УДК 004.02

**Галенко Н.І., Фісун М.Т.**

# МОДЕЛІ РИЗИКІВ У ПРОЕКТАХ ЗІ СТВОРЕННЯ ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ

*У статті розглянута можливість створення моделі ризику на базі методів імітаційного моделювання, апарату нечіткої логіки і байєсівської мережі довіри. Розроблені моделі проектного ризику. Зроблено висновок про ефективність застосування таких моделей ризиків у задачах ризик-менеджменту.*

***Ключові слова:*** *модель ризику, нечітка логіка, байєсівська мережа довіри, модель проектного ризику.*

*В статье рассмотрена возможность создания модели риска на базе методов иммитационного моделирования, апарата нечеткой логики и байесовской сети доверия. Разработаны модели проектного риска. Сделан вывод об эффективности применения таких моделей рисков в задачах риск-менеджмента.*

***Ключевые слова:*** *модель риска, нечеткая логика, байесовская сеть доверия, модель проектного риска.*

*The possibility of risk models creation on the basis of methods of imitational modeling, apparatus of fuzzy logic and Bayes Believe Network is considered in the article. The risk-model of project risk is developed. The conclusion about effectivety of using of such models in the tasks of risk-management is done.*

***Key words:*** *model is a hyphen, fuzzy logic, bayesian network, model project hyphen.*

## *ВСТУП*

Розробка і підтримка програмного забезпечення (ПЗ) – це складний високотехнологічний процес, який вимагає використання новітніх технологій, сучасного обладнання, висококваліфікованих розробників та ін. Використання таких складових надає конкурентоспроможності створюваному ПЗ, але і додає в проект з розробки ПЗ невизначеності. Невизначеності, якщо ними не керувати, трансформуються в ризики, які в свою чергу впливають на головні показники проекту з розробки ПЗ і, навіть, створюють загрозу для його успішного завершення. Невизначеності і ризики не слід уникати, тому що вони можуть створювати можливості для розвитку проекту та отримання додаткового прибутку, якщо процес ризик-менеджменту організовано раціонально і він відбувається на усіх етапах розробки і підтримки ПЗ.

Дослідження в області управління проектами зі створення програмного забезпечення показали, що в результаті неефективного управління 18 % проектів завершуються невдачею, 53 % потребують додаткових витрат фінансів і часу. В той же час в проектах реалізується тільки 69 % заявленої функціональності [1].

Наведені вище данні доводять актуальність задачі управління ризиками в проектах. Важливо збільшити кількість проектів, які не виходять за рамки бюджету, закінчуються вчасно і реалізують весь заявлений функціонал.

В роботах з ризик-менеджменту виділяють наступні етапи: визначення, аналіз, планування пом’якшення результатів матеріалізації, впровадження плану пом’якшення та моніторинг ризиків.

Аналіз ризиків є важливою складовою управління ризиками в проекті, від вдалого аналізу ризиків залежить якість системи ризиків для проекту в цілому.

В процесі аналізу ризиків виконують якісну оцінку ризиків, ранжування ризиків та кількісний аналіз.

Для отримання вихідних даних в кількісному аналізі використовують методи: опитування, дослідження ймовірностей виникнення ризиків, експертну оцінку.

Проекти з розробки програмного забезпечення більшою мірою уніфіковані [1], тому в процесі життєдіяльності кожного проекту виникають подібні невизначеності й ризики, на які накладається специфіка предметної області та локації виконання. Тому представляється важливою автоматизація процесу аналізу ризиків та створення системи ризиків, шляхом створення їх моделей для певної предметної області та місцезнаходження розробника.

### ОГЛЯД ПУБЛІКАЦІЙ

Управління ризиками є достатньо дискусійною темою і широко розглядається в багатьох літературних джерелах як вітчизняних [3, 4, 6], так і закордонних [1, 2, 9, 12]. Ці публікації призначені для широкого спектру читачів: як для професіоналів з ризик-менеджменту та розробки ПЗ [1, 5, 9, 12, 13], так і для широкого загалу [1, 2, 6, 7].

В даній статті кількісний аналіз ризиків розглядається в контексті управління ризиками [4, 6].

В [4, 5, 7, 9] розглянуті теоретичні основи імітаційного моделювання та створення імітаційних моделей на базі метода Монте-Карло. В [8, 10] розглянуті теоретичні основи апарату нечіткої логіки, в [11, 12, 13,14] наводяться теоретичні основи БМД та приклади їх використання.

Проте всі наведені вище джерела не враховують економічних та локаційних особливостей розробки ПЗ на Україні. Крім того, описані методології не розглядаються в якості інструмента для створення моделей ризиків в проектах з розробки ПЗ.

### ПОСТАНОВКА ЗАДАЧІ

У рамках даної роботи планується розв’язання наступних задач: (1) сформулювати системний підхід до створення моделі ризику в проектах з розробки ПЗ; (2) розробити моделі ризику, використовуючи імітаційне моделювання (метод Монте-Карло), апарат нечіткої логіки та графо-аналітичний метод (Байєсівські мережі довіри (БМД)); (3) визначити можливість та ефективність застосування запропонованих методів на практиці на прикладі дослідження ризику «Недостатня кваліфікація розробника».

### ДОСЛІДЖЕННЯ

**Імітаційне моделювання.** Суть методу Монте-Карло полягає в наступному: потрібно знайти значення $a$ деякої величини, що вивчається. Для цього вибирають таку випадкову величину $X$, математичне очікування якої дорівнює $a$:

$$M(X) = a. \tag{1}$$

Дисперсією випадкової величини $X$ буде число $D(X)$, яке є математичним очікуванням квадрата відхилення випадкової величини $X$ від її середнього значення $M(X)$.

$$\sigma(X) = M\,[(X–M(X))^2]. \tag{2}$$

Після виконання $n$ дослідів, в результаті яких отримають $n$ можливих значень $X$, обчислюють їх середнє значення (3) і приймають $x$ в якості оцінки (наближеного значення) $a^*$ шуканого числа $a$.

$$\bar{x} = \frac{\sum x_i}{n} \tag{3}$$

Теорія цього методу Монте-Карло вказує, як найдоцільніше обрати випадкову величину $X$, як знайти її можливі значення. Зокрема, розробляються способи зменшення дисперсії випадкових величин, що використовуються, внаслідок чого зменшується помилка, яка допускається при заміні шуканого математичного очікування $a$ його оцінкою $a^*$ [7].

Ґрунтуючись на положеннях методу Монте-Карло, сформулюємо алгоритм визначення коєфіцієнта впливу причинної події на кількісну оцінку матеріалізації ризику (R) таким чином:

- задати функції відносної ймовірності $f_i(x)$ для кожної з причин ризику, виходячи з припущення, що структура ризику була отримана на основі з якісного аналізу ризиків;

- задати точність оцінювання впливу складових на кількісну оцінку ризику проекту $\varepsilon$ і визначити кількість ітерацій N. Величина N може бути оцінена, виходячи із співвідношення

$$\varepsilon = \frac{3\sigma}{\sqrt{N}}, \tag{4}$$

де $\sigma$ – стандартне відхилення випадкової величини [7];

- визначити для кожної j-ї реалізацій випадкових величин $\{c_i^j\}, j = \overline{1, N}$ відносної ймовірності (нормально розподілених на інтервалі [0, 1]) для кожної з складових структури ризику значення коефіциента витрат (U) $\{U_i^j\}, j = \overline{1, N}$; цей крок повторюється N разів, знайдені значення зберігаються;
- визначити значення сумарних значень коефіцієнтів витрат;

$$U_{\Sigma}^{j} = \sum_{i=1}^{M} U_i^j \tag{5}$$

- на підставі отриманих значень $U_{\Sigma}^{j}$ можна побудувати гістограму для функції $R(x)$, яка дозволить визначити сумарне значення втрат від ризику та оцінити його вплив на проект [4, 5, 7].

На рис. 1 представлено графічне зображення моделі ризику, створеної на базі методу імітаційного моделювання Монте-Карло.

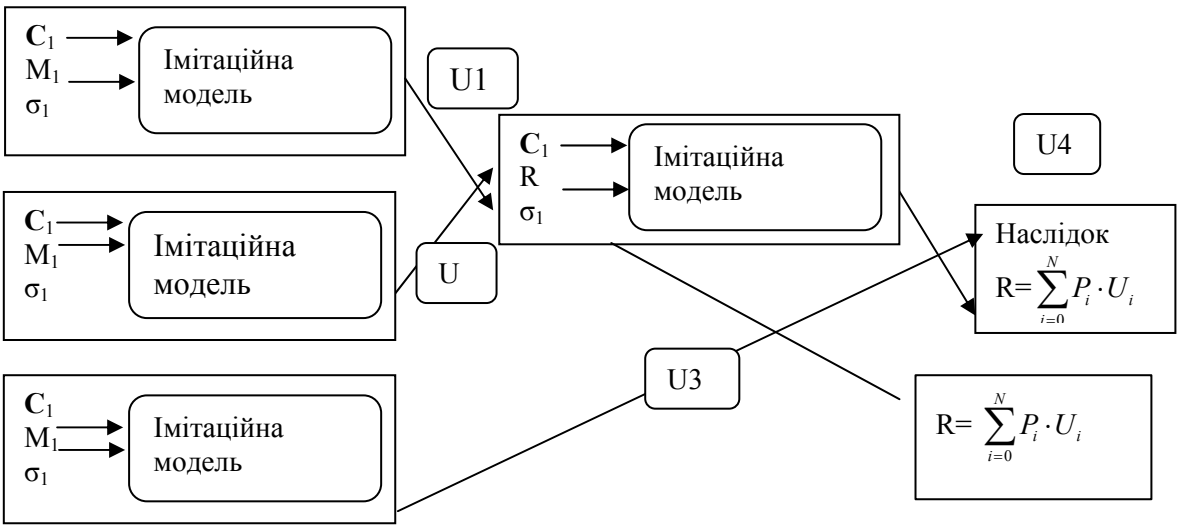



**Рис 1. Графічне зображення моделі ризику, побудованої за методом Монте-Карло:**
**C** – подія-причина, M – математичне очікування, $\sigma_1$ – стандартне відхилення,
R – кількісна оцінка ризику, U – коефіцієнт збільшення впливу причини ризику

**Нечітка модель ризику.** Апарат нечіткої логіки оперує лінгвістичними змінними, яким складно дати точну (об'єктивну) кількісну оцінку. Визначення кількісного значення ризику – завдання, що складно формалізується, значно зручніше оперувати складовими ризиків за допомогою лінгвістичних змінних.

Як інструмент перекладу лінгвістичних змінних на математичну мову для подальшого застосування методу використовуються функції належності $\mu_A(X)$. Функцією належності є математична функція, що задає ступінь або упевненість, з якою елементи деякої множини $X$ належать заданій нечіткій множині $A$ [4].

Далі пропонується використовувати алгоритм Мамдані для створення системи нечіткого виводу і визначення кількісного значення ризику.

Основними етапами нечіткого висновку за алгоритмом Мамдані є:

1. Формування бази правил системи нечіткого висновку.
2. Фазифікация вхідних параметрів.
3. Агрегація.

4. Активізація підумов в нечітких правилах продукції.
5. Дефазифікация.

*Формування бази правил системи нечіткого виводу*. Процес формування бази правил нечіткого виводу є формальним представленням емпіричних знань експерта в тій або іншій проблемній області.

*Фазифікация вхідних параметрів.* Фазифікацією, або введенням нечіткості, називається процес знаходження функції належності нечітких множин на основі звичайних початкових даних. На даному етапі встановлюється відповідність між чисельним значенням вхідної змінної системи нечіткого виводу і значенням функції належності лінгвістичної змінної, яка їй відповідає.

*Агрегування.* Метою даного етапу є визначення ступеня істинності кожного з підвисновків по кожному з правил системи нечіткого виводу.

*Активізація підумов в нечітких правилах продукцій.* Нечіткі підмножини, призначені для кожної вихідної змінної, об'єднуються разом, щоб сформувати одну нечітку підмножину для кожної змінної.

*Дефазифікация.* Отримані результати всіх вихідних змінних на попередніх етапах нечіткого висновку перетворяться в звичайні кількісні значення кожній з вихідних змінних. Дефазифікация нечіткої множини

$$C = \int_{[\underline{x};\overline{x}]} \mu_A(x)/x \qquad (6)$$

за методом центру ваги здійснюється по формулі:

$$c = \frac{\int_{\underline{x}}^{\overline{x}} x \times \mu_A(x)dx}{\int_{\underline{x}}^{\overline{x}} \mu_A(x)dx} \text{ [5]}; \qquad (7)$$

На рис. 2 представлено графічне зображення моделі ризику, створеної з застосуванням апарату нечіткої логіки.

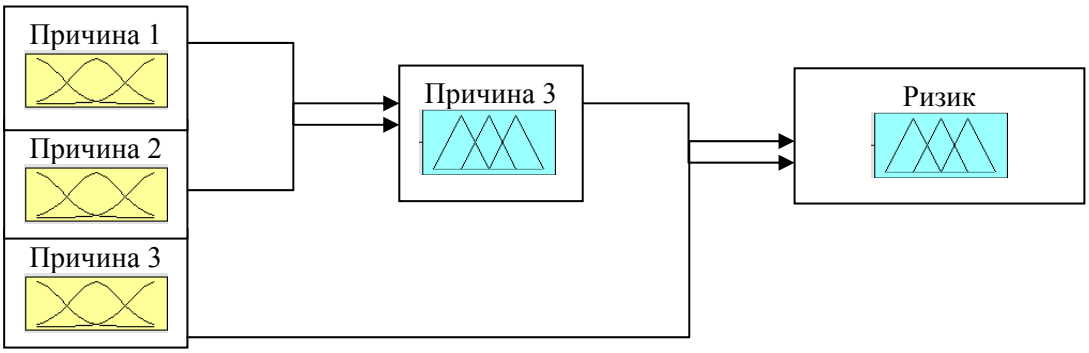



**Рис 2. Графічне зображення моделі ризику, побудованої на базі апарату нечіткої логіки**

**Байєсівські мережі довіри.** Байесівські мережі надають зручний апарат для дослідження ризиків, створення їхньої моделі й кількісної оцінки [11].

Байєсівська мережа (БМ) складається з множини випадкових змінних і направлених зв'язків між ними. Кожна змінна може приймати одне з кінцевої множини взаємовиключаючих значень. Змінні разом із зв'язками створюють орієнтований граф без циклів. Кожній змінній-нащадку $A$ із змінними-предками $B_1,\ldots, B_n$ приписується таблиця умовних ймовірностей $P(A|B_1,\ldots B_n)$ [11].

Поняття умовної ймовірності $P(A|B) = x$ складає основу байєсівського підходу до аналізу невизначеності.

Сумісна ймовірність настання подій $A$ і $B$ визначається формулою повної ймовірності:

$$P(A, B) = P(A \mid B)\cdot P(B) \qquad (8)$$

Маючи в розпорядженні інформацію про залежні змінні (наслідки), можна визначити порівняльні ймовірності вихідних змінних (причин) за допомогою теореми Байєса.

Нехай маємо умовну ймовірність $P(A|B)$ настання деякої події $A$ за умови, що відбулася подія $B$. Теорема Байєса дає рішення для зворотної задачі: яка ймовірність настання більш ранньої події $B$, якщо відомо, що більш пізня подія $A$ відбулася.

Нехай $A_1,\ldots A_n$ – повна група несумісних взаємовиключаючих подій (альтернативних гіпотез). Тоді апостеріорна ймовірність $P(A_j \mid B)$ кожної з подій $A_j$ за умови, що відбулась подія $B$, виражається через апріорну ймовірність $P(A_j)$:

$$P(A_j \mid B) = \frac{P(A_j) \cdot P(B \mid A_j)}{P(B) = \sum_{j=1}^{n} P(A_j) \cdot P(B \mid A_j)} \tag{9}$$

Зворотна ймовірність $P(B \mid A_j)$ називається *правдоподібністю*, а знаменник $P(B)$ в формулі Байєса – *свідоцтвом* [12].

Сумісна ймовірність є найбільш повним статистичним описом даних, за якими спостерігаємо. Сумісний розподіл є функцією багатьох змінних у задачі. В загальному випадку цей опис вимагає задавання ймовірностей всіх припустимих конфігурацій значень всіх змінних, що складно застосовувати в реальних задачах. В байєсівських мережах, за умови, коли маємо додаткову інформацію про ступінь залежності чи незалежності ознак, ця функція факторизується на функції меншої кількості змінних [13]:

$$P(A_1,\ldots A_n) = \prod_j P[A_j \mid pa(A_j)] \tag{10}$$

де $pa(A_j)$ – стан всіх змінних-предків для змінної $A_j$.

Виходячи з раніше перелічених особливостей БМ, а також використовуючи (1), (2) і (3), можна ефективно застосовувати апарат БМ у задачах розробки моделі ризику, кількісної оцінки ймовірності його виникнення та матеріалізації, а також впливу його на інші ризики в системі.

На рис. 3 зображено модель ризику, побудовану на базі Байєсівської мережі довіри.

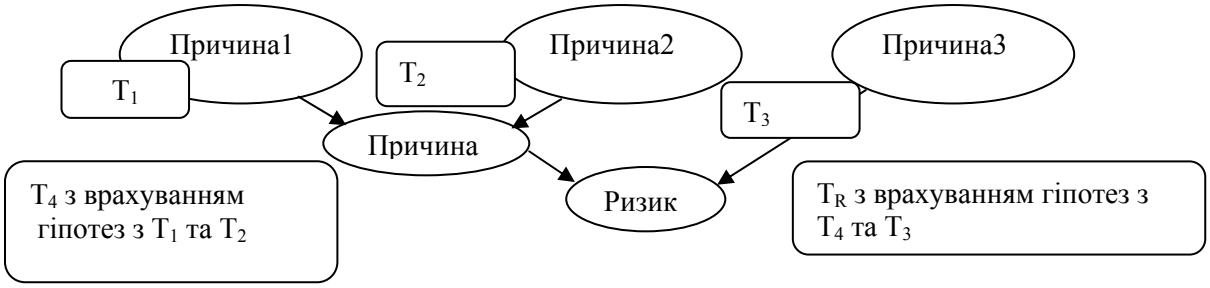



**Рис 3. Графічне зображення моделі ризику, побудованої у вигляді БМД:**
T*i* – таблиця умовної ймовірності

На рис. 4 зображено структуру ризику «Недостатня кваліфікація розробників».

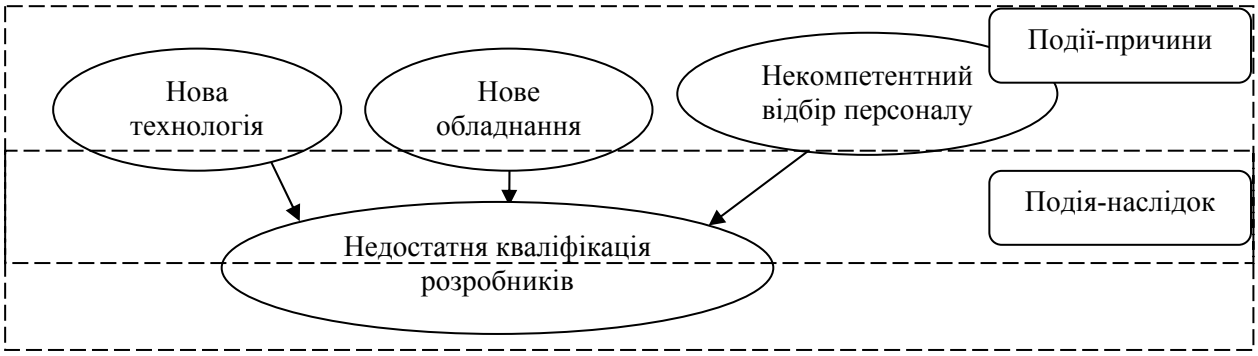
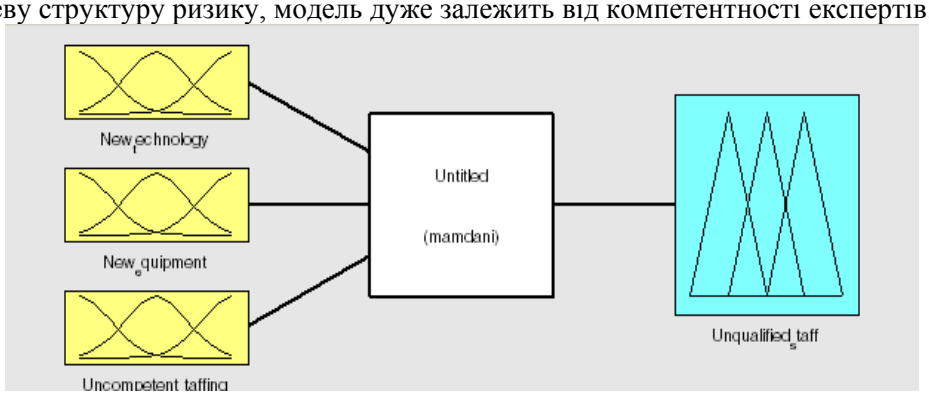


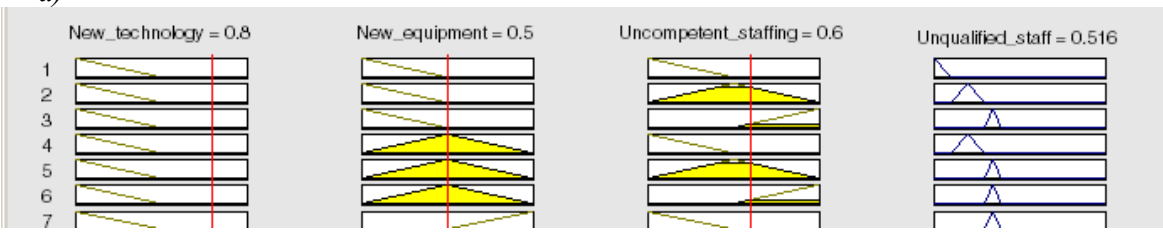

**Рис 4. Структура ризику «Недостатня кваліфікація розробників»**

На рис. 5 зображено модель ризику, реалізовану за допомогою апарату нечіткої логіки. Події-причини в цій моделі виступають у вигляді лінгвістичних вхідних змінних, подію-наслідок чи ризик представлено у вигляді вихідної лінгвістичної змінної. В якості ймовірнісних

складових вхідних і вихідного параметрів моделі використовувались данні з джерела [3], які були отримані в ході багаторічних спостережень за проектами з різних галузей з метою створення емпіричної шкали рівня ризиків. Для створення та компіляції моделі використовувався пакет Matlab. За допомогою такої моделі легко визначити вплив складових ризику. Як бачимо з рис. 5b, при ймовірностях подій-причин: *P*(Нова технологія)=0.8, *P*(Нове обладнання)=0.5, *P*(Некомпетентний відбір персоналу)=0.6, ймовірності матеріалізації ризику «Недостатня компетентність розробників» складе 0.516. Менеджер проекту може дослідити, яким чином впливають зміни ймовірностей подій-причин на ймовірність виникнення ризику, визначити, яке найбільше чи найменше значення може прийняти ризик і визначити подальші кроки стосовно ризик-менеджмента проекту. Проте ця модель має ряд недоліків: кількість вхідних лінгвістичних змінних значною мірою ускладнює базу правил, складно реалізувати багаторівневу структуру ризику, модель дуже залежить від компетентності експертів.




*a)*




*b)*

**Рис. 5. *a)* Модель ризику, реалізована за допомогою апарату нечіткої логіки; *b)* Результат дефазифікації лінгвістичних змінних**

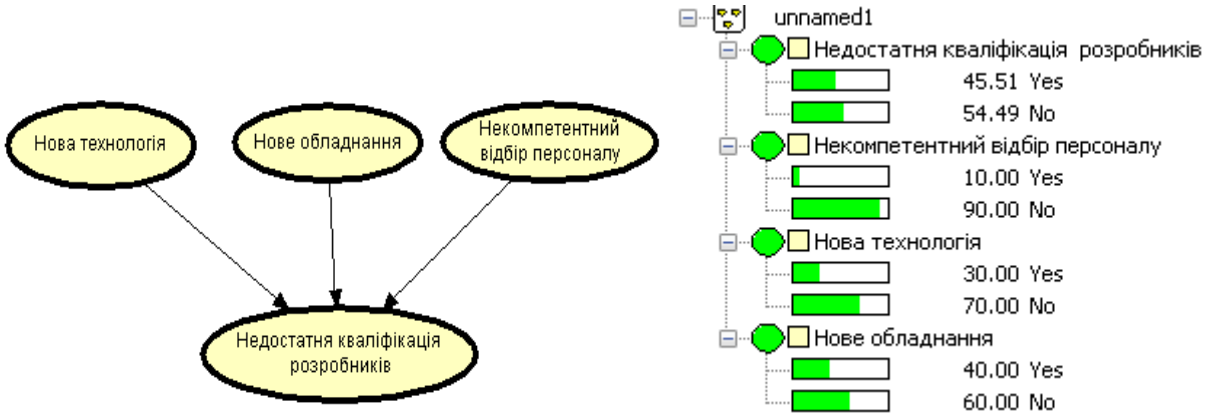
На рис. 6 зображено модель ризику, створену за допомогою БМД. Важливою особливістю для цього методу є те, що він вимагає представлення моделі у вигляді орієнтованого графа, а це дозволяє максимально точно дослідити складові і включити якомога більшу кількість факторів, які впливають на виникнення ризику.

Для кожної складової ризику були створені таблиці умовних ймовірностей. При створенні цих таблиць використовувались данні, що отримані з досвіду виконання і дослідження аналогічних проектів [3].

На рис. 6*b* бачимо, що ймовірність виникнення ризику з врахуванням умовних ймовірностей причин дорівнює 45,5 %.

Модель, що базується на БМД, динамічна, вона дозволяє легко додавати до системи свідоцтва та перераховувати всі ймовірності відповідно до них. Користуючись такою моделлю, можна реалізовувати складні багаторівневі структури, вона дозволяє використовувати ймовірнісні вхідні данні чи такі, що базуються на експертній оцінці.

Незручністю в використанні цього методу є залежність від суб’єктивної оцінки експерта та від якості розробки орієнтованого графа, на якому базується модель.



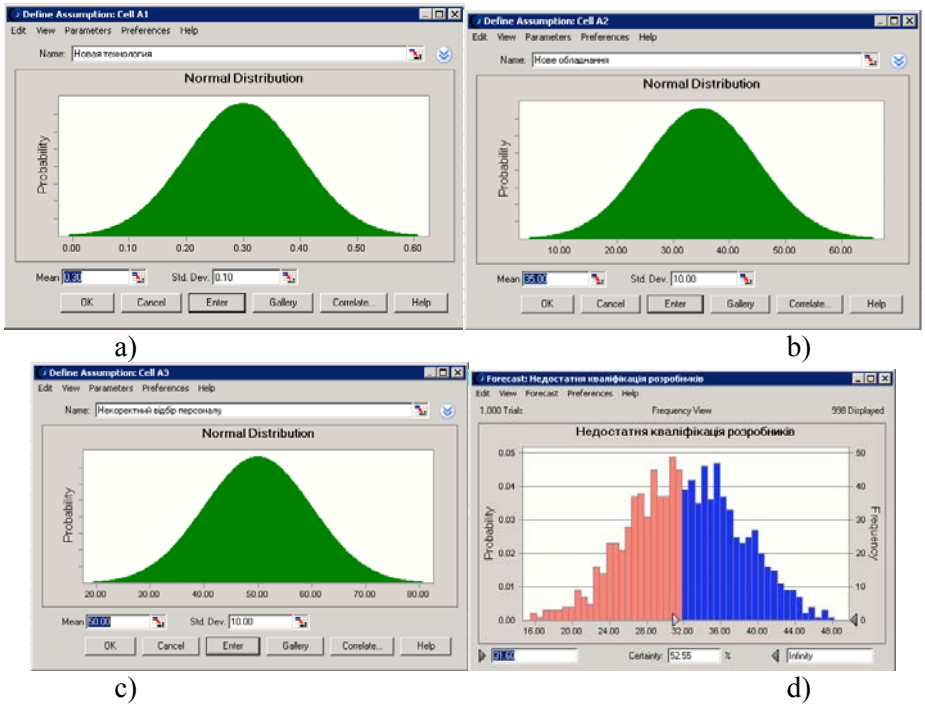

*a)* *b)*

**Рисунок 6.** ***a)*** **БМД для ризику «Недостатня кваліфікація розробників»;**
***b)*** **Ймовірнісні оцінки моделі ризику (розрахунок розповсюдження сум)**

В процесі створення імітаційної моделі ризику для представлення складових ризику використаємо нормальну форму розподілу ймовірностей (рис. 7*a*). Нехай кожна складова впливає на кількісну оцінку ризику, цей вплив виражається у вигляді втрати. Витрати будемо виміряти в умовних одиницях. В реальному проекті це можуть бути витрати часу, грошей чи іншого ресурсу. Сумарну витрату обчислимо за формулою (5). Математичне очікування, дисперсія та ймовірність виникнення кожної подій-причини було визначено, ґрунтуючись на базі проектів, реалізованих в організації [3].

Після виконання симуляції моделі в програмному пакеті CrystalBall, отримаємо результуючу гістограму (рис. 7*b*), користуючись якою можна визначити, які втрати зазнає проект у разі матеріалізації ризику і з якою ймовірністю.

Метод зручно використовувати, коли маємо недостатньо значень вхідних параметрів, отриманих досвідним шляхом. Він дозволяє створювати моделі зі складною структурою. Проте успішність симуляції і достовірність отриманих результатів залежить від обрання відповідної функції розподілу та розрахункової формули для визначення агрегованого впливу. Крім цього, метод імітаційного моделювання передбачає використання великої кількості ітерацій для досягнення припустимого рівня точності.




a) b)

c) d)

**Рисунок 7.** ***a-c)*** **Функції розподілу ймовірностей;** ***d)*** **Результати симуляції**

### ВИСНОВОК

В роботі було розглянуто три методології створення моделі ризику в проектах з розробки програмного забезпечення. Було доведено, що ці методи можуть бути використані для рішення задачі дослідження. За допомогою кожного з методів було створено модель реального проектного ризику та досліджено її.

Кожен з розглянутих методів має свої переваги і недоліки. Рішення про застосування того чи іншого метода приймає менеджер проекту, ґрунтуючись на попередньому досвіді виконання таких проектів і особливостях предметної області.

### ЛІТЕРАТУРА:


1. Guide to the Software Engineering Body of Knowledge – Washington IEEE, 2004. – 324 с.
2. ДеМарко Т. Вальсируя с медведями: управление рисками в проектах по разработке программного обеспечения / ДеМарко Т., Листер Т. – М.: Компания p.m. Office. М., 2005. – 190 с.
3. Максимов В.И. Моделирование риска и рисковых ситуаций: учебное пособие / Максимов В.И., Никонов О.И. – Екатеринбург: ГОУ ВПО УГТУ – УПИ, 2004. – 82 с.
4. Липаев В.В. Анализ и сокращение рисков проектов сложных программных средств / Липаев В.В. – М.: Синтег, 2005. – 208 с.
5. Голенко Д.И. Статистические методы сетевого планирования и управления / Голенко Д.И. – М. : Наука, 1968. – 400 с.
6. Куликова Е.Е. Управление рисками. Инновационный аспект / Куликова Е.Е. – М. : Бератор-паблишинг., 2008. – 224 с.
7. Соболь И.М. Метод Монте-Карло / Соболь И.М. – Москва: Наука, 1972. – 68 с.
8. Борисов А.Н. Принятие решений на основе нечетких моделей: примеры использования / Борисов А.Н., Крумберг О.А., Федоров И.П. – Рига: Зинатне, 1990. – 184 с.
9. Lawrence I. Goldman. Crystal ball professional introductory tutorial / Lawrence I. Goldman. – Denver: Decisioneering, Inc, 2002. – 68 с.
10. Леоненков А. Нечеткое моделирование в среде MATLAB и fuzzyTech / Леоненков А. – СПб.: БХВ-Петербург, 2003. – 736 с.
11. Терехов С.А. Введение в байесовы сети / Терехов С.А. – М.: МИФИ, 2003. – 188 с.
12. David Heckerman. / David Heckerman. – [A Tutorial on Learning With Bayesian Networks]. – : Microsoft Research. Technical Report, 1995. – 124 с
13. Jensen F. V. Bayesian networks basics / Jensen F. V. : Denmark, 1996.
14. Система для построения Байесовых систем. – http://www.hugin.com/.


Рецензенти: д.т.н., проф. Коваленко І.І.
д.т.н., проф. Казарєзов А.Я.

© Галенко Н.І., Фісун М.Т., 2009

*Стаття надійшла до редколегії 18.02.09*