

TRƯỜNG ĐẠI HỌC BÁCH KHOA HÀ NỘI
KHOA CÔNG NGHỆ THÔNG TIN

— * —

ĐỒ ÁN
TỐT NGHIỆP ĐẠI HỌC
NGÀNH CÔNG NGHỆ THÔNG TIN

LƯU TRỮ THÔNG TIN HƯỚNG DẪN LIỆU

Sinh viên thực hiện : **Lương Hữu Long**
Lớp KSTN CNTT- K49
Giáo viên hướng dẫn: TS. **Nguyễn Khanh Văn**

HÀ NỘI 6-2009

PHIẾU GIAO NHIỆM VỤ ĐỒ ÁN TỐT NGHIỆP

1. Thông tin về sinh viên

Họ và tên sinh viên: Lương Hữu Long

Điện thoại liên lạc: 01686669625

Email: huulongit@gmail.com

Lớp: KSTN CNTT – K49

Hệ đào tạo: Chính quy

Đồ án tốt nghiệp được thực hiện tại: Bộ môn Công nghệ phần mềm

Thời gian làm ĐATN: Từ ngày 15 / 01 / 2009 đến 25 / 05 / 2009

2. Mục đích nội dung của ĐATN

Tìm hiểu, nghiên cứu các vấn đề lý thuyết liên quan đến lưu trữ thông tin hướng dữ liệu (Data-centric Storage). Tìm hiểu, khai phá kiến thức về NS2 – một sản phẩm phần mềm lớn, chuyên dụng cho mô phỏng mạng; xây dựng hệ thống lưu trữ thông tin hướng dữ liệu (DCS) cơ bản mang tính mô phỏng trên NS-2.

3. Các nhiệm vụ cụ thể của ĐATN

1. Tìm hiểu lý thuyết tổng quan về vấn đề lưu trữ thông tin hướng dữ liệu (Data-centric Storage)
2. Nghiên cứu, đánh giá và so sánh hai thuật toán tìm đường cơ bản được sử dụng trong DCS là GPSR – giao thức định tuyến tiêu biểu áp dụng cho mạng không dây và HVGR – thuật toán sử dụng mô hình tọa độ ảo phân cấp trong đồ thị Voronoi (điểm mốc - landmark).
3. Tìm hiểu, nghiên cứu về vấn đề cân bằng tải trong mạng sensor nói chung và trong DCS nói riêng
4. Tìm hiểu, nghiên cứu về NS2 và xây dựng một hệ thống mạng đơn giản sử dụng DCS

4. Lời cam đoan của sinh viên:

Tôi – *Lương Hữu Long* – cam kết ĐATN là công trình nghiên cứu của bản thân tôi dưới sự hướng dẫn của *TS. Nguyễn Khanh Văn*.

Các kết quả nêu trong ĐATN là trung thực, không phải là sao chép toàn văn của bất kỳ công trình nào khác.

Hà Nội, ngày 25 tháng 05 năm 2009

Tác giả ĐATN

Lương Hữu Long

5. Xác nhận của giáo viên hướng dẫn về mức độ hoàn thành của ĐATN và cho phép bảo vệ:

Hà Nội, ngày tháng năm

Giáo viên hướng dẫn

TS. Nguyễn Khanh Văn

TÓM TẮT NỘI DUNG ĐỒ ÁN TỐT NGHIỆP

Công nghệ không dây ra đời từ thế kỷ 19 dẫn tới sự hình thành, phát triển của mạng không dây. Mạng không dây, mà điển hình là mạng cảm biến không dây (WSN), do đặc điểm của mạng được cấu tạo từ các thiết bị nhỏ có khả năng lưu trữ, xử lý thông tin thấp, quá trình hoạt động lại phụ thuộc vào nguồn cung cấp năng lượng hạn chế nên việc truy xuất dữ liệu trong mạng luôn gặp nhiều khó khăn. Vấn đề đặt ra ở đây là cần xây dựng một hệ thống mạng phân phối dữ liệu hiệu quả, hạn chế các tình trạng quá tải, tắc nghẽn có thể xảy ra trong quá trình truy xuất dữ liệu trong mạng sensor.

Đồ án “**Lưu trữ thông tin hướng dữ liệu trong mạng sensor**” được thực hiện nhằm tìm hiểu và nghiên cứu những vấn đề cơ bản của phương pháp lưu trữ thông tin hướng dữ liệu – một phương pháp phân phối dữ liệu hiệu quả, các kỹ thuật được sử dụng trong DCS; các thuật toán tìm đường cơ bản được sử dụng phổ biến trong DCS và vấn đề cân bằng tải cho DCS trong mạng sensor. Cuối cùng kết hợp những kiến thức tìm hiểu về công cụ NS2 và kiến thức về hệ thống DCS, đồ án sẽ xây dựng một hệ thống DCS đơn giản mô phỏng trong NS2 làm tiền đề cho việc xây dựng hệ thống DCS thực trong mạng sensor.

ABSTRACT OF THESIS

Wireless technology, born in 19th century, is the foundation for the establishment and development of wireless networks. Because wireless networks, particularly sensor wireless networks, consist of tiny devices with low storage and information processing, and their performance depends on limited power supply, so data manipulation encounters many problems. The problem is that we need construct a efficient data distribution network system, prevent system overload or bottleneck from occuring in data manipulation process in sensor network

The purpose of my thesis titled “**Data – centric storage**” is to study fundamental aspects of Data–centric storage, mechanisms for Data-centric storage, routing agorithms used in finding routes for DCS and load balancing in sensor networks; then, to assess the advantages, disadvantages and the feasibility of applying these algorithms. Lastly, on the basis of the integration between the knowledge of data-centric storage and that of NS2 tool, a simple DCS system are established so as to be used in NS2.

LỜI CẢM ƠN

Trước tiên tác giả xin gửi lời cảm ơn chân thành và sâu sắc nhất tới thầy TS. Nguyễn Khanh Văn – Chủ nhiệm bộ môn Công nghệ phần mềm, người đã tận tình chỉ bảo, hướng dẫn và giúp đỡ em trong suốt quá trình nghiên cứu làm đồ án này.

Xin được gửi lời cảm ơn tới tất cả các thầy, các cô trường Đại học Bách khoa Hà Nội nói chung và các thầy, cô khoa Công nghệ thông tin trong 5 năm qua đã tận tình dạy bảo, truyền đạt không chỉ những kiến thức khoa học mà cả những bài học về cuộc sống, làm hành trang cho em trong tương lai. Tác giả xin xin được gửi lời cảm ơn đến bạn bè lớp KSTN CNTT K49 về sự giúp đỡ trong quá trình học tập đã qua.

Cuối cùng con xin được gửi lời cảm ơn chân thành nhất tới bố mẹ. Sự quan tâm, tình thương của bố mẹ sẽ mãi là chỗ dựa vững chắc cho con trong hiện tại và trong cuộc sống sau này.

MỤC LỤC

DANH MỤC CÁC HÌNH MINH HỌA	8
DANH MỤC CÁC TỪ VIẾT TẮT	9
LỜI MỞ ĐẦU	10
CHƯƠNG 1 : ĐẶT VẤN ĐỀ	11
2.1. Tổng quan về mạng sensor	Error! Bookmark not defined.
2.1.1. Giới thiệu chung	Error! Bookmark not defined.
2.1.2. Đặc trưng và cấu hình mạng sensor	Error! Bookmark not defined.
2.1.3. Một số chuẩn mạng sensor	Error! Bookmark not defined.
2.1.4. Ứng dụng của mạng sensor	Error! Bookmark not defined.
2.2. Tổng quan về Data – centric storage (DCS)	Error! Bookmark not defined.
2.2.1. Khái quát về lưu trữ thông tin hướng dữ liệu (Data-centric storage)	Error! Bookmark not defined.
2.2.2. Vấn đề tìm đường (routing) trong DCS	Error! Bookmark not defined.
2.2.3. Các yêu cầu cần thiết trong DCS	Error! Bookmark not defined.
3.1. Các phương thức phân phối dữ liệu trong mạng sensor ..	Error! Bookmark not defined.
3.1.1. Các phương thức mẫu	Error! Bookmark not defined.
3.1.2. Chi phí giao tiếp gần đúng	Error! Bookmark not defined.
3.1.3. Các phương thức phân phối bổ sung	Error! Bookmark not defined.
3.2. Bản chất của DCS	Error! Bookmark not defined.
3.3. Các kỹ thuật được sử dụng trong DCS	Error! Bookmark not defined.
3.3.1. Thiết kế chuẩn cho DCS : Robustness ⁽¹⁾ và Scalability ⁽²⁾	Error! Bookmark not defined.
3.3.2. GHT : A Geographic Hash Table	Error! Bookmark not defined.
3.3.3. Grid GLS : Grid Geographic Location Server	Error! Bookmark not defined.
3.4. Thuật toán tìm đường trong DCS : Greedy Perimeter Stateless Routing (GPSR)	Error! Bookmark not defined.
3.4.1. Thuật toán GPSR	Error! Bookmark not defined.
3.4.2. Nút nhà (home node) và đường bao nhà (home perimeter)	Error! Bookmark not defined.
3.4.3. Giao thức làm mới đường bao (PRP)	Error! Bookmark not defined.
3.4.4. Sự tái tạo có cấu trúc	Error! Bookmark not defined.
3.5. Thuật toán tìm đường trong DCS : Hierarchical Voronoi Graph Routing (HVGR)	Error! Bookmark not defined.
3.5.1. Giới thiệu chung	Error! Bookmark not defined.
3.5.2. Thuật toán tìm đường cơ bản	Error! Bookmark not defined.
3.5.3. Xây dựng cấu trúc phân cấp	Error! Bookmark not defined.
3.5.4. Lựa chọn điểm mốc	Error! Bookmark not defined.
3.5.5. Cân bằng tải lưu trữ	Error! Bookmark not defined.
3.5.6. Điều khiển các thay đổi tự động	Error! Bookmark not defined.
3.6. So sánh GPSR và HVGR	Error! Bookmark not defined.
3.6.1. Về thuật toán	Error! Bookmark not defined.
3.6.2. Về cài đặt mô phỏng	Error! Bookmark not defined.
3.6.3. Về khả năng hỗ trợ cho DCS	Error! Bookmark not defined.

CHƯƠNG 4. CÂN BẰNG TẢI (LOAD – BALANCING) CHO DCS TRONG MẠNG SENSOR	Error! Bookmark not defined.
4.1. Khái niệm và ứng dụng cân bằng tải trong mạng sensor.....	Error! Bookmark not defined.
4.1.1. Khái niệm cân bằng tải	Error! Bookmark not defined.
4.1.2. Ứng dụng cân bằng tải trong các mạng sensor	Error! Bookmark not defined.
4.2. Thuật toán cân bằng tải	Error! Bookmark not defined.
4.2.1. Thuật toán cơ bản.....	Error! Bookmark not defined.
4.2.2. Cải tiến thuật toán	Error! Bookmark not defined.
4.3. Cân bằng tải cho DCS.....	Error! Bookmark not defined.
CHƯƠNG 5: CÔNG CỤ MÔ PHỎNG MẠNG SENSOR NS-2..	Error! Bookmark not defined.
5.1. Tổng quan về NS – 2	Error! Bookmark not defined.
5.1.1. Giới thiệu chung về NS – 2	Error! Bookmark not defined.
5.1.2. Sự liên kết OTcl (OTcl linkage)	Error! Bookmark not defined.
5.1.3. Lớp Simulation.....	Error! Bookmark not defined.
5.1.4. Quá trình chuyển tiếp gói và các nút.....	Error! Bookmark not defined.
5.1.5. Các Agent	Error! Bookmark not defined.
5.1.6. Các liên kết : liên kết đơn.....	Error! Bookmark not defined.
5.1.7. Gói dữ liệu (Packet)	Error! Bookmark not defined.
5.1.8. Tổng kết	Error! Bookmark not defined.
5.2. Kết quả mô phỏng bước đầu của hệ thống DCS trong NS-2..	Error! Bookmark not defined.

DANH MỤC CÁC HÌNH MINH HỌA

Hình 1.1. Một ví dụ về mô hình mạng sensor.....	11
Hình 1.2. Truy xuất dữ liệu tại một trạm cơ sở.....	13
Hình 1.3. Ví dụ topology mạng sử dụng DCS để lưu trữ thông tin về các loài vật..	15
Hình 2.1 : Sơ đồ mạng sensor	Error! Bookmark not defined.
Hình 3.1. Mô hình mạng sử dụng phương thức ES ..	Error! Bookmark not defined.
Hình 3.2. Mô hình mạng sử dụng phương thức LS ..	Error! Bookmark not defined.
Hình 3.3 : Mô hình mạng sử dụng DCS	Error! Bookmark not defined.
Hình 3.4 : Ví dụ về DCS lưu trữ các dữ liệu về các loài hoa trong vườn hoa...	Error! Bookmark not defined.
Hình 3.5 : Ví dụ về chuyển tham lam : x chuyển tới y , nút kề với nó gần D nhất	Error! Bookmark not defined.
Hình 3.6 : Ví dụ không thể thực hiện chuyển tham lam: x không có nút kề nó gần D hơn	Error! Bookmark not defined.
Hình 3.7 : Ví dụ về quy tắc bàn tay phải : gói tin chuyển theo chiều kim đồng hồ quanh một miền đóng	Error! Bookmark not defined.
Hình 3.8 : Khóa được lưu ở vị trí L, nút nhà a, các nút mô hình d và e trên đường bao nhà.....	Error! Bookmark not defined.
Hình 3.9 : T_1 sau khi nút a lỗi, nút mô hình d khởi tạo một gói làm mới cho L	Error! Bookmark not defined.
Hình 3.10 : Nút f trở thành nút nhà mới và tuyển các nút mô hình b, c, d và e	Error! Bookmark not defined.
Hình 3.11 : Ví dụ về sự tái tạo có cấu trúc với độ phân tích mức 2	Error! Bookmark not defined.
Hình 3.12 : tìm đường hướng miền.....	Error! Bookmark not defined.
Hình 3.13 : Tọa độ ảo phân cấp	Error! Bookmark not defined.
Hình 3.14 : Phân cấp điểm mốc	Error! Bookmark not defined.
Hình 4.1. Cây tìm đường đi ngắn nhất không cân bằng và cây cân bằng mức đỉnh	Error! Bookmark not defined.
Hình 4.2. Cây cân bằng tải trên mô hình lưới.....	Error! Bookmark not defined.
Hình 4.3. Cây cân bằng tải mức đỉnh trên một mạng lưới 10x10..	Error! Bookmark not defined.
Hình 4.4. a) Số các lân cận không đánh dấu của mỗi nút	Error! Bookmark not defined.
b) Không gian phát triển của mỗi nút.....	Error! Bookmark not defined.
Hình 4.5. Kết quả so sánh cân bằng tải giữa GHT và lưu trữ thông tin dựa trên điểm mốc	Error! Bookmark not defined.
Hình 5.1. Cây phân cấp của lớp TclObject.....	Error! Bookmark not defined.
Hình 5.2. Nút đơn hướng và nút đa hướng	Error! Bookmark not defined.
Hình 5.3. Đường kết nối	Error! Bookmark not defined.
Hình 5.4. Khuôn dạng gói dữ liệu trong NS2	Error! Bookmark not defined.
Hình 5.5. Dữ liệu được gửi (put) từ node 0 đến node 1	Error! Bookmark not defined.
Hình 5.6. Nút 2 yêu cầu get dữ liệu từ nút 1.....	Error! Bookmark not defined.
Hình 5.7. Dữ liệu được nút 1 trả về cho nút 2	Error! Bookmark not defined.

DANH MỤC CÁC TỪ VIẾT TẮT

STT	Từ viết tắt	Từ được viết tắt	Giải nghĩa
1	DCS	Data – Centric Storage	Lưu trữ thông tin hướng dữ liệu
2	RF	Radio Frequency	Tần số vô tuyến
3	ES	External Storage	Lưu trữ ngoài
4	LS	Local Storage	Lưu trữ cục bộ
5	GPS	Global Positioning System	Hệ thống định vị toàn cầu
6	PRP	Planar Face Traversal	Thuật toán duyệt miền phẳng
7	LBLSP	Load Balanced Local Shortest Path	Đường đi ngắn nhất cục bộ cân bằng tải
8	WSN	Wireless Sensor Network	Mạng cảm biến không dây

LỜI MỞ ĐẦU

Công nghệ không dây ra đời từ thế kỷ 19, bắt đầu từ sự phát triển điện tín không dây của Marconi. Từ đó công nghệ không dây ngày càng phát triển và đóng vai trò quan trọng trong kỹ thuật cũng như trong đời sống xã hội. Ngày nay có nhiều các kỹ thuật/công nghệ không dây (wireless) khác nhau với các thuộc tính, ứng dụng và khả năng hỗ trợ khác nhau.

Mạng sensor là một mạng không dây rộng lớn, gồm các nút cảm biến liên kết với nhau bằng sóng vô tuyến. Nó ngày càng được ứng dụng nhiều trong quân sự, công nghiệp và cuộc sống hằng ngày. Các hạn chế cơ bản của mạng sensor là bộ nhớ và năng lượng tại các nút sensor. Điều này đã làm ảnh hưởng lớn tới vấn đề truy xuất dữ liệu trong mạng sensor.

Đối với mạng sensor, việc truy xuất dữ liệu sẽ được thực thi tại các nút sensor. Nếu sử dụng phương pháp truy xuất dữ liệu tập trung tại một số nút cơ sở hay tại trạm cơ sở sẽ dễ dẫn đến việc quá tải, tắc nghẽn mạng. Điều này sẽ làm tiêu hao nhiều năng lượng và bộ nhớ của các nút sensor. Điều này đặt ra yêu cầu hệ thống mạng sensor với dữ liệu phân tán tại các nút. Chính vì thế, các giải pháp truy xuất dữ liệu phân tán đã được đưa ra trong đó lưu trữ thông tin hướng dữ liệu (data-centric storage) là một hệ thống điển hình cho phương pháp này trong mạng sensor.

Đề án “**lưu trữ thông tin hướng dữ liệu (data-centric storage DCS)**” tập trung tìm hiểu và nghiên cứu các vấn đề lý thuyết liên quan đến lưu trữ thông tin hướng dữ liệu sử dụng hai phương thức cơ bản là **put** và **get** theo từ khóa; kết hợp với đó là các thuật toán tìm đường cơ bản trong DCS. Ngoài ra, đề án còn đi vào tìm hiểu vấn đề cân bằng tải trong mạng sensor nói chung và trong DCS nói riêng. Cuối cùng, sẽ đi vào xây dựng một mô phỏng hệ thống DCS đơn giản trong NS-2, một công cụ mô phỏng mạng hiệu quả cao, qua đó sẽ phát triển hệ thống DCS thực tế với mạng sensor trong tương lai.

Trong chương đầu tiên của đề án, người viết đề án (NVĐA) xin được trình bày một cách tổng quan về các vấn đề trong mạng sensor, đặc biệt là vấn đề truy xuất dữ liệu và mục tiêu của đề án. Phần cuối chương, NVĐA sẽ trình bày các nội dung cụ thể trong từng chương của đề án.

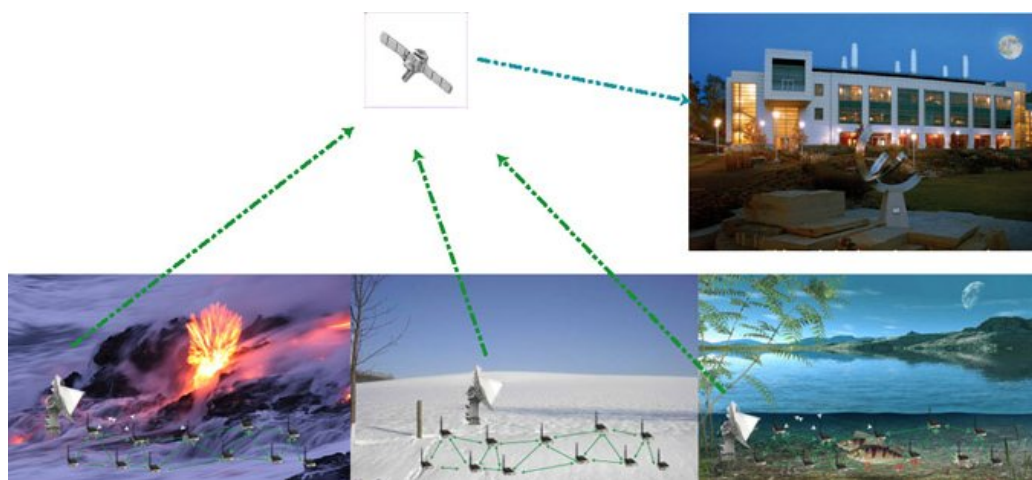
CHƯƠNG 1 : ĐẶT VẤN ĐỀ

Mạng sensor – các vấn đề và ứng dụng

Mạng không dây đang ngày càng đóng vai trò quan trọng trong xã hội với các ứng dụng trong các lĩnh vực di động, cảm biến ... Công nghệ không dây ra đời từ thế kỷ 19, bắt đầu từ sự phát triển điện tín không dây của Marconi. Từ đó đến nay, nó ngày càng trở nên phổ biến do độ tin cậy và tốc độ được nâng cao trong khi giá thành giảm nhiều đối với mọi thành phần người sử dụng. Ngày nay có nhiều các kỹ thuật/công nghệ không dây (wireless) khác nhau với các thuộc tính, ứng dụng và khả năng hỗ trợ khác nhau.

Mạng không dây đang thực sự trở thành công nghệ mà mọi người dùng đều nghĩ tới khi thiết lập một mạng máy tính mới hay nâng cấp hệ thống mạng máy tính cũ hoặc chỉ đơn giản là muốn kết nối Internet tốc độ cao mà không cần dây dẫn. Đặc biệt, với sự phát triển mạnh mẽ của thị trường điện thoại di động thì mạng không dây càng cho thấy tầm ảnh hưởng lớn lao của nó đối với xã hội. Nhu cầu sử dụng thiết bị di động và các ứng dụng trên điện thoại di động của con người ngày càng cao chính là động lực để các nhà nghiên cứu phát triển các hệ thống mạng không dây trong tương lai.

Mạng sensor là một mạng không dây phân tán rộng lớn, bao gồm các nút cảm biến được kết nối với nhau bằng sóng vô tuyến (RF connection). Trong đó, các nút mạng thường là các thiết bị đơn giản, nhỏ gọn, giá thành thấp ... và có số lượng lớn, được phân bố ngẫu nhiên (non-topology) trên một diện tích rộng (phạm vi hoạt động rộng). Các node cảm biến sử dụng nguồn năng lượng hạn chế (pin), nhưng cần thời gian hoạt động lâu dài (vài tháng đến vài năm) và có thể hoạt động trong môi trường khắc nghiệt (chất độc, ô nhiễm, nhiệt độ ...).



Hình 1.1. Một ví dụ về mô hình mạng sensor

Ứng dụng của mạng sensor : Mạng được ứng dụng nhiều trong các lĩnh vực quân sự, công nghiệp và cuộc sống hằng ngày. Ứng dụng đầu tiên của mạng sensor là trong lĩnh vực quân sự như việc phát hiện mìn, chất độc, hay sự di chuyển của quân địch... Cùng với sự phát triển mạnh mẽ của ngành công nghiệp và nhu cầu ngày càng cao của con người, mạng sensor cũng được ứng dụng nhiều trong công nghiệp điều khiển tự động, robotic, các thiết bị thông minh, y tế, môi trường... Qua đây, có thể thấy, mạng sensor tuy là một hệ thống mới trong lĩnh vực Telecom nhưng các ứng dụng của nó thì ngày càng lớn và quan trọng trong mọi lĩnh vực khoa học, kỹ thuật và đời sống hằng ngày.

Mạng sensor đang ngày càng phát triển và được mở rộng để đáp ứng được các nhu cầu về ứng dụng của con người. Nhiều công nghệ, phần cứng, các giao thức, phương thức, chuẩn lần lượt ra đời và đang tiếp tục được nghiên cứu và phát triển rộng rãi. Có thể kể ra một số phương thức chuẩn của mạng sensor đã được tìm ra như External Storage (ES), Local Storage (LS), Data-centric Storage (DCS), trong đó, mỗi phương thức có những ưu nhược điểm khác nhau, bổ sung cho nhau trong việc hỗ trợ phát triển các ứng dụng mạng sensor.

Các ưu điểm : Mạng sensor có tính linh hoạt cao, không bị ràng buộc cố định về phân bố địa lý. Trong mạng, ta có thể dễ dàng bổ sung hoặc thay thế các thiết bị tham gia vào mạng (như thêm nút, thay một nút bị lỗi hoặc nạp năng lượng cho nút) mà không phải cấu hình lại toàn bộ topology của mạng.

Các nhược điểm cần khắc phục : Mạng sensor vẫn còn tồn tại những hạn chế cơ bản đó là : Tốc độ đường truyền chưa cao; khả năng bị nhiễu và mất thông tin trên các vùng sensor có địa hình xấu là lớn; bộ nhớ lưu trữ tại các nút cảm biến bị hạn chế; và đặc biệt là hạn chế về năng lượng sử dụng trong mạng. Để khắc phục được hạn chế về năng lượng, các nhà nghiên cứu đã đưa ra các phương pháp lưu trữ và phân phối dữ liệu trong mạng sensor làm tăng hiệu quả truy xuất dữ liệu và đạt được cân bằng tải tốt hơn. Đây cũng là chủ đề chính xuyên suốt luận văn này.

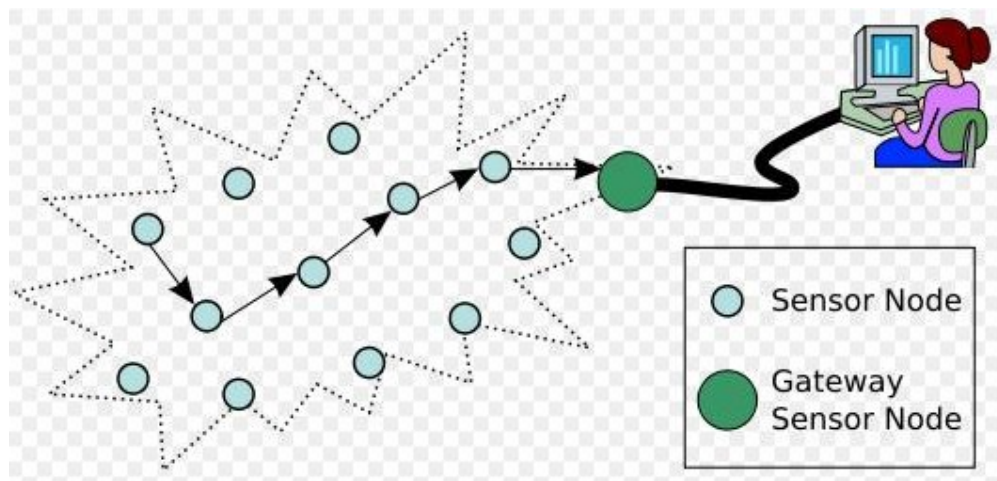
Mục tiêu chính của luận văn là tìm hiểu, nghiên cứu phương pháp lưu trữ và phân phối dữ liệu trong mạng sensor. Trong đó, NVĐA xin tập trung vào tìm hiểu một **hệ thống lưu trữ thông tin hướng dữ liệu (data-centric storage DCS)**. Ngoài ra, bài luận văn sẽ đi sâu vào tìm hiểu một số thuật toán tìm đường được áp dụng cho các hệ thống lưu trữ và phân phối dữ liệu mà điển hình là DCS và vấn đề cân bằng tải trong mạng sensor. Cuối cùng là giới thiệu một công cụ mô phỏng mạng sensor nói riêng và mạng thông tin nói chung NS-2, và một chương trình mô phỏng bước đầu trong hệ thống lưu trữ thông tin hướng dữ liệu DCS.

Các vấn đề truy xuất dữ liệu trong mạng sensor

Trong bất kỳ hệ thống mạng nào, vấn đề truy xuất dữ liệu luôn được đặt lên hàng đầu. Trong mạng sensor, nó càng đóng vai trò quan trọng. Ngày nay, khi các

mạng sensor mở rộng kích thước, sẽ có một số lượng lớn dữ liệu được tạo sẵn. Ngoài ra, nhu cầu về thông tin của con người ngày càng phong phú, nên dữ liệu thêm vào trong quá trình truy xuất cũng càng tăng lên. Khối lượng dữ liệu lớn, cộng thêm việc hạn chế về bộ nhớ và năng lượng trên các nút sensor chính là một vấn đề khó khăn đang thách thức các nhà nghiên cứu. Với hệ thống mạng sensor sử dụng phương pháp lưu trữ và phân phối dữ liệu, thì việc truy xuất dữ liệu tại các nút đã khắc phục được rất nhiều những hạn chế của mạng. Cụ thể, trong luận văn này chúng ta sẽ tập trung tìm hiểu về DCS và việc truy xuất dữ liệu trong DCS.

Ngoài DCS, có một phương pháp truy xuất dữ liệu đã được sử dụng là truy xuất tại các *trạm cơ sở (base station)* : Với phương pháp này, trong mạng sensor sẽ xây dựng một số trạm cơ sở (ví dụ với mạng 500 nút thì sẽ có 20 nút là trạm cơ sở). Ở đây, các trạm cơ sở có thể không phải là các nút sensor mà là một server với đầy đủ tài nguyên cần thiết cho việc truy xuất dữ liệu. Các base station này sẽ làm nhiệm vụ như một server lưu trữ dữ liệu cho một miền trong mạng. Khi đó, mọi truy vấn dữ liệu từ các nút sensor trong cùng một miền đều được gửi tới trạm cơ sở của miền đó thông qua các lân cận của nó.



Hình 1.2. Truy xuất dữ liệu tại một trạm cơ sở

Với phương pháp này, gánh nặng về bộ nhớ lưu trữ tại các nút sensor sẽ được giảm khi mà các base station có đầy đủ các tài nguyên cần thiết để lưu trữ một lượng dữ liệu rất lớn. Tuy nhiên, nó vẫn còn nhiều điểm hạn chế cần khắc phục như chi phí, xây dựng ..., trong đó, nhược điểm chính là dễ gây ra quá tải, tắc nghẽn xung quanh các base station (tại các lân cận của nó) ...

Thứ nhất, cần xây dựng các trạm cơ sở trên các miền trong mạng với đầy đủ tài nguyên như bộ nhớ, năng lượng, khả năng truyền dữ liệu... đều phải đạt yêu cầu với một lượng lớn dữ liệu truy xuất. Tất nhiên, các trạm này cũng phải có khả năng

cảm ứng để liên kết với các nút sensor lân cận trong miền. Các trạm cơ sở này có thể là một hệ thống máy tính hoặc cũng có thể chỉ là một kho lưu trữ dữ liệu lớn có gắn các thiết bị cảm ứng để giao tiếp... Như thế, chi phí cho một trạm cơ sở là rất lớn và việc xây dựng một base station trong mỗi miền cũng khá phức tạp.

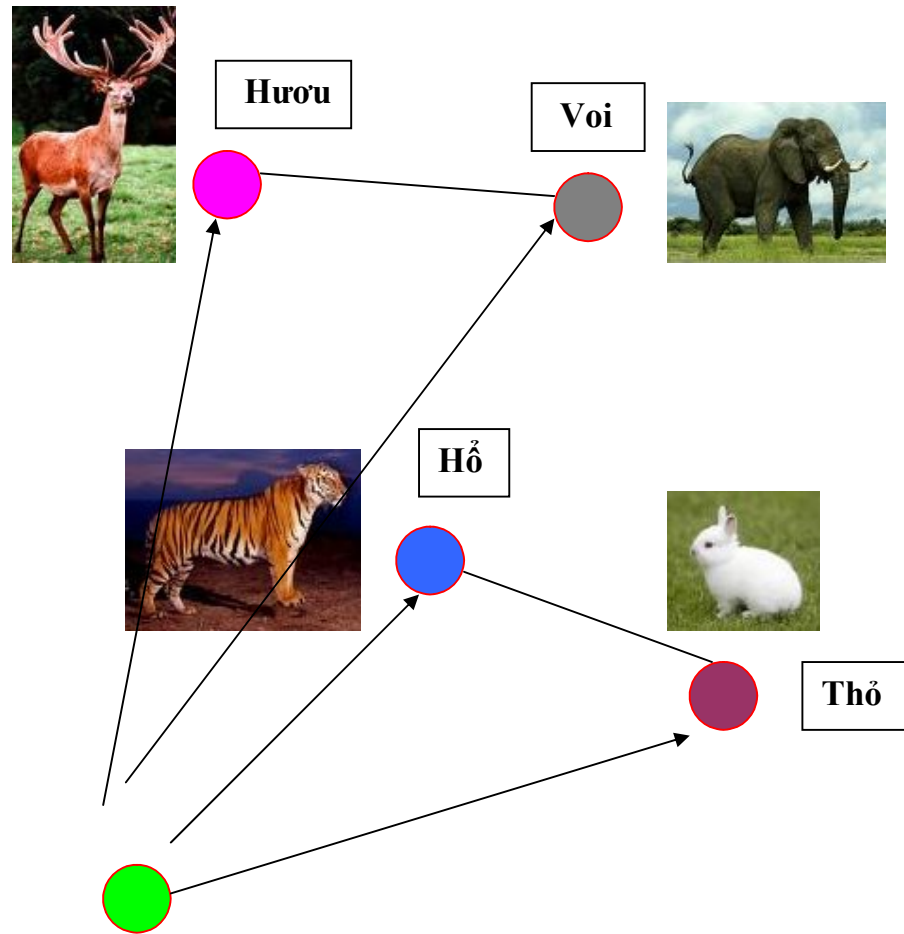
Thứ hai, và cũng là nhược điểm cơ bản của phương pháp truy vấn dữ liệu này là dễ gây ra quá tải và tắc nghẽn mạng tại các lân cận xung quanh trạm cơ sở. Với phương pháp truy vấn này, thì tại mỗi miền, việc truy xuất dữ liệu đều gửi yêu cầu truy vấn về trạm cơ sở của nó. Muốn đến được các trạm cơ sở, các truy vấn này đều phải qua các nút lân cận quanh nó. Trong một hệ thống mạng sensor lớn, thì các miền được chia cũng sẽ lớn với nhiều nút mạng. Điều này dẫn đến nhiều truy vấn sẽ được gửi tới các lân cận của trạm cơ sở. Do đó, việc quá tải và tắc nghẽn tại các lân cận này là điều không thể tránh khỏi.

Với các nhược điểm của phương pháp truy vấn trạm cơ sở, các nhà nghiên cứu đã đưa ra một phương pháp mới, hạn chế tối đa các nhược điểm trên, đó là phương pháp phân phối dữ liệu trên các nút sensor. Trong đó, điển hình là hệ thống lưu trữ thông tin hướng dữ liệu (DCS). Trong phương pháp này, yêu cầu quan trọng nhất là kỹ thuật phân tán dữ liệu trên các nút sensor. Trong mạng, dữ liệu được chia thành nhiều loại, mỗi loại dữ liệu được lưu trữ tại một nút mạng sensor khác nhau. Đây là phương pháp truy xuất được sử dụng khá phổ biến trong mạng sensor bởi tính hiệu quả và khả năng cân bằng tải rất tốt của nó. Để sử dụng kiểu truy xuất này, đã có nhiều giải pháp phân tán dữ liệu được đưa ra nhưng nhìn chung các giải pháp này đều mang tính hướng dữ liệu hơn. Có hai lớp các phương thức hướng dữ liệu phổ biến được đưa ra là: vận chuyển hướng dữ liệu (data-centric conveyance) và lưu trữ thông tin hướng dữ liệu (data-centric storage). Phương thức phân phối dữ liệu phụ thuộc vào bản chất của mạng sensor, môi trường triển khai nó và tải làm việc mong muốn. Trong đó, lớp các phương thức lưu trữ thông tin hướng dữ liệu thường sử dụng kho lưu trữ thông tin nội mạng thông minh giúp cho việc tìm dữ liệu hiệu quả hơn lớp các phương thức vận chuyển dữ liệu.

Lưu trữ thông tin hướng dữ liệu (data – centric storage DCS) là một giải pháp quan trọng và được sử dụng phổ biến trong các mạng sensor. Mục tiêu chính của DCS là lưu trữ các dữ liệu liên quan trên cùng một nút thông qua một từ khóa key. Khi đó, mỗi nút trên mạng sensor là một kho lưu trữ thông tin của cùng một loại dữ liệu.

Hình 1.3 là một ví dụ minh họa cho phương pháp phân phối dữ liệu. Ở đó, dữ liệu về các loài vật được chia riêng. Mỗi nút lưu trữ thông tin về một loài vật riêng biệt. Các truy vấn dữ liệu sẽ được chuyển đến các nút này từ các nút lân cận trung gian như hình vẽ. Đối với hệ thống DCS, cần phải gán một từ khóa key cho từng loại dữ liệu về loài vật, ví dụ như dùng các từ khóa thỏ, hổ tương ứng với các dữ liệu về loài thỏ, loài hổ ... Từ khóa này sẽ được ánh xạ tới nút lưu trữ thông tin về

loài vật đó để thực hiện việc truy xuất dữ liệu. Kỹ thuật ánh xạ chúng ta sẽ nhắc đến trong các phần sau của luận văn.



Hình 1.3. Ví dụ topology mạng sử dụng DCS để lưu trữ thông tin về các loài vật

Khi DCS ra đời, có nhiều thuật toán tìm đường cũng ra đời đáp ứng những đòi hỏi của nó. Bởi với cách lưu trữ dữ liệu trên từng nút theo tên trong DCS thì việc tìm đường tới một nút tương ứng với một tên là một yêu cầu cần thiết để **get** và **put** dữ liệu trên nút đó. Chính bởi nguồn năng lượng hạn chế trong mạng sensor mà cụ thể là năng lượng trên từng nút, mà nó đòi hỏi rất cao ở các thuật toán tìm đường để đạt được hiệu suất cao, cân bằng tải trong cả tìm đường và lưu trữ. Vậy bản chất và việc sử dụng DCS trong mạng sensor như thế nào. Với việc lưu trữ dữ liệu như thế thì tìm đường trong DCS sẽ có những khó khăn và thuận lợi gì, mang lại những ưu điểm và nhược điểm nào cho mạng sensor. Chúng ta sẽ lần lượt đi vào trả lời các vấn đề này trong các chương sau của luận văn này.

Nội dung và các vấn đề trong luận văn

Với các vấn đề truy xuất dữ liệu và những hạn chế về năng lượng trong mạng sensor đã trình bày ở trên thì trong luận văn này, người viết luận văn (NVĐA) xin tập trung vào tìm hiểu về bản chất cũng như các vấn đề trong lưu trữ thông tin hướng dữ liệu (DCS), một giải pháp hiệu quả và quan trọng trong mạng sensor. Nội dung LV được trình bày bao gồm : hai chương đầu đưa ra cái nhìn tổng quan về mạng không dây và mạng sensor, đồng thời chỉ ra các vấn đề hạn chế trong mạng sensor trong việc truy xuất dữ liệu và giải pháp lưu trữ thông tin hướng dữ liệu (DCS); hai chương tiếp theo sẽ đi vào tìm hiểu bản chất của DCS và vấn đề cân bằng tải trong mạng sensor nói chung, trong DCS nói riêng; và cuối cùng sẽ là một công cụ mô phỏng cho mạng sensor NS – 2 với kết quả mô phỏng bước đầu về hệ thống DCS :

Chương 1. Đặt vấn đề : Giới thiệu chung về mạng không dây, mạng sensor, các vấn đề truy xuất dữ liệu trong mạng sensor và giải pháp DCS.

Chương 2. Giới thiệu chung về mạng sensor và lưu trữ thông tin hướng dữ liệu DCS : Gồm 2 phần chính :

- +) Trình bày tổng quan cấu trúc mạng sensor và các ứng dụng của nó
- +) Trình bày tổng quan về DCS và vấn đề tìm đường trong DCS

Chương 3. Lưu trữ thông tin hướng dữ liệu (DCS) : Bao gồm hai vấn đề chính là :

- +) Trình bày bản chất của DCS
- +) Các thuật toán tìm đường cơ bản trong DCS

Chương 4. Cân bằng tải cho DCS trong mạng sensor : Trình bày các vấn đề cân bằng tải trong mạng sensor nói chung và trong DCS nói riêng.

Chương 5. Công cụ mô phỏng mạng sensor NS – 2 : Bao gồm:

- Trình bày tổng quan về hệ thống mô phỏng NS-2 với một số lớp cơ sở và việc giao tiếp giữa hai ngôn ngữ C++ và Tcl script trong NS-2
- Mô phỏng một hệ thống DCS cơ bản làm tiền đề để xây dựng hệ thống DCS trong mạng sens lớn hơn.

Cuối cùng, phần **Kết luận** sẽ tổng kết những kết quả, kiến thức mà NVĐA đã thu nhận được trong quá trình nghiên cứu làm đồ án và nêu ra hướng phát triển tiếp theo của đồ án.