ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ, СИСТЕМНИЙ АНАЛІЗ ТА КЕРУВАННЯ

DOI: 10.20535/1810-0546.2018.2.128989
УДК 519.766.4, 519.25

П.І. Бідюк*, Н.В. Кузнєцова
КПІ ім. Ігоря Сікорського, Київ, Україна

ЙМОВІРНІСНО-СТАТИСТИЧНИЙ МЕТОД ОЦІНЮВАННЯ РИЗИКУ ФІНАНСОВИХ ВТРАТ

Проблематика. Фінансові ризики, що виникають у різних сферах діяльності людини, пов'язані з великою кількістю невизначеностей, нечіткістю, неповнотою та неточністю даних. Для прогнозування фінансових втрат компаній необхідно опрацьовувати дані такого типу, тому актуальною є задача розробки нового методу, який дасть можливість здійснювати фільтрацію вхідних даних і прогнозувати фінансові втрати.
Мета дослідження. Розробити метод оцінювання ризику можливих фінансових втрат і запропонувати комплексну ймовірнісно-статистичну модель на його основі.
Методика реалізації. Комплексно застосовано: оптимальний фільтр для попередньої обробки даних та їх підготовки до побудови моделей, регресійну модель для формального опису і прогнозування умовної дисперсії та ймовірнісну модель у формі байєсівської мережі для оцінювання ймовірності можливих втрат.
Результати дослідження. Запропонований метод апробовано на задачі оцінювання фінансового ринкового ризику проведення операцій із фондовим ринком. Статистичні дані, які було використано, описують еволюцію цін на акції для відомих компаній. У результаті виконання обчислювальних експериментів було встановлено, що якість короткострокових прогнозів щодо волатильності поліпшується від 7,1 до 50 % завдяки оптимальній фільтрації даних. Застосування побудованої моделі у формі байєсівської мережі надало можливість подальшого удосконалення ймовірнісного оцінювання можливих фінансових втрат при здійсненні торговельних операцій на фондовому ринку акцій.
Висновки. Оцінювання ризику фінансових втрат є актуальною задачею, що може розв'язуватись різними методами. Високоефективним виявився запропонований ймовірнісно-статистичний метод для ймовірнісного оцінювання можливих фінансових втрат при здійсненні торговельних операцій на фондовому ринку акцій, тому в подальшому перспективним буде розширення його застосування на інші види фінансових ризиків.
Ключові слова: фінансові ризики; фільтр Калмана; мережа Байєса; регресійна модель; оцінювання фінансових втрат.

Вступ

Високої актуальності набули сьогодні задачі математичного моделювання й оцінювання фінансових ризиків. До ризиків такого типу належать, зокрема, ринкові та ризики неповернення кредитів з боку фізичних і юридичних осіб, інвестиційні та страхові ризики, операційні ризики, притаманні фінансовим організаціям будь-яких типів, та багато інших [1–4]. Для математичного моделювання ризиків фінансових втрат застосовують різні методи залежно від типу ризику, історичних даних стосовно розвитку фінансових процесів і методів їх обробки, наявності необхідних інструментальних засобів для моделювання, оцінювання і прогнозування фінансових процесів та досвіду використання тих чи інших методів і підходів. Широке застосування для розв'язання задач моделювання і оцінювання ризиків фінансових втрат знаходять ймовірнісно-статистичні методи і моделі, які ґрунтуються на комбінуванні різних підходів [3–5], що дає можливість використовувати переваги ідеологічно різних методів та отримувати оцінки можливих втрат у формі точкових оцінок та ймовірностей настання ризикових ситуацій.

Для оцінювання кредитоспроможності фізичних осіб застосовують лінійну і нелінійну регресію (моделі логіт і пробіт), дерева рішень, нейронні та байєсівські мережі, нейронечіткі моделі, методи кластерного аналізу, моделі на основі нечіткої логіки та інші. Високоякісні результати аналізу кредитоспроможності отримані, зокрема, за допомогою нелінійної регресії, дерев рішень, кластерного аналізу, нейронних і байєсівських мереж [4–7]. Сучасні дослідження концентруються на використанні методів інтелектуального аналізу даних, нелінійних регресійних моделей, кластерного аналізу, а також на комбінуванні вказаних методів у одній системі аналізу статистичних даних та експертних оцінок [8]. На основі таких моделей та сучасних інформаційних технологій будуються скорингові системи, які надають фінансовим підприємствам можливість виконання високоякіс-

*corresponding author: pbidyuke_00@ukr.net

ного аналізу даних, оцінювання потенційно можливих втрат, прогнозування ризиків та прийняття обґрунтованих управлінських рішень.

Слід зазначити, що наявні статистичні дані потребують, як правило, застосування методів попередньої обробки, спрямованих на заповнення пропусків вимірів, нормування, фільтрації даних з метою зменшення впливу похибок вимірів і випадкових зовнішніх збурень, бутстреп-аналізу тощо. Математичні моделі, які будуються на основі статистичних даних, потребують адаптації у процесі надходження нових вимірів і застосування нетрадиційних оптимізаційних процедур для оцінювання їх параметрів. Розв’язання цих задач можливе за допомогою сучасних систем підтримки прийняття рішень (СППР). Відомі розробки і практичне застосування СППР, які є надійним інструментом генерування об’єктивних, обґрунтованих альтернативних рішень і вибору кращих із них за допомогою множини відповідних критеріїв якості [7–10]. Методика розробки та реалізації СППР із використанням сучасних інформаційних технологій реалізована сьогодні на високому рівні.

Розробка, запропонована у цьому дослідженні, призначена для використання як складової в СППР, вона відповідає основним принципам системного аналізу, на основі яких проектуються сучасні СППР.

Постановка задачі

Метою дослідження є розв’язання таких задач: побудова комбінованої математичної моделі на основі застосування оптимального фільтра, регресійної моделі та байєсівського підходу; використання побудованої моделі для оцінювання ринкового фінансового ризику на основі статистичних даних стосовно біржової еволюції цін акцій відомих компаній; виконання порівняльного аналізу із застосуванням інших ймовірнісно-статистичних моделей до розв’язання задачі аналізу ринкового фінансового ризику.

Комбінована математична модель і метод оцінювання ризику фінансових втрат

Пропонується побудова комбінованої моделі на основі оптимального фільтра, регресійної моделі та байєсівської мережі. На цій моделі ґрунтується метод оцінювання ризику можливих фінансових втрат. Структура ймовірнісно-статистичного методу оцінювання ризику фінансових втрат подана на рис. 1.

Метод оцінювання фінансових втрат реалізується у вигляді кроків, поданих нижче.

Крок 1. Вхідні дані, які характеризують можливу появу фінансових втрат, можуть бути різного типу і різних форматів. Статистичні дані характеризуються, як правило, неточністю, неповнотою, нечіткістю інформації, наявністю похибок вимірів і впливом випадкових зовнішніх збурень. Тому виникає необхідність здійснити фільтрацію даних. Для цього можна застосувати цифрові або оптимальні фільтри, наприклад фільтр Калмана. Застосування цифрових фільтрів передбачає знання смуги частот, у якій знаходяться корисні складові даних, і смуги, де знаходяться згадані шкідливі випадкові впливи, які фільтр не повинен пропускати. На основі цієї інформації проектується цифровий фільтр у вигляді лінійних рівнянь авторегресії (АР) або авторегресії з ковзним середнім (АРКС), які будуть використовуватись для попередньої обробки даних згідно зі схемою, поданою на рис. 1.

Фільтр Калмана (ФК) у загальному випадку призначений для розв’язання таких задач: зменшення впливу випадкових зовнішніх збурень і похибок вимірів на статистичні дані (фільтрація), короткострокове прогнозування вектора стану на 1–3 кроки, заповнення невеликої кількості пропусків вимірів, оцінювання компонент вектора стану, що не вимірюються приладами, оцінювання деяких параметрів математичної моделі у просторі станів, що використовується фільтром для розв’язування вказаних задач [11]. Таким чином, основною умовою коректного функціонування алгоритму фільтрації є побудова математичної моделі даних

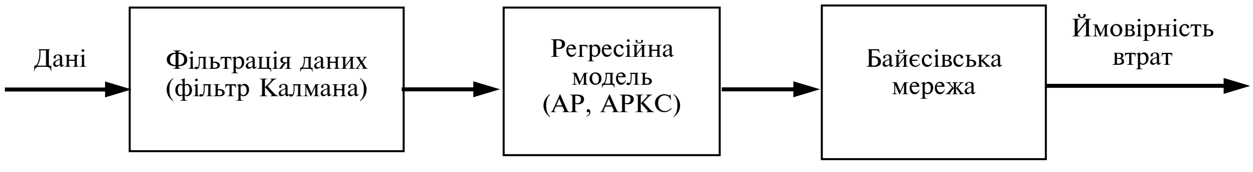

Рис. 1. Структурна схема ймовірнісно-статистичного методу оцінювання ризику фінансових втрат

та її представлення у просторі станів. Зазначимо, що для попередньої обробки даних можна використовувати спрощені моделі, наприклад векторну авторегресію на основі рівнянь низького порядку. Розв'язання такої задачі не потребує великих зусиль, але дає можливість скористатись перевагами застосування оптимального фільтра. Розглянемо детальніше рівняння оптимального фільтра.

Статистичні дані, що описують фінансовий ризик за різними показниками (ліквідності, обслуговування боргів тощо), можуть бути подані у табличному вигляді, а змінна або вектор змінних стану, що характеризують можливі фінансові втрати в часі, можуть бути описані рівнянням динаміки у просторі станів:

$$\mathbf{x}(k+1) = \mathbf{F}\mathbf{x}(k) + \mathbf{w}(k), \tag{1}$$

$$\mathbf{z}(k) = \mathbf{H}\mathbf{x}(k) + \mathbf{v}(k), \tag{2}$$

де $\mathbf{x}(k)$ — вектор стану досліджуваного процесу; $\mathbf{F}$ — перехідна матриця станів; $\mathbf{w}(k)$ — векторний процес зовнішніх випадкових збурень з коваріаційною матрицею $E[\mathbf{w}(k)\mathbf{w}^T(j)] = \mathbf{Q}(k)\delta_{kj}$; $\mathbf{z}(k)$ — вектор вимірів; $\mathbf{H}$ — матриця коефіцієнтів вимірів; $\mathbf{v}(k)$ — вектор похибок вимірів з коваріаційною матрицею $E[\mathbf{v}(k)\mathbf{v}^T(j)] = \mathbf{R}(k)\delta_{kj}$. У класичній постановці задачі випадкові процеси $\mathbf{w}(k)$ і $\mathbf{v}(k)$ не корельовані між собою та вектором стану та мають нульове середнє і сталі скінченні коваріації. Система матричних рівнянь (1), (2) — це модель даних у просторі станів, яка для спрощення аналізу не враховує можливих керуючих впливів. Початковим станом системи $\mathbf{x}_0$ будемо вважати випадкові змінні з відомими статистиками:

$$E[\mathbf{x}_0] = \bar{\mathbf{x}}_0;\ E[\mathbf{x}_0\mathbf{x}_0^T] = \mathbf{M};\ E[\mathbf{w}(k)\mathbf{x}_0^T] = 0,\ \forall k.$$

Оптимальна оцінка стану $\hat{\mathbf{x}}(k)$ повинна обчислюватися як найкраща за мінімумом середнього значення суми квадратів похибок оцінок вектора стану. Іншими словами, оцінка повинна бути такою, щоб

$$E[(\hat{\mathbf{x}}(k) - \mathbf{x}(k))^T(\hat{\mathbf{x}}(k) - \mathbf{x}(k))] = \min_{\mathbf{K}}, \tag{3}$$

де $\mathbf{x}(k)$ — точне значення вектора стану, яке може бути обчислене за допомогою детермінованої складової математичної моделі процесу; $\mathbf{K}$ — оптимальний матричний коефіцієнт фільтра, який необхідно обчислити в результаті розв'язання оптимізаційної задачі. Основне рівняння фільтрації має такий вигляд:

$$\hat{\mathbf{x}}(k) = \mathbf{F}(k)\hat{\mathbf{x}}(k-1) + \\ + \mathbf{K}(k)[\mathbf{z}(k) - \mathbf{H}(k)\mathbf{F}(k)\hat{\mathbf{x}}(k-1)]. \tag{4}$$

На основі рівняння (1) можна записати функцію прогнозування на один крок як умовне математичне сподівання вектора стану:

$$\hat{\mathbf{x}}(k+1 \mid k) = E_k[\mathbf{x}(k+1)] = \mathbf{F}\mathbf{x}(k), \tag{5}$$

де $\hat{\mathbf{x}}(k+1)$ — прогноз на один крок на основі інформації на момент k включно. Функцією (5) можна скористатись для обчислення прогнозу на довільну кількість кроків. Так, прогноз на два кроки має вигляд

$$\hat{\mathbf{x}}(k+2) = \mathbf{F}\hat{\mathbf{x}}(k+1) = \mathbf{F}\cdot\mathbf{F}\mathbf{x}(k) = \mathbf{F}^2\mathbf{x}(k),$$

і на довільну кількість кроків s:

$$\hat{\mathbf{x}}(k+s) = \mathbf{F}^s\mathbf{x}(k).$$

Очевидно, що дисперсія похибки прогнозу буде зростати пропорційно кількості кроків s. Так, похибка прогнозу на один і два кроки становить

$$\mathbf{e}_f(1) = \mathbf{x}(k+1) - \hat{\mathbf{x}}(k+1) = \\ = \mathbf{F}\mathbf{x}(k) + \mathbf{w}(k) - \mathbf{F}\mathbf{x}(k) = \mathbf{w}(k),$$

і

$$\mathbf{e}_f(2) = \mathbf{x}(k+2) - \hat{\mathbf{x}}(k+2) = \mathbf{F}\mathbf{x}(k+1) + \mathbf{w}(k+1) - \\ - \mathbf{F}[\mathbf{F}\mathbf{x}(k)] = \mathbf{F}\mathbf{x}(k+1) + \mathbf{w}(k+1) - \mathbf{F}[\mathbf{x}(k+1) - \\ - \mathbf{w}(k)] = \mathbf{w}(k+1) + \mathbf{w}(k).$$

Тобто дисперсію оцінок прогнозів на s кроків можна обчислити так:

$$\mathrm{Var}[\mathbf{e}_f(s)] = E\{[\mathbf{w}(k+s-1) + \\ + \mathbf{w}(k+s-2) + \ldots + \mathbf{w}(k)] \times [\mathbf{w}^T(k+s-1) + \\ + \mathbf{w}^T(k+s-2) + \ldots + \mathbf{w}^T(k)]\} = s\mathbf{Q}_w.$$

Оцінювання компонент вектора стану, які неможливо виміряти за допомогою приладів (або зареєструвати вручну), виконується таким чином. У випадку наявності невимірюваних компонент розмірність вектора вимірів $\mathbf{z}(k)$ є меншою вимірності вектора стану $\mathbf{x}(k)$, тобто $\dim[\mathbf{z}] < \dim[\mathbf{x}]$.

Оптимальний коефіцієнт фільтра обчислюється за рівнянням

$$\mathbf{K}(k) = \mathbf{P}'(k)\,\mathbf{H}^T[\mathbf{H}\mathbf{P}'(k)\,\mathbf{H}^T + \mathbf{R}]^{-1},$$

де $\dim[\mathbf{P}'(k)] = [n \times n]$; $\dim[\mathbf{H}^T] = [n \times r]$ за означенням; n — розмірність вектора стану і $\dim[\mathbf{H}\mathbf{P}'(k)\,\mathbf{H}^T + \mathbf{R}]^{-1} = [r \times r]$. Таким чином, $\dim[\mathbf{K}(k)] = [n \times r]$. Вектор нев'язок $\gamma(k)$ у рівнянні оцінювання (4)

$$\gamma(k) = \mathbf{z}(k) - \mathbf{H}\,\mathbf{F}\,\hat{\mathbf{x}}(k)$$

має розмірність $[r \times 1]$, а розмірність добутку $\dim[\mathbf{K}(k)\,\nu(k)] = [n \times 1]$. Наприклад, якщо $\dim[\mathbf{x}] = [3 \times 1]$, а $\dim[\mathbf{z}] = [2 \times 1]$, то добуток $\mathbf{K}(k)\gamma(k)$ має вигляд

$$\mathbf{K}(k)\,\gamma(k) = \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \\ K_{31} & K_{32} \end{bmatrix} \begin{bmatrix} \gamma_1 \\ \gamma_2 \end{bmatrix},$$

і оцінка вектора стану визначається за рівняннями:

$$\hat{x}_1(k) = \hat{x}_1(k, k-1) + K_{11}\gamma_1 + K_{12}\gamma_2,$$

$$\hat{x}_2(k) = \hat{x}_2(k, k-1) + K_{21}\gamma_1 + K_{22}\gamma_2,$$

$$\hat{x}_3(k) = \hat{x}_3(k, k-1) + K_{31}\gamma_1 + K_{32}\gamma_2.$$

Тобто невимірювана третя компонента вектора стану може бути оцінена, якщо матриці $\mathbf{P}'(k), \mathbf{P}(k)$ і $\mathbf{K}(k)$ мають ненульові відповідні елементи. Таким чином, фільтр дає можливість обчислити оцінки змінних, які є оптимальними у сенсі мінімуму функціоналу (3) і знаходяться під значно меншим впливом випадкових збурень стану і похибок вимірів.

Крок 2. На цьому кроці будується регресійна модель певної структури, яка здійснює прогнозування показників фінансового ризику втрат (фінансової стійкості) на наступний момент часу. Як правило, існує можливість побудови деякої множини моделей-кандидатів у вигляді авторегресії (векторної авторегресії) або авторегресії з ковзним середнім із різними параметрами для авторегресії та ковзного середнього, які будуть здійснювати прогнозування значень фінансових змінних на 1, 2, 3, ... кроки (періоди дискретизації) вперед. Цим самим забезпечується оцінювання прогнозних значень можливих втрат. При цьому для зручності обчислення прогнозів доцільно будувати і використовувати функції прогнозування на основі розв'язків рівнянь [11]. Так, для моделі авторегресії першого порядку функція прогнозування має вигляд

$$\hat{y}(k+s) = E_k[y(k+s)] =$$
$$= a_0\left(\sum_{i=0}^{S-1} a_1^i\right) + a_1^S y(k) = a_0\sum_{i=0}^{S-1} a_1^i + a_1^S y(k),$$

де $\hat{y}(k+s)$ — оцінка прогнозу основної (цільової) змінної на s кроків; E_k — оператор умовного математичного сподівання стосовно моменту часу k; a_0, a_1 — параметри регресійної моделі. Для моделі АРКС(2,1) побудована функція прогнозування на три кроки має такий вигляд:

$$\hat{y}(k+3) = E_k[y(k+3)] =$$
$$= a_0 + a_1 E_k[y(k+2)] + a_2 E_k[y(k+1)] =$$
$$= a_0(1 + a_1 + a_1^2 + a_2) + (a_1^3 + 2a_1 a_2)y(k) +$$
$$+ (a_1^2 a_2 + a_2^2)y(k-1) + \beta_1(a_1^2 + a_2)\,\varepsilon(k).$$

Табулювання функцій прогнозування дає можливість використовувати їх повторно.

Похибка прогнозування $f_k(s)$ для моделей типу АР та АРКС визначається випадковою складовою $\varepsilon(k)$ і оцінюється так:

$$f_k(s) = \varepsilon(k+s) + a_1\varepsilon(k+s-1) +$$
$$+ a_1^2\varepsilon(k+s-2) + \ldots + a_1^{S-1}\varepsilon(k+1). \quad (6)$$

З урахуванням того, що $E[f_k(s)] = 0$, оцінка прогнозу, яка обчислюється за виразом (6), є незміщеною, дисперсія похибки прогнозування визначається за виразом

$$\mathrm{Var}[f_k(s)] = \sigma^2[1 + a_1^2 + a_1^4 + a_1^6 + \ldots + a_1^{2(S-1)}],$$

тобто дисперсія є функцією кількості кроків прогнозування s. Асимптотичне значення дисперсії похибки прогнозу для стаціонарного процесу є константою:

$$\lim_{S\to\infty} \mathrm{Var}\,[f_k(s)] = \frac{\sigma^2}{1-a_1^2},$$

де a_1^2 — знаменник геометричної прогресії.

У випадку наявності пояснювальних змінних лінійна регресійна модель може мати таку структуру:

$$y(k) = a_0 + \sum_{i=1}^{p} a_i y(k-i) + b_1 x_1(k-d_1) + \ldots + $$
$$+ a_m x_m(k-d_m) + \varepsilon(k), \qquad (7)$$

де x_i, $i=1,\ldots,m$, — пояснювальні змінні (предиктори); d_i, $i=1,\ldots,m$, — дискретний час запізнення для пояснювальних змінних. Зазначимо, що незалежні змінні, які входять у модель із фактичним часом запізнення d, називають провідними індикаторами. Очевидно, що структура моделі (7) може містити складові, які описують тренд, сезонні ефекти, ковзне середнє та нелінійні ефекти. Оцінювання структури моделі здійснюється на основі кореляційного аналізу наявних даних, застосування статистичних тестів на інтегрованість, гетероскедастичність, наявність нелінійності та експертних оцінок. Як правило, для подальшого аналізу використовують кілька моделей-кандидатів.

Після вибору (оцінювання) структури моделі виконується оцінювання числових значень параметрів знайдених моделей-кандидатів. Найбільш поширеними методами оцінювання параметрів моделі є звичайний метод найменших квадратів (МНК) та його модифікації, метод максимальної правдоподібності (ММП), метод допоміжної (інструментальної) змінної (МДЗ), нелінійний метод найменших квадратів (НМНК), метод Монте-Карло для марковських ланцюгів (МКМЛ) та їх рекурсивні версії [7].

Далі аналізується якість моделі, тобто виконується перевірка оцінених кандидатів на адекватність процесу. Діагностика побудованих моделей може здійснюватися такими способами: візуальне дослідження графіка похибок; аналіз характеристик похибок; перевірка, чи корелюють похибки моделі між собою (критерій Дарбіна–Уотсона), чи коефіцієнт множинної детермінації R^2 прямує до 1 ($R^2 \to 1$, коли дисперсії вимірів змінної та оцінок цієї ж змінної, отриманих за моделлю, збігаються) і чи є мінімальною сума квадратів похибок для вибраної моделі, тобто

$$\sum_{k=1}^{N} e^2(k) = \sum_{k=1}^{N} [\hat{y}(k) - y(k)]^2 \to \min_{\hat{\theta}}.$$

Допоміжними критеріями вибору кращої моделі з оцінених кандидатів є інформаційний критерій Акайке:

$$AIC = N \ln\left(\sum_{k=1}^{N} e^2(k)\right) + 2n,$$

та критерій Байєса–Шварца:

$$BSC = N \ln\left(\sum_{k=1}^{N} e^2(k)\right) + n\ln(N),$$

де $n = p + q + 1$ — кількість параметрів моделі, які оцінюються за допомогою статистичних даних (p — кількість параметрів авторегресійної частини моделі; q — кількість параметрів ковзного середнього; “1” з’являється тоді, коли оцінюється зміщення (або перетин, тобто a_0)).

При використанні методів рекурсивного оцінювання для лінійної моделі 2-3-го порядку оцінки параметрів повинні збігатися до усталених значень після 30–40 ітерацій алгоритму оцінювання. Якщо кількість ітерацій набагато перевищує вказані значення, то це свідчить про те, що процес може бути нестаціонарним та/або нелінійним. Узагальнена методика побудови регресійних моделей та оцінювання прогнозів на їх основі докладно розглянута в [11].

К р о к 3. На третьому кроці здійснюється ймовірнісне оцінювання втрат на основі адаптивної динамічної мережі Байєса (МБ). МБ — ймовірнісно-статистична модель у формі спрямованого графа, вершинами якого є вибрані змінні, а дуги вказують на наявність причинно-наслідкових зв’язків між змінними. Кожній батьківській вершині ставиться у відповідність таблиця безумовних ймовірностей, а кожній залежній вершині — таблиця умовних ймовірностей. Числові значення в таблицях (параметри моделі) вказують на ймовірність набуття змінною конкретних значень. Сьогодні МБ — це потужний інструмент ймовірнісно-статистичного моделювання, який ґрунтується на спеціальних числових методах, які забезпечують оцінювання альтернативних структур і параметрів моделей, а також формування точного або наближеного ймовірнісного висновку — остаточного результату застосування моделі [12]. Побудова і застосування МБ надає такі переваги: можливість створювати моделі високої розмірності (десятки і сотні змінних); використання статистичних даних і експертних оцінок в одній моделі; можливість формування висновку в “прямому” та “зворотному” напрямках; побудова складних комбінованих ймовірнісно-статистичних моделей, які включають у себе МБ, регресійні та інші типи моделей. Сфери застосування МБ такі: розпізнавання ситуацій; прогнозування динаміки розвитку процесів; побудова діагностичних систем у медицині, економіці, техніці; підтримка прийняття управлінських рішень тощо [12–14].

Адаптація МБ полягає в тому, що структура і параметри моделі оцінюються повторно з надходженням нових даних. При цьому часовий інтервал, через який виконується повторне оцінювання, визначається емпірично для конкретної динаміки досліджуваного процесу і статистичних даних конкретного типу.

У комбінованій моделі для оцінювання втрат пропонується застосовувати динамічну мережу Байєса [12]. Для побудови цієї моделі використовуємо виходи вибраних на попередньому кроці кращих моделей авторегресії (або інших структур), які подаються на вхід стаціонарного шару ймовірнісної моделі.

Динамічні мережі Байєса – це розширення статичних байєсівських мереж для моделювання спільних розподілів ймовірностей на множині випадкових змінних $Z_1, Z_2, \ldots$. Як правило, змінні розбиваються на трійки $Z_t = (U_t, X_t, Y_t)$, що позначають множину змінних вхідного, прихованого та вихідного шарів моделі у просторі станів [12]. Надалі будемо розглядати лише моделі стохастичних процесів із дискретним часом, тобто індекс t буде зростати з появою кожного нового спостереження.

За означенням, динамічна мережа Байєса є парою $(B_1, B_{\rightarrow})$, де B_1 – МБ, що визначає апріорну ймовірність $P(Z_1)$, а $B_{\rightarrow}$ – двошарова МБ, що визначає $P(Z_t | Z_{t-1})$ за допомогою спрямованого ациклічного графа таким чином:

$$P(Z_t | Z_{t-1}) = \prod_{i=1}^{N} P(Z_t^i | Pa(Z_t^i)),$$

де Z_t^i – i-й вузол у момент часу t, що може бути компонентою X_t, Y_t або U_t, а $Pa(Z_t^i)$ – батьки Z_t^i на графі. Вузли першого шару двошарової МБ не мають жодних параметрів, що з ними асоціюються, але кожен вузол другого шару має зв’язаний із ним розподіл умовної ймовірності, що визначають $P(Z_t^i | Pa(Z_t^i))$ для всіх $t > 1$.

Батьки вершини $Pa(Z_t^i)$ можуть бути або ж у тому самому, або у попередньому часовому шарі. Втім це суто загальноприйняте спрощення: не існує суворих математичних обмежень стосовно того, що батьківські вершини знаходяться не далі, ніж у сусідньому шарі, і не можуть знаходитися, скажімо, через один шар. Дуги між шарами спрямовуються зліва направо, що позначає напрямок протікання часу. Якщо існує дуга від Z_{t-1}^i до Z_t^i, то ця вершина називається сталою. Дуги в межах одного шару є умовними, оскільки в цілому динамічна мережа є спрямованим ациклічним графом. У межах одного часового шару, як виняток, дозволяється використання неспрямованих дуг, що позначають сильні кореляції між змінними або деякі обмеження.

Оскільки на попередньому кроці пропонованого методу було отримано авторегресійну модель (або авторегресію з ковзним середнім), то перший (статичний) шар динамічної мережі Байєса (рис. 2) буде складатись зі змінних, що входять у рівняння авторегресії:

$$Z(t) = a_0 + a_1 U_1(t) + a_2 U_2(t) + \ldots + a_p U_p(t) + \\ + \beta_1 \varepsilon_1(t) + \beta_2 \varepsilon_2(t) + \ldots + \beta_q \varepsilon_q(t),$$

де p – порядок авторегресії, отриманої на другому кроці, а q – порядок ковзного середнього. Якщо порядок ковзного середнього становить 1, то модель вироджується у звичайну авторегресію порядку p.

На виході будемо отримувати ймовірність набуття вихідними змінними значення в часовому просторі. Для оцінювання фінансових ризиків це може бути ймовірність настання певного ризику або ймовірність фінансових втрат у часі (наприклад, через 30 днів). Мережа Байєса також може бути використана для короткострокового прогнозування ймовірності втрат у певному діапазоні та певному інтервалі значень на момент часу t. Метод може бути узагальнений на різні типи можливих розподілів.

Ілюстративний приклад. Для ілюстрації застосування запропонованого методу оцінювання втрат розглянемо задачу оцінювання ринкового ризику, який виникає при виконанні операцій з фінансовими активами. У цьому випадку вхідними даними комбінованої моделі є ціни акцій різних компаній, зокрема Microsoft, Google і Cisco.

Часові ряди даних, які характеризують ціни акцій при закритті торгів на фондовій біржі, взяті з історичної бази даних фінансової системи Yahoo Finance. Запропонований метод реалізовано в СППР, схему якої подано на рис. 3. Послідовність виконання обчислювальних операцій стосовно обробки даних, побудови математичних моделей, прогнозування волатильності (умовного стандартного відхилення) та оцінювання можливого ризику втрат відбувається за ієрархічною схемою зверху-вниз. Результати короткострокового прогнозування волатильності прибутку за акціями Microsoft подані в таблиці.

У таблиці подані статистичні показники точності короткострокового прогнозування волатильності на навчальній та перевірочній вибірках для досліджених моделей авторегресії з умовною гетероскедастичністю (АРУГ), узагальненої АРУГ (УАРУГ), експоненційної УАРУГ (ЕУАРУГ) і моделі стохастичної волатильності (МСВ). Значення середньої абсолютної похибки (САП) і середньої абсолютної похибки у відсотках (САПП), отримані на навчальній вибірці для моделей УАРУГ НВ, ЕУАРУГ НВ, МСВ НВ, показують менші значення похибок прогнозування, ніж оцінки прогнозів моделей на перевірочній вибірці (УАРУГ ПВ, ЕУАРУГ

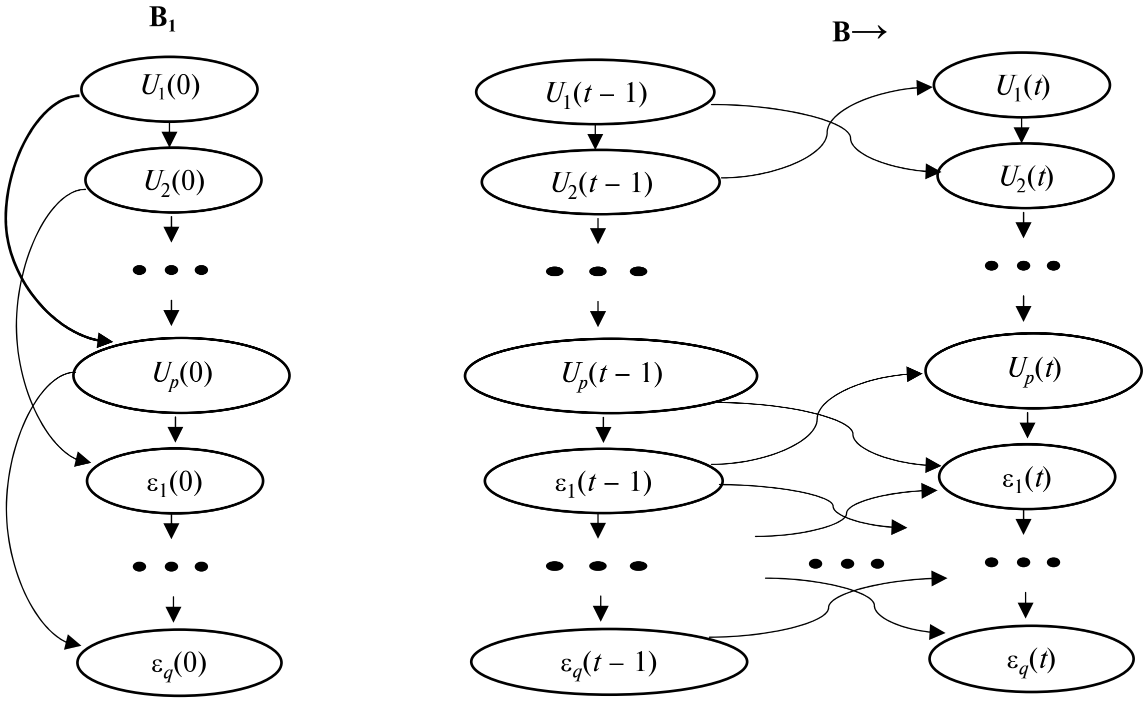

Рис. 2. Структура динамічної мережі Байєса

Таблиця. Прогнозування волатильності прибутку акцій Microsoft

Модель	САП без ФК	САПП без ФК	САПП з ФК
АРУГ НВ	0,000359	9454,4	9188,7
УАРУГ НВ	0,0000791	45,960	36,270
ЕУАРУГ НВ	0,4353	4,9400	3,7530
МСВ НВ	0,64	7,5000	4,9023
АРУГ ПВ	0,00041	5123,5	2494,4
УАРУГ ПВ	0,00013	51,993	28,396
ЕУАРУГ ПВ	0,5053	5,93	4,0710
МСВ ПВ	0,84	10,90	6,9590

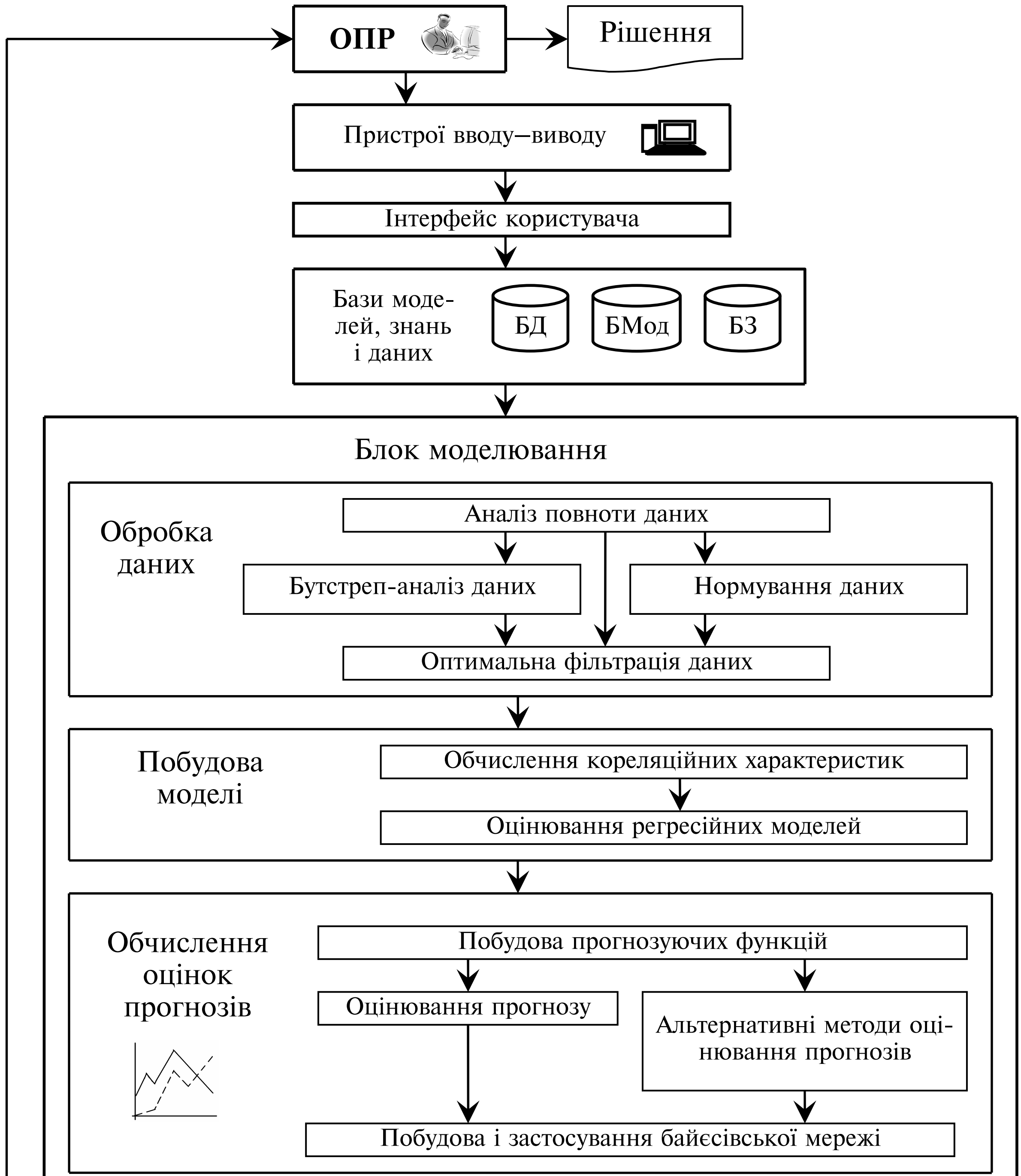

Рис. 3. Реалізація комбінованої моделі в архітектурі системи підтримки прийняття рішень

ПВ та МСВ ПВ), що й очікувалось. Зазначимо, що модель АРУГ демонструє велику міру неточності оцінок прогнозів як на навчальній вибірці, так і поза її межами, що пояснюється простотою її структури, яка не відображає реальної взаємодії змінних.

На третьому кроці здійснювалось ймовірнісне оцінювання втрат. Для цього побудована динамічна мережа Байєса на рис. 4, за якою спрогнозовано ймовірність фінансових втрат на наступні кроки.

Фінансові втрати за акціями в цілому за день оцінювались за формулою

$$Losses(k) =$$

$$= (Forecast\_Close(k) - Price(k)\_Close) * Volume(k).$$

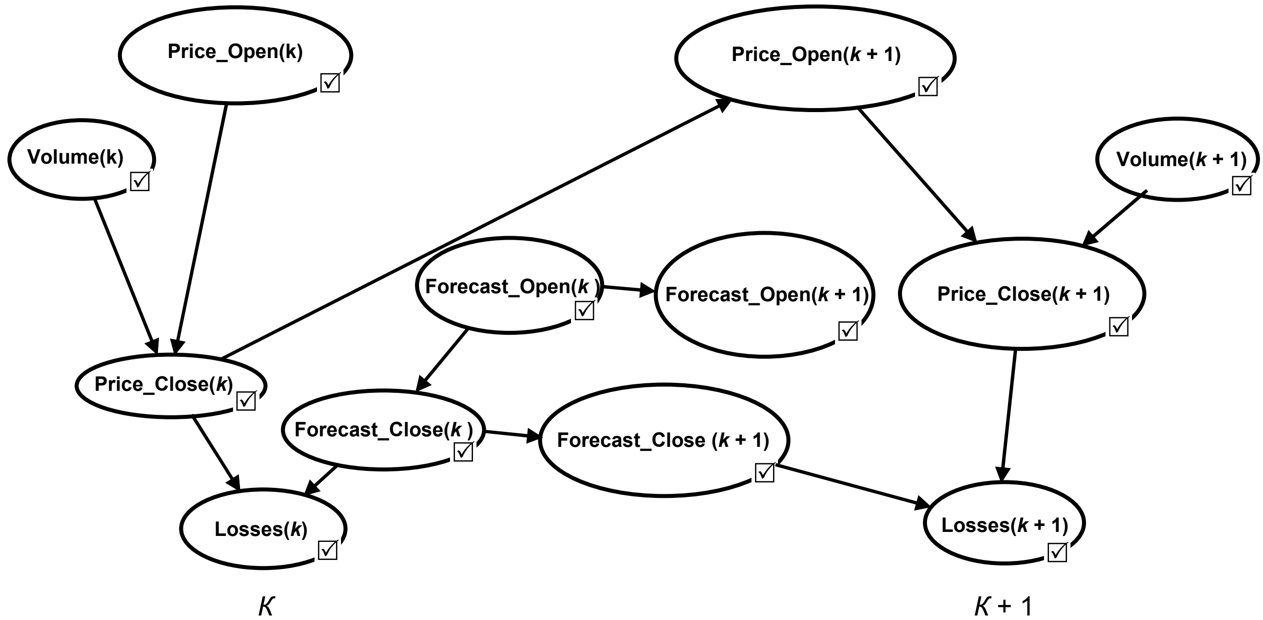

Рис. 4. Динамічна мережа Байєса для ілюстративного прикладу

Застосування фільтра Калмана (ФК) для попередньої фільтрації даних дало можливість підвищити якість оцінок прогнозів волатильності (без моделі АРУГ) від 7,1 до 50 %. Подібні результати моделювання і прогнозування волатильності отримані для цін акцій компаній Google і Cisco.

Оцінки прогнозів волатильності використані для обчислення величини можливих втрат за методикою Value-at-Risk (VaR). Оскільки методика VaR дає значення можливих втрат у вибраному довірчому інтервалі, то уточнення оцінки ймовірності настання втрат виконано за допомогою мережі Байєса, яка побудована на таких змінних: значення оцінок можливих втрат, обчислені на попередніх трьох інтервалах (періодах дискретизації даних); прогнози значень волатильності і фактичні втрати на попередніх десяти інтервалах. Ймовірності можливих втрат, отримані за допомогою байєсівської мережі, практично повністю збіглися з фактичними ймовірностями втрат, обчисленими частотним методом на основі наявних історичних даних. Характерні значення обчислених ймовірностей знаходяться в діапазоні 0,01–0,13. Таким чином, застосування запропонованої комплексної моделі дало можливість підвищити якість оцінок прогнозів волатильності (на використаних статистичних даних) і уточнити ймовірність настання ризикових ситуацій.

Висновки

Запропоновано метод оцінювання ризику можливих фінансових втрат і комплексну ймовірнісно-статистичну модель на його основі. Метод ґрунтується на комплексному застосуванні оптимального фільтра для підготовки статистичних даних до побудови математичної моделі волатильності (умовного стандартного відхилення або моделі для іншої змінної чи параметра), регресійної моделі для формального опису і прогнозування умовної дисперсії та ймовірнісної моделі у формі байєсівської мережі для оцінювання ймовірності можливих втрат. Виконано апробацію комплексної моделі на прикладі оцінювання ринкового фінансового ризику при виконанні операцій на фондовій біржі з використанням статистичних даних стосовно еволюції цін для відомих компаній. У результаті виконання обчислювальних експериментів із використанням фактичних статистичних даних встановлено, що якість оцінок прогнозів волатильності покращилась у середньому в 1,5 разу завдяки оптимальній фільтрації даних. Використання моделі у формі мережі Байєса дало можливість уточнити ймовірність настання можливих фінансових втрат при виконанні торговельних операцій з акціями компаній. Очевидно, що на першому етапі можна фільтрувати будь-які дані у формі часових рядів, а оцінювання ринкового фінансового ризику виконано для прикладу.

У подальших дослідженнях доцільно продовжити застосування запропонованого комплексного методу оцінювання втрат до множини альтернативних вибірок даних з метою аналізу інших типів фінансових ризиків. Розроблену комп'ютерну СППР доцільно розширити математичними моделями інших типів для опису фінансових ризиків та методами генерування альтернативних управлінських рішень стосовно менеджменту ризиків.

References

[1] M. Neil *et al.*, "Using Bayesian networks to model expected and unexpected operational losses", *Risk Analysis*, vol. 25, no. 4, pp. 963–972, 2005. doi: 10.1111/j.1539-6924.2005.00641.x

[2] *Handbook of Credit Scoring*, E. Mays, Ed. Chicago: Glenlake Publishing Company, Ltd., 2001.

[3] O.M. Trofymchuk *et al.*, "Decision support system for implementing systemic approach to forecasting", *Int. J. Comp. Technol.*, vol. 14, no. 5, pp. 5769–5778, 2015.

[4] N. Kuznietsova and P. Bidyuk, "Modeling financial risks in telecommunication field", *Naukovi Visti NTUU KPI*, no. 5, pp. 51–58, 2017. doi: 0.20535/1810-0546.2017.5.110338

[5] A.J. McNeil *et al.*, *Quantitative Risk Management*. Princeton: Princeton University Press, 2005, 538 p.

[6] P.I. Bidyuk and O.Ye. Matros, "Models for clients credit risk estimation", *Cybernetics and Computations*, vol. 153, pp. 87–95, 2008.

[7] W.R. Gilks *et al.*, *Markov Chain Monte Carlo in Practice*. New York: Chapman & Hall/CRC, 2000, 486 p.

[8] N.V. Kuznietsova, "Information technologies for clients' database analysis and behaviour forecasting", in *Selected Papers of the XVII Int. Sci. Practical Conf. Inform. Technol. Security (ITS 2017)*, 2017, vol. 2067, pp. 56–62 [Online]. Available: http://ceur-ws.org/Vol-2067

[9] F. Burstein and C.W. Holsapple, *Handbook of Decision Support Systems*. Berlin: Springer-Verlag, 2008, 908 p.

[10] C.W. Hollsapple and A.B. Winston, *Decision Support Systems*. Saint Paul: West Publishing Company, 1996, 860 p.

[11] P.I. Bidyuk *et al.*, *Methods of Forecasting*. Lugansk: Alma Mater, 2008, 608 p.

[12] K. Murphy. (2007). *A Brief Introduction to Graphical Models and Bayesian Networks* [Online]. Available: http://www.cs. ubc. ca/~murphyk/Bayes/bayes.html

[13] U. Kjærulff, "dHugin: A computational system for dynamic time-sliced Bayesian networks", *Int. J. Forecasting*, vol. 11, no. 1, pp. 89–111, 1995. doi: 10.1016/0169-2070(94)02003-8

[14] W.M. Bolstad, *Understanding Computational Bayesian Statistics*. Hoboken: John Wiley & Sons, Ltd, 2010, 334 p.

П.И. Бидюк, Н.В. Кузнецова

ВЕРОЯТНОСТНО-СТАТИСТИЧЕСКИЙ МЕТОД ОЦЕНКИ РИСКА ФИНАНСОВЫХ ПОТЕРЬ

Проблематика. Финансовые риски, возникающие в различных сферах деятельности человека, связаны с большим количеством неопределенностей, нечеткостью, неполнотой и неточностью данных. Для прогнозирования финансовых потерь компаний необходимо обрабатывать даже такие данные, поэтому актуальной является задача разработки нового метода, который будет осуществлять фильтрацию входных данных и прогнозировать финансовые потери.

Цель исследования. Разработать новый метод оценки риска возможных финансовых потерь и предложить комплексную вероятностно-статистическую модель на его основе.

Методы реализации. Комплексно применены: оптимальный фильтр для предварительной обработки данных и подготовки к построению моделей, регрессионная модель для формального описания и прогнозирования условной дисперсии и вероятностная модель в виде байесовской сети для оценки вероятности возможных потерь.

Результаты исследования. Предложенная модель была применена для оценки финансового рыночного риска проведения операций с фондовым рынком. Статистические данные, которые были использованы, описывают эволюцию цен на акции для известных компаний. В результате выполнения вычислительных экспериментов было установлено, что качество краткосрочных прогнозов волатильность улучшается от 7,1 до 50 % благодаря оптимальной фильтрации данных. Применение построенной модели в форме байесовской сети позволило дальнейшее совершенствование вероятностной оценки возможных финансовых потерь при осуществлении торговых операций на фондовом рынке акций.

Выводы. Оценка риска финансовых потерь является актуальной задачей, которая может решаться различными методами. Эффективным оказался предложенный вероятностно-статистический метод для вероятностной оценки возможных финансовых потерь при осуществлении торговых операций на фондовом рынке акций, поэтому в дальнейшем перспективным будет расширение его применения на другие виды финансовых рисков.

Ключевые слова: финансовые риски; фильтр Калмана; сеть Байеса; регрессионная модель; оценки финансовых потерь.

P.I. Biduyk, N.V. Kuznietsova

PROBABILISTIC-STATISTICAL METHOD FOR RISK ASSESSMENT OF FINANCIAL LOSSES

Background. Financial risks various fields of human activities may experience are associated with a large number of uncertainties, fuzziness, incompleteness and inaccuracy of data. To predict financial losses of companies, even such data need to be processed, so the task of developing a new method, which will filter incoming data and predict financial losses, is quite relevant.

Objective. The aim of the paper is to develop a new method for assessing the risk of potential financial losses and propose a comprehensive probabilistic-statistical model on its basis.

Methods. The following comprehensive methods were applied: optimal filter, for data pre-processing and preparation for model construction, regression model for formal description and prediction of conditional variance and probabilistic model in the form of Bayesian network for estimation of probability for possible losses.

Results. Proposed model was used to assess financial market risk of transactions with the stock market. The used statistical data describes evolution of stock prices for well-known companies. As a result of computational experiments it was found that the quality of short-term forecasts of volatility improves by an average from 7.1 to 50 % due to optimal data filtering. Application of the model constructed in the form of Bayesian network provided an opportunity for further improvement of probabilistic estimates for possible financial losses in the course of trading transactions at the stock market.

Conclusions. The risk assessment of financial losses is an urgent task that can be solved by different methods. The proposed probabilistic statistical method for the probabilistic estimation of possible financial losses in the course of trading operations in the stock market was effective, therefore, in the future it will be expanded to cover its application to other types of financial risks.

Keywords: financial risks; Kalman filter; Bayesian network; regression model; assessment of financial losses.

Рекомендована Радою
Навчально-наукового комплексу
“Інститут прикладного системного
аналізу” КПІ ім. Ігоря Сікорського

Надійшла до редакції
5 березня 2018 року

Прийнята до публікації
29 березня 2018 року