

І. М. Найдонов, аспірант

e-mail: samogot@gmail.com

Київський національний університет імені Тараса Шевченка,
вул. Володимирська, 60, м. Київ, 01033, Україна

ПРОБЛЕМА ГОЛОСОВОЇ ВЗАЄМОДІЇ В ЗАДАЧАХ УПРАВЛІННЯ ДИСТРИБУЦІЄЮ

Стаття присвячена проблемі автоматизації голосової взаємодії в задачах управління дистрибуцією. Проведено аналіз вітчизняних та зарубіжних літературних джерел для пошуку шляхів використання можливостей голосового управління для оптимізації процесів дистрибуції за трьома напрямками: системи управління дистрибуцією, системи розпізнання голосу та системи голосової взаємодії. Пропонується інтегративна модель, в якій поєднується принцип написання дерева можливих сценаріїв взаємодії та рефлекторна система голосового управління.

Ключові слова: голосове управління, голосова взаємодія, управління дистрибуцією, розпізнання мови.

Вступ. Дистрибуція – це діяльність, пов'язана з отриманням продукції, її зберіганням до моменту отримання замовлення і наступної доставки до клієнтів. Управління дистрибуцією включає в себе планування, організацію та контроль.

Інформаційні технології в управлінні дистрибуцією вже достатньо розроблені для забезпечення етапів отримання продукції та її збереження, отже нині найбільш інтенсивно відбувається розвиток етапу доставки продукції до кінцевих клієнтів. Зокрема розробляються системи автоматизації побудови планових маршрутів руху автотранспорту [1], системи керування транспортним парком (TMS) та моніторинг доставок у реальному часі.

Велику роль в управлінні дистрибуцією відіграють процеси голосової взаємодії, які наразі активно автоматизуються для підвищення ефективності, збереження ресурсів тощо. Голосова взаємодія поділяється на безпосередню та із залученням інформаційних технологій. Інформаційні технології в цьому контексті можуть слугувати лише засобом забезпечення зв'язку, що саме по собі може давати ефект, але найкращий результат можна отримати, якщо внести певні автоматизації голосової взаємодії.

Аналіз літературних даних та постановка проблеми. Голосове управління вже має певну історію використання в транспортній сфері. Передові автоконцерни світу, такі як Ford Motor Company, BMW AG, Daimler AG, прагнуть підвищити безпеку та комфорт водія, тому створюють можливість керування бортовою електронікою за допомогою голосу

[2]. Перша подібна система, що мала назву Linguatronic, була представлена інженерами Mercedes в автомобілі S-класу в 1996 р. [3]. Вона реалізувала голосове управління функціями вбудованого телефону та адресної книги, радіо та CD-програвача, а також кондиціонера. Fiat, працюючи з Microsoft, розробив систему з ініціацією водієм Blue&Me, в якій перед початком мовної команди треба було натиснути кнопку на кермі. Інженери BMW також розробили систему з ініціацією водієм, що була інтегрована з їх системою управління бортовою електронікою iDrive. Honda, використовуючи систему розпізнання мови IBM ViaVoice, надала можливість керування GPS навігацією для голосового вказування адреси прибуття [4].

Крім того, за наявності достатньо потужної системи голосову взаємодію з водієм можна використовувати для підтримання діалогу під час руху вночі, аби не дати водію заснути [5].

Також проводилися дослідження з розробки систем мовного управління бортовим обладнанням літаків, але через високі вимоги до швидкості та якості розпізнання, особливо за наявності потужних шумів та перешкод, вони ще досі не були запроваджені [6].

Використання таких часткових функцій голосового управління, які підвищують комфорт водія, також повинно мати певний позитивний ефект. Проте ці функції не забезпечують оптимізації саме процесів дистрибуції.

У сучасних системах автоматизації дистрибуції, доставки і керування транспортним парком достатньо розробленим є й процес

автоматизації побудови планових маршрутів руху автотранспорту [1]. Він включає в себе складові врахування топології, часових параметрів точки доставки (часові вікна доступності та час, необхідний на обслуговування точки), завантаженість автомобіля, кількість доступного транспорту тощо. Проте проблема вчасного корегування маршруту у випадках, коли реальний стан справ перестає відповідати запланованому маршруту, викликає достатньо великі витрати часу на комунікацію. Якби ці функції реалізувалися за допомогою автоматизованої голосової взаємодії, це дало б максимальний ефект для покращення управління дистрибуцією.

Мета та задачі дослідження. Метою дослідження є пошук шляхів використання можливостей голосового управління для оптимізації процесів дистрибуції, які не тільки підвищують комфорт водія, а й, фактично, дають змогу автоматизувати певні процеси управління дистрибуцією, покращуючи сервіс та знижуючи витрати.

Задача полягає в тому, щоб виокремити підходи, проаналізувати міжнародні здобутки, які вже отримані в різних галузях, для того щоб побудувати модель їх використання в автоматизації управління дистрибуцією на етапі доставки.

Матеріали та методи досліджень. Для вирішення поставлених завдань було проведено теоретичний аналіз зарубіжних та вітчизняних досліджень і використано такі методи: логічного узагальнення, аналогій, порівняльного зіставлення та мисленнєвого моделювання.

1. Шляхи використання можливостей голосового управління для оптимізації процесів дистрибуції

1.1. Системи управління дистрибуцією. Перший етап дослідження полягає в огляді сучасних інструментів управління дистрибуцією та визначенні їх недоліків.

Для управління доставкою вантажів у дистрибуції вкрай важливим є етап моніторингу руху автомобілів у режимі реального часу. Це дозволяє аналізувати ефективність водія, а також передбачати певні небажані інциденти. Для такого моніторингу використовують GPS дані руху автомобіля [7, 8]. На жаль, лише GPS треку недостатньо для однозначного розуміння стану справ. З треку тільки й видно, що водій був біля точки доставки, але не зрозуміло, чи виконана доставка, чи з якихось причин відмінена. З треку ясно, що за поточ-

ної швидкості водій відстає від плану та не встигає на наступну точку, але не зрозумілі причини відставання та чи має водій можливість надолужити втрачений час. Для отримання цієї інформації необхідна додаткова комунікація водія з диспетчером. Але телефонний дзвінок чи, ще гірше, комунікація через якийсь візуальний інтерфейс у смартфоні забирає певний час та знижує концентрацію уваги водія на дорозі, що може спричинити ДТП. Тому потрібна система, яка б дала змогу виявляти необхідну інформацію в голосових даних водія і відправляти її диспетчеру у формалізованому вигляді.

Найбільш подібна до цього система PickbyVoice [9]. Це система, що використовується в іншій сфері управління дистрибуцією – управлінні складськими процесами. Pick-by-Voice дає змогу водію по черзі отримувати голосові команди у вигляді: де, що і в якій кількості треба відібрати, а також у формі діалогу повідомляти про необхідність повторити завдання чи переходити до наступного тощо. Така система дає можливість звільнити руки та очі водія і в цілому збільшити його ефективність на 35 % [10].

На жаль, для управління транспортними доставками потрібна більш складна система, ніж наявні можливості Pick-by-Voice, адже вона повинна мати суттєво більший спектр необхідних для розпізнання команд. У передових системах управління міськими доставками вантажів (urban freight distribution) важливим параметром є часові вікна доставки [11]. Такий параметр зразу вводить цілу низку додаткової інформації, яку треба передати водію до диспетчера – наскільки вчасно були виконані доставки, скільки часу було витрачено на кожну з них, відставання від плану внаслідок пробок або інших непередбачуваних обставин тощо. Більше того, система повинна забезпечувати взаємодію з диспетчером у режимі реального часу, а не відтворювати заданий заздалегідь перелік завдань.

Таким чином, очевидною стає необхідність упровадження системи голосової взаємодії між водієм та диспетчером для отримання необхідної інформації від водія в мовній формі та автоматизації управління дистрибуцією.

1.2. Системи розпізнання голосу. Другий етап дослідження полягає в тому, щоб проаналізувати сучасні досягнення в системах

розпізнання мови і розглянути можливість їх використання в управлінні дистрибуцією.

Сучасні системи розпізнання мови в більшій своїй частині базуються на статистичних методах, використовують потужний апарат теорії ймовірностей та математичної статистики, що дає змогу суттєво підвищити якість розпізнання. Основні методи розпізнання мови – це приховані марківські моделі та штучні нейронні мережі [12, 13]. Але у сучасних системах більш поширеними є моделі на нейронних мережах, оскільки вони мають більшу швидкодію та стійкість до шумів [14].

Звісно, на вхід до нейронної мережі не подають «сирий» звук – амплітуду коливань по часу, адже це не дуже інформативна форма представлення акустичного сигналу для аналізу. Більш інформативним є спектр сигналу, але на практиці найчастіше використовується мелперетворення, в якому звуковий сигнал нарізується на фрейми розміром 20–40 мс, спектр кожного з яких масштабується через банк фільтрів та логарифмується для отримання даних, найбільш наближених до людського сприйняття [15].

Існує досить багато таких систем і вони доволі якісно виконують свою задачу. Але в більшості своїй ці системи розраховані на роботу в приміщеннях без сильних шумів, залучення дикторів з чіткою вимовою і використання потужних комп'ютерів або віддалених серверів, як, наприклад, Google Voice Search. Та й ці системи неідеальні, в них не вирішені проблеми фільтрації шумів та розпізнання великих обсягів даних, обмежені можливості налаштування під різні умови та різних дикторів [16]. Так, наприклад, в системах мовного управління бортовим обладнанням літаків мінімальна можлива якість становить 95 %, а час розпізнавання не повинен перевищувати 0,2 с при темпі мовлення близько 100 слів за хвилину [17]. Навіть у рамках однієї системи розпізнавання ці параметри можуть змінюватися залежно від багатьох факторів, у тому числі таких, що визначаються умовами польоту, різними акустичними перешкодами, впливом пілотажних перевантажень тощо [6].

Фактично, ті досягнення, які сьогодні здобуті в традиційних системах розпізнання мови, вже можуть бути використані для забезпечення голосової взаємодії в управлінні дистрибуцією. Проте вони не в змозі повністю забезпечити автоматизацію голосової взаємодії

в задачах управління дистрибуцією, тому що немає можливості ані встановити в кабіні водія потужне обладнання, ані забезпечити стабільний та швидкісний доступ до Інтернету. Крім того, кабіна водія – це неконтрольоване акустичне середовище з високим рівнем шуму. Проблема багатодикторності також актуальна для дистрибуції, адже в таких компаніях зазвичай працює від кількох десятків до кількох сотень і навіть тисяч водіїв, в яких можуть буди дефекти вимови, різноманітні акценти та інші індивідуальні особливості мовлення.

1.3. Системи голосової взаємодії. Третій етап дослідження полягає в пошуку інших ідей і методів, які дають можливість обійти обмеження класичних систем розпізнання мови. Аналіз розробок в галузі робототехніки та технологій управління дав змогу виявити такі перспективні напрями.

Японський дослідник з університету Осаки Ішігуро Хіроші з колегами вивчали різні аспекти комунікації та інформаційно-комунікаційних технологій, як, наприклад, використання комунікації з людиноподібними роботами як терапевтичної дії для людей похилого віку [18], аутистів [19] чи просто замкнених у собі людей, педагогічної дії щодо дітей та немовлят [20] тощо. Зокрема він проводив дослідження голосової комунікації двох людей опосередковано через комп'ютер [21].

У цьому дослідженні пара спілкувалася на загальні теми, обираючи варіанти своєї репліки із заздалегідь написаного дерева варіантів, свого роду сценарію. Жодному з партнерів не потрібно було нічого промовляти вголос: людині надавався набір з варіантів реплік на вибір, потрібно було лише натиснути на ту з них, яку б вона хотіла промовити, і ця репліка лунала з динаміків. Залежно від використаної репліки програма вибирала з дерева сценаріїв можливі варіанти відповідей і надавала їх на вибір співрозмовнику. Співрозмовник, у свою чергу, чуючи репліку першої людини, обирав свою з наданих варіантів. Це дослідження було спрямоване на подолання сором'язливості при спілкуванні з особами протилежної статі (що є особливо актуальним для Японії). Але такий підхід заздалегідь написаного дерева сценаріїв комунікації можна використовувати і в інших сферах.

У доповіді на світовому психологічному конгресі 2016 р. професор Ішігуро демонстрував використання цього сценарного підходу для роботів на виставках та в музеях. Щоб

уникнути необхідності розпізнавання голосу в шумному середовищі, поряд з експонатом ставиться людиноподібний робот та монітор, на якому показані варіанти запитань. Натискаючи на різні репліки, відвідувач може спілкуватися з роботом за заздалегідь написаним деревом сценаріїв, розпитуючи його про експонат, а робот буде відповідати голосом.

На жаль, для управління дистрибуцією постає зворотне завдання – водій має повідомити певну інформацію в систему і при цьому не повинен відволікатися на натискання кнопок на екрані. Тому пряме використання такої технології неможливе. Але застосування підходу описання всіх можливих сценаріїв комунікації залежно від контексту дозволить знизити кількість інформації, яку треба розпізнати, а отже й підвищити якість.

Наразі існує новий підхід до голосового управління, що базується на теорії несилової взаємодії [22], – рефлекторна система голосового управління [23]. Ідея, покладена в основу цього підходу, полягає в тому, щоб замість переведення голосової інформації в текстову репрезентацію, аналізувати безпосередньо інформаційну складову сказаного, визначаючи, яку з відомих реакцій потрібно виконати. «Традиційні системи розпізнавання мови базуються на принципі: «усна мова» → «репрезентація мови набором лінгвістичних конструкцій» → «розуміння мови». На основі теорії несилової взаємодії може бути запропонована інша модель розпізнавання природної мови: «усна мова» → «розрахунок несилової (інформаційної) взаємодії на реакції» → «реакція (розуміння чи поведінка)» [24].

Така модель розпізнавання називається рефлекторною, оскільки побудована за аналогією зі структурою умовного рефлексу, в якому виділяються афектори, центральний компонент та ефектори. Така модель може бути добре поєднана з ідеєю використання дерева сценаріїв, оскільки сценарії також складаються з реакцій, і одиницею моделювання стає не лінгвістична особливість мовлення, а реакція (або команда), яка може бути врахована автоматизованою системою розрахування маршрутів. Тобто, суть цього підходу полягає в тому, щоб перейти до іншої одиниці розпізнавання мови. У психології дискурсивного мислення і рефлексивній психології також накопичено досвід аналізу мови через виокремлення інших одиниць – функціональних висловлювань [25].

Оскільки в такій системі не потрібні словники, складні інтелектуальні моделі аналізу тексту та граматики, вони мають низку переваг порівняно з традиційними системами: багатодикторність, варіабельність природної мови, можливість обробки команд офлайн прямо на пристрої, робота в умовах шумів (не контрольованого акустично середовища), простота алгоритмів і менші складність та ціна реалізації [26].

У загальному випадку система рефлекторного голосового управління складається з трьох компонентів:

1. Фонемний стенограф

Відповідає за перетворення відцифрованого вхідного звукового сигналу, що поміщає усну мову в набір фонем або слів.

2. Ядро системи

Здійснює моделювання системи голосового управління. Містить програмну реалізацію всіх моделей, методів та алгоритмів системи, набір команд, протокол роботи, налаштування тощо.

3. База даних розпізнавання мови з можливістю навчання

Забезпечує зберігання інформаційної бази розпізнавання мови та виділення керуючого впливу. У базі даних зберігається статистика вхідних впливів та відповідних їм вихідних реакцій системи.

Варто зазначити, що модуль фонемного стенографа може бути забезпечений різними програмними засобами, що робить систему гнучкішою та більш адаптивною до умов середовища. А. В. Єгорченков [23] наводить перелік можливих модулів фонемного стенографа. Провівши порівняльне зіставлення стенографів з цього переліку, можна стверджувати, що найбільш прийнятним для задач дистрибуції є фонетичний стенограф на основі Julius speech recognition tool [27], тому що він у своїй роботі не потребує ні доступу до Інтернету, ні великих словникових баз, а дає на виході «сирі» фонем (наприклад, «п ь й! А м а» для слова «Прямо»), які можуть навіть краще сприйматися рефлекторною системою для подальшого визначення інформаційного компонента, ніж розпізнані слова, через вищу (в останньому випадку) ймовірність помилки.

2. Модель голосової взаємодії в задачах управління дистрибуцією. У результаті аналізу виявлено два найбільш перспективні напрями, поєднання яких дає змогу запропонувати нове принципове рішення і побудува-

ти рефлекторну модель голосової взаємодії в задачах управління дистрибуцією. В основу моделі покладено логічні сценарії взаємодії на тему управління дистрибуцією, які мають враховувати параметри основних причин невідповідності реальної ситуації запланованому маршруту, наприклад, запізнення або відмови обслуговування на точці доставки тощо. Це дає змогу отримати інформацію для прийняття 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.

Звичайно, абсолютно всі причини та параметри не можуть бути враховані заздалегідь, але пророблення й урахування основної типології дозволить приймати базові рішення і вдаватися до безпосереднього зв'язку з диспетчером лише у складних випадках, що розвантажуватиме водія та канали комунікації і дасть змогу підвищити загальну ефективність дистрибуції.

Найбільш ефективним шляхом пророблення дерева сценаріїв рефлекторної взаємодії є використання вхідних параметрів вже створеної системи автоматизації дистрибуції, у тому числі автоматичної побудови маршрутів [28], яка нині проходить широку експериментальну апробацію. Інтеграція модуля голосової взаємодії з цією системою буде значно спрощена, що сприятиме отриманню кращого економічного ефекту.

Виходячи з наявної логіки побудови маршрутів, вже зараз можна назвати принципові блоки сценаріїв, які потрібно буде розробити. Першим етапом, на якому можуть виникнути проблеми розбіжності плану і факту, є етап завантаження на складі (якщо, наприклад, буде виявлено невраховане перевантаження або недовантаження машини, будуть відсутні необхідні товари чи працівники складу не встигнуть їх вчасно відібрати, або навіть виявиться, що машина не здатна вийти на маршрут (наприклад, не заводиться на морозі)).

Другий етап сценаріїв голосової взаємодії визначають проблеми, що можуть виникнути в дорозі до певної точки доставки, як, наприклад, ремонт в дорозі за маршрутом руху або зміни в правилах руху на деяких вулицях, які ще не відбиті в алгоритмах прокладення маршруту (нові заборони поворотів чи односторонній рух), проблеми з автомобілем на

дорозі, які призводять до зниження швидкості або відмови в подальшому русі за маршрутом, або найбільш поширена проблема заторів на дорогах.

Третій етап сценаріїв голосової взаємодії викликаний можливими невідповідностями між планом та фактом в обслуговуванні на точці доставки. Це можуть бути як проблеми з боку клієнта («нікого немає вдома», клієнт не має грошей, клієнт відмовляється від замовлення чи стверджує, що він замовляв щось інше), так і проблеми з боку водія (запізнення на точку доставки, тобто непотрапляння в заплановане дозволене часове вікно доступності, пошкодження товару тощо). Найбільш поширеною є ситуація, коли водій проводить у точці доставки більше часу, ніж заплановано, що призводить до проблем на всьому подальшому маршруті.

Ці та інші інциденти на всіх зазначених етапах, що зображені в моделі на схемі (рис. 1), потребують вирішення із залученням диспетчера для вибору найкращої стратегії і мінімізації втрат через проблему. Відповідно дерево сценаріїв голосової взаємодії повинне відбивати всі три етапи та типові відомі проблеми і способи їх розв'язання.

Модель голосової взаємодії суб'єктів дистрибуції (рис. 1) складається з трьох етапів, два останніх з яких можуть циклічно повторюватися при наявності кількох точок доставки в маршруті. Стрілками позначено процеси голосової взаємодії, які можуть розгортатися на кожному з етапів при невідповідності плану та факту. У верхній та нижній частинах схеми показані принципові типи результатів голосової взаємодії для кожного з суб'єктів (диспетчера та водія). Ця логічна схема визначає принциповий алгоритм побудови дерева сценаріїв голосової взаємодії.

Висновки. Проведено аналіз вітчизняних і зарубіжних літературних джерел для пошуку шляхів використання можливостей голосового управління з метою оптимізації процесів дистрибуції за трьома напрямками: системи управління дистрибуцією, системи розпізнання голосу та системи голосової взаємодії.

Виявлено обмеження інструментів GPS контролю, які використовуються в системах управління дистрибуцією на етапі моніторингу доставки, але не відбивають причин відхилення реальної ситуації від запланованого маршруту. Саме на подолання цих обмежень має бути спрямована система голосової взаємодії в задачах управління дистрибуцією.

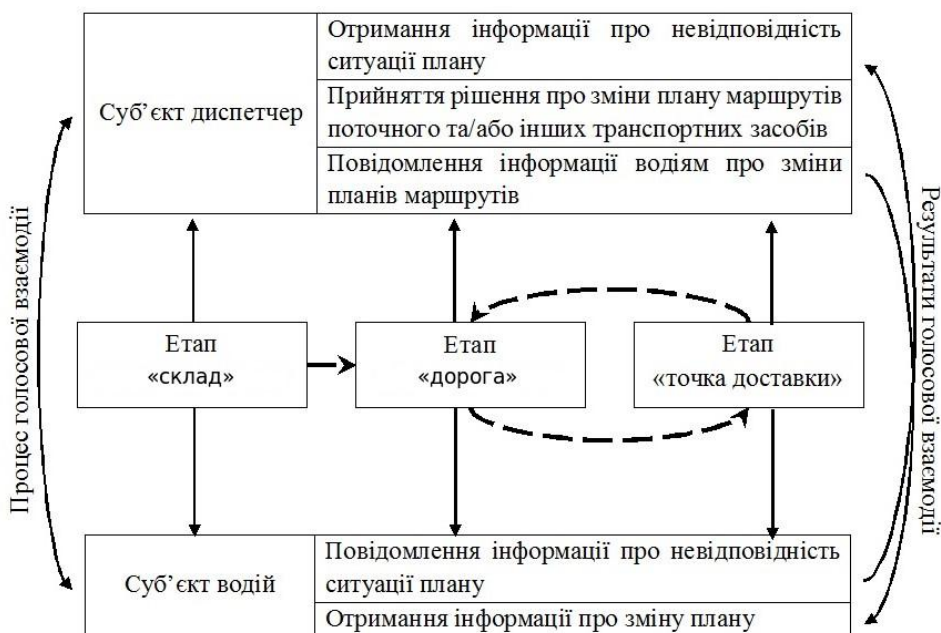


Рис. 1. Модель голосової взаємодії суб'єктів дистрибуції

Найвні традиційні системи розпізнання голосу на основі нейронних мереж та прихованих марківських моделей не забезпечують необхідного для задач дистрибуції рівня стійкості до шумів, багатодикторності та потребують потужного обладнання або стабільного доступу до Інтернету.

Нові принципи розв'язання проблеми – перехід до іншої одиниці розпізнання мови (рефлексу або команди) та побудова дерева можливих сценаріїв взаємодії для зменшення кількості необхідних до розпізнання команд залежно від контексту ситуації.

Таким чином, пропонується інтегративна рефлекторна модель голосової взаємодії в задачах управління дистрибуцією, в якій поєднуються два сформульовані принципи: написання дерева можливих сценаріїв взаємодії та рефлекторна система голосового управління.

Перспективою подальшого дослідження є розробка програмного забезпечення, яке реалізуватиме запропоновану модель та інтегруватиметься з існуючою авторською системою автоматизації побудови маршрутів і управління дистрибуцією.

Список літератури

1. Найдъонов І. М. Топологічна евристика в розв'язанні проблеми маршрутизації транспортних засобів / І. М. Найдъонов // ScienceRise. – 2015. – Т. 11, № 6. – С. 52–58. – DOI: 10.15587/2313-8416.2015.44381.

- Кравченко А. П. Автоматизированная компьютерная система голосового управления автомобилем / Кравченко А. П., Крамарь Н. М., Морозов И. В. // Автомобильный транспорт. – 2009. – № 25. – С. 44–47.
- Heisterkamp P. Linguatronic product-level speech system for mercedes-benz cars / P. Heisterkamp // Proceedings of the first international conference on human language technology research. – San Diego: Association for Computational Linguistics, 2001. – С. 1–2. – (HLT '01). – DOI: 10.3115/1072133.1072199.
- Jonsson I.-M. Social and emotional characteristics of speech-based in-vehicle information systems: impact on attitude and driving behaviour / I.-M. Jonsson. – Linköping: Linköping University Electronic Press, 2009. – 193 p.
- Кравченко А. П. Автоматизированная компьютерная система для интерактивного общения с водителем / А. П. Кравченко, Т. М. Терещенко, И. В. Морозов // Автомобильный транспорт. – 2012. – № 30. – С. 54–58.
- Экспериментальное исследование влияния акустических помех разных видов на результаты автоматического распознавания речевых команд / О. Н. Корсун, А. А. Яцко, И. М. Финаев, В. Я. Чучупал // Наука и образование: науч. изд-е МГТУ им. Н. Э. Баумана. – 2013. – Т. 01. – С. 12.

7. GPS-based data production in urban freight distribution / J. Gonzalez-Feliu, P. Pluvinet, M. Serouge, M. Gardrat. – Nova Science. – 06.2013.
8. Comendador J. A. A GPS analysis for urban freight distribution / J. A. Comendador, M. E. López-Lambas, A. Monzón de Cáceres // The Seventh International Conference on City Logistics. – E.T.S.I. Caminos, Canales y Puertos (UPM), 06.2012. – DOI: 10.1016/j.sbspro.2012.03.127.
9. Pick-to-Voice. Increased productivity, accuracy and flexibility [Electronic resource] / Dematic Global Website. – Mode of access: <http://www.dematic.com/en/supply-chain-solutions/bytechnology/voice-and-light-systems/pick-to-voice/>
10. Baumann H. Order picking supported by mobile computing / H. Baumann. – Bremen : University of Bremen, 2012. – 190 p.
11. Quak H. Urban distribution: the impacts of different governmental time-window schemes / H. Quak, R. de Koster; Erasmus Research Institute of Management (ERIM). – 09.2006. – 15 p.
12. Маковкин К. А. Гибридные модели: скрытые марковские модели и нейронные сети, их применение в системах распознавания речи / К. А. Маковкин // Модели, методы, алгоритмы и архитектуры систем распознавания речи. – М., 2006.
13. Гефке Д. А. Применение скрытых марковских моделей для распознавания звуковых последовательностей / Д. А. Гефке, П. М. Зацепин // Известия Алтайского государственного университета. – Барнаул, 2012.
14. Deep neural networks for acoustic modeling in speech recognition / G. Hinton, L. Deng, D. Yu et al. // IEEE, SIGNAL PROCESSING MAGAZINE. — 2012.
15. Saini P. Automatic speech recognition: a review / P. Saini, P. Kaur // International Journal of Engineering Trends and Technology. – 2013. – Т. 4, № 2.
16. Волков А. В. Анализ существующих методов распознавания на инвариантность к фоновым помехам и дикции диктора / А. В. Волков // Известия Тульского государственного университета. Технические науки. – 2014. – Т. 9–2. – С. 6. – DOI: 10.14489/vkit.2016.05.pp.021-028.
17. Бондарос Ю. Г. Система распознавания команд речевого интерфейса пилота для интегрированной модульной авионики / Ю. Г. Бондарос, К. А. Маковкин, В. Я. Чучупал // Вестник компьютерных и информационных технологий. – 2007. – Т. 4. – С. 2–13.
18. Nishio S. Using androids to provide communication support for the elderly / S. Nishio, T. Minato, H. Ishiguro // New Breeze. – 2015, Oct. – Т. 27, № 4. – P. 14–17.
19. Robot-mediated interventions for social anxiety in individuals with autism spectrum disorder / H. Kumazaki, Y. Yoshikawa, Y. Matsumoto et al. // 2016 International Meeting for Autism Research. – 05.2016. – P. 270–271. – (IMFAR 2016).
20. Do infants consider a robot as a social partner in collaborative activity? / Y. Park, S. Itakura, A. M. Henderson et al. // Proceedings of the 3rd International Conference on Human-Agent Interaction. – Daegu, Kyungpook, Republic of Korea : ACM, 10.2015. – P. 91–95. – (HAI '15). – DOI: 10.1145/2814940.2814953.
21. Ishiguro H. Adaptation to teleoperated robots / H. Ishiguro // International Journal of Psychology. – 2016. – Т. 51. – P. 10–10. – DOI: 10.1002/ijop.12361.
22. Тесля Ю. М. Введение в информатику природы / Ю. М. Тесля. – К. : Маклаут, 2010. – 255 с.
23. Егорченков А. В. Прикладное применение рефлекторной системы голосового управления / А. В. Егорченков // Управление развитием складных систем. – К., 2016. – № 25. – С. 103–107.
24. The non-force interaction theory for reflex system creation with application to TV voice control / I. Teslia, N. Popovych, V. Pylypenko, O. Chornyi // Proceedings of the 6th International Conference on Agents and Artificial Intelligence. – 2014. – С. 288–296. – DOI: 10.5220/0004754702880296.
25. Найдьонов М. І. Формування системи рефлексивного управління в організаціях / М. І. Найдьонов. – К. : Міленіум, 2008. – 484 с.
26. Тесля Ю. М. Рефлекторная система голосового управления техническими устройствами (РСГУ) / Ю. М. Тесля, О. Чорний // Управление развитием складных систем. – К., 2013. – № 15. – С. 105–110.
27. Пилипенко В. В. Распознавание ключевых слов в потоке речи при помощи фонетического стенографа / В. В. Пилипенко // Речевые технологии. – 2009. – № 1. – С. 75–79.

28. Математичний сервер автоматичної побудови логістичних маршрутів: а.с. 60788 Україна / І. М. Найдьонов. – № 61210 ; заявл. 25.05.2015 ; зарег. 23.07.2015 ; опубл. 30.10.2015, Бюл. № 38. – 2 с.

References

- Nayd'onov, I. M. (2015) Topologic euristics in solving the problem of vehicles routing. *ScienceRise*, 11 (6), pp. 52–58, DOI:10.15587/2313-8416.2015.44381 [in Ukrainian].
- Kravchenko, A. P., Kramar, N. M. and Morozov, I. V. (2009) Automated computer system of vehicle voice control. *Avtomobil'nyj transport*, (25), pp. 44–47 [in Russian].
- Heisterkamp, P. (2001). Linguatronic product-level speech system for mercedes-benz cars. In: *Proceedings of the first international conference on human language technology research*, (HLT '01). San Diego: Association for Computational Linguistics, pp. 1–2, DOI:10.3115/1072133.1072199.
- Jonsson, I.-M. (2009) Social and emotional characteristics of speech-based in-vehicle information systems: impact on attitude and driving behaviour. Linköping: Linköping University Electronic Press, 193 p.
- Kravchenko, A. P., Tereshhenko, T. M. and Morozov, I. V. (2012) Automated computer system for interactive communication with a driver. *Avtomobil'nyj transport*, (30), pp. 54–58 [in Russian].
- Korsun, O. N., Jacko, A. A., Finaev, I. M. and Chuchupal, V. Ja. (2013) Experimental research of the impact of acoustic noises of different kinds on the results of automated recognition of voice commands. *Nauka i obrazovanie: nauch. izd-e MGTU im. N. E. Bauman*, (01), p. 12 [in Russian].
- Gonzalez-Feliu, J., Pluvinet, P., Serouge, M. and Gardrat, M. (2013) GPS-based data production in urban freight distribution. *Nova Science*.
- Comendador, J. A., López-Lambas, M. E. and Monzón de Cáceres, A. (2012) A GPS analysis for urban freight distribution. In: *The seventh international conference on city logistics*. E.T.S.I. Caminos, Canales y Puertos (UPM). DOI:10.1016/j.sbspro.2012.03.127.
- Pick-to-Voice. Increased productivity, accuracy and flexibility, available at: <http://www.dematic.com/en/supply-chain-solutions/by-technology/voice-and-light-systems/pick-to-voice/>
- Baumann, H. (2012) Order picking supported by mobile computing. Bremen: University of Bremen, 190 p.
- Quak, H. and Koster, R. de. (2006) Urban distribution: the impacts of different governmental timewindow schemes. Erasmus Research Institute of Management (ERIM), 15 p.
- Makovkin, K. A. (2006) Hybrid models: latent markovian models and neural networks, their application in voice recognition systems. *Modeli, metody, algoritmy i arhitektury sistem raspoznavanija rechi*. Moscow [in Russian].
- Gefke, D. A. and Zacepin, P. M. (2012) Application of latent markovian models for recognition of voice sequences. *Izvestija Altaj. gos. un-ta*. Barnaul [in Russian].
- Hinton, G., Deng, L., Yu, D. et al. (2012) Deep neural networks for acoustic modeling in speech recognition. *IEEE, SIGNAL PROCESSING MAGAZINE*.
- Saini, P. and Kaur, P. (2013) Automatic speech recognition: a review. *International Journal of Engineering Trends and Technology*, 4 (2).
- Volkov, A. V. (2014) The analysis of existing methods for invariance recognition to background noise and speaker's diction. *Izvestija Tul'skogo gosudarstvennogo universiteta. Tehnicheskie nauki*, (9–2), p. 6, DOI:10.14489/vkit.2016.05.pp.021-028 [in Russian].
- Bondaros, Ju. G., Makovkin, K. A. and Chuchupal, V. Ja. (2007) The system of recognition of pilot's speech interface for integrated modular avionics. *Vestnik kompjuternyh i informacionnyh tehnologij*, (4), pp. 2–13 [in Russian].
- Nishio, S., Minato, T. and Ishiguro, H. (2015) Using androids to provide communication support for the elderly. *New Breeze*, 27 (4), pp. 14–17.
- Kumazaki, H., Yoshikawa, Y., Matsumoto, Y. et al. (2016) Robot-mediated interventions for social anxiety in individuals with autism spectrum disorder. In: *2016 International Meeting for Autism Research (IMFAR 2016)*, pp. 270–271.
- Park, Y., Itakura, S., Henderson, A. M. et al. (2015) Do infants consider a robot as a social partner in collaborative activity? In: *Proceedings of the 3rd International*

- Conference on Human-Agent Interaction (HAI '15)*, pp. 91–95. Daegu, Kyungpook, Republic of Korea: ACM. DOI:10.1145/2814940.2814953.
21. Ishiguro, H. (2016) Adaptation to teleoperated robots. *International Journal of Psychology*, (51), p. 10–10. DOI:10.1002/ijop.12361.
22. Teslia, Yu. M. (2010) Introduction in nature informatics. Kyiv: Maklout, 255 p. [in Ukrainian].
23. Egorchenkov, A. V. (2016) Applied use of reflex system of voice control. *Upravlinnya rozvytkom skladnyh system*, (25), pp. 103–107 [in Russian].
24. Teslia, I., Popovych, N., Pylypenko, V. and Chorni, O. (2014) The non-force interaction theory for reflex system creation with application to TV voice control. In: *Proceedings of the 6th International Conference on Agents and Artificial Intelligence*, pp. 288–296. DOI: 10.5220/0004754702880296.
25. Nayd'onov, M. I. (2008) Formation of reflex control system in organizations. Kyiv: Milenium, 484 p. [in Ukrainian].
26. Teslja, Ju. M. and Chornyj, O. (2013) Reflex system of voice control of technical devices (RSGU). *Upravlinnya rozvytkom skladnyh system*, (15), pp. 105–110 [in Russian].
27. Pilipenko, V. V. (2009) Recognition of key words in the flow of speech with the help of phonetic stenographer. *Rechevyve tehnologiyi*, (1), pp. 75–79 [in Russian].
28. Nayd'onov, I. M. (2015) Mathematical server for automatic construction of logistic routes. No. 60788. Ukraine [in Russian].

I. M. Naydonov, postgraduate student
e-mail: samogot@gmail.com

Taras Shevchenko National University of Kyiv
Volodymyrska str., 60, Kyiv, 01033, Ukraine

THE PROBLEM OF VOICE INTERACTION IN DISTRIBUTION MANAGEMENT TASKS

The problem of automation of voice interaction in distribution management tasks is analyzed in the article. The main goal is to analyze domestic and foreign experience and single out ways for the use of voice control capabilities to optimize distribution processes in three areas: distribution management systems, voice recognition systems and voice interaction systems. The limits of GPS monitoring systems, which are used to control delivery processes, but do not reflect the causes of real situation deviation from the planned route, are discovered. To overcome these limits, the direction of voice interaction system for management tasks in distribution is offered. Available traditional voice recognition systems based on neural networks and hidden Markov models do not provide the required level for distribution because of resistance to noise, they require powerful hardware or stable access to the Internet. Two new principles of resolving the problem – the transition to another speech recognition unit and building a tree of possible scenarios of interaction to reduce the amount necessary to recognize commands, depending on the context of the situation, are offered. Thus, integrative model of reflex voice interaction in distribution management, which combines two articulated principles, is offered. The prospect of further research is to develop a software, which will be realized in proposed model, and to integrate it with existing authoring system of automation routing.

Keywords: voice control, voice interaction, distribution management, speech recognition.

Рецензенти: А. О. Білощицький, д.т.н., професор,
Ю. В. Триус, д.пед.н., к.ф.-м.н., професор