УДК 004.75; 004.657

И. П. Карпова, канд. техн. наук, доц., e-mail: karpova_ip@mail.ru,
Национальный исследовательский университет "Высшая школа экономики",
Национальный исследовательский центр "Курчатовский институт", Москва

Хранение и обработка распределенных данных в группе мобильных роботов

Рассматривается способ организации обработки распределенных данных, носителями и потребителями которых являются мобильные роботы. Введены определения неточных и противоречивых данных применительно к данным, которыми обмениваются роботы в группе. Предложен способ работы с неточными и противоречивыми данными, основанный на мультимножествах, нечетких множествах и оценке степени достоверности данных, полученных от разных роботов.

Ключевые слова: *групповая робототехника, статический рой, база данных, обработка неточных и противоречивых данных*

Введение

Основная идея групповой робототехники заключается в том, что группа относительно простых роботов может решать сложные задачи за счет объединения своих усилий [1]. Под "простотой" понимается ограниченность сенсорных, вычислительных и когнитивных возможностей робота. Одним из механизмов, позволяющих роботам согласовывать свои действия, является обмен данными. Он может быть явным, когда роботы обмениваются сообщениями, и неявным, когда взаимодействие осуществляется посредством влияния на окружающую среду или наблюдением за действиями друг друга. Нас в первую очередь интересует явный обмен данными как наиболее эффективный способ коммуникации.

Каждый робот обладает некоторыми данными, которые могут включать в себя: показания датчиков робота; сведения о задачах, которые он решал или решает в настоящий момент; описание текущего состояния робота и т. п. Объединение этих данных можно рассматривать как общую (коллективную) базу данных (БД) группы роботов.

Общая постановка задачи формирования и функционирования базы данных в группе роботов приведена в работе [2]. Она включает в себя необходимость решения таких вопросов, как определение структуры БД, разработка формата хранения данных, формулирование правил организации запросов к БД и создание протокола обмена сообщениями.

Цель данной работы — рассмотреть структуру общей БД для группы роботов, логическую организацию запросов и теоретические основы обработки неточных и противоречивых данных, которые могут быть получены в ответ на запрос робота к общей БД.

Структура статьи следующая. Сначала кратко описывается понятие статического роя, приводится перечень отличий общей базы данных для группы роботов от распределенной БД в классическом представлении, приводятся сведения о работах по аналогичной тематике. Затем описывается структура хранимых данных и подход к логической организации запросов в общей базе данных группы роботов. Далее вводятся понятия неточных и противоречивых данных и предлагается подход, позволяющий на основе таких данных получить единственное значение с учетом особенностей рассматриваемой предметной области.

Общая база данных в статическом рое

Группу роботов будем рассматривать как статический рой [2]. Под статическим роем понимается однородная группа роботов, для которой выполняются два условия: 1) каждый робот имеет фиксированное число коммуникационных портов и может общаться с ограниченным числом своих ближайших соседей; 2) связи между членами группы — роя — могут считаться постоянными на некотором интервале времени.

Таким образом, взаимное расположение роботов определяет также и схему их соединения, которую можно рассматривать как сеть.

Общая база данных группы роботов существенно отличается от распределенной БД в классическом понимании. Она не требует глобального каталога данных, в котором хранится информация о местоположении фрагментов БД, так как на каждом узле (роботе) структура фрагмента БД одинакова. Эта структура может быть отражена в справочной таблице, которая загружается в память робота при его инициализации и фактически является каталогом данных. Таким образом, обращение к данным других узлов может происходить на основе локального каталога. Также не требуется поддержка распределенных транзакций. Процессы записи данных в общей БД носят локальный характер, так как в БД каждого робота хранятся только те данные, которые робот собрал самостоятельно, и неважно каким образом: считал с датчиков или получил в результате запроса к другому узлу и сохранил в локальной БД.

Очевидно, что нецелесообразно использовать для организации БД в группе роботов обычные СУБД, поддерживающие распределенные базы данных, так как они требуют большого объема памяти, надежных и высокоскоростных каналов связи и т. д. Но существует специализированный программный инструмент для хранения и обмена данными в распределенной БД: СУБД TinyDB, которая создана специалистами Калифорнийского университета в Беркли [3]. Она работает под управлением операционной системы TinyOS [4], предназначена для организации обмена данными в беспроводных сенсорных сетях (БСС) и адаптирована к потребностям и ограничениям БСС, в том числе по объему памяти на устройствах и по затратам энергии при передаче данных. TinyDB настраивается на датчики, которые находятся на устройстве, расположенном в узле сети. Недостатком этой системы является централизованный характер сбора и обработки данных: по запросу центрального узла TinyDB собирает с узлов данные и передает их на центральный узел. К сожалению, по условиям нашей задачи применение TinyDB невозможно, так как в однородной группе роботов никакой централизации или иерархии быть не должно, и инициатором запроса может быть любой робот.

Из работ по аналогичной тематике стоит остановиться на работе [5], в которой предлагается подход, подразумевающий наличие общей базы данных в группе роботов (агентов). Эта БД содержит прецеденты — наборы ситуаций, в которых побывал агент, выполненных в каждой ситуации действий и полученных за эти действия оценок. Ситуация характеризуется набором наблюдаемых объектов, представленным в виде мультимножества, а действия агента определяются стохастическими векторами. Агенты могут обмениваться этими прецедентами, усваивая чужой "жизненный опыт". Агент-новичок сравнивает ситуацию, в которой он оказался, с ситуациями из баз прецедентов, сформированных другими агентами, используя для качественного сравнения мультимножеств метрику Хэмминга. Если при сравнении выявлены противоречивые совпадения, т. е. текущее мультимножество обнаружено в базах нескольких агентов, но экстремумы стохастических векторов найденных прецедентов не совпадают, то сравнивается весь "жизненный опыт" нового агента с "жизненным опытом" тех агентов, с которыми были совпадения. Новым агентом будут использованы действия того агента, с чьей базой прецедентов было выявлено больше совпадающих или похожих ситуаций.

Но перед нами стоит иная задача — предложить структуру для хранения данных более общего вида и разработать способы работы с такими данными, возможно неточными и противоречивыми, которые могут поступить в ответ на запрос одного робота к другим.

Логическая организация общей базы данных и запросов к ней

В самом общем виде данные, хранящиеся в БД робота, могут быть представлены в виде множества элементов следующей структуры:

(<имя факта> <значение> <атрибуты>).

Именем факта является название параметра (например, имя соответствующего датчика), а атрибут может содержать любую дополнительную информацию: состояние узла, временну́ю метку и др. [2]. Атрибут представляет собой пару (*<имя атрибута>*, *<значение>*).

Концептуальная схема общей БД приведена на рис. 1.

Справочная таблица содержит перечень показателей, в первую очередь, датчиков, которыми оснащен робот. Каждый датчик имеет название, для него определены единица измерения, погрешность измерения и границы допустимых значений. В качестве первичного ключа справочной таблицы целесообразно использовать число-

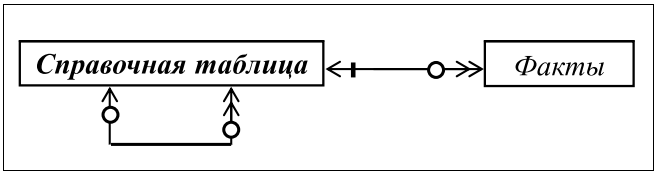

Рис. 1. Концептуальная схема общей БД

вой идентификатор, а не название, так как это сократит издержки памяти для таблицы фактов.

Структура факта и структура атрибута похожи, поэтому имеет смысл хранить общие данные об атрибутах в той же справочной таблице, а связь атрибута с показателем отражается с помощью внешнего ключа из справочной таблицы на саму себя. Это даст возможность и для атрибута (при необходимости) указать единицу измерения, погрешность измерения и границы допустимых значений.

Обсуждение возможных физических форматов хранения данных выходит за рамки данной статьи, так как конкретная реализация схемы хранения зависит от характера данных и решаемых задач. Можно сделать только одно замечание: в условиях ограничений на объем памяти, выделенной под хранение БД, память целесообразно организовать как кольцевой буфер, способ записи — стек. Это позволит начинать поиск с последних добавленных значений, которые, естественно, востребованы чаще других.

С учетом приведенных выше ограничений на организацию каналов связи в статическом рое и принципиально низкую пропускную способность этой сети в общей БД должны поддерживаться элементарные запросы. Это запросы на получение некоторого факта по его имени и, в общем случае, по атрибуту. При выполнении запросов могут возникнуть следующие ситуации: робот, запросивший данные, 1) не получил ответа; 2) получил один ответ; 3) получил несколько ответов, и полученные данные одинаковы; 4) получил несколько ответов, и полученные данные разные.

Для начала определим, что будет подразумеваться под понятиями *одинаковые* и *разные* данные. Данными в первую очередь являются значения, полученные с датчиков (сенсоров) робота, т. е. скалярные величины. Далее речь пойдет именно о них. Каждый датчик имеет погрешность измерения ε, поэтому одинаковыми будут считаться значения x_i и x_j, отличающиеся друг от друга не более чем на эту погрешность:

$$\left|x_i - x_j\right| \leqslant \varepsilon. \qquad (1)$$

В противном случае данные считаются разными.

Упорядочим по значениям данные, полученные от N роботов. Если для упорядоченного множества значений $X = \{x_i\}$ выполняются условия

$$\begin{gathered} \left|x_i - x_{i+1}\right| \leqslant \varepsilon, \;\; i = \overline{1, N-1} \\ \exists i, j : \left|x_i - x_j\right| > \varepsilon \end{gathered}, \qquad (2)$$

то (2) описывает отношение нестрогого равенства.

В соответствии с особенностями функционирования группы роботов каждое значение может иметь временную отметку и привязку к координатам. Эти данные также невозможно определить с абсолютной точностью. Следовательно, и для этих показателей необходимо определить окрестности (будем также называть это погрешностью ε), в пределах которых временные отметки и координаты считаются совпадающими. При определении этих окрестностей необходимо учитывать как физические параметры роботов, так и степень достоверности используемых ими методов определения собственных координат и времени. Эти показатели зависят от конкретных робототехнических устройств и от решаемых ими задач, поэтому они не всегда могут быть заданы заранее и должны являться параметрами запросов. Но для упрощения запросов необходимо предусмотреть некоторые значения по умолчанию, которые сохраняются в справочной таблице БД.

Теперь рассмотрим, каковы могут быть действия узла (робота), запрашивающего данные, в вышеприведенных ситуациях.

1. Робот не получил ответа за некоторый промежуток времени — тайм-аут. Значение тайм-аута зависит от времени обмена данными и от того периода, на протяжении которого связи между членами группы — роя — могут считаться постоянными. Эти значения могут быть определены только экспериментально, так как зависят от физических характеристик конкретных роботов и условий, в которых они функционируют. Но можно привести некоторые общие соображения относительно дальнейших действий робота в этой ситуации. Они зависят от того, насколько критичным для робота является отсутствие запрашиваемых данных. Например, если эти данные требуются для уточнения принимаемого решения, то он может начать выполнять принятое решение, не дожидаясь ответа, а при его получении скорректировать свои действия. Если робот не может принять решение без запрашиваемых данных, то по истечении тайм-аута робот может послать повторный запрос. Дальнейшие действия робота, например, если он и на повторный запрос не получил ответа, определяются общим алгоритмом его поведения.

2. Робот получил один ответ. Этот вариант тривиален.

3. Робот получил несколько ответов, и полученные данные одинаковы (с учетом погрешности измерения). Этот вариант также не требует действий, касающихся дополнительной обработки данных и/или обмена данными.

4. Робот получил несколько ответов, и полученные данные разные. В этом случае речь идет о работе с неточными или противоречивыми данными.

Обработка неточных и противоречивых данных

Рассмотрим наиболее распространенную ситуацию, при которой из всех полученных роботом разных значений определенного *r*-го показателя он должен выбрать одно наиболее правдоподобное (достоверное) значение Z^r.

Существуют различные подходы к обработке неточных и противоречивых данных. В области БД наиболее частой причиной появления таких данных являются синтаксические или семантические ошибки (при ручном вводе, при некачественной автоматизации ввода и обработки данных и т. д.) [6]. Проблема обработки неточных данных в БД чаще всего решается путем использования усредненных значений. Последовательность обработки противоречивой информации в БД обычно такая: сначала ошибки необходимо выявить, затем либо исключить из обработки, либо исправить. Один из подходов, основанный на удалении противоречивых данных, описан в работе [7], а обзор методов по очистке данных можно найти в работе [8]. Эти подходы подразумевают знание специфики предметной области и чаще всего предполагают использование различных статистических методов, позволяющих обнаружить ошибки путем поиска отклонений в данных, затем игнорировать ошибочные значения или, если это возможно, исправить их. Для нашей задачи этот подход не даст удовлетворительных результатов, так как его точность при небольшом объеме выборки невелика, а объем выборки определяется ограничениями статического роя, в котором каждый робот может общаться только со своими ближайшими соседями.

Вместе с тем большое внимание проблеме работы с противоречивыми данными уделяется в системах поддержки принятия решений (СППР). Например, в работе [9] авторы предлагают использовать для работы с такими данными аппарат аргументации, а в работе [10] — подход на основе использования мультимножеств и систем нечеткого вывода. Но в СППР в качестве противоречивых данных выступают субъективные оценки критериев или альтернатив, которые дают эксперты или лица, принимающие решения.

При решении проблемы противоречивых данных применительно к групповой робототехнике надо учитывать следующее. Во-первых, нет возможности воспользоваться экспертным мнением для выявления более правдоподобного значения показателя. Во-вторых, метод обработки противоречивых данных должен быть вычислительно простым, так как робот работает в режиме реального времени, а длительные вычисления могут приводить к недопустимым задержкам.

Будем считать, что при записи значений показателей в БД проверяются ограничения целостности, которым должны удовлетворять эти показатели. Например, если датчик температуры имеет рабочий диапазон от −60 до +100 °C, то значения, выходящие за этот диапазон, будут признаны ошибочными и не будут записаны в БД. Перечень ограничений целостности должен храниться в справочной таблице на каждом узле (роботе), поэтому ограничения целостности могут быть разными на разных узлах.

Неточные данные

Для начала рассмотрим самый простой вариант. Обозначим полученные в ответ на запрос данные как $U^r = \{x_i^r\}$, $0 < i < N$, где x_i^r — значения *r*-го показателя, снятые в одно время в одном и том же месте датчиками *N* разных роботов. Естественно, время и место определяются также с учетом погрешностей, которые даны в справочной таблице.

Пусть для $U^r = \{x_i^r\}$ выполняются следующие условия:

$$\begin{aligned} &\text{для } \forall i \ \ \exists j : \left|x_i^r - x_j^r\right| \leqslant \varepsilon_{\max}^r ;\ i = \overline{1,N},\ j = \overline{1,N}, \\ &\text{и } \exists m,l : \left|x_m^r - x_l^r\right| > \varepsilon_{\max}^r;\ m = \overline{1,N};\ l = \overline{1,N}, \\ &\varepsilon_{\max}^r = \max(\varepsilon_1^r,\ldots,\varepsilon_N^r), \end{aligned} \qquad (3)$$

где ε_n^r — погрешность *r*-го показателя, полученного от *n*-го робота; $\varepsilon_{\max}^r$ — максимальная погрешность измерения *r*-го показателя среди *N* роботов. Будем называть данные, удовлетворяющие условию (3), *неточными*, так как разброс значений превышает погрешность измерения. Все результаты измерения конкретного показателя можно представить как точки на числовой оси. Эти точки x_i^r, $1 \leqslant i \leqslant N$, можно рассматривать как реализации дискретной случайной величины *X* и использовать методы математической статистики для определения центра группировки и разброса точек. При вычислении математического ожидания $M[X] = \sum_{i=1}^{N} x_i^r p_i$ возникает вопрос с определением вероятностей p_i. Естественным образом в отсутствие дополнительных данных можно предположить, что все полученные значения равновероятны, и тогда математическое ожидание превращается в среднее значение. В подобных задачах для оценки достоверности полученного значения обычно вычисляют среднее квадратическое отклонение или используют метод доверительных интервалов. Но этот метод дает удовлетворительные по точности результаты при наличии выборки хотя

бы в 20—30 значений [11], а в наших условиях этого нельзя гарантировать ввиду специфики предметной области.

Откажемся от предположения о равновероятности полученных значений x_i^r и оценим степень достоверности данных, полученных от разных роботов. В соответствии с ГОСТ Р 51170—98 под достоверностью подразумевается свойство данных не иметь скрытых ошибок. Здесь можно провести аналогию с хорошо известным методом Дельфи, который используется для определения значения прогнозируемой величины на основе многократного опроса экспертов. Одним из этапов применения этого метода является расчет коэффициента компетентности эксперта с точки зрения всей группы:

$$w_i = \frac{Z_i - \bar{Z}}{\bar{Z}},$$

где $\bar{Z}$ — среднее значение оценок прогнозируемой величины, а Z_i — значение прогнозируемой величины, которое дал i-й эксперт. Этот коэффициент предъявляется эксперту при повторных опросах для того, чтобы он мог изменить свое мнение относительно прогнозируемой величины, если ранее данное им значение отличается от мнения других экспертов. В некоторых модификациях метода Дельфи (например, [12]) коэффициент компетентности интерпретируется как вес эксперта, и среднее значение $\bar{Z}$ рассчитывается как взвешенная сумма оценок всех экспертов, деленная на их число. Фактически, на основе достоверности ранее полученных данных можно рассчитать аналогичный по своей сути коэффициент и учитывать его при определении результирующего значения Z^r.

Будем учитывать опыт предыдущих обменов данными и рассчитаем *степень достоверности* данных W_n^r для n-го узла (робота), с которого поступает r-й показатель. Предлагается определять W_n^r для n-го робота как отношение числа достоверных значений r-го показателя к общему числу полученных от него значений r-го показателя:

$$W_n^r = \frac{I_n^{r+}}{I_n^r}. \quad (4)$$

Значение r-го показателя считается достоверным, если оно отличается от значения Z^r не более чем на значение погрешности измерения r-го показателя $\varepsilon_{\max}^r$. (Заметим, что значение $\varepsilon_{\max}^r$ — переменная величина, которая рассчитывается заново для каждого запроса, на который пришло несколько ответов). Проблема заключается в том, что при первом получении значения r-го показателя от n-го узла степень достоверности по формуле (4) не может быть вычислена, так как I_n^{r+} еще неизвестно. Кроме того, если полученное значение не будет признано достоверным, то степень достоверности для n-го робота будет равна 0, и его показания просто не будут учитываться. Поэтому предлагается оценивать степень достоверности W_n^r следующим образом:

$$W_n^r = \frac{1 + I_n^{r+}}{1 + I_n^r}. \quad (5)$$

То есть изначально будем считать, что полученные от узла данные абсолютно достоверны ($W_n^r = 1$). Окончательно значение Z^r при условии (3) можно определить так:

$$Z^r = \sum_{i=1}^{N} x_i^r w_n^r, w_n^r = \frac{W_n^r}{\sum_{n=1}^{N} W_n^r}, \quad (6)$$

где w_n^r — коэффициент достоверности значения r-го показателя, полученного с n-го узла.

Противоречивые данные

Более сложным является вариант, при котором элементы в упорядоченном по значениям множестве $U^r = \{x_i^r\}$, $1 \leqslant i \leqslant N$, не образуют компактную группу точек, т. е. существуют хотя бы два соседних элемента, разница между которыми больше погрешности. Пусть для U^r выполняются следующие условия:

$$\exists m, l: \left|x_m^r - x_l^r\right| > \varepsilon_{\max}^r;\ m = \overline{1, N};\ l = \overline{1, N},$$
$$\text{и } \forall j: \left|x_m^r - x_j^r\right| > \varepsilon_{\max}^r \text{ или } \left|x_l^r - x_j^r\right| > \varepsilon_{\max}^r;\ j = \overline{1, N}, \quad (7)$$
$$\varepsilon_{\max}^r = \max(\varepsilon_1^r, \ldots, \varepsilon_N^r).$$

Тогда множество U^r можно разбить на два подмножества, причем разность значений любых двух элементов разных подмножеств превысит максимальную погрешность измерений. Будем называть данные, удовлетворяющие условию (7), *противоречивыми*. Рассмотрим вариант, при котором упорядоченное по значениям элементов множество U^r можно разбить на подмножества $U_1^r = \{x_1, \ldots, x_d\}$ и $U_2^r = \{x_{d+1}, \ldots, x_N\}$ следующим образом (рис. 2):

$$|x_1 - x_d| \leqslant \varepsilon_{\max}^r;$$
$$|x_{d+1} - x_N| \leqslant \varepsilon_{\max}^r;$$
$$|x_{d+1} - x_d| > \varepsilon_{\max}^r.$$

Так как среди элементов этих подмножеств могут быть равные значения, их можно рассматривать как мультимножества или множества с повторяющимися элементами [13]. Более того, мы

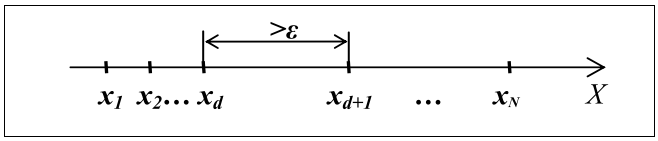

Рис. 2. Пример разбиения множества значений на два подмножества

договорились считать два значения одинаковыми, если разность между ними не превышает погрешности измерений. Если в упорядоченном подмножестве значений U_i^r разница между двумя крайними элементами не превышает погрешности, то элементы этого подмножества неразличимы. Величины d и $(N - d)$ являются мощностью мультимножеств U_1^r и U_2^r соответственно (обозначим d как d_1 и $(N - d)$ как d_2). При условии, что каждое из этих мультимножеств состоит из одинаковых элементов, их размерность (т. е. число различных элементов) равна 1, а d_1 и d_2 являются как кратностью вхождения элементов в соответствующие мультимножества ($k(U_1^r) = d_1$, $k(U_2^r) = d_2$), так и их высотой h: $h\ U_1^r = \max k(U_1^r) = d_1$, $h\ U_2^r = \max k(U_2^r) = d_2$. При этом возможны следующие варианты:

1) $d_1 = 1$; $d_2 = 1$;
2) $d_1 > 1$; $d_2 = 1$ или $d_2 > 1$; $d_1 = 1$;
3) $d_1 > 1$; $d_2 > 1$.

В случае варианта 1 рациональным является выбор того значения, достоверность которого выше. Тривиальной является ситуация 2, когда мощность одного подмножества равна 1, а мощность другого — больше 1. Например, если $d_1 = 1$ и $d_2 > 1$, то можно считать, что противоречивое значение x_1 возникло вследствие ошибки измерения или ошибки передачи данных, и игнорировать это значение.

Если мощность обоих мультимножеств больше 1, то нельзя просто выбрать в качестве Z^r значение из того подмножества, чья мощность (высота) больше. Необходимо учитывать наличие противоречивых данных, поэтому предлагается ввести дополнительную характеристику W_{Ui}^r — общую степень достоверности подмножества U_i^r. Рассмотрим степень достоверности W_n^r (5) как функцию принадлежности нечеткого множества. Если считать каждый из элементов подмножества свидетельством в пользу того, что данный параметр принадлежит этому подмножеству, то общую степень достоверности данных подмножества U_i^r можно рассчитать, например, как алгебраическую сумму W_n^r. Для случая, когда подмножество U_i^r состоит из двух значений, алгебраическая сумма вычисляется так:

$$W_{Ui}^r = W_1^r + W_2^r - W_1^r \cdot W_2^r. \qquad (8)$$

Далее Z^r будет рассчитываться по формуле (6) на основе данных того мультимножества, для которого величина W_{Ui}^r больше.

Аналогичным образом проводят расчеты в случае, когда число подмножеств больше двух.

Неточные и противоречивые данные

Противоречивые данные также могут быть неточными, если разница между двумя крайними значениями подмножества U_i^r превышает максимально допустимую погрешность (условие (3)). Высота такого подмножества будет меньше его мощности, так как не все элементы подмножества являются одинаковыми в соответствии с (1). Здесь для выбора подмножества также может быть рассчитана общая степень достоверности (8). Если же оба подмножества имеют одинаковую общую степень достоверности, то в качестве основного целесообразно брать подмножество с большей высотой, а при одинаковых высотах — с большей мощностью.

В том случае, если полученные по запросу данные таковы, что оба подмножества имеют одинаковые общую степень достоверности, мощность и высоту, то возможны как минимум два варианта действий:

1) послать повторный запрос для уточнения данных;

2) отказаться от возможности получить единственное значение вместо множества и применить, например, метод недоопределенных вычислений.

Об особенностях применения метода недоопределенных вычислений в робототехнике более подробно рассказано в работе [14].

Аналогичные рассуждения могут быть применены к ситуации, когда данные разбиваются более чем на два подмножества.

Таким образом, в работе были предложены подходы к решению проблемы обработки неточных и противоречивых данных в группе мобильных роботов с помощью оценки степени достоверности данных, которая основана на учете ранее полученных значений. При этом степень достоверности (8) может выступать в качестве атрибута значения Z^r и в дальнейшем использоваться роботом при принятии решений.

Заключение

Важная особенность предложенного подхода к работе с неточными и противоречивыми данными заключается в том, что он является вычислительно простым и не требует большого объема

памяти у робота для хранения вспомогательной информации.

Недостатком этого подхода является возможная несогласованность сведений о достоверности данных, полученных от одних и тех же источников разными роботами. Это связано с тем, что степень достоверности вычисляется локально и на основе тех данных, которые запросил и получил конкретный узел (робот). Никакой процедуры согласования не предусматривается, так как это сильно увеличит трафик и, кроме того, это сложно выполнить технически. Группа роботов время от времени может распадаться на несвязанные подгруппы, между которыми прерывается обмен данными; в этой ситуации согласование невозможно в принципе.

В ходе дальнейших исследований планируется разработать необходимое алгоритмическое и программное обеспечение и провести эксперименты на реальных робототехнических устройствах.

Работа выполнена при частичной финансовой поддержке грантов РФФИ 16-29-04412 (метод работы с неточными и противоречивыми данными) и РНФ 16-01-00018 (выявление отличительных характеристик общей базы данных для группы роботов и описание структуры хранимых данных).

Список литературы

1. **Каляев И. А., Гайдук А. Р., Капустян С. Г.** Модели и алгоритмы коллективного управления в группах роботов. М.: Физматлит, 2009. 280 с.

2. **Карпов В. Э.** Управление в статических роях. Постановка задачи // "Интегрированные модели и мягкие вычисления в искусственном интеллекте". Сб. научных трудов VII-й Международной научно-практической конференции (Коломна, 20—22 мая 2013). В 3 томах. М.: Физматлит, 2013. Т. 2. С. 730—739.

3. **Madden S., Hellerstein J., Hong W.** TinyDB: In-Network Query Processing in TinyOS: Version 0.4. 2003. 46 p.

4. **TinyOS** Documentation Wiki [Electronic resource]. URL: http://tinyos.stanford.edu/tinyos-wiki/index.php/TinyOS_Documentation_Wiki.

5. **Воробьев В. В., Паршикова Е. А.** Применение мультимножеств для оценки ситуации мобильным агентом // Информационные технологии. 2015. Т. 21, № 6. С. 421—426.

6. **Singh R., Singh K.** A Descriptive Classification of Causes of Data Quality Problems in Data Warehousing // IJCSI Int. J. Comput. Sci. Issues. 2010. Vol. 7, N. 2. P. 41—50.

7. **Горбатков С. А., Белолипцев И. И., Мурзина Е. А.** Удаление противоречивых наблюдений как процедура предрегуляризации нейросетевой модели налогового контроля // Вестник УГАТУ. 2013. Т. 17, № 5. С. 110—114.

8. **Rahm E., Do H.** Data cleaning: Problems and current approaches // IEEE Data Eng. Bull. 2000. Vol. 23, N 4. P. 3—13.

9. **Вагин В. Н., Фомина М. В.** Аргументация в индуктивном формировании понятий // Образовательные ресурсы и технологии. 2014. Вып. 2, № 5. С. 34—39.

10. **Гусева М. В., Демидова Л. А.** Генерирование решающих правил классификации инвестиционных проектов на основе систем нечеткого вывода и мультимножеств // Системы управления и информационные технологии. 2006. № 4. С. 46—53.

11. **Вентцель Е. С.** Теория вероятностей: учеб. для вузов. 6-е изд. стер. М.: Высш. шк., 1999. 576 с.

12. **Беляевский И. К.** Маркетинговое исследование: информация, анализ, прогноз. М.: Финансы и статистика, 2014. 320 с.

13. **Петровский А. Б.** Пространства множеств и мультимножеств. М.: Едиториал УРСС, 2003. 248 с.

14. **Карпов В. Э.** О некоторых особенностях применения недоопределенных моделей в робототехнике // Международная научно-практическая конференция "Интегрированные модели и мягкие вычисления в искусственном интеллекте" (28—30 мая 2009) Сб. научных трудов. Т. 1. М.: Физматлит, 2009. С. 520—532.

I. P. Karpova, Associate Professor, karpova_ip@mail.ru,
National Research University Higher School of Economics, Moscow, Russia,
National Research Center "Kurchatov Institute", Moscow, Russia

Distributed Data Storage and Processing for Mobile Robotic Groups

This paper discusses a problem of distributed data processing in mobile robot's group. The robot's group is typified as a static swarm. Static swarm is a model, which is characterized by the absence of a control center and is represented by the network with fixed topology at some time interval which consists of locally interacting agents. The main features that distinguish the robot's group general database from the classic distributed database are described. It is shown that the database does not require the global data dictionary storing information about the location of the database fragments. The data structure on each node is the same and can be described in the reference table. This table is loaded into robot's memory when robot is initialized and, in fact, is the data dictionary. This database does not require distributed transactions, because data is written in the general database locally. The approach of logical queries organization in general database is offered. Definitions of imprecise and inconsistent data conformably to data which robots in the group are exchanging with are given. The approach of processing imprecise and inconsistent data which come to robot is proposed. This approach is based on elements of multisets theory, fuzzy sets theory and on evaluation of data reliability degree. The reliability degree is based on the experience of the previous data exchanges between robots. An important feature of the proposed method of imprecise and inconsistent data processing is that it is computationally simple and does not require much memory to store auxiliary information.

Keywords: *group robotics, static swarm, database, processing of imprecise and inconsistent data*

References

1. **Kaljaev I. A., Gajduk A. R., Kapustjan S. G.** *Modeli i algoritmy kollektivnogo upravlenija v gruppah robotov*, (Models and algorithms of collective control in robots groups), Moscow, Fizmatlit, 2009. 280 p. (in Russian).
2. **Karpov V. E.** Upravlenie v staticheskih rojah. Postanovka zadachi (Control in static swarms. Problem statement), *Integrirovannye modeli i mjagkie vychislenija v iskusstvennom intellekte. Sb. nauchnyh trudov VII-j Mezhdunarodnoj nauchno-prakticheskoj konferencii* (Kolomna, 20.05—22.05.2013), In 3 vol. Moscow, Fizmatlit, 2013, vol. 2, pp. 730—739 (in Russian).
3. **Madden S., Hellerstein J., Hong W.** *TinyDB: In-Network Query Processing in TinyOS: Version 0.4.* 2003. 46 p.
4. **TinyOS** Documentation Wiki, available at: http://tinyos.stanford.edu/tinyos-wiki/index.php/TinyOS_Documentation_Wiki.
5. **Vorob'ev V. V., Parshikova E. A.** Primenenie mul'timnozhestv dlja ocenki situacii mobil'nym agentom (Application of multisets for assessment of the situation by the mobile agent), *Informacionnye tehnologii*, 2015, vol. 21, no. 6, pp. 421—426 (in Russian).
6. **Singh R., Singh K.** A Descriptive Classification of Causes of Data Quality Problems in Data Warehousing, *IJCSI Int. J. Comput. Sci. Issues.*, 2010, vol. 7, no. 2, pp. 41—50.
7. **Gorbatkov S. A., Belolipcev I. I., Murzina E. A.** Udalenie protivorechivyh nabljudenij kak procedura predreguljarizacii nejrosetevoj modeli nalogovogo kontrolja (Removal of inconsistent observations as the procedure of prerequisitie neural network model of tax control), *Vestnik UGATU*, 2013, vol. 17, no. 5, pp. 110—114 (in Russian).
8. **Rahm E., Do H.** Data cleaning: Problems and current approaches, *IEEE Data Eng. Bull.*, 2000, vol. 23, no. 4, pp. 3—13.
9. **Vagin V. N., Fomina M. V.** Argumentacija v induktivnom formirovanii ponjatij (The argument in the inductive formation of concepts), *Obrazovatel'nye resursy i tehnologii*, 2014, vol. 2, no. 5, pp. 34—39 (in Russian).
10. **Guseva M. V., Demidova L. A.** Generirovanie reshajushhih pravil klassifikacii investicionnyh proektov na osnove sistem nechetkogo vyvoda i mul'timnozhestv (Decision rules generating for the investment projects classification based on fuzzy inference systems and multisets), *Sistemy upravlenija i informacionnye tehnologii*, 2006, no. 4, pp. 46—53 (in Russian).
11. **Ventcel' E. S.** *Teorija verojatnostej* (Probability theory): Ucheb. dlja vuzov. 6-e izd. ster., Moscow, Vyssh. shk., 1999. 576 p. (in Russian).
12. **Beljaevskij I. K.** *Marketingovoe issledovanie: informacija, analiz, prognoz* (Marketing research: information, analysis, forecast). Moscow, Finansy i statistika, 2014, 320 p. (in Russian).
13. **Petrovskij A. B.** *Prostranstva mnozhestv i mul'timnozhestv* (Spaces of sets and multisets), Moscow, Editorial URSS, 2003. 248 p. (in Russian).
14. **Karpov V. E.** O nekotoryh osobennostjah primenenija nedoopredelennyh modelej v robototehnike (Some features of the application of subdefinite models in robotics), *Mezhdunarodnaja nauchno-prakticheskaja konferencija "Integrirovannye modeli i mjagkie vychislenija v iskusstvennom intellekte"* (28.05—30.05.2009), sb. nauchnyh trudov, vol. 1, Moscow, Fizmatlit, 2009, pp. 520—532 (in Russian).

УДК 004.94, 620.91

Д. Н. Кобзаренко, д-р техн. наук, зав. лаб., e-mail: kobzarenko_dm@mail.ru,
А. М. Камилова, вед. специалист, e-mail: anna702@mail.ru,
Н. Ш. Газанова, вед. специалист, e-mail: dyuzya@gmail.com,
Федеральное государственное бюджетное учреждение науки Институт проблем геотермии Дагестанского научного центра Российской академии наук, г. Махачкала, Россия,
А. М. Дадашев, начальник, e-mail: daggidromet@mail.ru,
Дагестанский центр по гидрометеорологии и мониторингу окружающей среды — филиал Федерального государственного бюджетного учреждения "Северо-Кавказское управление по гидрометеорологии и мониторингу окружающей среды", г. Махачкала, Россия

Применение непрерывного вейвлет-преобразования в изучении временных рядов ветромониторинга на примере Дагестана

Рассмотрены результаты частотно-временного анализа временных рядов — скоростей ветра с помощью информационных технологий и вейвлет-преобразования вейвлетом Morlet.

Идея пространственно-временного анализа данных ветромониторинга состоит в том, чтобы показать наличие или отсутствие закономерностей во временных рядах — скоростях и направлениях ветра. При этом анализ выполняется как по пространству (сопоставление данных ветромониторинга за один временной период), так и по времени (несколько лет).

Для выполнения анализа получены достоверные данные ветромониторинга, выполняемого в Дагестане метеорологическими станциями. Данные охватывают временной период 2011—2015 гг. и получены с метеостанций: "Ахты", "Дербент", "Махачкала" и "Кочубей". На основе исходных данных подготовлены по четыре временных ряда для каждой метеостанции. Каждый временной ряд охватывает два календарных года (2011—2012, 2012—2013, 2013—2014, 2014—2015). Данные представлены частотой 8 измерений в сутки.

Ключевые слова: *скорость ветра, ветроэнергетика, вейвлет-преобразование, частотно-временной анализ*

Введение

Исследования по применению вейвлет-преобразования для анализа данных метеорологических временных рядов известны как в России [1, 2], так и за рубежом [3]. Причем в работах [1, 2] приведены идеи о применении вейвлет-преобразования для восстановления промежуточных