

УДК 004.82:322(075):621.311.004 DOI: 10.20998/2411-0558.2018.24.09

А. А. АЗАРЯН, д-р техн. наук, проф., ДВНЗ "КНУ", Кривий Ріг,

І. А. КОТОВ, канд. техн. наук, проф., ДВНЗ "КНУ", Кривий Ріг

ОЦІНКА ЕФЕКТИВНОСТІ ТЕЗАУРУСА ПРОФЕСІЙНОЇ ЛЕКСИКИ В ОНТОЛОГІЯХ СИСТЕМ ПІДТРИМКИ РІШЕНЬ ДЛЯ КЕРУВАННЯ РЕЖИМАМИ ЕНЕРГОСИСТЕМ

Розглянуті основні типи тезаурусів і сфери їх застосування в системах професійних онтологій. Досліджені формальні моделі синтезу онтологій професійних областей, в тому числі, професійної галузі керування режимами енергосистем. Детально досліджена проблема інтегрованого подання знань в експертних системах підтримки рішень оперативного диспетчерського персоналу енергосистем. Поставлено завдання розробки и застосування методів оцінки ефективності тезауруса професійної лексики уніфікованої онтології подання знань. Результатами роботи є дослідження статистичних залежностей ефективності тезауруса професійної лексики від об'єму вибірки і ступеня її спеціалізованості, формальна система подання професійного тезауруса висловлювань галузі управління режимами енергооб'єктів. Іл.: 2. Бібліогр.: 15 назв.

Ключові слова: онтологія; тезаурус; лексика; висловлювання; енергосистема; керування; експертна система; диспетчерський персонал.

Постановка проблеми. В даний час оперативне керування складними технічними комплексами в режимі реального часу протікання технологічних процесів набуває особливої актуальності. Виникнення аварійних ситуацій значно ускладнює завдання оперативного управління. Особа, яка приймає управлінські рішення, не в змозі адекватно реагувати на розвиток нештатної ситуації через психофізіологічні обмеження і емоційне перевантаження внаслідок суттєвої відповідальності. Отже, проблема оперативного керування таким складним об'єктом як енергосистема, повинна бути розв'язана шляхом автоматизації засобів інтелектуальної підтримки рішень диспетчерського персоналу. Таким засобом автоматизації управління є спеціалізовані експертні системи підтримки рішень. Останні повинні інтегруватися в оперативні інформаційно-керуючі комплекси автоматизованої системи диспетчерського управління енергосистеми – SCADA. Центральною ланкою будь-яких інтелектуальних систем є бази знань і, разом з ними, моделі подання та обчислення професійної лексики. Уніфікованої моделлю подання професійних знань може служити онтологія, побудована на базі професійного тезауруса. Ефективність роботи системи підтримки рішень в значній мірі залежить від ефективності використання професійного тезауруса. Таким чином, завдання оцінки ефективності тезауруса професійної лексики в онтологіях системи підтримки рішень

для керування режимами енергосистем є актуальним.

Аналіз літератури. У науковій області теоретичного і практичного створення експертних систем, в тому числі, систем підтримки рішень працюють вітчизняні та зарубіжні вчені. Перш за все, слід відзначити роботи Башликова А.А., Вагіна В.Н., Венделіна А.Г., Геловані В.А., Евланова Л.Г., Заде Л.А., Крєтова В.С., Ларичева О.І., Любарського Ю.Я., Лорьєра Ж.Л., Люгєра Дж.Ф., Поспєлова Д.А., Саати Т., Самойлова В.Д., Таунсенда К., Терелянського П.В., Ульмана Дж., Уотермена Д., Фішберна П.К., Форсайта Ф., Чаплінського Ю.П. та багато інших. Окремо необхідно відзначити наукові результати в галузі інтелектуалізації управління режимами електроенергетичних систем – роботи Баркалова С.А., Баркан Я.Д., Бартоломея П.І., Башликова А.А., Бердникова Р. Н., Буткевича А.Ф., Глазунової А.М., Гришанова С.О., Любарського Ю. Я., Місріханова М.Ш., Панасецкого Д.А., Поспєлова Г.Є., Чебана В.М. та інших [1 – 7].

Проведений аналіз досліджень і публікацій показує актуальність і великий інтерес до проблеми побудови та використання баз знань при побудові АСДУ енергосистем. З іншого боку, стає очевидною відсутність єдиних підходів до вирішення даної проблеми. У роботі пропонується рішення на основі об'єднання (уніфікації) онтологій професійних знань і проводиться оцінка їх ефективності.

Мета статті – розробка математичної моделі професійного тезауруса і оцінка його ефективності в рамках загальної моделі онтології.

Багато існуючих технічних рішень по використанню різних форм подання знань представляють науковий інтерес, однак, не відрізняються значною ефективністю. Додаткові обмеження на використання тих чи інших форм подання знань вносять специфічні професійні області. Аналіз показав, що сьогодні не існує єдиної методології подання і використання знань в системах підтримки рішень [8]. На основі проведених досліджень було поставлено завдання розробки і впровадження в структуру підсистем SCADA енергосистем узагальнюючих інкорпорованих моделей подання професійних знань, реалізованих через формалізми онтологій [9, 10]. Природно, що при цьому виникає супутнє завдання оцінки ефективності використання професійного тезауруса.

Оцінка ефективності професійного тезауруса. Визначимо систему атомарних висловлювань та їх властивості як множину першого шару системи інкорпорації знань [8, 11, 12].

Задаємо безліч символів загальної лексики.

Множина цифр: $A_d = \{0, 1, \dots, 9\}$.

Множина символів спеціальної професійної лексики – спеціальні знаки: $A_p = \{\text{†}, \text{⊗}, \text{⌢}, \text{♠}, \text{⌘}, \text{⦶}, \text{⦿}, \text{⋈}, \text{⚡}, \text{⚙}, \text{⚖}, \text{⚗}, \text{⚛}, \text{⚞}, \text{⚟}, \text{⚝}, \text{⚜}, \text{⚛}, \text{⚞}, \text{⚟}, \text{⚝}, \text{⚜}\} \cup \{\square_p\}$, де $\{\square_p\}$ – множина інших символів спеціальної професійної лексики.

Множина символів спеціальних професійних термінів і сленгу: $A_{sl} = \{\text{блінкер} - \text{сленгове назва вказівного (сигнального) реле з ручним поверненням, квітіровать (приймати до відома, підтверджувати прийом інформації), посадка (характер з'єднання деталей, який визначається зазором або натягом), тепловідвід (компонент, який проводить і розсіює тепло від об'єкта), ключка (пристрій для ручного включення масляних вимикачів з електромагнітним приводом), коротун (коротке замикання), ЛЕП (лінія електропередачі), ПС (підстанція), адмін (системний адміністратор), апгрейд (збільшення продуктивності системи, шляхом заміни модулів або додавання додаткових елементів), глюк (якась помилка в апаратно-програмній системі)} \cup \{\square_{sl}\}$, де $\{\square_{sl}\}$ – множина інших символів спеціальних професійних термінів і сленгу.

Множина символів абревіатур професійної області: $A_{ab} = \{\text{АВР (Автоматичне включення резервного живлення і обладнання), АЛАР (Автоматика ліквідації асинхронного режиму), АПВ (Автоматичне повторне включення), АПНУ (Автоматика запобігання порушенню стійкості), АСКОЕ (Автоматизована система комерційного обліку електроенергії), АЧР (Автоматична частотна розвантаження), ВЛ (повітряна лінія), КВП (Контрольно-вимірювальні прилади і автоматика), ТСН (трансформатор власних потреб), ПРВВ (пристрій резервування відмови вимикача)\} \cup \{\square_{ab}\}$, де $\{\square_{ab}\}$ – множина інших символів абревіатур професійної області.

Множина символів, що утворюють нетермінальний алфавіт:

$$A_N = \{\text{буква, символ, знак, цифра, число, слово, лексема, висловлювання}\}.$$

Якщо прийняти, що всі атомарні висловлювання без дублікатів, то властивості шару атомарних висловлювань БЗ повинні бути наступними:

$$S^{c_i} = S_1^{c_i} \cup S_2^{c_i} \cup \dots \cup S_m^{c_i} \cup \dots \cup S_{n_g}^{c_i} = \bigcup_{m=1}^{n_g} S_m^{c_i}, \quad (1)$$

$$S_1^{c_i} \cap S_2^{c_i} \cap \dots \cap S_m^{c_i} \cap \dots \cap S_{n_g}^{c_i} = \bigcap_{m=1}^{n_g} S_m^{c_i} = \emptyset, \quad (2)$$

$$\forall S_m^{c_i} (S_m^{c_i} \subseteq S^{c_i}), \quad (3)$$

$$\forall S_p^{c_i} \forall S_q^{c_i} (S_p^{c_i} \neq S_q^{c_i}), \quad (4)$$

де $c_i \in C$, $i = \overline{1, n_c}$, C – множина контекстів; n_c – кількість контекстів; n_g – кількість висловлювань, які відносяться до контексту c_i .

Синтаксис атомарних висловлювань:

$\langle \text{вислів} \rangle ::= \langle \text{лексема} \rangle \mid \langle \text{вислів} \rangle \langle \text{лексема} \rangle$;

$\langle \text{лексема} \rangle ::= \langle \text{число} \rangle \mid \langle \text{слово} \rangle$;

$\langle \text{число} \rangle ::= \langle \text{число} \rangle \mid \langle + \rangle \langle \text{число} \rangle \mid \langle - \rangle \langle \text{число} \rangle$;

$\langle \text{число} \rangle ::= \langle \text{цифра} \rangle \mid \langle \text{число} \rangle \langle \text{цифра} \rangle \mid \langle . \rangle \langle \text{число} \rangle$;

$\langle \text{цифра} \rangle ::= 0 \mid 1 \mid 2 \mid 3 \mid 4 \mid 5 \mid 6 \mid 7 \mid 8 \mid 9$;

$\langle \text{слово} \rangle ::= \langle \text{буква} \rangle \mid \langle \text{символ} \rangle \mid \langle \text{знак} \rangle \mid \langle \text{слово} \rangle \langle \text{буква} \rangle$,

де $\forall \langle \text{буква} \rangle (\langle \text{буква} \rangle \in A_l)$;

$\forall \langle \text{символ} \rangle (\langle \text{символ} \rangle \in A_s)$;

$\forall \langle \text{знак} \rangle (\langle \text{знак} \rangle \in (A_p) \vee \langle \text{знак} \rangle \in (A_{sl}) \vee \langle \text{знак} \rangle \in (A_{ab}))$.

Всі висловлювання атомарні і не пов'язані один з одним. Тому вони не утворюють графової структури [13, 14] –

$$\forall s_k^{c_i}, k = \overline{1, n_s} (\bigvee_{k=1}^{n_s} s_k^{c_i} \in S_m^{c_i}). \quad (5)$$

Тоді мова атомарних висловлювань для предметної області і деякого m -го класу контексту визначиться в такий спосіб

$$L(G)^{c_i} = \langle \Sigma, N, P, S \rangle, \quad (6)$$

де G – формальна граматика;

Σ – основний кінцевий термінальний алфавіт;

N – допоміжний кінцевий нетермінальний алфавіт;

P – правила підстановки (продукції) формальної граматики:

$$\exists a, \exists b, (a, b) \in P : a \rightarrow b ;$$

S – стартовий нетермінальний символ граматики G ;

$$N \cap \Sigma = \emptyset \text{ и } P \subset ((N \cup \Sigma)^+ \times (N \cup \Sigma)^*).$$

Визначимо правила формальної граматики – P з урахуванням 8 нетерміналів в формальній мові $L(G)^{c_i}$:

$S \rightarrow \langle \text{лексема} \rangle$
 $S \rightarrow \langle \text{лексема} \rangle \langle \text{вислів} \rangle$
 $\langle \text{лексема} \rangle \rightarrow \langle \text{число} \rangle$
 $\langle \text{лексема} \rangle \rightarrow \langle \text{слово} \rangle$
 $\langle \text{число} \rangle \rightarrow \langle + \rangle \langle \text{число} \rangle$
 $\langle \text{число} \rangle \rightarrow \langle - \rangle \langle \text{число} \rangle$
 $\langle \text{число} \rangle \rightarrow \langle \text{цифра} \rangle$
 $\langle \text{число} \rangle \rightarrow \langle \text{число} \rangle \langle \text{цифра} \rangle$
 $\langle \text{слово} \rangle \rightarrow \langle \text{буква} \rangle$
 $\langle \text{слово} \rangle \rightarrow \langle \text{слово} \rangle \langle \text{буква} \rangle$
 $\langle \text{слово} \rangle \rightarrow \langle \text{символ} \rangle$
 $\langle \text{слово} \rangle \rightarrow \langle \text{слово} \rangle \langle \text{символ} \rangle$
 $\langle \text{слово} \rangle \rightarrow \langle \text{знак} \rangle$
 $\langle \text{слово} \rangle \rightarrow \langle \text{слово} \rangle \langle \text{знак} \rangle$

Модель професійної онтології для шару бази знань атомарних висловлювань – тезауруса професійної лексики визначиться трійкою:

$$O = \langle X, R, F \rangle. \quad (7)$$

Визначимо складові кортежу

$$X^{c_0} = S^{c_0} = \{s_k^{c_0} \mid k = \overline{1, n_s}\}$$

або

$$X^{c_0} = \{s_k^{c_0} \mid k = \overline{1, n_s}\},$$

де c_0 – один певний (будь-який) контекст, до якого відносяться висловлювання, які беруть участь у виразах. Так як розглядається приклад єдиного контексту для всіх висловлювань, він далі не вказується.

Так як для даного шару відносини між висловлюваннями відсутні (тобто – висловлювання використовуються ізольовано один від одного), то $R = \emptyset$.

Визначимо множину функцій інтерпретації F :

$$S = \{\{S_1\}, \{S_2\}\}, \quad (8)$$

де $S_1 \cup S_2 = S$, $S_1 \cap S_2 = \emptyset$.

Тоді

$$\exists (s_{1i} \in S_1, s_{21}, s_{22}, \dots, s_{2k} \in S_2), \quad (9)$$

$$s_{1i} = f(s_{21}, s_{22}, \dots, s_{2k}), f \in F.$$

У більш загальному випадку кількість тих, що інтерпретують і інтерпретованих підмножин може бути довільна. Тоді вони можуть бути розбиті на ситуаційні пари в залежності від специфіки модельованих завдань:

$$S = \{[\{S_{11}\}, \{S_{21}\}], [\{S_{12}\}, \{S_{22}\}], \dots, [\{S_{1j}\}, \{S_{2j}\}], \dots, [\{S_{1N}\}, \{S_{2N}\}]\}, \quad (10)$$

$$\bigcup_{j=1}^N [S_{1j} \cup S_{2j}] = S. \quad (11)$$

Для кожної j -ої пари отримаємо:

$$\exists (s_{1ij} \in S_{1j}, s_{21j}, s_{22j}, \dots, s_{2kj} \in S_{2j}),$$

$$s_{1ij} = f_j(s_{21j}, s_{22j}, \dots, s_{2kj}), f_j \in F.$$

Для реалізації функції інтерпретації конкретний підхід буде залежати від предметної області, вимог до бази знань і семантики концептуального шару. В якості одного з варіантів прийmemo

$$S_2 = \{(s_{2j}, I_j) \mid j = \overline{1, n}; n < N\}, \quad (12)$$

де I_j – індекс релевантності між підмножиною термінів S_{2j} і концептом s_{1j} ; N – кількість висловлювань, які відносяться до даного контексту.

Тоді функція інтерпретації

$$f_j : Op(\{(s_{2j}, I_j)\}) \rightarrow (s_{1j}, I_j), \quad (13)$$

де Op – операція агрегації термінів, наприклад, перерахування з упорядкуванням, конкатенація, вибір найбільш ймовірного, вибір найбільш значимого, вибір самого нового, логічні зв'язки, логічне слідування і т. і.

Таким чином, в результаті отримана формальна модель уніфікованої онтології для рівня 0 ієрархії професійних онтологій типу "активний словник" з наступною узагальненою структурою

$$O = \langle \bigcup_{j=1}^N [S_{1j} \cup S_{2j}], \emptyset, \{F\} \rangle. \quad (14)$$

Як приклад джерела знань професійної області використовуємо фрагменти з документа [15]. Введемо безлічі атомарних висловлювань професійного тезауруса із зазначенням індексу релевантності

$$S = \{[\{S_1\}, \{S_2\}]\}.$$

Очевидно, що професійний контекст може бути визначений таким чином

$$c_i \in C,$$

де $c_i = \{\text{"Дії оперативного персоналу при втраті власних потреб на ПС"}\}.$

Визначимо підмножину концептів.

$S_1 = \{s_{11} = \text{"номер розділу"}, 1; s_{12} = \text{"відключення робочого трансформатора власних потреб"}, 2; s_{13} = \text{"напруги немає на одній із секцій (можлива відмова у включенні від АВР секційного вимикача)"}, 3; s_{14} = \text{"після відновлення напруги на секціях власних потреб черговий зобов'язаний"}, 4; s_{15} = \text{"при відсутності резерву (перевірка показала, що відключення сталося не від внутрішніх пошкоджень, а внаслідок перевантаження, зовнішнього к. з. або несправності в ланцюгах захисту)"}, 5\}.$

Аналогічним чином визначаються інші елементи підмножини концептів.

Визначимо підмножину термінів.

$S_2 = \{s_{21} = \text{"12"}, 1; s_{22} = \text{"в разі"}, 2; s_{23} = \text{"необхідно перш за все перевірити"}, 2; s_{24} = \text{"відновлення напруги на секціях в результаті дії АВР 0,4 кВ"}, 2; s_{25} = \text{"якщо"}, 2; s_{26} = \text{"на секціях немає напруги"}, 2; s_{27} = \text{"черговий персонал зобов'язаний"}, 2; s_{28} = \text{"включити резервний тр-р і тим самим подати напругу на секцію 0,4 кВ"}, 2; s_{29} = \text{"слід"}, 3; s_{210} = \text{"відключити вступної автомат"}, 3; s_{211} = \text{"включити секційний вимикач ключем управління"}, 3; s_{212} = \text{"по блінкерам які спрацювали визначити, від якого захисту відключився тр-р"}, 4; s_{213} = \text{"провести огляд тр-ра і його приєднань і повідомити диспетчеру і керівництву підрозділу"}, 4; s_{214} = \text{"тр-р може бути включений повторно"}, 5\}.$

Аналогічним чином визначаються інші елементи підмножини термінів.

В якості класифікатора обрана одна ознака – семантична приналежність до документа [15].

Наведемо приклади деяких функцій інтерпретації в рамках отриманої онтології:

$$f_1 : Op(\{(s_{21}, 1)\}) \rightarrow (s_{11}, 1);$$

$$f_2 : Op(\{(s_{23}, 2), (s_{24}, 2), (s_{25}, 2), (s_{26}, 2), (s_{27}, 2), (s_{28}, 2)\}) \rightarrow (s_{12}, 2);$$

$$f_3 : Op(\{(s_{29}, 3), (s_{210}, 3), (s_{211}, 3)\}) \rightarrow (s_{13}, 3);$$

$$f_4 : Op(\{(s_{212}, 3), (s_{213}, 3)\}) \rightarrow (s_{14}, 4);$$

$$f_5 : Op(\{(s_{214}, 5)\}) \rightarrow (s_{15}, 5).$$

Аналогічним чином визначаються інші елементи підмножини функцій інтерпретації F .

У роботі був побудований більш детальний тезаурус професійної лексики.

Прийmemo, що потужність множини тезауруса професійної лексики не перевищує 0xFFFF. Тоді зберігання кодів висловлювань тезауруса можна робити, наприклад, в даних типу unsigned short int. При цьому були отримані наступні співвідношення обсягів вихідних висловлювань професійної лексики, кодів тезауруса онтології і ефективності тезауруса, характерних для розглянутої фрагментарною моделі професійної області, які наведені в табл.

У табл. використовуються наступні позначення: n – кількість розділів (по наростаючій); V_P – обсяг вибірки професійної лексики; V_S – обсяг висловлювань тезауруса; ΔV – текстова надмірність; V_T – обсяг кодів тезауруса; $E_T = \frac{V_P - V_T}{V_P} * 100\%$ – ефективність тезауруса.

Графіки, що ілюструють показники ефективності використання тезауруса онтологічної моделі професійної області диспетчерського управління енергосистемою (рис. 1, рис. 2), наведені нижче.

Ефективність графіків на рис. 1 і рис. 2 полягає в можливості наочної оцінки залежності зростання ефективності тезауруса розробленої моделі онтології від початкового об'єму професійної лексики бази знань. Крім цього, проілюстроване зростання ліквідованої надмірності вихідного професійного тексту по відношенню до обсягу тезауруса.

Таблиця

Кількісні характеристики фрагменту тезауруса

n	V_P , байт	V_S , байт	ΔV , байт	V_T , байт	E_T , %
1	226	198	28	8	12,39
2	312	274	38	12	12,18
3	508	428	80	16	15,75
4	834	648	186	18	22,30
5	1256	872	384	26	30,57
6	1612	1088	524	28	32,51
7	2164	1354	810	28	37,43
8	3224	1536	1688	32	52,36
9	4178	1682	2496	32	59,74
10	5854	1728	4126	34	70,48

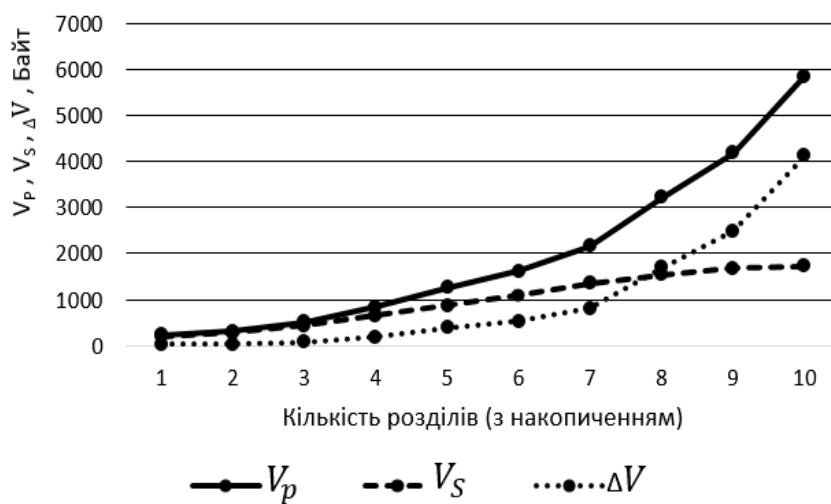


Рис. 1. Обсяги вибірки професійного тезауруса

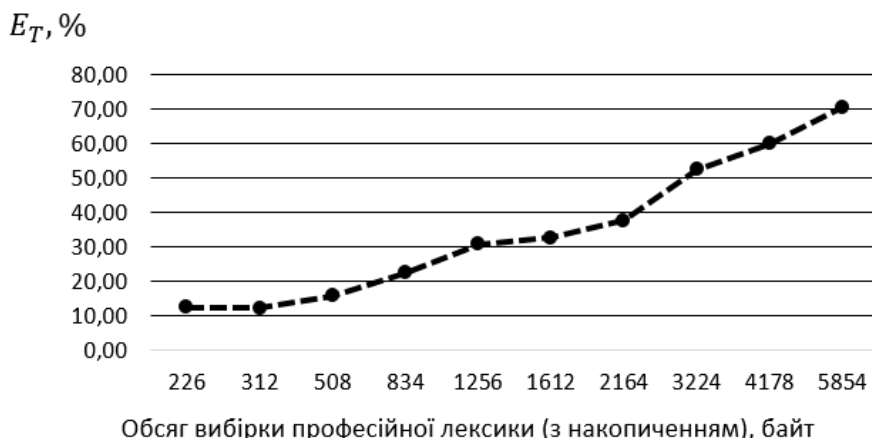


Рис. 2. Оцінка ефективності професійного тезауруса в залежності від обсягу вибірки

Продемонструємо практичну корисність отриманих результатів для трьох перших рядків табл.

Вибірка професійної лексики для розділу 1:

"12.1. У разі відключення робочого трансформатора власних потреб необхідно перевірити відновлення напруги на секціях в результаті дії АВР 0,4 кВ. Якщо на секціях немає напруги, черговий персонал зобов'язаний включити резервний тр-р і подати напругу на секцію 0,4 кВ".

Обсяг вибірки – 226.

Тезаурус:

"відключення робочого трансформатора власних потреб
перевірити відновлення напруги на секціях
дії АВР 0,4 кВ
черговий персонал зобов'язаний включити резервний тр-р
подати напругу на секцію
на секціях немає напруги
необхідно
якщо".

Обсяг тезауруса: 198. Текстова надмірність $226 - 198 = 28$. Кількість необхідних кодів: 8. Ефективність тезауруса в процентах: $28/226 \cdot 100 = 12,39\%$.

Вибірка професійної лексики для розділів 1 і 2:

"12.1. У разі відключення робочого трансформатора власних потреб необхідно перевірити відновлення напруги на секціях в результаті дії АВР 0,4 кВ. Якщо на секціях немає напруги, черговий персонал зобов'язаний включити резервний тр-р і подати напругу на секцію 0,4 кВ.

12.2. Якщо напруги немає на одній із секцій відключити вступної автомат і включити секційний вимикач".

Обсяг вибірки – 312.

Тезаурус:

"відключення робочого трансформатора власних потреб
перевірити відновлення напруги на секціях
дії АВР 0,4 кВ
черговий персонал зобов'язаний включити резервний тр-р
подати напругу на секцію
на секціях немає напруги
необхідно
якщо
на одній із секцій
відключити вступної автомат
включити секційний вимикач
напруги немає".

Обсяг тезауруса: 274. Текстова надмірність $312 - 274 = 38$. Кількість необхідних кодів: 12. Ефективність тезауруса в процентах: $38/312 \cdot 100 = 12,18\%$.

Вибірка професійної лексики для розділів 1, 2 і 3:

"12.1. У разі відключення робочого трансформатора власних потреб необхідно перевірити відновлення напруги на секціях в результаті дії АВР 0,4 кВ. Якщо на секціях немає напруги, черговий персонал зобов'язаний включити резервний тр-р і подати напругу на секцію 0,4 кВ.

12.2. Якщо напруги немає на одній із секцій відключити вступної автомат і включити секційний вимикач.

12.3. Після відновлення напруги на секціях власних потреб черговий зобов'язаний по блінкеру визначити, від якого захисту відключився тр-р, провести огляд тр-ра і його приєднань і повідомити диспетчеру і керівництву підрозділу".

Обсяг вибірки – 508.

Тезаурус:

"відключення робочого трансформатора власних потреб
перевірити відновлення напруги на секціях
дії АВР 0,4 кВ
черговий персонал зобов'язаний включити резервний тр-р
подати напругу на секцію
на секціях немає напруги
необхідно
якщо
на одній із секцій
відключити вступної автомат
включити секційний вимикач
напруги немає
після відновлення напруги на секціях власних потреб
черговий зобов'язаний по блінкеру визначити
провести огляд тр-ра і приєднань
повідомити диспетчеру і керівництву підрозділу".

Обсяг тезауруса: 428. Текстова надмірність $508 - 428 = 80$. Кількість необхідних кодів: 16. Ефективність тезауруса в процентах: $80/508 \cdot 100 = 15,75\%$.

З наведених прикладів видно, що для професійної лексики є доцільною побудова онтологій на основі тезаурусів термінів і лексичних концептів. Причому з ростом обсягу бази знань V_p значно збільшується обсяг ліквідованої текстової надмірності ΔV , і зростає ефективність тезауруса E_T . Ця обставина дозволить на практиці будувати компактні, уніфіковані бази знань професійних областей і ефективні системи підтримки рішень.

Висновки. Наукова новизна полягає в наступному:

1. Розроблено модель онтології атомарних висловлювань, яка може використовуватися в якості базису для подання практично будь-яких форм знань без додаткових змін (відповідно до моделі, запропонованої автором в статті [8]).

2. На основі отриманих результатів можна зробити висновок, що для вузько спеціалізованих лексичних областей є доцільним перехід від алфавітного до професійно-ієрогліфічного подання інформації.

3. Запропоновано будувати тезауруси професійної лексики на основі лінгвістичних змінних, змістом яких є переважно професійний сленг (з урахуванням їх можливої фазифікації).

4. Обґрунтовано необхідність врахування поняття і моделі контексту в формальних моделях мови тезауруса (атомарних

висловлювань) і онтології професійної лексики, так як інтерпретація висловлювань тезауруса залежить від навколишнього контексту і його потужності (в теоретико-множинному сенсі).

5. Досліджено ефективність використання тезауруса спеціальної лексики в вузько професійної області. Показане пропорційне зростання ефективності використання тезауруса від ступеня специфічності (спеціалізованості) професійної області та обсягу лексичного матеріалу.

Практична корисність отриманих результатів полягає в наступному:

1. Обсяг тезауруса набагато менше обсягу вихідного лексичного матеріалу, тому побудова, подання, сприйняття і розуміння (інтерпретація) висловлювань на основі тезауруса значно спрощується.

2. Використання тезауруса стандартизує комп'ютерну автоматизацію побудови і використання баз знань в інтелектуальних системах.

3. Розроблений підхід дає можливість будувати уніфіковані системи підтримки рішень для різних професійних областей.

4. Як наслідок п. 3, запропоновані моделі дозволяють скоротити час розробки систем підтримки прийняття рішень та суттєво знизити їх вартість.

Можна зробити висновок про те, що використання моделі онтології по відношенню до професійної області є доцільним внаслідок специфічної структури професійної лексики. Застосування онтологічних моделей для різних професійних областей дозволяє уніфікувати форми подання знань і на їх основі будувати ефективні інтелектуальні системи підтримки рішень. Отримані залежності показують, що з ростом обсягу задіяної професійної лексики текстова надмірність зростає швидше, ніж обсяг тезауруса, чим пояснюється зростання ефективності останнього. Очевидно, можна очікувати, що для загальноживаної лексики ефект використання онтологічних моделей буде значно нижче внаслідок необхідності використання тезауруса набагато більшого обсягу. Тому онтологічні моделі доцільно застосовувати в специфічних професійних галузях, що відрізняються наявністю широкого спектру професійної термінології (професійного сленгу). Використовуючи отримані результати дослідження, в подальшому планується побудова онтологічних моделей не тільки на базі тезауруса атомарних висловлювань професійної лексики, але також на семантичних і продукційних онтологічних моделях.

Список літератури:

1. *Башильков А.А.* Основы конструирования интеллектуальных систем поддержки принятия решений в атомной энергетике: учебник / *А.А. Башильков, А.П. Еремеев.* – М.: ИНФРА-М – 2017 – 350 с.
2. *Daniel J. Power* Support, Analytics And Business Intelligence, 3rd Edition / *J. Power*

Daniel, Heavin Decision Ciara. – NY: Business Expert Press, LLC. – 2017. – 182 p.

3. Глазунова А.М. Решение задач диспетчерского управления интеллектуальными электроэнергетическими системами на базе методов оценивания состояния / А.М. Глазунова, И.Н. Колосок // Сб. докл. Всерос. научно-практической конф. "Энергетика России в XXI веке. Инновационное развитие и управление". – Иркутск. – 2015. – С. 1-8.

4. Negnevitsky M. Intelligent Approach for Preventing Large-Scale Emergencies in Electric Power Systems / M. Negnevitsky, N. Tomin, D. Panasetsky, V. Kurbatsky // IEEE International Conference on Electric Power Engineering PowerTech. – 2013. – Grenoble, France, 16-20 June. – 2013. – P. 1-6.

5. Гришанов С.О. Реалізація експертної системи для діагностування генераторів теплових електричних станцій / С.О. Гришанов, М.О. Канашиевич // Збірник наукових праць X Міжнародної науково-технічної конференції молодих учених і спеціалістів у місті (Кременчук 28-29 березня 2012р.). – Кременчук: КНУ, 2012. – С. 305-306.

6. Баркалов С.А. Введение в системное проектирование интеллектуальных баз знаний / С.А. Баркалов, А.В. Душкин, С.А. Колодяжный. – М.: Горячая линия-Телеком. – 2017. – 108 с.

7. Антамошин А.Н. Интеллектуальные системы управления организационно-техническими системами / А.Н. Антамошин, О.В. Близнова, А.А. Большаков и др. – М.: Горячая линия – Телеком. Издание 2-е, стер. – 2016. – 160 с.

8. Котов І.А. Автоматизація інтелектуальних систем підтримки рішень оперативного керування шляхом інкорпорації професійних онтологій / І.А. Котов // Вісник НТУ "ХПІ". Серія: Інформатика і моделювання. – Харків: НТУ "ХПІ". – 2016. – № 44 (1216). – С. 63-76

9. Величко В.Ю. Комплексные инструментальные средства инженерии онтологий / В.Ю. Величко, К.С. Малахов, В.В. Семенов, А.Е. Стрижак // International Journal "Information Models and Analyses". – Vol. 3. – № 4. – 2014. – С. 336-360.

10. Самойлов Ю.В. Обзор программных приложений для реализации онтологического подхода к управлению знаниями / Ю.В. Самойлов // Сборник статей VI Международной научно-практической конференции "Инновационные научные исследования". – М.: МЦНС "Наука и просвещение". – 2017. – С. 82-86

11. Морозова В.А. Представление знаний в экспертных системах: учебное пособие / В.А. Морозова, В.И. Паутов. – Екатеринбург: Изд-во Урал. ун-та. – 2017. – 120 с.

12. Кравченко Ю.А. Онтологический подход формирования информационных ресурсов на основе разнородных источников знаний / Ю.А. Кравченко, В.В. Марков // Известия ЮФУ. Технические науки. – 2013. – № 7 (144). – С. 116-120.

13. Зыков А.Г. Математическая логика / А.Г. Зыков, В.И. Поляков, В.И. Скорубский. – СПб: НИУ ИТМО, 2013. – 131 с.

14. Колмогоров А.Н. Математическая логика. Введение в математическую логику / А.Н. Колмогоров, А.Г. Драгалін. – М.: Ленанд. – 2017. – 240 с.

15. Инструкция ДС-8 по предотвращению и ликвидации технологических нарушений в электрической части электростанций и электрических сетей региона Днепровской ЭС / Министерство топлива и энергетики Украины. Государственное предприятие Национальная энергетическая компания "Укрэнерго". – Запорожье: Днепровская электроэнергетическая система. – 2008. – 67 с.

References:

1. Bashlykov, A.A., and Ereemeev, A.P. (2017), *Fundamentals of designing intelligent decision support systems in nuclear power engineering: textbook*, Moskow, INFRA-M, 350 p.
2. Daniel, J. Power, and Ciara Heavin Decision (2017), *Support, Analytics And Business*

Intelligence, 3rd Edition, NY, Business Expert Press, LLC, 182 p.

3. Glazunova, A.M., and Kolosok, I.N. (2015), "Solving the problems of dispatching control of intelligent electric power systems on the basis of methods for estimating the state", *The collection of reports of the All-Russian scientific-practical. Conf. "Energy of Russia in the XXI century. Innovative Development and Management"*, Irkutsk, pp. 1-8.
4. Negnevitsky, M., Tomin, N., Panasetsky, D., and Kurbatsky, V. (2013), "Intelligent Approach for Preventing Large-Scale Emergencies in Electric Power Systems", *IEEE International Conference on Electric Power Engineering PowerTech*, Grenoble, France, 16-20 June, pp. 1-6.
5. Grishanov, S.O. (2012), "Implementation of an expert system for the diagnosis of generators of thermal power plants", *The collection of scientific works X International scientific and technical conference young students and specialists at the Kremenchuk 28-29 March 2012 p.*, Kremenchuk, KNU, pp. 305-306.
6. Barkalov, S.A., Dushkin, A.V., and Kolodyazhny, S.A. (2017), *Introduction to the system design of intellectual knowledge bases*, Moscow, Hot line-Telecom, 108 p.
7. Antamoshin, A.N., Bliznova, O.V., and Bolshakov, A.A. (2016), *Intellectual systems of management of organizational and technical systems*, Moscow, Hot line-Telecom, Edition 2nd, 160 p.
8. Kotov, I.A. (2016), "Automation of intelligent systems of decision support for operational control by incorporating professional ontologies", *Bulletin of NTU "KhPI". Series: Informatics and Modeling*, Kharkiv, NTU "KhPI", No. 44 (1216), pp. 63-76.
9. Velychko, V.Yu., Malakhov, K.S., Semenov, V.V., and Strizhak, A.E. (2014), "Integrated instrumental tools for ontology engineering", *International Journal "Information Models and Analyzes"*, Vol. 3, No. 4, pp. 336-360.
10. Samoilov, Yu.V. (2017), "An overview of software applications for the implementation of an ontological approach to knowledge management", *A collection of articles of the VI International Scientific and Practical Conference "Innovative Scientific Research"*, Moscow, ICPS "Science and Education", pp. 82-86.
11. Morozova, V.A., and Pautov, V.I. (2017), *"Representation of knowledge in expert systems: training manual"*, Yekaterinburg, News of the Ural University, 120 p.
12. Kravchenko, Yu.A., and Markov, V.V. (2013), "Ontological approach of formation of information resources on the basis of heterogeneous sources of knowledge", *Izvestiya SFU. Technical sciences*, No. 7 (144), pp. 116-120.
13. Zykov, A.G., Polyakov, V.I., and Skorubsky, V.I. (2013), *Mathematical logic*, SPb, NIU ITMO, 131 p.
14. Kolmogorov, A.N., and Dragalin, A.G. (2017), *Mathematical logic. Introduction to Mathematical Logic*, Moscow, Lenand, 240 p.
15. *Instruction DC-8 on prevention and liquidation of technological violations in the electric part of power plants and electric networks of the region of the Dnieper Electric Power System* (2008), Ministry of Fuel and Energy of Ukraine. State Enterprise National Energy Company "Ukrenergo", Zaporozhye: Dnieper Electric Power System, 67 p.

Статтю представив д-р техн. наук, проф. ДВНЗ "Криворізький національний університет", зав. кафедрою автоматизованих електромеханічних систем в промисловості та транспорті Сінчук О.М.

Поступила (received) 22.03.2018

Повторно 23.04.2018

Kotov Igor Anatoliyovich, Cand. Tech Sciences, Associate Professor
State Higher Educational Institution "Kryvy Rih National University"
Str. Vyacheslav Chornovil, 41, fl. 23, Kryvy Rih, Ukraine, 50093
Tel.: (067) 970-17-17, e-mail: rioexito@gmail.com
ORCID ID: 0000-0003-2445-6259

Azaryan Albert Aramaysovich, Doctor of Techn. Sciences, Prof.
State Higher Educational Institution "Kryvy Rih National University"
Str. Krivbasovskaya, 54, fl. 12, Kryvy Rih, Ukraine, 50093
Tel. (067) 391-05-43, e-mail: aazaryan325@gmail.com
ORCID ID: 0000-0002-1381-579X

УДК 004.82:322(075):621.311.004

Оцінка ефективності тезауруса професійної лексики в онтологіях систем підтримки рішень для керування режимами енергосистем / Азарян А.А., Котов І.А. // Вісник НТУ "ХПІ". Серія: Інформатика і моделювання. – Харків: НТУ "ХПІ". – 2018. – № 24 (1300). – С. 100 – 116.

Розглянуті основні типи тезаурусів і сфери їх застосування в системах професійних онтологій. Досліджені формальні моделі синтезу онтологій професійних областей. Детально досліджена проблема подання знань в експертних системах підтримки рішень оперативного диспетчерського персоналу енергосистем. Результатами роботи є дослідження статистичних залежностей ефективності тезауруса професійної лексики від об'єму вибірки і ступеня її спеціалізованості, формальна система подання тезауруса висловлювань галузі управління режимами енергооб'єктів. Іл.: 2. Табл.: 1. Бібліогр.: 15 назв.

Ключові слова: онтологія; тезаурус; лексика; висловлювання; енергосистема; керування; експертна система; диспетчерський персонал.

УДК 004.82: 322 (075): 621.311.004

Оценка эффективности тезауруса профессиональной лексики в онтологиях систем поддержки решений для управления режимами энергосистем / Азарян А.А., Котов И.А. // Вестник НТУ "ХПИ". Серія: Інформатика и моделирование. – Харьков: НТУ "ХПИ". – 2018. – № 24 (1300). – С. 100 – 116.

Рассмотрены основные типы тезаурусов и сферы их применения в системах профессиональных онтологий. Исследованы формальные модели синтеза онтологий профессиональных областей. Подробно исследована проблема представления знаний в экспертных системах поддержки принятия решений оперативного диспетчерского персонала энергосистем. Результатами работы является исследование статистических зависимостей эффективности тезауруса профессиональной лексики от объема выборки и степени ее специализированности, формальная система представления тезауруса высказываний в области управления режимами энергообъектов. Ил.: 2. Табл.: 1. Библиогр.: 15 назв.

Ключевые слова: онтология; тезаурус; лексика; высказывания; энергосистема; управление; экспертная система; диспетчерский персонал.

UDC 004.82: 322 (075): 621.311.004

Assessment of the effectiveness of the thesaurus of professional vocabulary in ontologies of decision support systems for managing power system regimes / Azaryan A.A., Kotov I. A. // Herald of the National Technical University "KhPI". Subject issue: Information Science and Modelling. – Kharkov: NTU "KhPI". – 2018. – №. 24 (1300). – P. 100 – 116.

The main types of thesauri and their application in professional ontology systems are considered. Formal models for the synthesis of ontologies of professional fields. The problem of representation of knowledge in expert decision support systems for operational dispatch personnel of power systems has been thoroughly investigated. The task is to develop and apply methods for assessing the effectiveness of the thesaurus of professional vocabulary of the unified ontology of knowledge representation. The results of the research are the statistical dependencies of the vocabulary vocabulary effectiveness on the sample size and the degree of its specialization, the formal system of representation of the thesaurus of the utterances of the management area of the modes of power facilities. Figs.: 2. Tabl.: 1. Refs.: 15 titles.

Keywords: ontology; thesaurus; vocabulary; utterances; power system; management; expert system; dispatcher system.