

**ПОБУДОВА ДИНАМІЧНИХ МЕРЕЖ БАЙЄСА**

*Анотація. Запропонована методика побудови динамічних мереж Байєса на основі статистичних даних, яка складається з двох етапів: побудова статичної структури мережі та динамічної структури мережі, що включає визначення зв'язків між двома сусідніми інтервалами часу. Отриману структуру використовують для формування висновку на кожному часовому інтервалі. Наведено приклад застосування методики до фактичних даних. Ключові слова. Динамічні байєсовські мережі, методологія розробки, статистичні данні.*

**Вступ**

Байєсові мережі – перспективний ймовірнісний інструментарій для моделювання складних ієрархічних процесів (статичних і динамічних) з невизначеностями довільного характеру. Байєсова мережа складається з множини випадкових змінних (вузлів графа) і спрямованих зв'язків між змінними, які разом утворюють орієнтований граф. Кожній змінній ставиться у відповідність таблиця умовних ймовірностей, яка характеризує ймовірність прийняття вузлом (змінною) того чи іншого значення за умови, що зв'язані з ним вузли також приймають конкретні значення.

Мережі Байєса (МБ) ефективно використовуються для розв'язання досить складних задач у різних областях, таких як медична і технічна діагностика, обробка зображень та відеосигналів в різних галузях науки і техніки, аналіз і прогнозування фінансово-економічних процесів. Ймовірнісні моделі у вигляді спрямованих ациклічних графів надають широкі можливості стосовно включення в них відносно великої дискретних і безперервних змінних та формування висновку стосовно будь-якої вибраної змінної.

**1.1. Використання мереж Байєса у задачах моделювання**

Мережі Байєса можна використовувати для моделювання складних процесів різної природи – в техніці, біології, медицині, промислових технологіях, економіці, фінансах і т. ін. Так у роботі [1] розглядається застосування МБ для моделювання поведінки робота, а точніше вивчення і аналізу машиною можливостей, що надає оточу-

юче середовище – еффорданс (від англ. affordance – те, що навколишній світ дозволяє чи надає можливість здійснити індивідууму чи машині).

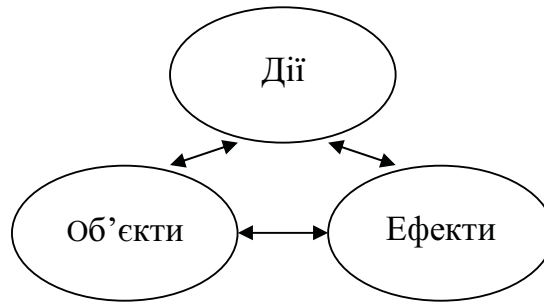


Рисунок 1 – Еффорданс, як відношення між діями (Д), об'єктами (О) та ефектами (Е), може використовуватися для вибору дії, об'єкта чи передбачення результатів дії на об'єкт

Таблиця 1

Входи і виходи для прикладу з роботом

Входи	Виходи	Функція
(О, Д)	Е	Передбачення ефекту
(О, Е)	Д	Розпізнавання дії і планування
(Д, Е)	О	Розпізнавання об'єкту і вибір

Запропонована структура мережі для даного простого прикладу:

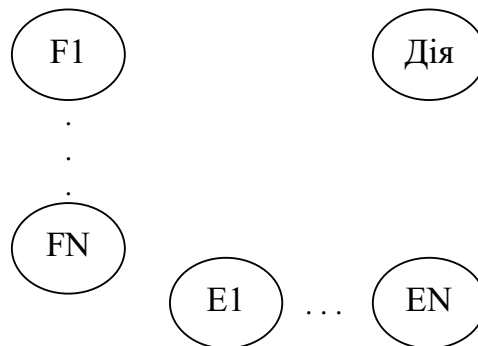


Рисунок 2 – Модель Байєсівської мережі для моделювання еффордансу. Вузли представляють дії Д, властивості об'єктів F(1)...F (n) та ефекти отримані в результаті дій E(1)...E(m)

У розглянутій в роботі [1] моделі робот наділений певними навичками, він здатний сприймати близькі об'єкти і вимірювати їх основні параметри, а також взаємодіяти з оточуючим середовищем через певний набір дій. Набір вузлів (змінних) для цієї ситуації,  $X = \{A, Fr, Fo, E\}$ , має чотири типи складових (дискретних випадкових

змінних):  $A = \{a_i\}$  – дія робота;  $Fr = \{Fr(1), \dots, Fr(nr)\}$  – характеристики робота (наприклад, можливі положення руки);  $Fo = \{Fo(1), \dots, Fo(no)\}$  – властивості об'єкта  $O$ ;  $E = \{E(1), \dots, E(ne)\}$  – ефекти, що спостерігаються в результаті виконання роботом певних дій.

В роботі наведено результати експерименту, виконаного за умов, наведених у таблиці 2.

Таблиця 2

Умови виконання експерименту з роботом

Змінна	Опис	Значення
A	Дія	Схопити, вдарити, доторкнутися
C	Колір	Зелений1, зелений2, жовтий, блакитний
Sh	Форма	М'яч, куб
S	Розмір	Малий, середній, великий
V	Швидкість об'єкта	Мала, середня, велика
HV	Швидкість руки робота	Мала, велика
Di	Відстань від руки до об'єкта	Мала, середня, велика
Ct	Довжина контакту	Короткий, довгий

За результатами експерименту із застосуванням методу максимальної правдоподібності побудовано таку мережу:

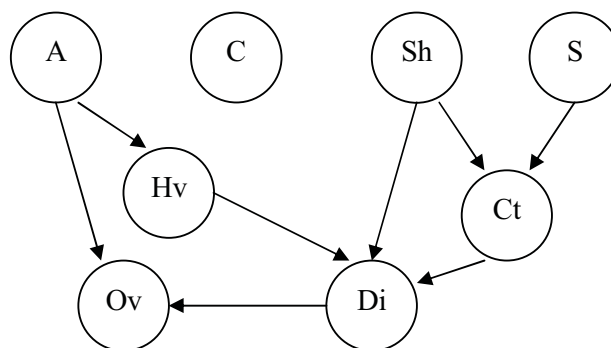


Рисунок 3 – Модель Байєсової мережі

Представлена модель дозволяє описати правдоподібну модель знання про взаємодію робота з оточуючим середовищем, що, в свою чергу, може бути використано при розв'язанні задач опису поведінки робота методами штучного інтелекту.

## 1.2. Використання мереж Байєса у задачах прогнозування

У роботі [2] представлена розробка, спрямована на покращення зручності введення тексту до мобільного телефону. Байєсівська мережа використана для автоматичного підбору символів на основі попередньо введеного тексту і натиснутої кнопки. Запропоновано методику, яку автори назвали ВАРТІ (Bayesian Predictive Text Input).

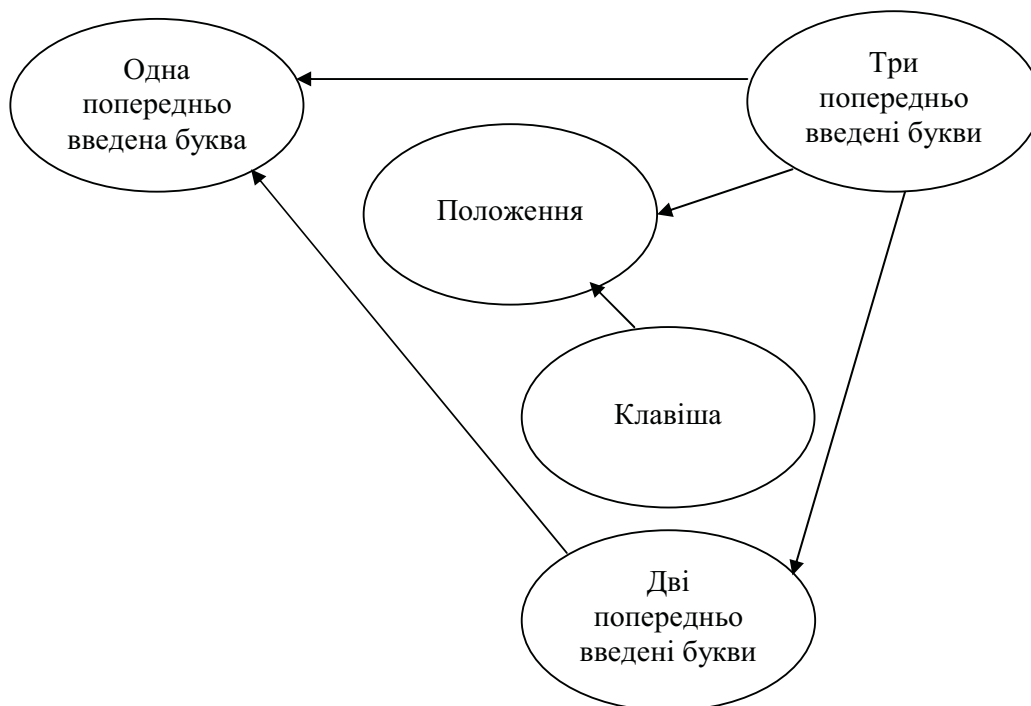


Рисунок 4 – Мережа Байєса для прогнозування букви

На рис. 4 показано побудовану мережу, що враховує три попередньо введені літери та натиснуту клавішу. Вузол „Клавіша” символізує натиснуту клавішу і приймає значення від 1 до 9. Вузол „Положення” визначає порядок літери на клавіші, що обирається і приймає значення від 1 до 3.

Отримані авторами дослідження результати свідчать про те, що використання МБ дозволяє значно пришвидшити введення тексту. В середньому кількість натискання клавіш зменшується на 37.4%. У 91.2% випадків ВАРТІ правильно прогнозує букву. Перевагою даного методу, у порівнянні з іншими, є те, що використання запропонованої методики не потребує зберігання чітко визначеного словника.

## 1.3. Використання мереж Байєса у задачах класифікації

Мережа Байєса може бути використана для розв’язання задач класифікації. Для цього необхідно вибрати вузол, що представляє со-

бою значення результату класифікації, а інші вузли – це параметри, за якими відбувається класифікація. Для кожного стану обраного вузла обраховується умовна вірогідність, як наведено у формулі:

$$P(x_1 = X_1, \dots, x_n = X_n) = \prod_{i=1}^n p(x_i = x_i / \pi_i - \Pi_i)$$

У роботі [3] пропонується структура МБ, яка б найкраще відповідала потребам задач класифікації – часткова мережа Байєса (Partial Bayesian Network). Процедура побудови такої мережі складається з трьох кроків. На першому кроці перевіряють до якої групи відноситься вузол  $x_i \in Z - \{x_c\}$  – батьків чи нащадків вузла  $x_c$  – тобто вузла-класифікатора. Якщо вузол  $x_i$  додається як батьківський для  $x_c$ , то ймовірність мережі змінюється на:

$$\delta_p = \frac{g(c, \pi_c \cup \{i\})}{g(c, \pi_c)} .$$

Якщо ж  $x_i$  – нащадок  $x_c$ , то ймовірність розпізнавання змінюється на:

$$\delta_c = \frac{g(i, \pi_i \cup \{c\})}{g(i, \pi_i)} .$$

Перевіряючи умови  $\delta_p > \delta_c$ , вузол  $x_i$  додається до батьківської множини вузлів  $x_c$ , або до множини його нащадків. Якщо ж виконується умова  $\max(\delta_p, \delta_c) < 1$ , то зв'язок не утворюється.

На другому кроці додаються батьки, що вибираються з множини батьківських та незалежних вузлів для вузлів з множини нащадків. На третьому кроці визначаються зв'язки між вузлами з множини нащадків.

В роботі наведені результати експериментів для двох множин тестових даних, що свідчать про досить високу ефективність запропонованого методу. На першій, простішій множині, результати всіх методів приблизно однакові, але на ускладненій множині даних метод на сонові МБ показав найкращі результати.

#### 1.4. Типи мереж Байєса та їх характеристики

Існує декілька різних типів мереж Байєса, які відрізняються за типами змінних. Виділяють такі основні типи мереж Байєса: дискретні МБ, динамічні МБ, неперервні МБ та гібридні МБ [4].

Дискретні МБ – це мережі, у яких змінні вузлів – дискретні величини. Дискретні МБ мають такі властивості: (1) кожна вершина представляє собою подію, яка описується випадковою величиною, що може мати декілька станів; (2) всі вершини, пов'язані з "батьківським" вузлом, визначаються таблицею умовних ймовірностей або функцією умовних ймовірностей; (3) для вершин без "батьків" ймовірності їх станів є безумовними (маргінальними).

Динамічні МБ – мережі, у яких значення вузлів змінюється з часом. Динамічні МБ ідеально підходять для моделювання змінних у часі процесів. Їх перевага полягає у тому, що вони використовують табличне представлення умовних ймовірностей, яке полегшує представлення різних нелінійних явищ. Найпростіший тип динамічної МБ – це прихована модель Маркова, у якої в кожному шарі є один дискретний прихований вузол і один дискретний або неперервний вузол, що спостерігається.

Безперервні МБ – ті, в яких змінні вузлів мережі є безперервними величинами. У багатьох випадках події можуть приймати будь-які значення стани із деякого діапазону (області визначення). Тобто змінна  $X$  буде безперервною випадковою величиною, простором можливих станів якої буде весь діапазон допустимих її значень,  $X = \{x | a \leq x \leq b\}$ , що містить нескінченну безліч точок. При цьому вже не можна говорити про ймовірності окремого стану, оскільки при їх нескінченно великій кількості вага кожного з них буде наближатись до нуля. Тому розподіл ймовірностей для безперервних випадкових величин визначається інакше ніж у дискретному випадку і для їх опису використовуються функції розподілу ймовірностей та щільності розподілу ймовірностей. Безперервні МБ використовують для моделювання стохастичних процесів у просторі станів з безперервним часом.

Гібридні МБ – мережі, що містять вузли з дискретними та безперервними змінними. При використанні таких МБ існують такі обмеження:

1. Дискретні змінні не можуть мати безперервних батьків;

2. Безперервні змінні повинні мати нормальний закон розподілу, умовний на значеннях батьків;

3. Розподіл безперервної змінної  $X$  з дискретними батьками  $Y$  та безперервними батьками  $Z$  є нормальним розподілом  $P(X|Y=y, Z=z) = N(\mu_x(\mu_y, \mu_z), \sqrt{\sigma_x(\sqrt{\sigma_y})})$ , де  $\mu_x(\mu_y, \mu_z)$  – математичні сподівання;  $\sigma_x, \sigma_y$  – дисперсії;  $\sqrt{\sigma_x}, \sqrt{\sigma_y}$  – середньоквадратичні відхилення. Значення  $\mu_x$  лінійно залежить від безперервних батьків, а  $\sigma_x$  взагалі не залежить від безперервних батьків. Однак, обидва параметри ( $\mu_x$  і  $\sigma_x$ ) залежать від дискретних батьків. Це обмеження гарантує можливість формування точного виводу.

### 1.5. Методика побудови динамічної мережі Байєса

Динамічні мережі Байєса — це спосіб розширити мережі Байєса для моделювання ймовірнісних розподілів за напів-нескінченну кількість випадкових величин  $X_1, X_2, \dots$ . Ми розглядаємо випадкові процеси лише з дискретним часом, тому ми збільшуємо індекс  $t$  на одиницю кожного разу, коли надходить нове спостереження. Зазначимо, що термін "динамічний" означає, що ми моделюємо динамічну систему, а не те, що мережа змінюється з часом.

Динамічна мережа Байєса визначається як пара,  $B_0, B_1$ , де  $B_0$  – структура статичної мережі Байєса, що визначає апріорну ймовірність  $P(X_1)$ , а  $B_1$  – перехідна (транзитивна) мережа, що складається з двох часових зрізів мережі Байєса, які визначають спільним розподілом  $P(X_t, X_{t-1})$ .

При моделюванні динаміки процесу еволюцію його стану розглядають у послідовні моменти часу. При цьому структура мережі залишається, як правило, незмінною.

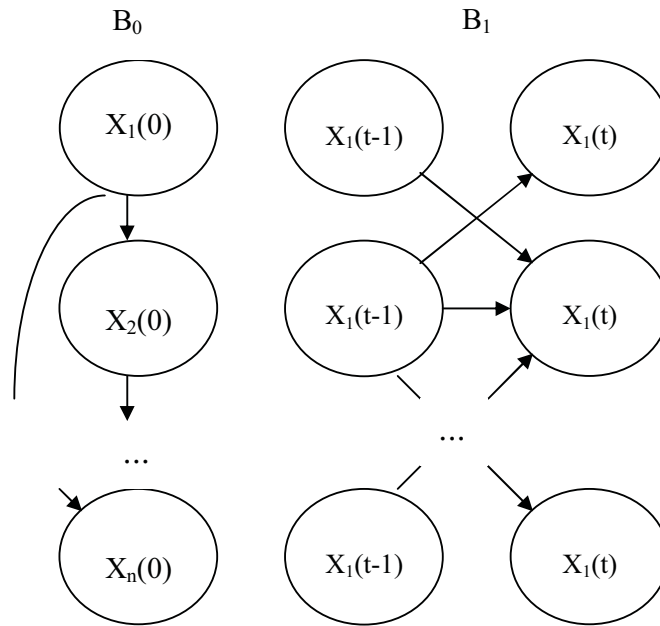


Рисунок 5 – Структура динамічної мережі Байєса

Побудова структури ДМБ виконується за два етапи:

1. Побудова статичної структури мережі. Тобто, це побудова структури тієї частини мережі, що буде повторюватися на кожному інтервалі часу.

Вхідні дані. Навчальна вибірка  $D = \{d_1, \dots, d_n\}$ ,  $d_i = \{x_i^{(1)} x_i^{(2)} \dots x_i^{(N)}\}$  (нижній індекс – номер спостереження, а верхній – номер змінної),  $n$  – число спостережень;  $N$  – число вершин (змінних).

Перший етап. Для всіх пар вершин обчислюємо значення взаєм-

ної інформації  $Set\_MI = \left\{ MI(x^i, x^j); \forall i, j \right\}$ . Після цього елементи

множини  $Set\_MI$  упорядковуємо за спаданням:

$$Set\_MI = \{MI(x^{m_1}, x^{m_2}), MI(x^{m_3}, x^{m_4}), MI(x^{m_5}, x^{m_6}), \dots\}.$$

Другий етап

Крок 1. З множини значень взаємної інформації  $Set\_MI$  вибираємо перші два максимальних значення  $MI(x^{m_1}, x^{m_2})$  и  $MI(x^{m_3}, x^{m_4})$ .

За отриманим значенням  $MI(x^{m_1}, x^{m_2})$  и  $MI(x^{m_3}, x^{m_4})$  будуємо множину моделей  $G$  вигляду:



$\{ (m_1 \rightarrow m_2; m_3 \rightarrow m_4), (m_1 \rightarrow m_2; m_3 \leftarrow m_4), (m_1 \leftarrow m_2; m_3 \leftarrow m_4), (m_1 \leftarrow m_2; m_3 \rightarrow m_4), (m_1 \leftarrow m_2; m_3 \text{ не залежить від } m_4), (m_1 \rightarrow m_2; m_3 \text{ не залежить від } m_4), (m_1 \text{ не залежить від } m_2; m_3 \rightarrow m_4), (m_1 \text{ не залежить від } m_2; m_3 \leftarrow m_4), (m_1 \text{ не залежить від } m_2; m_3 \text{ не залежить від } m_4) \}$ .

Запис вигляду  $m_i \rightarrow m_j$  означає, що вершина  $x^{m_i}$  є предком вершини  $x^{m_j}$ .

Крок 2. Виконуємо пошук серед множини моделей  $G$ . В параметрі  $g^*$  зберігаємо оптимальну мережеву структуру. Оптимальною структурою буде та, яка буде мати найменше значення деякого функціоналу. Наприклад,  $L(g, x^n)$  – опис мінімальної довжини (ОМД) структури моделі при заданій послідовності із  $n$  спостережень  $x^n = d_1 d_2 \dots d_n$ .

1.  $g^* \leftarrow g_0 (\in G)$ ;
2. для  $\forall g \in G - \{g_0\}$ : якщо  $L(g, x^n) < L(g^*, x^n)$  то  $g^* \leftarrow g$ ;
3. на виході  $g^*$  – шукане рішення.

Крок 3. Після того як знайдено оптимальну структуру  $g^*$  із  $G$ , з множини значень взаємної інформації  $Set - MI$  вибираємо максимальне значення:  $MI(x^{i\_next}, x^{j\_next})$ . За отриманим значенням  $MI(x^{i\_next}, x^{j\_next})$  і структурою (структурами)  $g^*$  будуємо множину моделей  $G$  вигляду:  $\{ (g^*; i\_next \rightarrow j\_next), (g^*; i\_next \leftarrow j\_next), (g^*; i\_next \text{ не залежить від } j\_next) \}$ . Переходимо на крок 2.

Умова закінчення процедури пошуку. Евристичний пошук продовжується до тих пір, поки не буде виконано аналіз визначеного

$$\frac{N \cdot (N - 1)}{2}$$

числа елементів множини або ж всіх елементів множини  $Set - MI$ . Як показує практика, у більшості випадків немає сенсу ви-

конувати аналіз більше половини (тобто  $\frac{N \cdot (N-1)}{4}$ ) елементів множини  $Set\_MI$ .

Вихід: оптимальна структура (структури)  $\mathcal{G}^*$ .

2. Побудова динамічної структури мережі. На цьому етапі будуємо структуру, що визначає зв'язки між двома сусідніми часовими зрізами, тобто  $t$  і  $t+1$ .

Побудова виконується за алгоритмом побудови структури статичної мережі, але розглядається набір вершин для двох сусідніх інтервалів часу, тобто множина вершин, між якими шукають зв'язки, подвоюється.

Розглядаємо тільки ті пари, що складаються з вершин, які знаходяться в різних зрізах. Щоб отримати дані про змінні на сусідньому часовому зрізі, зсуваємо вибірку даних на одну позицію вперед (для отримання минулого зрізу) або назад (для отримання наступного).

Вхідні дані. Навчальна вибірка  $D_t = \{d_1, \dots, d_n\}$ ,  $D_{t+1} = \{d_2, \dots, d_{n+1}\}$ .

Перший етап. Для всіх пар вершин із сусідніх часових зрізів об-

числюємо значення взаємної інформації

$$Set\_MI = \left\{ MI(x_t^i, x_{t+1}^j); \forall i, j \right\}$$

Після цього елементи множини  $Set\_MI$  упорядковуємо за спаданням:

$$Set\_MI = \{MI(x_t^{m_1}, x_{t+1}^{m_2}), MI(x_t^{m_3}, x_{t+1}^{m_4}), MI(x_t^{m_5}, x_{t+1}^{m_6}), \dots\}$$

Другий етап

Крок 1. Із множини значень взаємної інформації  $Set\_MI$  вибираємо перші два максимальних значення  $MI(x_t^{m_1}, x_{t+1}^{m_2})$  и  $MI(x_t^{m_3}, x_{t+1}^{m_4})$ . Будуємо множину моделей  $G$  вигляду:

$\{ (m_1 \rightarrow m_2; m_3 \rightarrow m_4), (m_1 \rightarrow m_2; m_3 \text{ не залежить від } m_4), (m_1 \text{ не залежить від } m_2; m_3 \rightarrow m_4), (m_1 \text{ не залежить від } m_2; m_3 \text{ не залежить від } m_4) \}$ . На відміну від алгоритму побудови статичної мережі з роз-

гляду виключаються всі моделі зі зворотнім зв'язком, тобто ті у яких  $x_t^{m_i}$  є предком вершини  $x_{t+1}^{m_i}$ .

Крок 2. Виконується пошук серед множини моделей  $G$ . В параметрі  $g^*$  зберігаємо оптимальну мережеву структуру. Оптимальною буде та структура, у якій буде найменше значення функції  $L(g, x^n)$  (ОМД структури моделі).

Крок 3. Після того як знайдено оптимальну структуру  $g^*$  з  $G$ , з множини значень взаємної інформації  $Set\_MI$  вибираємо максимальне значення:  $MI(x_t^{i\_next}, x_{t+1}^{j\_next})$ . За отриманим значенням і структурою  $g^*$  будемо множини моделей  $G$  вигляду:  $\{ (g^*; i\_next \rightarrow j\_next), (g^*; i\_next \text{ не залежить від } j\_next) \}$ . Переходимо на крок 2.

Умова закінчення процедури пошуку. Евристичний пошук продовжується до тих пір, поки не буде виконано аналіз визначеного числа елементів множини або ж всіх  $N^2$  елементів множини  $Set\_MI$ . Як показує практика, у більшості випадків немає сенсу виконувати аналіз більше половини (тобто  $\frac{N \cdot N}{2}$ ) елементів множини  $Set\_MI$ .

Вихід: оптимальна структура (структури)  $g^*$ .

### 1.6 Приклад побудови динамічної мережі Байєса

Наведемо приклад побудови ДМБ. В даному прикладі використані дані «Augsburg Indoor Location Tracking Benchmarks» [5]. В них зафіксовано переміщення чотирьох осіб в будівлі; дані отримані в період з липня 2003 по січень 2004 року у рамках проекту «Smart Doorplate». Мета: визначити присутність чи відсутність особи у конкретному приміщенні.

На основі цих даних будемо вибірку з 4-х дискретних змінних:

## Вихідні дані

№	Змінна	Позначення	Можливі значення
1	Кімната	ROOM	402, 403, 404, 405, 406, 407, 408, 409, 410, 411, 412, corridor, printer, kitchen, restroom, away
2	Час перебування	DUR	<0.5hr, 0,5-1hr, 1-2hr, 2-3hr, >3hr
3	День тижня	DAY	1, 2, 3, 4, 5
4	Час дня	TIME	9-10, 10-11, 11-12, 12-13, 13-14, 14-15, 15-16, 16-17, 17-18

**I. Побудова статичної структури мережі**

Обчислюємо значення взаємної інформації між вершинами та ОМД для 3 варіантів структури мережі. Структура, для якої значення ОМД мінімальне, є найкращою. Повторюємо цей етап для всіх вершин у порядку спадання взаємної інформації. В результаті отримуємо статичну структуру мережі, зображену на рис. 6:

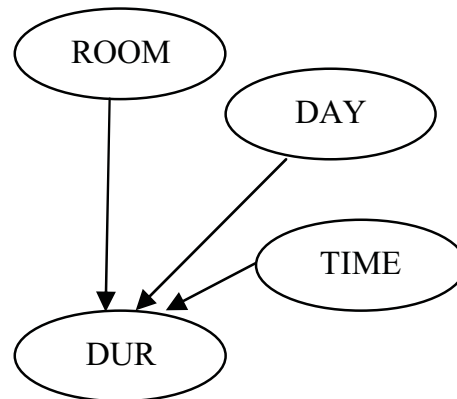


Рисунок 6 – Структура статичної мережі

**II. Побудова динамічної структури мережі**

Спочатку сформуємо дані для двох часових зрізів  $t$  і  $t+1$ . Для  $D_t$  виключаємо з вибірки записи з TIME = «17-18» – будемо вважати, що останнє місце перебування не вплине на перше місце наступного робочого дня. Для  $D_{t+1}$  виключаємо записи з TIME = «9-10», оскільки на них не впливає попереднє місце перебування працівника. Таким чином, отримуємо дані, наведені у табл. 4.

Вихідні дані після попередньої обробки

T				t+1			
ROOM	DUR	DAY	TIME	ROOM	DUR	DAY	TIME
402	<0.5hr	1	9-10	402	1-2hr	1	10-11
402	1-2hr	1	10-11	405	0,5-1hr	1	11-12
405	0,5-1hr	1	11-12	407	0,5-1hr	1	12-13
407	0,5-1hr	1	12-13	kitchen	0,5-1hr	1	13-14
kitchen	0,5-1hr	1	13-14	407	<0.5hr	1	14-15
407	<0.5hr	1	14-15	402	0,5-1hr	1	15-16
402	0,5-1hr	1	15-16	402	1-2hr	1	16-17
402	1-2hr	1	16-17	402	2-3hr	1	17-18

Обчислюємо взаємну інформацію між вершинами сусідніх зрізів та для всіх пар вершин в порядку спадання величини взаємної інформації, перевіряємо критерій ОМД. В результаті отримуємо структуру динамічної мережі Байєса, наведену на рис. 7.

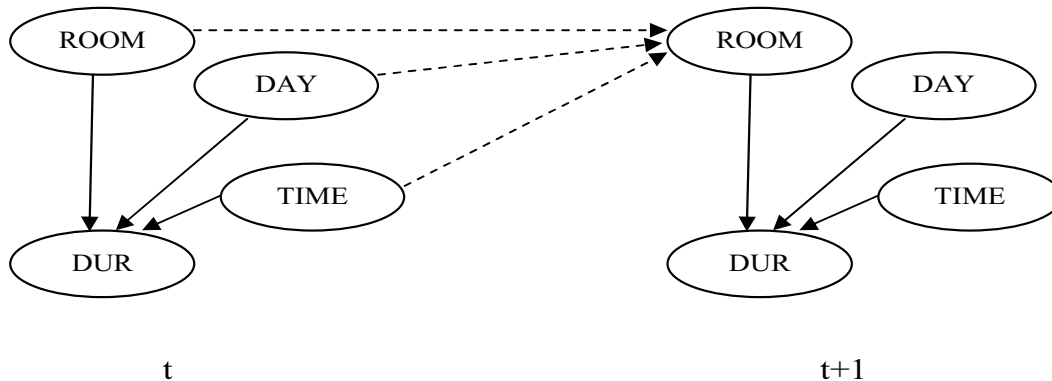


Рисунок 7 – Структура двох сусідніх часових зрізів динамічної мережі

Отримана структура може бути використана для формування ймовірнісного висновку стосовно ймовірностей перебування осіб у вибраному приміщенні.

### Висновки

В роботі розглянуто використання мереж Байєса на конкретних прикладах, зокрема, моделювання, прогнозування та класифікації. Наведені приклади свідчать, що мережі Байєса – сучасний підхід до обробки даних та моделювання процесів різної природи і складності. Також розглянуто основні види мереж Байєса і методика побудови

динамічних мереж. Динамічні мережі Байеса описують динаміку процесу і призначені для прийняття рішень стосовно оцінок його стану в умовах наявності невизначеностей. Для простоти представлення моделі у більшості випадків приймають, що кількість змінних та зв'язки між ними повторюються в кожний наступний момент часу, а властивості ДМБ в цілому відповідають марковському процесу першого порядку.

Побудова динамічної мережі Байеса складається з таких етапів: (1) – побудова статичної структури мережі (що повторюється на кожному інтервалі часу) та (2) – побудова динамічної структури мережі, що включає визначення зв'язків між двома сусідніми інтервалами часу. Отриману структуру використовують для формування висновку на кожному часовому інтервалі.

Отже, динамічні мережі Байеса описують динаміку процесу і призначені для прийняття рішень щодо його поточного стану в умовах наявності невизначеностей. В подальших дослідженнях необхідно застосувати наведену вище методику до фактичних даних та порівняти отримані результати з іншими методами.

## ЛІТЕРАТУРА

1. Montesano L., Lopes M., Bernardino A., Jose Santos-Victor. Modeling Affordances using Bayesian networks / IEEE/RSJ International Conference on Intelligent Robots and Systems, San Diego, USA, 2007, 12 p.
2. Maragoudakis M., Tselios N.K., Fakotakis N., Avouris N.M. Improving SMS usability using Bayesian Networks / Wire Communications Laboratory, Technical Report, 2005, 45 p.
3. Madden M.G. A New Bayesian Network Structure for Classification Tasks. – Berlin: Springer, 2002. – pp. 183-197.
4. Бидюк П.И., Терентьев А.Н. Построение и методы обучения байесовских сетей // Таврический вестник информатики и математики. – Симферополь: КНЦ НАНУ, 2004, № 2, с. 139-153.
5. <http://www.informatik.uniaugsburg.de/en/chairs/sik/research/finished/ailtbenchmarks/>.