УДК 519.217

Г.А. Степанкова, І.В. Баклан

ПРИХОВАНІ МАРКОВСЬКІ МОДЕЛІ: КЛАСИФІКАЦІЯ

Анотація. В статті запропонований погляд на класифікацію прихованих марковських моделей та їх місце в загальній класифікації марковських моделей. Були визначені основні класифікаційні ознаки, за якими і був виконаний класифікаційний аналіз.

Ключові слова: марковські моделі, приховані марковські моделі, гібридні моделі, класифікація марковських моделей.

Вступ

Імовірнісні моделі, такі як приховані марковські моделі (ПММ) та баєсові мережі, на сьогодні є одними з найпопулярніших математичних формалізмів, які використовуються у сучасних інтелектуальних системах прийняття рішень. Тому дуже важливо мати повну класифікацію різновидів цих ймовірнісних моделей та їх застосування.

Постановка проблеми

Теорія прихованих марковських моделей не нова. Її основи опублікував Баум і його колеги в кінці 60-х, на початку 70-х років. Тоді ж, на початку 70-х років, Бейкер і Джелінек з колегами із IBM застосували ПММ в розпізнаванні мови [1, 2].

Тим не менш, широке розповсюдження ПММ отримала зовсім недавно. Основи теорії ПММ були опубліковані у журналах для математиків, які не дуже популярні серед інженерів, що займаються розпізнанням мови. Опублікована теорія не містила відповідних навчальних матеріалів, яки б пояснювали можливості та способи застосування ПММ у різних прикладних галузях. Це призвело до того, що до цього часу не створена цілісна теорія ПММ та не визначене їх місце у теорії імовірнісних моделей. Зокрема, не була створена загальна класифікація ПММ.

Аналіз досліджень і публікацій

Реальним початком імовірнісного моделювання були роботи монаха Томаса Баєса та, особливо, його теорема про умовні ймовірно-

 ISSN 1562-9945

сті у його великій роботі «Нариси до вирішення проблеми доктрини шансів».

З імовірнісним моделюванням пов'язані імена таких видатних світових математиків, як А.А. Марков [3], А.М. Колмагоров [4] та інших.

На сьогодні існують цілі теорії імовірнісного моделювання, які використовуються для вирішення задач теорії масового обслуговування, розпізнання образів, біоінформатики тощо.

Роботи останніх років показали, що в економіці, генетиці, розпізнаванні образів знайшли місце для вирішення цих задач такі математичні апарати, як баєсові мережі та, особливо, приховані марковські мережі (ПММ).

Мета роботи

Метою даної статті є ознайомлення зі створеною нами найбільш повною класифікацією марковських моделей, та, зокрема, встановлення місця прихованих марковських моделей у загальній класифікації імовірнісних моделей.

Загальна класифікація імовірнісних моделей

У загальному вигляді класифікація імовірнісних моделей представлена на рисунку 1.

Імовірнісні моделі (стохастичні моделі) – це моделі, які на відміну від детермінованої моделі містять випадкові елементи (величини). Таким чином, при завданні на вході моделі деякої сукупності значень, на її виході можуть бути результати, які, залежно від дії випадкового фактора, розрізняються між собою.

Класифікація марковських моделей

В свою чергу, ми підрозділяємо марковські моделі на марковські ланцюжки, марковські логічні мережі та приховані марковські моделі.

Марковські ланцюжки

Якщо випадкова послідовність має марковську властивість, то вона зветься ланцюжком Маркова. Тобто марковським ланцюжком називають таку послідовність випадкових подій, в якій ймовірність кожної події залежить тільки від стану, в якому процес знаходиться в поточний момент та не залежить від більш ранніх станів. Скінченний дискретний ланцюжок визначається:

– *множиною станів* $S = \{s_1, \ldots, s_n\}$, подією є перехід з одного стану в інший в результаті випадкового випробування;

– *вектором початкових ймовірностей* (початковим розподілом) $p^{(0)} = \{p^{(0)}(1), \ldots, p^{(0)}(n)\}$, що визначає імовірності $p^{(0)}(i)$ того, що в початковий момент часу $t = 0$ процес знаходився в стані s_i;

– *матрицею перехідних ймовірностей* $P = \{p_{ij}\}$, що характеризує імовірність переходу процесу з поточним станом s_i в наступний стан s_j, при цьому сума ймовірностей переходів з одного стану дорівнює 1: $\sum_{j=1 \ldots n} p_{ij} = 1$.

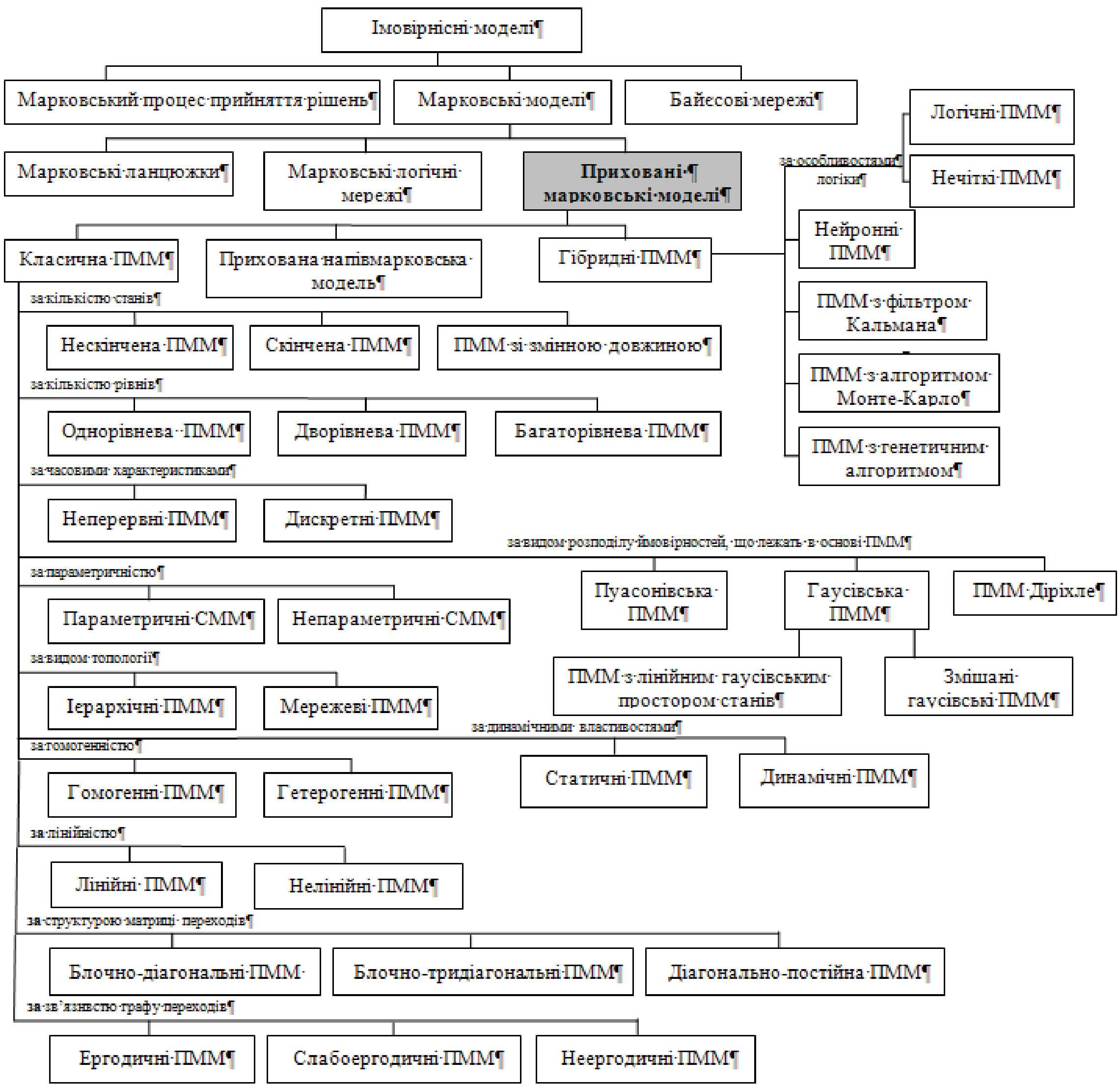

Рисунок 1 - Загальна класифікація імовірнісних моделей марковського типу

Марковські логічні мережі

Марковська логічна мережа – це ймовірнісна логіка, яка застосовує ідею марковської мережі до логіки першого порядку [5].

Імовірнісні графічні моделі дозволяють нам достатньо ефективно вручну представлять невизначеність. Логіка першого порядку дозволяє компактно представляти обширні та різноманітні знання. Марковська логічна мережа [5], що є досить простим методом, поєднує в собі ймовірність та логіку першого порядку без будь-яких обмежень, крім реального домену.

Класифікація прихованих марковських моделей

Тепер перейдемо безпосередньо до класифікації ПММ. У широкому сенсі ПММ можна розділити на наступні загальні класи:

- прихована напівмарковська модель;
- класична ПММ;
- гібридні марковські моделі.

Приховані напівмарковські моделі

Приховані напівмарковські моделі є узагальненням прихованих марковських моделей. Вони дають значну гнучкість часовим розподілам, які беззаперечно виконують геометричний розподіл у разі прихованих марковських ланцюжків. Прихована напівмарковська модель є узагальнення ПММ, що дозволяє використовувати більш загальний розподіл часу перебування.

Класифікація класичних прихованих марковських моделей

Тепер розглянемо класифікацію безпосередньо класичних ПММ. Класичні ПММ розділяються за наступними основними класифікаційними ознаками:

– за кількістю станів на скінчені та нескінчені ПММ;

– за кількістю рівнів на однорівневі, дворівневі та багаторівневі ПММ;

– за часовими характеристиками на неперервні та дискретні ПММ;

– за видом розподілу ймовірностей, що лежать в основі ПММ, на пуасонівські, гаусівські ПММ та ПММ Діріхле;

– за параметричністю на параметричні та непараметричні ПММ;

– за видом топології на мережеві та ієрархічні ПММ;

– за гомогенністю на гомогенні та гетерогенні ПММ;

– за лінійністю на лінійні та нелінійні ПММ;

– за динамічними властивостями на статистичні та динамічні ПММ;

– за структурою матриці переходів на блочно-діагональні, блочно-тридіагональні та діагонально-постійні ПММ;

– за зв'язністю графу переходів на ергодичні, слобоергодичні та неергодичні ПММ.

Класифікація ПММ за кількістю станів

Основним припущенням моделювання класичної ПММ є те, що дані породжуються деяким дискретним станом змінної, яка може приймати одне з кількох значень, що є необґрунтованим для більшості реальних задач. Нескінчені ПММ – це ПММ, у яких зліченне нескінченне число прихованих станів. Відповідно, скінчені ПММ – це моделі, в яких скінченне число прихованих станів. ПММ зі змінною довжиною – це ПММ, в яких в кожний момент часу різне число прихованих станів .

Класифікація ПММ за кількістю рівнів

Моделювання ПММ передбачає рознесення її елементів за відповідними ієрархічними рівнями. Ідея рівнів передбачає ієрархічну будову ПММ, домінування одних елементів над іншими та, навпаки, підпорядкування одних елементів іншим.

Елементи нижнього рівня входять як інтегранти в елементи більш високих рівнів.

Класифікація ПММ за часовими характеристиками

У неперервних ПММ фігурують неперервні змінні. Це стосується і такої незалежної змінної, як час. У дискретних ПММ змінні, в тому числі й час, дискретні, тобто для них визначена деяка множина дозволених значень, в окремому випадку їх всього два (виконавчі змінні) [6].

Класифікація ПММ за видом розподілу ймовірностей, що лежать в основі ПММ

Пуасонівські приховані марковські моделі (ППММ) є різновидом ПММ [7, 8]. ППММ є дискретними часовими стохастичними процесами $\{(X_t, Y_t)\}_{t\in N}$, такими що $\{X_t\}_{t\in N}$ є недосліджуваною нескінченою множиною марковських ланцюжків та $\{Y_t\}_{t\in N}$ є досліджуваною

 ISSN 1562-9945

послідовністю випадкових змінних, залежних від $\{X_t\}_{t\in N}$. Це змодельоване припущення про те, що умовний розподіл кожного досліджуваного Y_t задається послідовністю $\{X_t\}_{t\in N}$ та залежить виключно від одночасно недосліджуваних X_t. Більш того, задані $\{X_t\}_{t\in N}$ та $\{Y_t\}_{t\in N}$ є послідовностями умовно незалежних спадкових величин. Якщо ми припустимо, що для кожного t, Y_t, який задає множину X_t, є пуасонівською випадковою величиною, то будемо мати так звані ППММ. В цьому випадку X_t визначає паусонівський параметр, який використовується для генерації Y_t.

Гаусівські ПММ є різновидом ПММ. Гаусівські ПММ поділяються на ПММ з лінійним гаусівським простором станів та змішані гаусівські ПММ.

Класифікація ПММ за параметричністю

ПММ діляться на параметричні та непараметричні. В параметричних ПММ передбачається, що характеристики досліджуваних об'єктів описуються у вигляді розподілів, що залежать від (одного чи декількох) числових параметрів. Параметризація дозволяє за короткий час «програти» (з допомогою зміни параметрів) різні ситуації та уникнути принципових помилок.

Непараметричні ПММ не пов'язані зі специфікацією параметричного сімейства для розподілу досліджуваних характеристик. На відміну від параметричної ПММ для непараметричних ПММ не робиться ніяких припущень про нормальність вибірок. Це істотно розширює коло розглянутих завдань.

Класифікація ПММ за видом топології

У мережевих ПММ компактно відображаються найбільш суттєві відносини між елементами. Зазвичай мережеві моделі зображуються в явному графічному вигляді. Поширеною формою подання мережевих моделей є графи.

Якщо в ПММ елементи, що входять до неї, розглядаються, в свою чергу, як ПММ, то сама ПММ зветься ієрархічною ПММ.

Класифікація ПММ за гомогенністю

Гомогенна ПММ - однорідна модель, тобто модель, яка складається з однотипних елементів. Наприклад, гомогенною ПММ буде ПММ з однаковим розподілом імовірності переходів між усіма станами. Гетерогенна ПММ - неоднорідна модель, що складається з однорідних частин (фаз), розділених між собою. Однорідні частини (фази) можуть відрізнятися одна від одної за складом та властивостями. Тобто гетерогенні ПММ – це моделі, в яких різні розподіли імовірності переходів між різними станами.

Класифікація ПММ за лінійністю

Лінійна залежність однієї величини від іншої - це пропорційність їх приростів. Аналогічно визначається поняття лінійної ПММ. Будемо вважати, що на початку відліку входи та виходи вибрані так, що нульовому входу відповідає нульовий вихід. Тоді ПММ називається лінійною, якщо в ній виконаний принцип суперпозиції (накладення), тобто при складанні входів складаються і виходи, а при множенні входу на будь-яке число вихід множиться на те ж число. Якщо цей принцип не виконаний, ПММ називається нелінійною ПММ. Лінійні моделі зазвичай описуються лінійними неоднорідними рівняннями - алгебраїчними, диференціальними тощо, в яких неоднорідний член відповідає входу, а рішення – виходу.

Класифікація ПММ за динамічними властивостями

Статичні ПММ – це моделі, в яких система представляється незмінною в часі. Такі моделі зручні, коли потрібно описати структуру системи, тобто з яких об'єктів вона складається, як ці об'єкти пов'язані один з одним, і які властивості цих об'єктів. Образно кажучи, статична модель є як би «фотографією» істотних властивостей системи в певний момент часу.

Динамічні ПММ – це моделі, які містять інформацію про поведінку системи та її складових частин. Для опису поведінки зазвичай використовуються записані у вигляді формул, схем чи комп'ютерних програм співвідношення, що дозволяють обчислити параметри системи та її об'єктів, як функції часу.

Класифікація ПММ за структурою матриці переходів

Блочно-діагональні ПММ – це ПММ, в якій матриця переходів P за структурою є блочно-діагональною. *Блочно-діагональна матриця* – це блочна квадратна матриця, в якій блоки також є квадратними матрицями, а блоки поза основною діагоналлю є нульовими матрицями. Блочно-тридіагональні ПММ – це ПММ, в яких матриця переходів P за структурою є блочно-тридіагональною. Діагонально-постійні ПММ – це ПММ, в яких матриця переходів P за структурою є діагонально постійною.

 ISSN 1562-9945

Гібридні приховані марковські моделі

Нейронні ПММ

Приховані марковські моделі та штучні нейронні мережі (ШНМ) взаємно доповнюють один одного та компенсують властиві їм недоліки. Це призвело до ідеї комбінувати ці структури в рамках однієї нової моделі, яку можна визначити як гібридну ПММ/ШНМ модель чи гібридну нейронну ПММ. Гібридна модель дозволяє ефективно об'єднати переваги марківських моделей та нейронних мереж, тобто ПММ забезпечує можливість моделювання довготривалих залежностей, а ШНМ забезпечує непараметричну універсальну апроксимацію, оцінку імовірності, алгоритми дискримінантного навчання, зменшення числа параметрів для оцінки, які зазвичай потрібні для стандартних ПММ. Результатом використання таких гібридних структур явилося значне підвищення якості розпізнавання в порівнянні із стандартними методами [9 – 12].

ПММ з фільтром Калмана

Неперервне навчання ПММ віконується доволі повільно. Це серйозне обмеження можна обійти, розглянувши навчання ПММ з учителем як завдання оптимальної фільтрації, рішення якої рекурсивно використовує інформацію, що міститься в даних навчання, неявно повертаючись до першої ітерації процесу навчання. Ця ідея лежить в основі ПММ з фільтром Калмана [7]. Серед переваг ПММ з фільтром Калмана слід виділити ефективне використання інформації, що міститься у вхідних даних.

ПММ з алгоритмом Монте-Карло

Гібридна ПММ з алгоритмом Монте-Карло (ПМММК) являє собою алгоритм навчання для ПММ з неперервними простором станів та простором спостережень. ПМММК працює з неперервним простором станів та спостережень. Цей підхід використовує методи Монте-Карло для апроксимації великих непараметричних класів функцій щільності. Також в ПМММК передбачається, що всі простори та випадкові змінні (наприклад, стан, спостереження) вимірні. Передбачається також, що якщо інше не зазначено, всі розподіли ймовірностей неперервні та мають неперервні функції щільності.

Класифікація гібридних ПММ за особливостями логіки

Побудова гібридної моделі на основі апарату нечітких множин та прихованих марковських моделей отримала назву нечіткої прихованої марковської моделі (НПММ). Для НПММ застосовуються вже відомі принципи гібридизації різноманітних моделей з прихованими марковськими моделями [13, 14]. Особливістю цього підходу є застосування алгоритму нечіткого очікування в ПММ. Основною перевагою НПММ у порівнянні зі звичайними ПММ є вища продуктивність НПММ у порівнянні зі звичайними ПММ.

Висновки

На сьогодні не існувало повної класифікації ПММ, які були визначені у багатьох наукових статтях. Тому ціллю нашого дослідження було створення такого роду класифікації.

Слід зазначити, що нами була отримана найбільш повна класифікація сучасних ПММ та галузей їх застосування.

Головним результатом дослідження стала класифікація прихованих марковських моделей. Також були розглянуті основи теорії прихованих марковських моделей, а саме: марковський процес, загальна класифікація імовірнісних моделей, класифікація марковських моделей, класифікація прихованих марковських моделей, класифікація класичних ПММ, гібридні ПММ, а також класифікація типів задач, які розв'язуються за допомогою ПММ.

ЛІТЕРАТУРА

1. Baker J. K. The DRAGON system / Baker J. K. – IEEE Trans. Acoust., Speech Signal Processing, vol. ASSP-23, Feb. 1975. – pp. 24–29.
2. Baker J. Trainable grammars for speech recognition / Baker J. – Boston, MA.: In Speech communication paper presented at th 97th Meeting of the Acoustical Society of America, 1979. – pp. 547–550.
3. Марков А.А. Теория чисел. Теория вероятностей / Марков А.А. – М.: Изд-во АН СССР, 1951. – (Избранные труды).
4. Колмогоров А. Н. Основные понятия теории вероятностей / Колмогоров А. Н. – [Изд. 2-е]. – М.: Наука, 1974. – 120 с.
5. Richardson M. Markov logic networks / Richardson M., Domingos P. – Dept. Comp. Sci. & Eng., Univ. Washington, Seattle, 2004.
6. Fraser A. M. Hidden Markov models and dynamical systems / Fraser Andrew M. – Philadelphia: Society for Industrial and Applied Mathematics. 2008.
7. Баклан І.В. Пуасонівські приховані марковські моделі / Баклан І.В., Степанкова Г.А. / Сучасні інформаційні та інноваційні техно-

логії на транспорті: матеріали Міжнародної науково-практичної конференції. – Херсон: Видавництво Херсонського державного морського інституту, 2010. Т. 1. – С.122 – 126.

8. Баклан І.В. Дослідження властивостей пуасонівських прихованих марковських моделей для прогнозування часових рядів / Баклан І.В., Степанкова Г.А. – Інтелектуальні системи прийняття рішень та проблеми обчислювального інтелекту: матеріали міжнародної наукової конференції. – Херсон: ХНТУ, 2010. Т. 1. – С.238 – 243.
9. Степанкова Г.А. Особливості гібридизації нейронних мереж та прихованих марківських моделей / Степанкова Г.А., Баклан І.В. – Искусственный интеллект. Интеллектуальные системы: материалы IX Международной научно-технической конференции. – Донецк: ИПИИ «Наука і освіта». 2008. Т.1. – С.328 – 333.
10. Степанкова Г.А. Побудова гібридних моделей на основі прихованих марківських моделей та нейронних мереж / Степанкова Г.А., Баклан І.В. – Автоматика. Автоматизация. Электротехнические комплексы и системы. 2008. №1(21). – С.5 – 11.
11. Баклан І.В. Застосування гібридних нейронних мереж для прогнозування фінансових часових рядів / Баклан І.В., Степанкова Г.А. – Системний аналіз та інформаційні технології: матеріали XI Міжнародної науково-технічної конференції (26-30 травня 2009р., Київ). – К.: ННК «ІПСА» НТУУ «КПІ», 2009. – С.261.
12. Баклан І.В. Гибридні моделі в статистичних методах розпізнавання образів / Баклан І.В., Рифа В.М. – Вісник Херсонського Державного Технічного Університету № 3 (19). - Херсон: ХДТУ. 2003. – С.26-28.
13. Баклан І.В., Гібридні технології в проектуванні інтелектуальних систем прийняття рішень / Баклан І.В. – Сучасні інформаційні та інноваційні технології на транспорті: матеріали Міжнародної науково-практичної конференції. – Херсон: Видавництво Херсонського державного морського інституту, 2009. Т.1 – С.32 –37.
14. Баклан І.В. Нечіткі приховані марковські моделі для прогнозування часових рядів / Баклан І.В., Степанкова Г.А. – Искусственный интеллект. Интеллектуальные системы ИИ-2010: материалы Международной научно-технической конференции. – Донецк: ИПИИ «Наука і освіта». 2010. Т.2. – С. 19 – 24.