

УДК 681.562

ПРИНЦИПИ ДІАГНОСТУВАННЯ КОМП'ЮТЕРНИХ МЕРЕЖ НА БАЗІ НЕЧІТКОЇ ЛОГІКИ

Б.М. Герасимов, І.В. Самойлов, С.В. Толюпа

Військовий інститут телекомунікацій та інформатизації

Національного технічного університету України "КПІ"

e-mail: samoilov_igor@bigmir.net

Сучасні інформаційні технології характеризуються широким використанням локальних обчислювальних мереж (ЛОМ) як засобів передачі інформації у всіх сферах використання комп'ютерної техніки. Розгалуженість, складність обчислювальних мереж створюють проблему забезпечення надійності їх функціонування. Однією з центральних задач діагностики ЛОМ є встановлення причин виникнення несправностей та їх усунення. Велика кількість різних факторів, вплив яких необхідно враховувати при встановленні цих причин, значно ускладнюють задачу діагностики [1]. Звичайно, кваліфіковані адміністратори мереж в змозі успішно вирішувати задачі діагностики, однак кількість таких компетентних експертів відносно мала, і їм важко передати досвід та інтуїцію молодим адміністраторам. В зв'язку з цим виникає необхідність в розробці комп'ютерних систем, які б забезпечували інтелектуальну підтримку прийняття діагностичних рішень адміністраторами середньої кваліфікації [2, 3].

З особливостей побудови структури комп'ютерних мереж визначимо риси мереж як об'єктів контролю та діагностування. До них віднесемо:

неможливість коректно змоделювати структуру ЛОМ і процеси, що відбуваються в ній;

неоднозначність діагностичної інформації;

наявність в окремих компонентах ЛОМ засобів вбудованого контролю;

наявність несправностей визначеного типу в залежності від видів використовуваних протоколів і конфігурації ЛОМ;

залежність прояву несправностей від щільності мережного трафіка і кількості підключених абонентів;

відсутність точних кількісних даних, які відображають зв'язок між зовнішніми проявами несправностей та причинами їх виникнення.

Перераховані та інші особливості не дають можливості в повній мірі використовувати існуючі методи діагностування. Тому виникає інтерес до застосування методів та засобів штучного інтелекту [4], які дозволяють формалізовувати і використовувати доступну лінгвістичну інформацію, яка відображає розуміння експертом причинно-наслідкових зв'язків. Одним з перспективних шляхів формалізації лінгвістичної інформації є теорія нечітких множин [5]. Але на сьогоднішній день в літературі не визначені методологічні принципи побудови систем діагностики комп'ютерних мереж на базі теорії нечітких множин. При формуванні цих принципів автори використовували ідеї лінгвістичного моделювання, які запропоновані в [6].

Принцип лінгвістичності входних та вихідних змінних. У відповідності з цим принципом, входи об'єкта та його вихід розглядаються як лінгвістичні змінні, які оцінюються якісними термами (терм. – від англ. term – називати).

Згідно Л. Заде [5], лінгвістичною змінною називається така змінна, значеннями якої є слова або речення природної мови, тобто якісні терми. Сукупність усіх можливих значень лінгвістичної змінної називається терм-множиною. В якості прикладу наведемо декілька входних лінгвістичних змінних і відповідні терм-множини, які можуть використовуватися в задачі діагностування комп'ютерних мереж:

ЯКЕ МЕРЕЖЕВЕ ОТОЧЕННЯ “БАЧИТЬ” РОБОЧА СТАНЦІЯ? – {не бачить жодного комп’ютера мережі; бачить всю мережу крім серверу; бачить всю мережу},

ДРУК ДОКУМЕНТІВ З КОМП’ЮТЕРІВ МЕРЕЖІ – {не виконується ні з серверу, ні з жодної з робочих станцій; не виконується з жодної робочої станції, але виконується з серверу; не виконується тільки з цієї робочої станції; виконується з усіх комп’ютерів мережі}.

Вихідною змінною в цій задачі діагностування може бути:

ПРИЧИНА ВИНИКНЕННЯ НЕСПРАВНОСТІ – {сервер друку; принтер; робоча станція; мережевий комутатор}.

Використовуючи поняття функції належності, кожен з термів, якими оцінюються лінгвістичні змінні, можна формалізувати у вигляді нечіткої множини, яка задана на відповідній універсальній множині. Згідно Заде [5], нечіткою множиною A в універсальній множині U називається сукупність пар $(\mu^A(u), u)$, де $u \in U$ та $\mu^A(u): U \rightarrow [0, 1]$ – функція належності, яка являє собою деяку суб’єктивну міру відповідності елемента $u \in U$ властивостям нечіткої множини A . $\mu^A(u)$ може приймати значення від нуля, який показує абсолютну не належність, до одиниці, яка, навпаки, говорить об абсолютній належності елемента u нечіткій множині A .

Якщо універсальна множина U складається з кінцевої кількості множин $u_1, u_2, \dots, u_n$, то нечітка множина A представляється у вигляді:

$$A = \mu^A(u_1) / u_1 + \mu^A(u_2) / u_2 + \dots + \mu^A(u_n) / u_n = \sum_{i=1}^n \mu^A(u_i) / u_i.$$

У випадку неперервної множини U використовується наступне позначення:

$$A = \int_U \mu^A(u) / u.$$

Принцип формування структури діагностичної залежності “входи-вихід” у вигляді нечіткої бази знань. Нечітка база знань являє собою сукупність правил ЯКЩО “входи”, ТО “вихід”, які відображають досвід експерта та його розуміння причинно-наслідкових зв’язків. Прикладом експертного правила з нечіткої бази знань для діагностування комп’ютерних мереж може служити наступне висловлювання:

ЯКЩО робоча станція не бачить жодного комп’ютера мережі,

ТА друк документів не виконується тільки з цієї робочої станції,

ТО причина виникнення несправності – робоча станція.

Особливість подібних висловлювань складається у тому, що їх адекватність не змінюється при незначних коливаннях умов експерименту. Тому формування нечіткої бази знань можна трактувати як аналог структурної ідентифікації [7], на якому будується груба модель діагностування з параметрами, які підлягають настроюванню. У даному випадку настроюванню підлягають форми функції належності нечітких термів, за допомогою яких оцінюються входи та вихід об’єкту.

Крім цього, сукупність правил ЯКЩО-ТО можна розглядати як набір експертних точок у просторі “входи-вихід”. Використання апарату нечіткого логічного висновку [5, 6] дозволяє відновлювати по цих точках багатовимірну поверхню, яка допомагає отримувати значення виходу при різноманітних комбінаціях вхідних змінних.

Принцип ієрархічності діагностичних знань. Використання цього принципу дозволяє подолати так зване “прокляття розмірності”. При великій кількості вхідних змінних побудова системи висловлювань о невідомій залежності “входи-вихід” стає досить складною. Це зумовлено тим, що в оперативній пам’яті людини одночасно утримується не більш 7 ± 2 понять-ознак. В зв’язку з цим доцільно провести класифікацію вхідних змінних та по ній побудувати дерево висновку, яке буде визначати систему

вкладених одне в одне висловлювань-знань меншої розмірності. Приклад такого дерева для одинадцяти вхідних змінних показано на рис. 1.

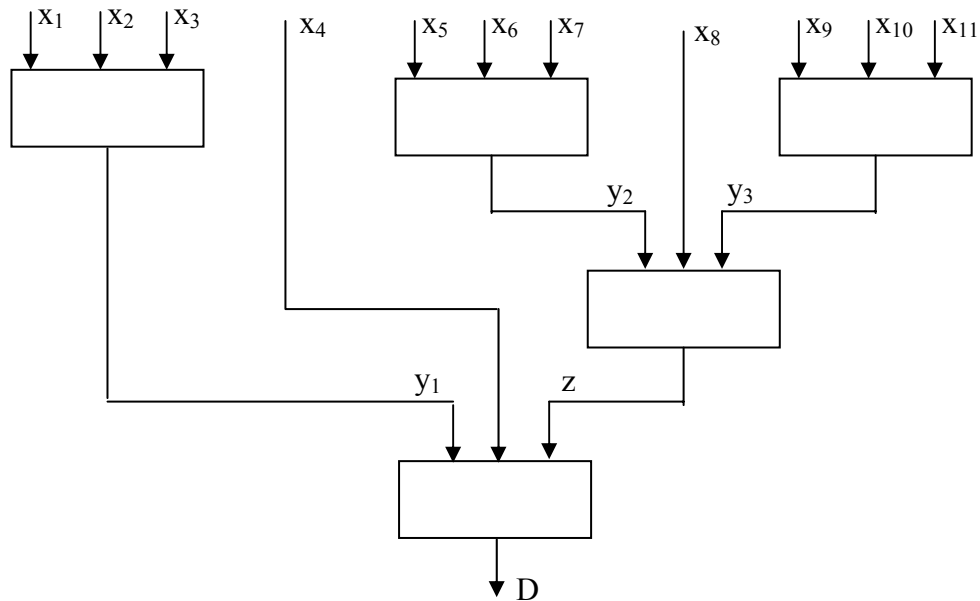


Рис. 1. Приклад дерева висновку

З цього прикладу видно, що знання виду

$$D = f(x_1, x_2, \dots, x_{11}),$$

які зв'язують входи $x_1 \div x_{11}$ з виходом D , замінюються послідовністю співвідношень:

$$D = f_1(y_1, x_4, z), y_1 = f_2(x_1, x_2, x_3), z = f_3(y_2, x_8, y_3), y_2 = f_4(x_5, x_6, x_7), y_3 = f_5(x_9, x_{10}, x_{11}),$$

де y_1, y_2, y_3, z – проміжні виходи, які розглядаються як лінгвістичні змінні.

За рахунок принципу ієрархічності можна враховувати практично необмежену кількість вхідних змінних, які впливають на оцінку вихідної змінної. При побудові дерева висновку необхідно намагатися, щоб кількість аргументів (вхідних стрілок) у кожному вузлі дерева не перевищувало 7 ± 2 . Доцільність порівневого представлення експертних знань обумовлена не тільки природною ієрархічністю об'єктів діагностування, але й необхідністю врахування додаткових змінних по мірі накопичення знань про об'єкт.

Принцип трьохваріантної оцінки значень вхідних змінних. Суть цього принципу полягає в тому, що експертна оцінка деякої вхідної змінної може бути здійснена по одному з трьох варіантів: числом, лінгвістичним термом та за принципом термометра. Перший варіант застосовується при можливості кількісної оцінки вхідної змінної та наявності інструментальних засобів для відповідних вимірювань. Другий варіант застосовується коли значення вхідної змінної задається лінгвістичною оцінкою з відповідною терм-множини. По третьому варіанту експертна оцінка деякої вхідної змінної здійснюється зафарбуванням частини шкали (рис. 2), ліва та права межа якої відповідають найменшому і найбільшому рівням змінної, що розглядається. Принцип термометра зручно застосовувати у тих випадках, коли експерт не в змозі оцінити деяку змінну ні числом, ні якісним термом, а лише інтуїтивно відчуває її рівень. Зручність такого підходу складається в тому, що він дозволяє розглядати різноманітні за природою лінгвістичні змінні на єдиній універсальній множині $U = [\underline{u}, \overline{u}]$.

Принцип двоетапного настроювання нечітких баз знань. У відповідності з цим принципом побудова моделі діагностування здійснюється у два етапи, які за аналогією з класичними методами [7] можна вважати етапами структурної та параметричної ідентифікації. Параметрами які підлягають настроюванню є: ваги нечітких правил ЯКЩО-ТО та форми функцій належності.



Рис. 2. Оцінка змінної за принципом термометра

Перший етап є традиційним для технології нечітких експертних систем. На цьому етапі здійснюється формування та груба настройка нечіткої бази знань за доступною експертною інформацією. Для грубої настройки ваг правил та форм функцій належності можна застосовувати модифікований метод парних порівнянь Сааті, який описаний в [6]. Чим вище професійний рівень експерта, тим краще адекватність нечіткої моделі діагностування комп'ютерних мереж, побудованої на першому етапі. Ця модель називається чистою експертною системою, так як для її побудови застосовується тільки експертна інформація. Однак, співпадання результатів нечіткого логічного висновку (теорія) та експериментальних даних, не є гарантованими. Тому необхідний другий етап, на якому здійснюється тонка настройка нечіткої моделі шляхом її навчання по експериментальним даним.

Суть другого етапу ідентифікації є підбір ваг нечітких правил з бази знань, та таких параметрів функцій належності, які мінімізують відхилення між експериментальними даними і результатами нечіткого логічного висновку. Тонка настройка нечіткої бази знань зводиться до рішення задачі нелінійної оптимізації, яка може бути вирішена різними методами. Однак, при великій кількості вхідних змінних та нечітких термів в базі знань, застосування традиційних технологій оптимізації (метод найшвидшого спуску, квазіньютонівські методи тощо) потребує великих затрат машинного часу. Тому на цьому етапі пропонується застосовувати генетичні алгоритми оптимізації. Ці алгоритми є аналогом випадкового пошуку, який ведеться одночасно з різних початкових точок, що скорочує час пошуку оптимальних параметрів нечіткої бази знань.

Запропоновані принципи дозволяють застосовувати в якості початкової інформації для побудови моделей діагностики, як експериментальні дані, так і лінгвістичні висловлювання-правила, які відображають розуміння експертом причинно-наслідкових зв'язків.

Література

1. Пархоменко П.П., Согомонян Е.С. Основы технической диагностики: Оптимизация алгоритмов диагностирования, аппаратные средства. – М.: Энергия, 1981. – 236с.
2. Тоценко В.Г. Експертні системи діагностики і підтримки рішення. К: Наукова думка, 2004. – 124с.
3. Герасимов Б.М., Дивизинюк М.М., Субач И.Ю. Системы поддержки принятия решений: проектирование, применение, оценка эффективности. – Севастополь: СНИЯЭиП, 2004. – 320с.
4. Поспелов Г.С. Искусственный интеллект – основа новой информационной технологии. – М.: Наука, 1988. – 220 с.
5. Заде Л. Понятие лингвистической переменной и её применение к принятию приближенных решений. – М.: Мир, 1976. – 167с.
6. Ротштейн А.П. Интеллектуальные технологии идентификации: нечеткие множества, генетические алгоритмы, нейронные сети. – Винница: УНІВЕРСУМ-Вінниця, 1999. – 320с.
7. Цыпкин Я.З. Основы информационной теории идентификации. – М.: Наука. – 1984. – 320с.