

Київський національний університет імені Тараса Шевченка
Міністерство освіти і науки України

НАЙДЬОНОВ ІВАН МИХАЙЛОВИЧ



УДК 656.073, 004.934

**ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ ФОРМАЛІЗАЦІЇ ГОЛОСОВОЇ
ІНФОРМАЦІЇ В СИСТЕМАХ ДИСПЕТЧЕРСЬКОГО КОНТРОЛЮ
ЗА РУХОМ АВТОТРАНСПОРТУ**

Спеціальність 05.13.06 — інформаційні технології

Автореферат
дисертації на здобуття наукового ступеня
кандидата технічних наук

Дисертацією є рукопис.

Роботу виконано на факультеті інформаційних технологій, кафедрі технологій управління Київського національного університету імені Тараса Шевченка Міністерства освіти і науки України.

Науковий керівник: доктор технічних наук, професор
Тесля Юрій Миколайович,
Національний авіаційний університет
МОН України, м. Київ,
проректор з молодіжної політики та інноваційного
навчання

Офіційні опоненти: доктор технічних наук, професор,
Зайченко Юрій Петрович,
Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського»
МОН України,
професор кафедри математичних методів системного аналізу

доктор технічних наук, професор,
Голуб Сергій Васильович,
Черкаський державний технологічний університет
МОН України,
професор кафедри математичного та програмного забезпечення автоматизованих систем

Захист відбудеться «23» вересня 2019 р. о 14:00 годині на засіданні спеціалізованої вченої ради Д26.001.51 в навчальному корпусі факультету інформаційних технологій Київського національного університету імені Тараса Шевченка за адресою: 04116, м. Київ, вул. Богдана Гаврилишина, 24, ауд. 310.

З дисертацією можна ознайомитись у науковій бібліотеці імені Михайла Максимовича Київського національного університету імені Тараса Шевченка за адресою: 01033, м. Київ, вул. Володимирська, 60, зала № 12.

Автореферат розіслано «17» серпня 2019 р.

Учений секретар спеціалізованої вченої ради
кандидат технічних наук

Т. В. Латишева

ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

Актуальність теми дослідження. Системи диспетчерського контролю за рухом автотранспорту, призначені ефективно коригувати відхилення від запланованих маршрутів при зіткненні з непередбачуваними обставинами, потребують ефективного обміну повідомленнями між водієм і диспетчером. Різні форми автоматизації диспетчерського контролю (GPS, додатки з сенсорним інтерфейсом, мобільний інтернет) на сьогодні не здатні замінити голосову взаємодію, в якій диспетчер отримує необхідну для прийняття рішень інформацію зокрема про характер і причини відхилень від плану.

Таким чином підвищення ефективності передачі повідомлень за рахунок формалізації голосової взаємодії між водієм та диспетчером є одним із перспективних напрямів удосконалення системи диспетчерського контролю, що робить тему дисертаційного дослідження інформаційних технологій формалізації голосової інформації в системах диспетчерського контролю за рухом автотранспорту **актуальною**.

Питаннями формалізації голосової взаємодії, побудови діалогових систем та сценаріїв голосової взаємодії займалися такі вчені, як Ishiguro H., Iosif E., Herbert D., Lopes J., Khouzaimi H., але їх дослідження не були спрямовані на сферу систем диспетчерського контролю за рухом автотранспорту.

Досліджували голосове управління в транспортних системах Кравченко А. П., Корсун О. Н., Heisterkamp P., Jonsson I.-M., проте результати цих досліджень переважно стосувалися автоматизації голосового управління бортовим обладнанням.

Загалом різні способи покращення диспетчерського контролю при виконанні доставок автотранспортом розглядали Prasanna K., Stopher P. R., Liu S., Govindan K., Papetti A., Gonzalez-Feliu J., Comendador J. A., Baumann H., Quak H., однак їхні роботи не мали на меті формалізацію голосової взаємодії.

Питання формалізації голосової інформації за рахунок переведення її в текст розглядає велика кількість учених, зокрема Пилипенко В. В., Людовик Т., Васильєва Н. Б., Womack B. D., Зірнєєва Г. В., Гладунов С. А., Робейко В. В., Abdel-Hamid O., Zhang Y., Sharma P., Єрмоленко Т. В., Не Y., насамперед шляхом переведення голосу в текст, який, внаслідок необхідності великих словників та доступу до інтернету, має певні обмеження для використання на мобільних пристроях.

У зв'язку з цим на даний час існує необхідність вирішення актуального наукового завдання розробки моделей і методів формалізації голосової інформації в системах диспетчерського контролю за рухом автотранспорту.

Вирішення зазначених суперечностей теорії та практики формалізації голосової інформації може спиратися на дослідження голосового управління, заснованих на теорії несилової взаємодії та рефлекторні системи голосового управління, що належать такими вченим, як Тесля Ю. М., Пилипенко В. В., Чорний О. Ю., Єгорченков А. В.

Зв'язок роботи з науковими програмами і планами.

Дисертаційна робота виконана відповідно до пріоритетного напрямку розвитку інформаційних та комунікаційних технологій, що визначені в Законі України «Про пріоритетні напрями розвитку науки і техніки» на період до 2020 року, та тематичного плану науково-дослідних робіт Київського національного університету імені Тараса Шевченка в рамках науково-дослідної роботи «Розробка теоретико-методологічних основ впровадження систем управління проектами для розвитку підприємств і організацій» (№ держреєстрації 0117U002694), в якій автор брав участь як виконавець, запропонувавши впровадження графу сценаріїв для систем формалізації голосового управління.

Метою дослідження є підвищення ефективності розпізнавання повідомлень у голосовій взаємодії водія з диспетчером на основі розробки та використання інформаційної технології формалізації голосової інформації в системах диспетчерського контролю за рухом автотранспорту.

Для досягнення сформульованої мети поставлено ряд часткових **завдань досліджень**:

- здійснити аналіз сучасних інформаційних систем обробки та формалізації голосової інформації;
- розробити метод формалізації голосової інформації в допоміжних системах диспетчеризації автотранспорту;
- удосконалити математичну модель голосової взаємодії водія та диспетчера в системах диспетчерського контролю за рухом автотранспорту та представити у вигляді повного графу сценаріїв усіх етапів процесу доставки «склад – дорога – точка доставки»;
- адаптувати метод структурної ідентифікації згорткових нейронних мереж для класифікації голосових команд для розпізнавання фонемного тексту;
- поєднати метод інтелектуальних рефлексорних систем з понятійним апаратом теорії нейронних мереж;
- провести експериментальні дослідження на основі математичного моделювання формалізації голосової інформації, отриманої від водіїв, для оптимізації диспетчерського контролю за рухом автотранспорту.

Об'єктом дослідження є процеси голосової взаємодії в системах диспетчерського контролю за рухом автотранспорту.

Предмет дослідження – моделі і методи формалізації голосової взаємодії в системах диспетчерського контролю за рухом автотранспорту.

Методи дослідження. Для досягнення поставленої мети в роботі використано: теорію графів — для опису моделі голосової взаємодії, теорію інформації та теорію несилової взаємодії — для вдосконалення методу інтелектуальних рефлексорних систем, теорію штучних нейронних мереж та методи обробки природної мови — для вдосконалення методу згорткових нейронних мереж.

Наукова новизна отриманих результатів полягає в тому, що в дисертаційній роботі:

- вперше розроблено метод формалізації голосової інформації в системах підтримки диспетчеризації автотранспорту, який на відміну від аналогів поєднує використання інтелектуальних рефлекторних систем та згорткових нейронних мереж, що дає змогу автоматизувати процес передачі голосової інформації;
- удосконалено математичну модель голосової взаємодії водія та диспетчера в системах диспетчерського контролю за рухом автотранспорту, яка на відміну від існуючих представлена у вигляді повного графу сценаріїв усіх етапів процесу доставки «склад – дорога – точка доставки», що дає змогу виділити контексти голосової взаємодії для підвищення точності подальшої формалізації;
- набув подальшого розвитку метод структурної ідентифікації згорткових нейронних мереж для класифікації голосових команд, в якому на відміну від існуючих ведеться розпізнання фонемного тексту, що уможлиблює класифікацію голосових команд без переведення голосу в лексичний текст;
- отримав подальший розвиток метод інтелектуальних рефлекторних систем, який відрізняється від існуючих поєднанням з понятійним апаратом теорії нейронних мереж, що дає можливість оптимізувати значення інформованості та визначеності шляхом навчання моделі методом зворотного розповсюдження помилки.

Практичне значення отриманих результатів полягає в тому, що з використанням наукових результатів закладається можливість підвищення точності та швидкості розпізнавання голосових повідомлень безпосередньо на мобільному пристрої, що покращує можливості диспетчерського контролю за рухом автотранспорту. Розроблені на базі запропонованих особисто автором моделей і методів програмні засоби становлять практичний результат, який впроваджений на підприємстві ТОВ «УІТ», м. Київ.

Особистий внесок здобувача. Наукові положення, розробки та висновки дисертаційної роботи є результатом самостійно проведеного дослідження здобувача. Основні наукові результати, представлені в дисертації, отримані здобувачем особисто.

Апробація результатів досліджень. Основні положення дисертаційної роботи були апробовані на 5-ти міжнародних науково-технічних конференціях, у тому числі:

- XVII Міжнародній науково-технічній конференції «Системний аналіз та інформаційні технології» (м. Київ, 22–25 червня 2015 р.),
- XII Міжнародній конференції «Управління проектами у розвитку суспільства», тема: «Комплексне управління проектами розвитку в умовах нестабільного оточення» (м. Київ, 21-23 травня 2015 р.),
- III Міжнародній науково-практичній конференції «Інформаційні технології та взаємодії» (м. Київ, 8–10 листопада 2016 р.),

- 16th EAGE International Conference on Geoinformatics - Theoretical and Applied Aspects (м. Київ, 15–17 травня 2017 р.),
- IV Міжнародній науково-практичній конференції «Інформаційні технології та взаємодії» (м. Київ, 8–10 листопада 2017 р.).

Публікації. За результатами дисертаційних досліджень опубліковано 11 наукових праць. Основні наукові положення викладено у 5 наукових статтях [1–5], серед яких [1–4] опубліковані у спеціалізованих фахових виданнях України, [5] опубліковано у закордонному науковому виданні. За матеріалами виступів на науково-технічних конференціях опубліковано 5 тез доповідей [6–10]. Додатково результати досліджень відображено в науковій статті [11].

Структура і обсяг роботи. Дисертаційна робота представлена на 246 сторінках друкованого тексту, включає 51 рисунок, 20 таблиць, які розташовані на 31 повній сторінці тексту. Робота складається з вступу, чотирьох розділів, висновків і списку використаних джерел із 142 найменуваннями, який розміщений на 21 сторінці. Основний текст викладений на 124 сторінках роботи.

ОСНОВНИЙ ЗМІСТ РОБОТИ

У **вступі** наведено загальну характеристику роботи, яка підкреслює її актуальність, відповідність науковим темам, наукову новизну та практичне значення, визначено предмет та об'єкт дослідження, сформульовано мету та завдання дослідження.

У **першому розділі** «Проблема формалізації голосової інформації в системах диспетчерського контролю за рухом автотранспорту» проведено аналіз сучасних інформаційних систем формалізації голосової інформації, уточнено проблему, гіпотезу та завдання дослідження.

У результаті аналізу проблеми формалізації голосової інформації в системах диспетчерського контролю за рухом автотранспорту виявлено, що інформаційні технології в управлінні дистрибуцією є достатньо розробленими для забезпечення етапів отримання продукції та її збереження, але недостатньо — для етапу доставки продукції до кінцевих клієнтів, особливо щодо проблеми «останньої милі». При цьому значну роль в управлінні дистрибуцією відіграють процеси голосової взаємодії, автоматизація яких здатна підвищити ефективність системи дистрибуції.

На сучасному етапі голосового управління в організаційно-технічних системах існує проблема своєчасного коригування в необхідних випадках планових маршрутів руху автотранспорту, що інколи призводить до досить великих витрат часу на комунікацію і, відповідно, є найбільш обґрунтованим напрямом формалізації голосової взаємодії.

Автоматизація голосової взаємодії для своєчасного коригування планових маршрутів руху автотранспорту повинна доповнити наявні засоби автоматизації

управління в системах дистрибуції, такі, зокрема, як відстеження руху автомобілів у режимі реального часу за допомогою GPS треку, позаяк існуючі в дистрибуції системи формалізації голосової взаємодії для управління зберіганням є занадто спрощеними для використання в задачах доставки.

Проведений аналіз сучасних інформаційних систем формалізації голосової інформації показав, що традиційно вони включають три етапи: попередня обробка голосової інформації з виділенням ознак, перетворення голосової інформації в текстову та виділення змісту з текстової інформації. Такий підхід не в змозі повністю забезпечити формалізацію голосової взаємодії в задачах диспетчерського контролю за рухом автотранспорту, оскільки переведення голосової інформації в лексичний текст потребує встановлення в кабіні водія потужного обладнання або забезпечення стабільного та швидкісного доступу до Інтернету. Висунуто ідею, що розроблення моделі голосової взаємодії без блоку переведення звуку голосу в текст може принципово покращити формалізацію голосової взаємодії в системах диспетчерського контролю.

Встановлено пріоритетність підходу до голосового управління, заснованого на теорії несилової взаємодії з рефлекторною системою голосового управління. Він передбачає безпосередній аналіз інформаційного повідомлення, вимовленого суб'єктом, з визначенням керуючого впливу з поміж відомих реакцій у певному предметному полі. У результаті проведеного аналізу сформульовано основну ідею, гіпотезу, мету та спрямування дослідження, поставлено завдання досліджень дисертаційної роботи. Визначено необхідність у розробці моделей і методів, що дадуть змогу формалізувати голосову інформацію в системах диспетчерського контролю за рухом автотранспорту.

У **другому розділі** «Підходи до формалізації голосової взаємодії в системі диспетчеризації» визначено науково-методологічні основи формалізації голосової взаємодії в системі диспетчеризації.

Запропоновано концепцію створення системи формалізації голосової взаємодії в задачах диспетчерського контролю за рухом автотранспорту, що має дві складові: (а) інтелектуальні рефлекторні системи голосового управління, що включають блок розпізнавання звукового сигналу та блок виділення його змісту; (б) модель сценаріїв голосової взаємодії на трьох етапах доставки (завантаження на складі, дорога до точки доставки, розвантаження в точці доставки).

Найбільш перспективним напрямом, який дає змогу запропонувати нове принципове рішення і побудувати рефлекторну модель голосової взаємодії в задачах диспетчерського контролю за рухом автотранспорту, є застосування моделей логічних сценаріїв голосової взаємодії, які мають враховувати параметри основних причин невідповідності реальної ситуації запланованому маршруту. При отриманні такої інформації приймається своєчасне рішення про повернення вантажу на склад, про відміну чи відкладення обслуговування однієї точки доставки, щоб мати можливість встигнути на іншу, більш важливу, про зміну маршруту для об'їзду затору або про утворення нового маршруту з резервною машиною тощо.

Структурування моделі за трьома етапами доставки дає змогу вирішувати проблемні моменти із залученням диспетчера для вибору найкращої стратегії і мінімізації втрат. Структурна модель є принциповим алгоритмом побудови графу сценаріїв голосової взаємодії для кожного з суб'єктів (диспетчера та водія).

Розроблена система автоматичного розрахунку планових маршрутів та практика її використання забезпечили накопичення параметрів непередбачуваних ситуацій на плановому маршруті доставки, що впливають на створення сценаріїв голосової взаємодії.

У результаті дослідження методів автоматизації руху автотранспорту в доставці вантажів автотранспортом та параметрів, що впливають на сценарії голосової взаємодії, розроблено евристичний алгоритм, який максимально враховує вимоги логістів та водіїв щодо оптимальності наповнення кожного конкретного маршруту, забезпечуючи глобальну оптимальність сукупності всіх маршрутів при високій швидкості обчислення.

Завдяки формалізації голосової взаємодії дані GPS доповнюватимуться додатковою інформацією щодо закінчення виконання однієї з точок у межах єдиної зупинки і початок виконання наступної точки. Для забезпечення зазначеного необхідно впровадити зручний інтерфейс, який не відволікатиме водія від основного завдання, тобто голосовий інтерфейс, який сприйматиме команди про початок та завершення виконання доставки.

Модель голосової взаємодії запропоновано будувати у вигляді орієнтованого графу, в якому вершини позначають стан системи та діалогові фрази, які буде озвучувати система, а ребра – репліки (стимули), які можуть бути сприйняті системою в кожній конкретній вершині, множина ж всіх ребер, що виходять з однієї вершини, позначатиме перелік стимулів розпізнання відповідно до її стану. У результаті для повноцінного опису запропоновано використовувати такі сутності: «контекст» (або «стан»), «стимул» (або «подія») та «реакція» системи відповідно до стимулу.

Принципи побудови рефлекторних систем на основі теорії несилової взаємодії адаптовано для формалізації голосової взаємодії в системах диспетчерського контролю за рухом автотранспорту.

Під час дослідження принципів побудови рефлекторної системи голосової взаємодії встановлено особливості застосування теорії несилової взаємодії як основи інтелектуальних рефлекторних систем, а також теоретично запропоновано використання рефлекторного методу для побудови рефлекторної системи голосової взаємодії в системах диспетчерського контролю за рухом автотранспорту.

Третій розділ «Методи формалізації голосової інформації в системах диспетчерського контролю за рухом автотранспорту» присвячено створенню моделі та методів формалізації голосової інформації в системах диспетчерського контролю за рухом автотранспорту.

Запропонована класифікація реакцій для водія на етапах «склад – дорога – точка доставки» базується на узагальненні зібраних статистичних даних що-

до процесу доставки різних вантажів автомобільним транспортом у провідних логістичних компаніях України.

Розроблено математичну **модель голосової взаємодії водія та диспетчера в системах диспетчерського контролю за рухом автотранспорту**, яку представлено у вигляді повного графу сценаріїв усіх етапів процесу доставки «склад – дорога – точка доставки» (рис. 1).

Ця модель може бути представлена у вигляді орієнтовного графу G , що складається з множини вершин V та множини ребер E :

$$G = \langle V, E \rangle;$$

$$E = \{ \langle v_i, v_j \rangle | v \in V \}.$$

При цьому існує відношення f_R множини ребер до множини реакцій та відношення f_C множини вершин до множини контекстів, такі, що:

$$f_R : E \rightarrow R, \quad R \subset \mathbb{N}, \quad |R| \leq |E|;$$

$$f_C : V \rightarrow C, \quad C \subset \mathbb{N}, \quad |C| \leq |V|;$$

$$R_V(v_i) = \{ f_R(e) | \forall j : e = \langle v_i, v_j \rangle \in E \};$$

$$\forall i, j : f_C(v_i) = f_C(v_j) \iff R_V(v_i) = R_V(v_j),$$

де $R_V(v_i)$ множина реакцій можливих у вершині v_i .

Для створення контекстної моделі голосової взаємодії в системах диспетчерського контролю за рухом автотранспорту було зібрано статистичні дані, зауваження та коментарі щодо процесу доставки різних вантажів автомобільним транспортом у провідних вітчизняних логістичних компаніях. Здійснено систематизацію та обробку зібраних оригінальних коментарів щодо статусу доставки, що використовуються в різних компаніях, за результатами яких розроблено граф сценаріїв взаємодії водія та диспетчера.

У повному графі сценаріїв усіх етапів доставки (рис. 1) враховано всі виявлені можливі причини затримки або невиконання етапів «склад», «дорога», «точка доставки», а для таких випадків існують вказівки (інструкції) диспетчера щодо подальших дій водія.

Виділено перелік унікальних контекстів голосової взаємодії, формалізація голосової інформації в яких може відбуватися незалежно один від одного, що дає змогу знизити кількість реакцій для розпізнання.

Адекватність розробленої моделі підтверджується її повною відповідністю статистичним даним інцидентів, зібраних за період впровадження системи, та експериментально - порівнянням результатів математичного моделювання з використанням моделі без неї.

Метод формалізації голосової інформації в системах підтримки диспетчеризації автотранспорту з використанням інтелектуальних рефлексорних

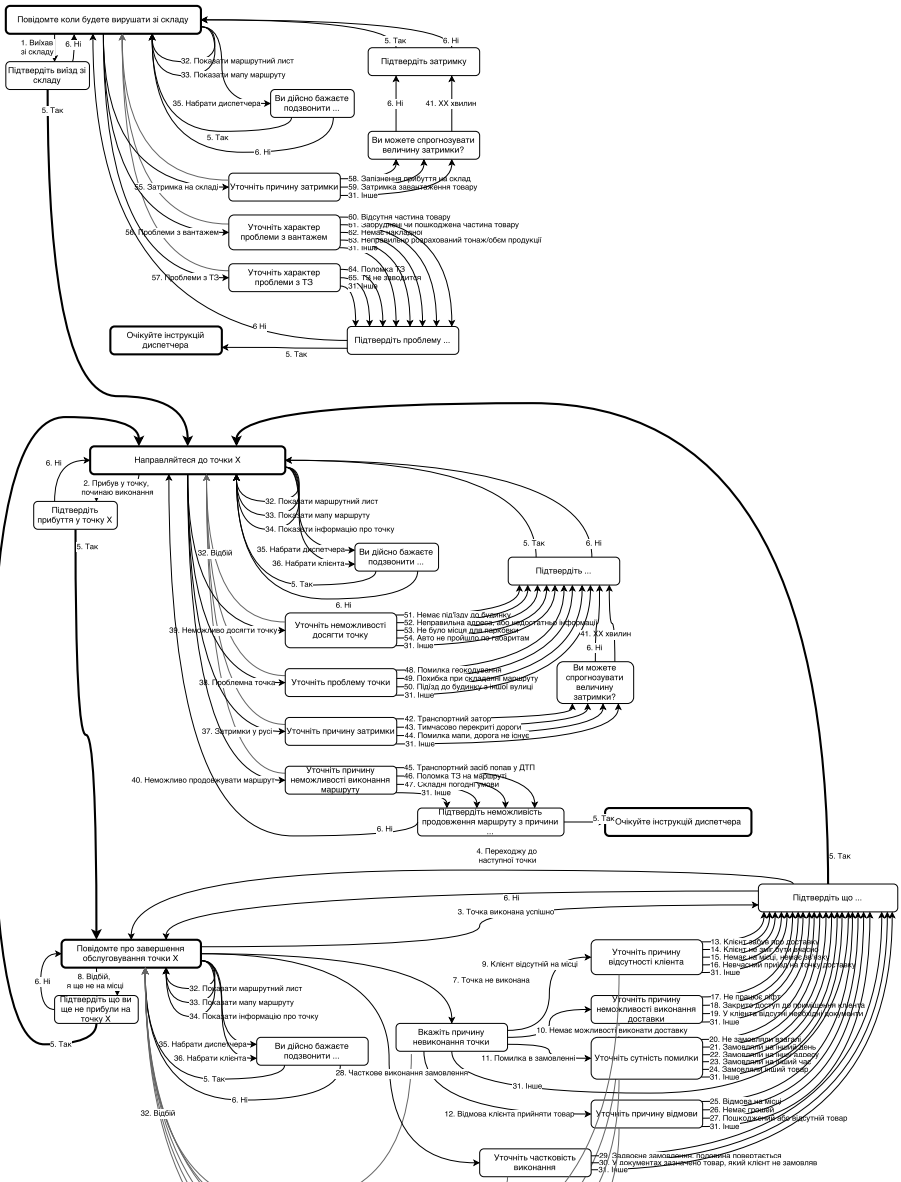


Рис. 1 — Повний граф сценаріїв голосової взаємодії

систем дає змогу автоматизувати процес передачі голосової інформації, уникнувши перевodu звукової інформації в лексичний текст за рахунок використання двох основних модулів - автоматичного фонетичного стенографа та ядра рефлекторної системи голосового управління. Розроблений метод можна представити наступною послідовністю операцій:

- запис фрази, вимовленої водієм: $A_i = \langle a_1, a_2, \dots, a_n \rangle; t = \frac{n}{s}; s = 16 \text{ (kHz)}$;
- перетворення записаної фрази на фонетичний текст за допомогою фонемного стенографа: $P_i = S(A_i); P = \langle p_1, p_2, \dots, p_k \rangle; p_i \in F$;
- класифікація фонемної репрезентації голосової команди: $y_i = C_c(P_i)$;
 - розбиття фонетичного тексту на N-грами фонем різної довжини;
 - розрахунок інтроформаційного впливу кожного N-граму фонем на можливі команди у вибраному контексті;
 - вибір команди з найбільшою ймовірністю;
- виконання відповідної реакції (озвучення відповіді, виконання команди та/або відправка структурованих даних диспетчеру);
- переключення контексту на новий, відповідний до вибраної реакції: $c_{i+1} = f(c_i, y_i)$;
- очікування та запис наступної фрази.

Де: A_i — цифровий аудіозапис команди водія тривалістю t секунд,

n — кількість семплів аудіосигналу,

s — частота дискретизації аудіосигналу,

P_i — представлення команди водія у вигляді фонемного тексту — кортежу фонем довжини k ,

F — множина фонем української мови,

S — фонемний стенограф,

y_i — реакція з моделі голосової взаємодії водія та диспетчера, що відповідає вимовленій команді,

C_c — класифікатор фонемної репрезентації голосових команд відповідно до поточного контексту c_i ,

f — функція визначення наступного контексту залежно від поточного контексту c_i та вибраної реакції y_i .

Таким чином, в систему обробки надходить увесь вхідний набір фонем без виділення слів, команд, пропозицій тощо. Як і в мозку людини при слуханні усного мовлення або читанні листа не розпізнаються букви і слова, а розпізнається сенс, так само відбувається і при використанні рефлекторної системи голосового управління: не потрібно створювати жодних словників, виконувати морфологічний, синтаксичний, семантичний аналіз тексту, а також виділяти слова і команди; робота здійснюється зі звуковим потоком, з якого система формалізації голосової інформації, як і людина, виділяє інформативну частину за максимальною визначеністю.

Метод формалізації голосової інформації в моделі голосової взаємодії водія при диспетчерському контролі за рухом автотранспорту реалізований в системі, що складається з двох основних модулів: автоматичного фонетичного

стенографа і ядра рефлекторної системи голосового управління, поточна реалізація яких визначає умови їх використання в моделі голосової взаємодії.

Для реалізації ядерного компонента рефлекторної системи голосового управління запропоновано дуальну систему класифікації голосових команд, яка може бути налаштована на предметну область і використовувати метод інтелектуальних рефлекторних систем або метод згорткових нейронних мереж залежно від того, який з них показує кращі результати.

Для підвищення ефективності розпізнавання запропоновано удосконалену схему системи формалізації голосової інформації в моделі голосової взаємодії водія та диспетчера (рис. 2), яка включає: (а) моделювання за кожним контекстом із моделі голосової взаємодії та (б) дуальну систему класифікації фонемної репрезентації голосових команд, що дає змогу вибрати кращий метод класифікації для предметної області.

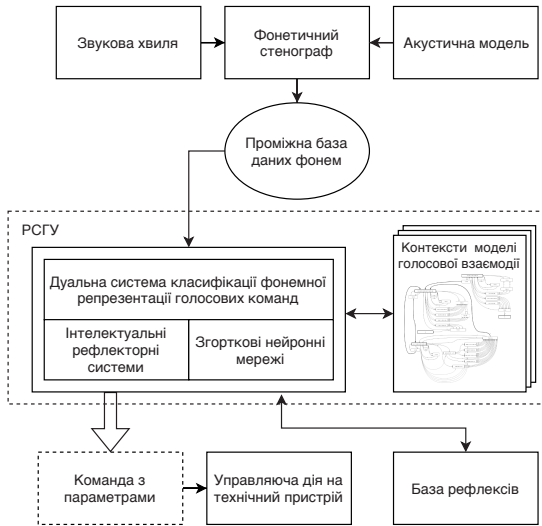


Рис. 2 — Удосконалена схема системи формалізації голосової інформації в моделі голосової взаємодії водія при диспетчерському контролі за рухом автотранспорту

Метод інтелектуальних рефлекторних систем, що використовується для класифікації голосових команд, поєднаний з понятійним апаратом теорії нейронних мереж та представлений у матричній формі, полягає в наступному:

1. Розрахунок визначеності для інтелектуальної системи відносно всіх вхідних N -грам фонем і можливих голосових команд:

$$\begin{aligned}
D_A &= \pm 0.5(P_A \odot (J_{1,p} - P_A) + (J_{1,p} - P_A) \odot P_A - 2J_{1,p})^{\circ \frac{1}{2}}; \\
D_{AB} &= \pm 0.5(P_{AB} \odot (J_{p,q} - P_{AB}) + (J_{p,q} - P_{AB}) \odot P_{AB} - 2J_{p,q})^{\circ \frac{1}{2}}; \\
I_A &= (D_A^{\circ 2} + J_{1,p})^{\circ \frac{1}{2}}; \quad I_{AB} = (D_{AB}^{\circ 2} + J_{p,q})^{\circ \frac{1}{2}},
\end{aligned}$$

- де: $p = |A|$ — потужність множини голосових команд;
 $q = |B| = \sum_{i=s_{\min}}^{s_{\max}} f^i$ — потужність множини N-грам фонем;
 f — кількість фонем в акустичній моделі;
 s_{\min} та s_{\max} — мінімальний та максимальний розміри N-грам;
 $J_{i,j}$ — матриця одиниць розміром $i \times j$;
 P_A — матриця безумовної ймовірності вибору команд з множини A (розмір матриці $1 \times p$);
 D_A — матриця визначеності щодо команд з множини A ;
 I_A — матриця інформованості щодо команд з множини A ;
 P_{AB} — матриця умовної ймовірності вибору команд з множини A при наявності впливу N-граму фонем з множини B (розмір матриці $p \times q$);
 D_{AB} — матриця визначеності щодо команд з множини A при наявності впливу N-граму фонем з множини B ;
 I_{AB} — матриця інформованості щодо команд з множини A при наявності впливу N-граму фонем з множини B ;
 \circ, \odot та \otimes — операції матричного поелементного добутку, піднесення до ступеня та ділення Адамара.

- Отримання додаткової визначеності, що є в N-грамів фонем відносно голосових команд:

$$D_{\Delta} = D_{AB} \circ (J_{p,1} I_A) - I_{AB} \circ (J_{p,1} D_A),$$

де D_{Δ} — матриця додаткової визначеності щодо команд з множини A , яку надає наявність N-граму фонем з множини B (розмір матриці $p \times q$).

- Розрахунок сумарного впливу на голосову команду, реакцію інтелектуальної системи всіх наявних N-грамів фонем:

$$D_{\Sigma} = X D_{\Delta}; \quad I_{\Sigma} = (D_{\Sigma}^{\circ 2} + J_{n,q})^{\circ \frac{1}{2}},$$

- де: n — кількість вхідних команд для розпізнання або навчання системи;
 X — вхідна матриця команд для розпізнання або навчання системи, представлена у форматі «мішок N-грам фонем», тобто матриці розміром $n \times p$, де $x_{ij} = 1$, якщо для відповідної голосової команди i існує N-грам фонем j , в інакшому випадку $x_{ij} = 0$;
 D_{Σ} — матриця сумарної додаткової визначеності щодо команд з множини A під впливом всіх N-грамів фонем з множини B (розмір матриці $n \times q$);

I_{Σ} — матриця сумарної додаткової інформованості щодо команд з множини A під впливом всіх N -грамів фонем з множини B .

4. Обчислення нової інформованості та визначеності голосової команди:

$$D_Y = D_{\Sigma} \circ (J_{n,1} I_A) - I_{\Sigma} \circ (J_{n,1} D_A); \quad I_Y = (D_Y^{\circ 2} + J_{n,q})^{\circ \frac{1}{2}},$$

де D_Y — матриця нової (вихідної) визначеності щодо команд з множини A під впливом всіх N -грамів фонем з множини B (розмір матриці $n \times q$); I_Y — матриця нової (вихідної) інформованості щодо команд з множини A під впливом всіх N -грамів фонем з множини B .

5. Обчислення сумісної умовної ймовірності команди A_i (при наявності всіх N -грамів фонем $B_j \in B$):

$$Y = P_Y = 0.5 J_{n,p} + D_Y \circledast 2I_Y,$$

де P_Y — матриця сумісної умовної ймовірності команд з множини A під впливом всіх N -грамів фонем з множини B .

Як альтернативний класифікатор фонетичного тексту голосової команди запропоновано використовувати згорткові нейронні мережі, що знайшли широке застосування в різноманітних задачах класифікації звукових даних та природномовних текстів. **Метод структурної ідентифікації згорткових нейронних мереж** для класифікації голосових команд з використанням розпізнання фонемного тексту в якості вхідних даних полягає в наступному:

1. Представлення кожної фонемі у вигляді one-hot вектору;
2. Розрахунок декількох одновимірних згорткових шарів з фільтрами розмірами 2, 3 та 4 і кроком 1;
3. Розрахунок агрегаційного шару виділенням максимального значення кожного фільтру;
4. Конкатенація результатів обрахунку всіх фільтрів;
5. Розрахунок повнозв'язного шару з функцією активації ReLU та нормалізацією Dropout;
6. Розрахунок точності та функції втрат.

Таким чином система формалізації голосової інформації в моделі голосової взаємодії водія при диспетчерському контролі за рухом автотранспорту містить фонемний стенограф та згорткову нейронну мережу, яка працює з фонемами. Реалізація ЗНМ виконана на мові Python з використанням TensorFlow. Нейронна мережа містить паралельні одновимірні шари з різними варіантами кроку фільтра з активаційною неспадаючою диференційованою функцією ReLU; фонемі представляються у вигляді one-hot вектору, фрази нормалізовані за максимальною довжиною з використанням вектора з усіма нулями в якості заповнювача. Для навчання використано Adam-алгоритм зворотного розповсюдження помилки із стохастичним градієнтним спуском, який дозволяє регулювати величину швидкості навчання залежно від параметрів. Для зниження ефекту перенавчання

ЗНМ використано Dropout шар. Також побудовано алгоритм зворотного поширення помилки ЗНМ.

Метод інтелектуальних рефлекторних систем представлено в термінах нейронних мереж, що дає можливість отримати оптимальні значення параметрів рефлекторних систем $p(A_i/B_j)$ та $p(A_i)$ не традиційним частотним методом обрахунку, а шляхом навчання методом зворотного розповсюдження помилки.

У **четвертому розділі** «Засоби формалізації голосової інформації в системах диспетчерського контролю за рухом автотранспорту» описано розроблені засоби формалізації голосової інформації в системах диспетчерського контролю за рухом автотранспорту, особливості їх використання, проведене експериментальне дослідження ефективності засобів та їх впровадження.

На основі запропонованих методів та моделей розроблено засоби формалізації голосової інформації, які дають змогу водію не відволікатись від управління автомобілем і слідкувати за дорожніми умовами та обстановкою, що сприяє прискоренню доставки продукції, а також підвищенню рівня безпеки.

Розглянуті особливості використання розроблених засобів формалізації голосової інформації в системах диспетчерського контролю за рухом автотранспорту показали, що водію автомобіля, який буде здійснювати доставку продукції і вперше зіштовхнеться із засобом формалізації голосової інформації, що діє в рамках системи диспетчерського контролю за рухом автотранспорту, необхідно попередньо перевірити і за потреби донавчити систему розпізнавати його голосові команди з відповідних контекстів.

Процес порівняння ефективності різних методів класифікації в дуальній моделі формалізації голосової взаємодії експериментально виконано за допомогою моделювання розпізнавання команд на основі ітеративного процесу збору даних та введення нових критеріїв оцінки, якщо попередні не дали достатньої точності оцінювання. Використано розширений набір метрик оцінки ефективності моделей класифікації, що крім оцінки точності включав ще й робастні метрики для незбалансованої вибірки (прецизійність, повноту, F-міру) та візуальний аналіз матриць помилок.

Оцінку ефективності дуальної системи формалізації голосової інформації проведено експериментальним шляхом у три етапи: на першому етапі первинного моделювання виявлено необхідність збільшення кількості вхідних даних; на другому перевірено гіпотезу про недостатність кількості вхідних даних; на третьому — гіпотезу про недостатню якість звукового сигналу. Прийнятний для практичного використання рівень точності в моделі, побудованій методом згорткових нейронних мереж, досягнуто на другому етапі моделювання, а в моделі, побудованій методом інтелектуальних рефлекторних систем — на третьому.

Результати третього етапу моделювання обома описаними методами наведено в таблиці 1. Розміри N-грам при моделюванні інтелектуальними рефлекторними системами були вибрані в діапазоні 2–4. Розміри згорткових фільтрів при моделюванні згортковими нейронними мережами також були взяті в діапазоні 2–4.

Табл. 1 — Порівняння двох методів формалізації ІРС і ЗНМ

№ Кон- тексту	Точність ІРС	F-міра ІРС	Точність ЗНМ	F-міра ЗНМ	Кількість стиму- лів	Кількість реакцій
1	0.850	0.590	0.900	0.900	100	2
3	0.866	0.862	0.997	0.997	350	7
4	0.870	0.867	0.990	0.990	200	4
5	0.830	0.830	0.967	0.966	300	6
6	0.835	0.829	0.975	0.975	200	4
7	0.774	0.775	0.948	0.948	500	10
8	0.857	0.851	0.983	0.983	300	6
9	0.884	0.737	0.976	0.976	250	5
10	0.864	0.865	0.960	0.960	250	5
11	0.824	0.814	0.952	0.951	250	5
12	0.751	0.753	0.951	0.950	450	9
13	0.813	0.620	0.927	0.926	150	3
14	0.730	0.734	0.950	0.950	300	6
15	0.780	0.778	0.923	0.922	300	6
16	0.840	0.832	0.968	0.968	250	5
17	0.646	0.641	0.951	0.952	350	7
18	0.740	0.739	0.928	0.928	250	5
19	0.825	0.821	0.980	0.980	200	4
По всій вибірці	0.637	0.628	0.890	0.890	3200	64

Впровадження протягом року в трьох логістичних компаніях підтвердило ефективність розробленої інформаційної технології формалізації голосової інформації: система підтримки диспетчеризації автотранспорту підвищує загальну ефективність процесу доставки за рахунок скорочення кількості необхідних транспортних засобів та збільшення кількості точок, які можуть бути обслуговувані одним транспортним засобом; запровадження голосового інтерфейсу може підвищити відсоток уникнення чи виправлення водіями інцидентів і відхилень від планового маршруту.

Для апробації розроблених засобів систему підтримки диспетчеризації автотранспорту було впроваджено на підприємстві ТОВ «Українські Інформаційні Технології» та досліджено її використання протягом року в трьох підприємствах-клієнтах.

Результати цієї апробації показали загальне підвищення ефективності процесу доставки на 14.5 %, у тому числі за рахунок скорочення кількості необхідних транспортних засобів — на 7.1 % та підвищення кількості точок, які можуть бути обслуговувані одним транспортним засобом, — у середньому на 9.4 %.

При цьому дослідження показали, що, залежно від навантаженості доби, від 5 % до 15 % точок доставки пов'язані з певними інцидентами та відхилення-

ми від плану, і лише 10 % з цих інцидентів вдається ліквідувати або надолужити. Статистичне моделювання, проведене на основі порівняння поведінки різних водіїв, показало, що водії, які вчасно повідомляють про можливі інциденти через додаток із сенсорним управлінням, можуть уникнути чи виправити до 50 % інцидентів, а отже запровадження голосового інтерфейсу може поширити ці результати на всіх водіїв.

ВИСНОВКИ

У дисертаційній роботі вирішено актуальне наукове завдання розробки моделей і методів формалізації голосової інформації в системах диспетчерського контролю за рухом автотранспорту. Загалом можна зробити такі висновки.

1. Дослідження теоретико-методологічних засад формалізації голосової інформації в системах дистрибуції показало, що значну роль в їх управлінні відіграють процеси голосової взаємодії особливо щодо своєчасного коригування планових маршрутів руху автотранспорту. Розроблення моделі голосової взаємодії без блоку переведення звуку голосу в текст може принципово покращити автоматизацію голосової взаємодії в системах контролю дистрибуції.

2. Запропонована система автоматичного розрахунку планових маршрутів та практика її використання забезпечили накопичення параметрів непередбачуваних ситуацій в процесі доставки, що впливають на створення сценаріїв голосової взаємодії, які представлено у вигляді орієнтованого графу та контекстів взаємодії. Принципи побудови рефлекторних систем на основі теорії несилової взаємодії адаптовано для формалізації голосової інформації в системах диспетчерського контролю за рухом автотранспорту.

3. Удосконалено математичну модель голосової взаємодії водія та диспетчера в системах диспетчерського контролю за рухом автотранспорту, яку представлено у вигляді повного графу сценаріїв усіх етапів доставки «склад – дорога – точка доставки». Виділено перелік унікальних контекстів голосової взаємодії, формалізація голосової інформації в яких може відбуватися незалежно, що дає можливість зменшити кількість реакцій для автоматизованого розпізнання.

4. Розроблено метод формалізації голосової інформації в системах підтримки диспетчеризації автотранспорту з використанням інтелектуальних рефлекторних систем, що дає змогу автоматизувати процес передачі голосової інформації, уникнувши переведення звукової інформації в лексичний текст завдяки використанню двох основних модулів – автоматичного фонетичного стенографа і ядра рефлекторної системи голосового управління. Для реалізації ядерного компонента запропоновано дуальну систему класифікації голосових команд, яка може використовувати метод інтелектуальних рефлекторних систем або метод згорткових нейронних мереж.

5. Метод структурної ідентифікації згорткових нейронних мереж для класифікації голосових команд адаптовано до розпізнавання фонемного тексту, що

дає змогу класифікувати голосові команди без переведення голосу в лексичний текст.

6. Метод інтелектуальних рефлекторних систем поєднано з понятійним апаратом теорії нейронних мереж, що дає змогу оптимізувати значення інформованості та визначеності шляхом навчання моделі методом зворотного розповсюдження помилки.

7. Результати математичного моделювання формалізації голосової інформації засвідчили підвищення ефективності розпізнавання повідомлень у голосовій взаємодії водія з диспетчером, а саме підвищення точності розпізнавання в середньому на 6.6 % для згорткових нейронних мереж і на 19.1 % — для інтелектуальних рефлекторних систем завдяки використанню моделі голосової взаємодії водія та диспетчера. Крім того, використання моделей на основі згорткових нейронних мереж дало підвищення швидкості розпізнавання на 15 % порівняно з інтелектуальними рефлекторними системами.

8. Результати досліджень упроваджені в ТОВ «УІТ», м. Київ (довідка від 4 січня 2019 р.) та протягом року використовувалися в трьох логістичних компаніях-клієнтах.

9. Мету досліджень щодо підвищення ефективності розпізнавання повідомлень у голосовій взаємодії водія з диспетчером досягнуто, всі часткові завдання вирішено повністю. Наукові результати досліджень є внеском у розвиток наукових і методологічних основ створення та застосування інформаційних технологій та інформаційних систем для автоматизованої переробки інформації й управління.

10. Перспективним напрямком подальших досліджень у зазначеній сфері може бути вивчення широкого кола питань, зокрема щодо розробки та дослідження інших реалізацій фонемного стенографа, використання розроблених методів та моделей класифікації фонемного тексту для роботи з лексичним текстом, а також створення моделей голосової взаємодії у вигляді графу сценаріїв для інших предметних областей.

СПИСОК ОПУБЛІКОВАНИХ ПРАЦЬ ЗА ТЕМОЮ ДИСЕРТАЦІЇ

1. Найдьонов І. М. Проблема голосової взаємодії в задачах управління дистрибуцією / І. М. Найдьонов // Вісник Черкаського державного технологічного університету. Серія: Технічні науки. – 2016. – № 3. – С. 63–71.

2. Найдьонов І. М. Модель голосової взаємодії водія в системах диспетчерського контролю за рухом автотранспорту / І. М. Найдьонов // Комп'ютерно-інтегровані технології: освіта, наука, виробництво. – 2018. – № 33. – С. 121–127.

3. Найдьонов І. М. Формалізація голосової інформації в системах диспетчерського контролю за рухом автотранспорту / І. М. Найдьонов // Наукові нотатки. – 2018. – № 64. – С. 135–140.

4. Найдьонов І. М. Порівняння ефективності двох методів формалізації голосової взаємодії / І. М. Найдьонов // Вісник НТУ «ХПІ», Серія: Нові рішення в сучасних технологіях. – 2018. – № 45 (1321). – С. 104–112.

5. Naydonov I. Comparison of efficiency of convolution neural networks and intelligent reflex systems for voice interaction formalization / I. Naydonov // American Journal of Engineering Research. – 2019. – Vol. 8, issue 5. – P. 99–108.

6. Найдьонов І. М. Побудова оптимальних маршрутів транспортних засобів з урахуванням топології дорожньої мережі / І. М. Найдьонов // Системний аналіз та інформаційні технології: матеріали 17-ї Міжнародної науково-технічної конференції SAIT 2015, Київ, 22-25 червня 2015 р. / ННК “ІПСА” НТУУ “КПІ”. – К. : ННК “ІПСА” НТУУ “КПІ”, 2015. – С. 89.

7. Найдьонов І. М. Аналіз інструментів голосової взаємодії в задачах управління дистрибуцією / І. М. Найдьонов // Тези доповідей XII Міжнародної конференції «Управління проектами у розвитку суспільства». Тема: Комплексне управління проектами розвитку в умовах нестабільного оточення / Відповідальний за випуск С. Д. Бушуев. – К. : КНУБА, 2015. – С. 175–176.

8. Найдьонов І. М. Шляхи використання можливостей голосового управління для оптимізації процесів дистрибуції / І. М. Найдьонов // III Міжнародна науково-практична конференція «Інформаційні технології та взаємодії». 8–10 листопада 2016. Тези доповідей. – К. : Київський університет, 2016. – С. 243–244.

9. Naydonov I. Geoinformation system of vehicle routing and parameters of voice interaction of subjects of logistics / I. Naydonov // 16th EAGE International Conference on Geoinformatics - Theoretical and Applied Aspects. – 2017. – P. 118–121.

10. Найдьонов І. М. Сценарії автоматизації голосової взаємодії суб'єктів дистрибуції / І. М. Найдьонов // IV Міжнародна науково-практична конференція «Інформаційні технології та взаємодії». 8–10 листопада 2017. Тези доповідей. – К. : Київський університет, 2017. – С. 177–178.

11. Найдьонов І. М. Топологічна евристика в розв'язанні проблеми маршрутизації транспортних засобів (VRP) / І. М. Найдьонов // ScienceRise. – 2015. – Т. 6, № 2(11). – С. 52–58.

АНОТАЦІЯ

Найдьонов І. М. Інформаційна технологія формалізації голосової інформації в системах диспетчерського контролю за рухом автотранспорту. — Кваліфікаційна наукова праця на правах рукопису.

Дисертація на здобуття наукового ступеня кандидата технічних наук за спеціальністю 05.13.06 – «інформаційні технології». — Київський національний університет імені Тараса Шевченка, Київ, 2019.

Дисертаційна робота присвячена вирішенню актуального наукового завдання — розробці моделей і методів формалізації голосової інформації в системах

диспетчерського контролю за рухом автотранспорту. Розроблено інформаційну технологію: модель голосової взаємодії водія та диспетчера (яка представлена у вигляді повного графу сценаріїв усіх етапів доставки «склад – дорога – точка доставки»); метод формалізації голосової інформації в системах підтримки диспетчеризації автотранспорту з використанням інтелектуальних рефлексорних систем; дуальна система класифікації голосових команд (яка налаштована на предметну область і використовує метод інтелектуальних рефлексорних систем або метод згорткових нейронних мереж у залежності від того, який з них ефективніший); засоби формалізації голосової інформації у вигляді мобільного додатку для системи Android. Використання експериментально перевіреної інформаційної технології здатне підвищити ефективність управління процесом дистрибуції.

Ключові слова: диспетчеризація автотранспорту, маршрути доставки, голосова взаємодія, розпізнавання мовлення, фонетичний текст, інтелектуальні рефлексорні системи, згорткові нейронні мережі.

АННОТАЦИЯ

Найдёнов И. М. Информационная технология формализации голосовой информации в системах диспетчерского контроля за движением автотранспорта. — Квалификационный научный труд на правах рукописи.

Диссертация на соискание ученой степени кандидата технических наук по специальности 05.13.06 – «информационные технологии». — Киевский национальный университет имени Тараса Шевченко, Киев, 2019.

Диссертация посвящена решению актуального научного задания — разработке моделей и методов формализации голосовой информации в системах диспетчерского контроля за движением автотранспорта. Разработана информационная технология: модель голосового взаимодействия водителя и диспетчера (которая представлена в виде полного графа сценариев всех этапов доставки «склад – дорога – точка доставки»); метод формализации голосовой информации в системах поддержки диспетчеризации автотранспорта с использованием интеллектуальных рефлексорных систем; дуальная система классификации голосовых команд (которая настроена на предметную область и использует метод интеллектуальных рефлексорных систем или метод сверточных нейронных сетей в зависимости от того, какой из них эффективнее); средства формализации голосовой информации в виде мобильного приложения для системы Android. Использование экспериментально проверенной информационной технологии способно повысить эффективность управления процессом дистрибуции.

Ключевые слова: интеллектуальные рефлексорные системы; сверточные нейронные сети; голосовое взаимодействие; распознавание речи; фонетический текст; системы дистрибуции; маршруты доставки; последняя миля.

ANNOTATION

Naydonov I. M. Information technology of the formalization of voice information in systems of dispatch control of vehicle traffic. — Manuscript.

Thesis for the degree of candidate of technical sciences in the specialty 05.13.06 – «information technologies». — Taras Shevchenko National University of Kyiv, Kyiv, 2019.

This research is devoted to the solution of the actual scientific task — the development of models and methods for the formalization of voice information in dispatch control systems of vehicle traffic. The information technology of the formalization of voice information in dispatch control systems of vehicle traffic was developed: the model of voice interaction of driver and dispatcher in dispatch control systems of vehicle traffic (which is presented as a complete communication script graph of all stages of the delivery «depot - road - delivery point»); the method of formalizing the voice information in vehicle dispatching support systems using intelligent reflex systems; dual system of classification of voice commands (which is tuned to the subject area and uses the method of intelligent reflex systems or the method of convolutional neural networks, depending on which of them is more effective); tool for formalizing voice information as a mobile application for the Android system. The use of experimentally proven information technology can improve the management of the distribution process.

The scientific novelty of the obtained results is that the scientific problem of integration of models and methods of formalization of voice information with the management of the distribution process has been solved for the first time in a single system of voice information formalization in dispatch control systems of vehicle traffic. Herewith, a model of voice interaction of distribution entities in dispatch control systems of vehicle traffic was developed for the first time, which is presented as a complete communication script graph of all stages of the distribution «depot - road - delivery point», which allows to narrow the scope of voice interaction to the boundaries of the subject area; a method for formalizing voice information in vehicle dispatching support systems using intelligent reflex systems was created for the first time, which allows automating voice interactions; the method of convolutional neural networks was further developed for the classification of voice commands in order to formalize the voice information in the systems of dispatch control of vehicle traffic, which was applied to phonemic text; the methods of constructing of intelligent reflex systems was further developed based on the use of neural networks to formalize the interaction processes, which makes it possible to obtain optimal values of the parameters of the reflex systems by trainig using the backpropagation method.

The developed system of automatic calculation of planned routes and the practice of its use ensured accumulation of parameters of unpredictable situations in the delivery process, which influence the creation of voice interaction scenarios, which are presented in the form of a targeted graph and interaction contexts. The principles of

constructing reflex systems based on the non-force interaction theory are adapted for the formalization of voice information in systems of dispatch control of vehicle traffic.

The evaluation of the effectiveness of the dual system for the formalization of voice information was conducted experimentally in three stages: on the first stage of the primary modeling, the need to increase the number of input data was identified; on the second stage, the hypothesis of insufficient number of input data is checked; on the third — the hypothesis of insufficient quality of the sound signal. Acceptable for practical use, the level of accuracy in the model built by the method of convolutional neural networks is achieved at the second stage of modeling, and in the model, built by the method of intelligent reflex systems — on the third.

Analysis of modeling results showed that the use of script graph and splitting a full set of commands to several contexts are appropriate, since the value of accuracy without contexts are the lowest for both methods of classification. Comparing the simulation results for different classification methods showed that both methods can be used in practice, and learning models using intelligent reflex systems much faster than convolution neural networks, but the actual recognition is faster using convolution neural networks. The accuracy of simulation is also higher using convolution neural networks.

The developed tool for formalizing voice information in the form of a mobile application for the Android system allows the driver to not distract from driving and monitor road conditions that can accelerate the delivery of products during the distribution, as well as increase the level of security.

Actual implementation during the year in three distribution companies confirmed the effectiveness of the developed information technology for the formalization of voice information: vehicle dispatching support system increases the overall efficiency of the delivery process by reducing the number of vehicles needed and increasing the number of points that can be serviced by one vehicle; the introduction of a voice interface can increase the percentage of avoiding or correcting incidents and deviations from the planned route by drivers.

Keywords: intelligent reflex systems; convolution neural network; voice interaction; speech recognition; phonetic text; distribution system; delivery routes; last mile

Підписано до друку 27.06.2019 р. Формат 60×90 1/16.

Папір офсетний. Друк - різнографія.

Ум. друк. арк. - 1,6. Обл.-видав. арк. - 2,1.

Наклад 150 прим.

Надруковано в _____

