

РОЗПОДІЛЕНА БАЗА ДАНИХ ВІРТУАЛЬНИХ КООРДИНАТ ВУЗЛІВ

¹Валуйський С.В., ¹Лукашук В.В. ²Чумаченко С.М.

¹ Інститут телекомунікаційних систем КПІ ім. Ігоря Сікорського, Україна

² Український науко-дослідний інститут цивільного захисту, Україна

E-mail: vas77794@gmail.com; sergiy23.chumachenko@gmail.com

Distributed data base of nodes virtual coordinates

The main purpose of this work was to denote main disadvantages and problems related with saving data about nodes in WSN. Estimate memory usage of distributed database based on hash tables using notion of virtual coordinates.

Хеш-таблиця являє собою узагальнення звичайного масиву, будучи асоціативною пам'яттю, в якій доступ до елементів здійснюється по ключу, а не за індексом. Хеш-таблиця дозволяє зберігати пари (ключ, значення) і підтримує три базові операції: додавання в хеш-таблицю значення пари ключ і значення, зчитування та видалення.

У безпроводних сенсорних мережах хеш-таблиці можуть використовуватися для зберігання різного роду інформації (наприклад, зібрані показники датчиків або звіти про зафіксовані події). Однак централізоване зберігання хеш-таблиці на деякому виділеному вузлу має ряд істотних недоліків, зокрема: низька стійкість до відмов, погана масштабованість і висока ймовірність перевантаження системи. Отже, більш доцільним рішенням є організація розподіленої хеш-таблиці на основі мережі з безлічі вузлів, які виступають в ролі елементів пам'яті. В цьому випадку завдання хеш-функції полягає у визначенні відповідності між ключем k і вузла, відповідального за зберігання значення d .

Розподілені географічні хеш-таблиці отримала свій розвиток в системі CellHashRouting [1] в якій площа покриття мережі розбивається на безліч кластерів, і між ними виконується розподіл зберігається інформації за допомогою GPSR. Однак географічні хеш-таблиці мають той ж недолік, що принцип маршрутизації який лежить в їх основі, тому в даній роботі пропонується оцінити використання пам'яті у системі розподілених хеш-таблиць на основі віртуальних координат.

Однією з проблем більшості алгоритмів геометричній маршрутизації є необхідність знати координати одержувача для початку процесу доставки пакета. Очевидно, що в більшості випадків відправник знає тільки адресу (ідентифікатор) одержувача, тому потрібна деяка мережева служба, що

дозволяє встановити відповідність між адресом вузла та його віртуальними координатами.

З деяких роботах з маршрутизації по віртуальним координатам дане питання розглядається тільки в [2] і [3], в яких пропонується використання опорних вузлів у якості сховищ координат усіх інших вузлів мережі, ґрунтуючись на тому факті, що вузли завжди знають як доставити пакет до опорного вузла. Якщо віртуальні координати вузла v змінилися, то він обчислює номер опорного вузла, відповідального за зберігання цієї інформації, за допомогою деякої хеш-функції H , яка перетворює адреса вузла a_v в один з індексів опорних вузлів, і передає йому значення v . Якщо деякому вузлу w необхідно передати пакет до v , то спочатку він запитує значення вектора v в опорного вузла з номером $H(a_v)$ і після отримання відповіді починає процес маршрутизації.

Такий даний підхід не відповідає вимогам масштабованості системи з наступних причин. Припустимо ідеальну ситуацію, в якій хеш-функція забезпечує рівномірний розподіл віртуальних координат n вузлів між n_L опорними вузлами, тобто в кожному опорному вузлі зберігаються n / n_L віртуальних координат. Враховуючи, що для кожного вектору віртуальних координат потрібно $\Theta(n_L)$ пам'яті, навантаження на кожен опорний вузол становитиме

$$M_L = \frac{n}{n_L} \Theta(n_L) \quad (1)$$

що робить цю схему непридатною в мережах великих розмірів.

Для вирішення цієї проблеми пропонується використовувати розподілену хеш-таблицю на основі віртуальних координат, в якій ключами є адреси вузлів, а значеннями - їх координати, тобто в таблиці розміщуються пари виду (a_v, \vec{v}) . При цьому кількість записів в хеш-таблиці і її ємність збігаються, тобто $n_k = n$. Отже, для організації подібної хеш-таблиці кожному вузлу мережі в середньому буде потрібно $M_{avg} = \Theta(n_L)$ пам'яті, а в найгіршому випадку – $M_{max} = \Theta(\rho n_L)$ що значно краще оцінки (1), враховуючи, що, як правило, $\rho = \text{const} \cdot n_L = \text{const}$, а також $n_L \ll n$.

Потрібно відмітити, що оцінки для витрат пам'яті мають місце при типі трафіку «багато-до-багатьох», коли кожен вузол мережі є потенційним одержувачем. Якщо ж умови прикладної завдання такі, що кількість одержувачів мала, то і витрати пам'яті будуть значно меншими. Мережева служба для співставлення віртуальних координат вузла на його адреси має виконати наступні дії:

1. Кожен вузол мережі v , який потенційно може бути одержувачем пакета у загальному випадку зазначені дії виконують всі вузли мережі), при зміні своїх віртуальних координат поміщає нове значення \vec{v} в розподілену хеш-таблицю, тобто відправляє відповідний пакет віртуальному вузлу з координатами $H(a_v) = \vec{h}(a_v)$.

2. Потенційні одержувачі пакетів також періодично оновлюють збережене в хеш-таблиці значення своїх віртуальних координат, навіть якщо координати не змінилися.

3. Якщо деякому вузлу w необхідно передати пакет вузлу v , то вузол w робить запит в хеш-таблицю, передаючи пакет з координатами $\vec{h}(a_v)$, і після отримання у відповідь вектору v починає процес маршрутизації. При наступній передачі пакета вузлу w немає необхідності повторно зчитувати координати з хеш-таблиці за рахунок кешування.

Топологія мережі може змінюватися з різних причин, тому фізичний вузол із сховище, з відповідними віртуальним координатам $\vec{h}(a_v)$, також може змінитися. Зазначене в пункті 2 примусове періодичне оновлення значень у хеш-таблиці встановить нову відповідність між точкою і віртуальним простором $\vec{h}(a_v)$ і найближчим до неї фізичним вузлом згідно поточної топології мережі. Час оновлення може мати різні значень, виходячи із швидкості змін, що відбуваються в топології мережі. При цьому період може бути як фіксованим (заданим на етапі проектування системи), так і динамічно налаштованому процесі роботи. Сказане також відноситься і до параметрів кешування у пункті 3.

Крім кешування можливі інші варіанти оптимізації запитів в базу даних координат, вибір найбільш доцільного з них залежить від умов експлуатації мережі і розв'язуваної прикладної завдання.

Було розглянуто основні недоліки і проблеми пов'язані із зберіганням даних в безпроводних сенсорних мережах, які використовують методи географічної маршрутизації. Запропонована оцінка використання пам'яті при використанні розподіленої бази даних з віртуальними координатами.

Література

1. CHR: a distributed hash table for wireless ad hoc networks / R Araujo [et al.] // Proceedings Of the 25th IEEE international conference of distributed computing systems workshops. Columbus (USA), 2005. Vol. 6, no. 10. P. 407-413.
2. Bea convector routing: scalable point-to-point routing in wireless sensor nets / R. Fonseca [etal.] // Proceedings in the 2nd symposium networked systems designand implementation. Boston (USA), 2005. P. 329-342.
3. Efficient hop ID based routing for spar sea doc networks / Y. Zhao [etal] // Proceedings of the 13th IEEE international conference of network protocols. Boston (USA), 2005. P. 179-190.