розвитку територій**

Нормативне обґрунтування складових підсистем є запорукою побудови узгодженої моделі даних в ДГІС Зонінг. Тепер з'являються два шляхи побудови БГД.

1. Процес, який передбачає використання матеріалів, нагромаджених за минулі роки.
2. Процес, заснований на уніфікації з ранжуванням, що зазначений вище.

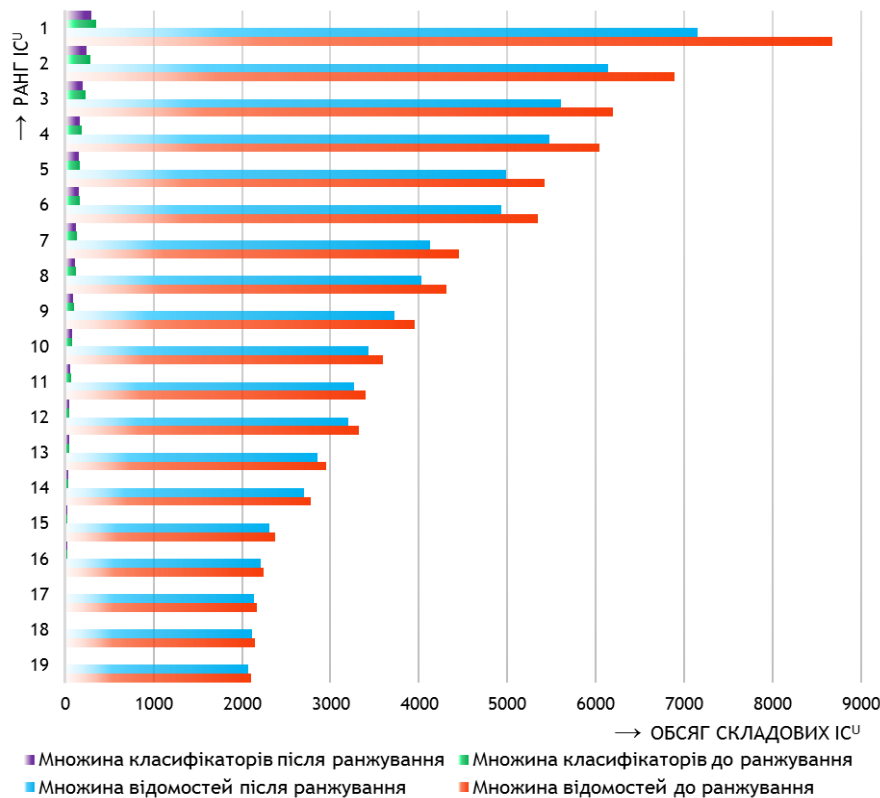


Рисунок 2 – Графік залежності кількості складових ІСᵁ до їх рангів у площині динамічної геоінформаційної системи Зонінг

Стосовно першого варіанту відомо, що весь набір даних формувалася за старими вимогами. Згідно директиви про інфраструктуру просторових даних Європейського союзу (INSPIRE) кількість класів нової класифікації геопросторових даних значно зменшена. Розробник набору на об'єкти робіт минулих років вимушений повторно організувати ті самі екземпляри класів але за новими стандартами.

Тому для узгодженого функціонування ДГІС Зонінг необхідна конвертація масиву даних зі старого класифікатора просторової інформації  $K^{OLD}$  до нового  $K^{NEW}$ .

В результаті проведеного дослідження запропоновано алгоритм автоматизованого конвертування геопросторових об'єктів до класифікаторів за оновленими стандартами, який дозволяє конвертувати будь-які диференційовані набори БГД у простір єдиного цифрового опису [19]. Реалізовано у ПЗ Digitalis [20]. Така конвертація тим самим частково вирішує завдання уніфікації.

При будь-якій побудові інформаційної системи можливі випадки пропущених даних або «missing data» [21], що в державному комплексі стандартів [22] визначається, як значення, що не є готовим для постачальника даних, більше того, його правильне значення може не існувати. Іншими словами, це неповні, невірні, відсутні, дубльовані, пошкодженні, знищені або застарілі дані.

Таким чином виникає необхідність верифікації та валідації даних у новостворених системах. Верифікація може бути застосована при конвертуванні архівів інформації. Валідація передбачає аналіз самих підсистем за нормативно-правовими правилами. За результатами

таких перевірок відбувається маркування пропущеної інформації в БГД, що дозволяє значно скоротити час опрацювання відомостей в підсистемах.

Для того щоб побудована модель ДГІС Зонінг була нормативно прийнятною, пропущені дані доцільно коригувати за допомогою продукційної моделі. Маркування пропущених даних на основі валідації допомагає визначити такі об'єкти для коригування [23]. Базові знання продукційної моделі закладені в нормативних документах, що описують відомості ДГІС Зонінг.

Зберігати та передавати весь контент БГД передбачається у XML-форматі. Така структура організації даних активно використовується в даному предметному середовищі і багато мов програмування підтримують цей формат.

Концептуально розгортання ДГІС Зонінг на великі території тягне за собою великі об'єми даних. Тому в дослідженні передбачено використання технології паралельних обчислень. На мові C# (бібліотека PLINQ) показано приклад автоматизації усіх етапів розпаралелювання обчислень, враховуючи поділ роботи на завдання, виконання цих завдань різними потоками і об'єднання результатів в одну вихідну послідовність.)

Також на мові C# було здійснено візуалізацію масивів геопросторових даних, які зберігаються у XML-форматі. Програма відкриває XML-файл з даними, визначає геометричні примітиви об'єктів у файлі, параметрично генерує ці об'єкти та пошарово візуалізує на екрані.

Кінцевим етапом побудови моніторингу регіональних ресурсів після валідації, відтворення пропущених даних та візуалізації є створення проектної документації.

Результати візуалізації у поєднанні з нормативним наповненням за допомогою програмного продукту дизайну, генерації та представлення звітів FastReport готові для подання до обговорення громадою. Така документація може бути перспективним планом розвитку або проектним рішенням для досліджуваної території.

В результаті дана методологія побудови моніторингу регіональних ресурсів дозволяє нормативно створювати варіанти моделей перспективного розвитку територій (рис. 3).



Рисунок 3 – Концептуальна модель реалізації динамічної геоінформаційної системи для здійснення моніторингу регіональних ресурсів

### Висновки

За результатами проведеного комплексного дисертаційного дослідження, виконано наступне:

– вперше запропоновано розуміти Зонінг в Україні як динамічна геоінформаційна система, а не просто містобудівна документація;

- для побудови моніторингу регіональних ресурсів вперше застосовано технологію уніфікації з ранжуванням (опираючись на нормативні положення);
- дістало подальшого розвитку застосування продукційної моделі для відтворення пропущених даних геопросторових об'єктів на основі уніфікації з ранжуванням;
- удосконалено побудову бази геопросторових даних за допомогою конвертування до оновлених стандартів, верифікації та валідації відомостей;
- дістало подальшого розвитку методологія паралельних обчислень для моніторингу та управління територіями;
- розширено та удосконалено застосування візуалізації зі створенням проектної документації на основі XML-файлів;
- вперше запропонована методологія побудови моніторингу регіональних ресурсів, яка включає: уніфікацію з ранжуванням, верифікацію, валідацію та відтворення даних.

### Список літератури

1. Про затвердження Положення про регіональні кадастри природних ресурсів : Постанова від 28.12.2001 № 1781. База даних «Законодавство України». КМ України. Дата оновлення: 19.06.2015. URL: <https://zakon.rada.gov.ua/rada/show/1781-2001-%D0%BF> (дата звернення: 14.02.2019).
2. Ахромкін Є. М. Концепт інноваційної моделі забезпечення ефективного використання регіональних ресурсів. Ефективна економіка. 2011. № 2. URL: <http://www.economy.nayka.com.ua/index.php?operation=1&iid=461> (дата звернення: 04.02.2019).
3. Мартин А. Реформування системи моніторингу земель в Україні: напрями та механізми. Землевпорядний вісник. 2017. № 11. С. 22–25.
4. Боровий В. О., Зарицький О. В. Застосування нової референційної системи координат УСК-2000. Проблеми та пропозиції. Землевпорядний вісник. 2017. № 4. С. 22–27. URL: <https://www.researchgate.net/publication/316915780> (дата звернення: 04.02.2019).
5. Боровий В. О., Зарицький О. В. Щодо точності визначення площ в новій референційній системі координат УСК-2000. Землевпорядний вісник. 2019. № 3. С. 28–32. URL: <https://www.researchgate.net/publication/332291662> (дата звернення: 18.04.2019).
6. NYC DCP Zoning And Land Use (ZoLa) [online]. NYC.gov. 2019. URL: <http://gis.nyc.gov/doitt/nycitymap/template?z=8&p=986992,205269&a=ZOLA&c=ZOLA> (accessed: 14.02.2019).
7. Боровий В. О., Зарицький О. В. ГІС-технології в геодезії та землеустрої : монографія. Київ: ТОВ "ВІСТКА", 2017. 252 с.
8. Про регулювання містобудівної діяльності : Закон від 17.02.2011 № 3038-VI. База даних «Законодавство України». ВР України. Дата оновлення: 22.05.2019. URL: <https://zakon.rada.gov.ua/laws/show/3038-17> (дата звернення: 29.05.2019).
9. Ничвид М. Р., Шейдик А. А. Міжнародний досвід в сфері 3D кадастру. 2016. URL: <https://dspace.uzhnu.edu.ua/jspui/handle/lib/9110> (дата звернення: 19. 06. 2019).
10. Про внесення змін до деяких законодавчих актів України щодо вирішення питання колективної власності на землю, удосконалення правил землекористування у масивах земель сільськогосподарського призначення, запобігання рейдерству та стимулювання зрошення в Україні : Закон від 10.07.2018 № 2498-VIII. База даних «Законодавство України». ВР України. URL: <https://zakon.rada.gov.ua/laws/show/2498-19> (дата звернення: 14.02.2019).
11. Земельний кодекс України : Кодекс від 25.10.2001 № 2768-III. База даних «Законодавство України». ВР України. Дата оновлення: 07.02.2019. URL: <https://zakon.rada.gov.ua/laws/show/2768-14> (дата звернення: 14.02.2019).
12. Craglia M. Contribution to the extended impact assessment of INSPIRE. Creator INSPIRE Framework definition support (FDS) working group Max Craglia–Sheffield University (contractor). Publisher Environment Agency for England and Wales, 2003. 65 с.
13. Національна інфраструктура геопросторових даних. 2019. URL: <http://nsdi.land.gov.ua> (дата звернення: 14.02.2019).

14. Відкрите просторове планування (PMAP). 2019. URL: <https://pmap.minregion.gov.ua/> (дата звернення: 19.06.2019).
15. Про затвердження Порядку інформаційної взаємодії між кадастрами та інформаційними системами : Постанова від 03.06.2013 № 483. База даних «Законодавство України». КМ України. Дата оновлення: 17.03.2017. URL: <https://zakon.rada.gov.ua/laws/show/483-2013-%D0%BF> (дата звернення: 04.02.2019).
16. Боровий В. О., Зарицький О. В. Роль земельно-кадастрової інвентаризації в зонуванні населених пунктів. Землевпорядний вісник. 2013. № 6. С. 33–36. URL: <https://www.researchgate.net/publication/311709266> (дата звернення: 04.02.2019).
17. Костенко О. Б., Зарицький О. В. Декомпозиція з ранжуванням інформаційних систем при інфологічному проектуванні. Математичні моделі та новітні технології управління економічними та технічними системами / за заг. ред. В. О. Тимофєєва, І. В. Чумаченко. Харків: ФОП Мезіна В. В., 2017. С. 247–260. URL: <https://www.researchgate.net/publication/319956142> (дата звернення: 04.02.2019).
18. Костенко О. Б., Буласенко М. В., Зарицький О. В. Аналіз математичних методів декомпозиції в системах відтворення втрачених даних. Математичні моделі та новітні технології управління економічними та технічними системами / за заг. ред. В. О. Тимофєєва, І. В. Чумаченко. Харків: ФОП Мезіна В. В., 2017. С. 261-267. URL: <https://www.researchgate.net/publication/319956215> (дата звернення: 04.02.2019).
19. Костенко О. Б., Буласенко М. В., Зарицький О. В. Автоматизація конвертування просторових об'єктів до стандартів оновлених класифікаторів. Математичні моделі та новітні технології управління економічними та технічними системами: колективна монографія. Харків: ФОП Панов А.М., 2018. С. 295-301. URL: <https://www.researchgate.net/publication/329339630> (дата звернення: 04.02.2019)
20. Приложение D «Digitals book». Введение в Digitals Script. Digitals. Использование в геодезии, картографии и землеустройстве. Д. В. Федоров. URL: <http://www.vinmap.net/book/apd.html> (дата обращения: 12.04.2019).
21. Codd E. F. Missing Information (Applicable and Inapplicable) in Relational Databases. SIGMOD Record. 1986. № 15. С. 53–78.
22. COY ISO 19136:2009 "Обмінний формат геопросторових даних на основі географічної мови розмітки GML (ISO 19136:2007)".
23. Zarytskyi O. V., Kostenko O. B., Bulaienko M. V. Marking of Incomplete Information in the Databases of Topographic Data Using Validation. Information systems and innovative technologies in project and program management / Collective monograph edited by I. Linde, I. Chumachenko, V. Timofeyev. Riga: ISMA, 2019. С. 185-197. URL: <https://www.researchgate.net/publication/336552874> (accessed 10.09.2019).

НАУКОВЕ ВИДАННЯ

**СУЧАСНІ ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ**  
**Матеріали засідань школи-семінару**  
**Випуск 4**

Відповідальний за випуск С.І. Шматков

Комп'ютерне складання та верстання В.Є. Стрілець

План 2019 р., поз. 15

Підп. до друку 09.12.2019. Формат 60x84 1/16. Папір офсетний.

Riso-друк. Гарнітура Таймс. Ум. друк. арк. 3,38.

Наклад 25 прим. Зам. № 09-12-2019. Ціна договірна

---

Друкарня "Технологічний Центр"

Свідоцтво про внесення суб'єкта видавничої справи ДК №4452 від 10.12.2012

Адреса: 61145, м. Харків, вул. Шатилова дача, 4

---