

THUẬT TOÁN KẾT HỢP PHÂN CỤM VÀ LẤY MẪU NÉN LÀM GIẢM NĂNG LƯỢNG TIÊU THỤ TRONG MẠNG CẢM BIẾN KHÔNG DÂY

Nguyễn Tuấn Minh*

Trường Đại học Kỹ thuật Công nghiệp - ĐH Thái Nguyên

TÓM TẮT

Trong nghiên cứu này, công nghệ lấy mẫu nén (Compressive Sensing - CS) được giới thiệu và áp dụng cho mạng cảm biến không dây (Wireless sensor networks - WSNs). Công nghệ CS và mạng cảm biến không dây đều có rất nhiều ứng dụng trong nhiều lĩnh vực khác nhau, dân sự cũng như quân sự. Dựa trên đặc tính của dữ liệu cảm biến có độ tương quan cao, một thuật toán mới về truyền dữ liệu trong mạng WSN được đề xuất nhằm giảm đáng kể năng lượng tiêu thụ cho toàn mạng. Bên cạnh đó, năng lượng tiêu thụ cho truyền dữ liệu trong mạng được lặp công thức, sau đó chạy mô phỏng để xác định tinh chính xác của công thức đã lập. Ngoài ra, các trường hợp hoạt động khác nhau của mạng được xem xét cùng với những đề xuất đưa ra các trường hợp tối ưu năng lượng tiêu thụ tương ứng.

Từ khóa: Mạng cảm biến không dây; công nghệ lấy mẫu nén; phân cụm nhóm; bộ cảm biến; năng lượng tiêu thụ; dữ liệu rỗng.

GIỚI THIỆU CHUNG

Mạng cảm biến không dây (viết tắt là WSN) đang mang lại rất nhiều ứng dụng không chỉ trong kỹ thuật quân sự mà còn trong cả dân sự [1]. Mạng là sự liên kết các bộ cảm biến bằng các đường truyền dữ liệu không dây như WiFi, Bluetooth, RF (radio frequency). Những bộ cảm biến này thường làm việc ở những nơi thường là khó xâm nhập hoặc con người không muốn vào vì các điều kiện khắc nghiệt như ở các môi trường hóa chất, phóng xạ. Do vậy vấn đề về việc sử dụng năng lượng dự trữ ở các bộ cảm biến, hay ở mạng cảm biến không dây luôn là chủ đề rất được quan tâm. Điều này trực tiếp ảnh hưởng tới thời gian hoạt động của toàn mạng trên.

Công nghệ lấy mẫu nén (Compressive sensing - CS) [2] đã được ứng dụng trong mạng cảm biến để làm giảm năng lượng tiêu thụ [3]. Trong các phương pháp này, chỉ một số lượng mẫu cảm biến nhất định được gửi về trung tâm xử lý dữ liệu để khôi phục toàn bộ dữ liệu thu được từ mạng cảm biến.

Các thuật toán phân cụm nhóm trong mạng cảm biến không dây (WSN) được coi là những thuật toán tiết kiệm năng lượng trong việc cân bằng năng lượng tiêu thụ trong mạng

cảm biến [4]. Mỗi một cụm có một trưởng nhóm (CH), đồng thời cũng là một bộ cảm biến trong nhóm. Một số các toán diễn hình như K-mean [5], LEACH [4], HEED [6], EEUC [7] được biết đến với vai trò làm giảm năng lượng tiêu thụ trong toàn mạng có phân cụm.

Các mạng cảm biến kết hợp các phương pháp thu nhận dữ liệu khác nhau ứng dụng công nghệ nén cảm biến. [10] đã đưa ra phương thức kết hợp (lai) giữa các cấu trúc mạng truyền thống để làm giảm dữ liệu truyền từ mỗi bộ cảm biến. Trong [11], một mạng WSN được chia nhỏ ra thành các cụm nhóm. Các dữ liệu cảm biến được gửi tới những nhóm trưởng và những bộ cảm biến này sẽ lại chuyển tiếp dữ liệu tới trạm gốc BS.

Trong nghiên cứu này, một thuật toán mới và một số kết quả mới được công bố. Khác với những kết quả nghiên cứu trước, tác giả đã đề xuất để những nhóm trưởng CH sẽ tạo ra các ma trận con rồi nhân với dữ liệu trong nhóm để tạo ra các mẫu cảm biến. Toàn bộ mẫu này sẽ được chuyển tới trạm gốc BS. Theo đó, thuật toán mới của nghiên cứu này (CCS) đã giảm đáng kể số lần truyền dữ liệu không chỉ ở mỗi bộ cảm biến mà còn ở các nhóm trưởng. Những kết quả này không chỉ cho thấy tiềm năng của việc ứng dụng công nghệ

CS vào WSNs mà còn cho thấy khả năng áp dụng cho các mạng cảm biến lưu động, mạng robot, mạng phương tiện với mục đích thu thập thông tin. Những đóng góp chính của nghiên cứu này được liệt kê như sau:

1. Thuật toán truyền dữ liệu trong mạng cảm biến không dây được thiết lập.
2. Công thức tính năng lượng tiêu thụ trong mạng được trình bày và mô phỏng.
3. Số lượng nhóm cần thiết cho việc tối ưu năng lượng tiêu thụ được đề xuất.

Phần còn lại của bài báo sẽ được trình bày như sau: Công nghệ nén cảm biến và thuật toán CCS sẽ được trình bày ở phần II và III. Phần V và VI sẽ phân tích năng lượng tiêu thụ và đưa ra các kết quả mô phỏng. Phần cuối cùng (VII) sẽ là kết luận và những gợi ý cho những nghiên cứu trong tương lai.

CÔNG NGHỆ LẤY MẪU NÉN

Xây dựng hệ thống mạng cảm biến

Giả sử có trong một mạng cảm biến bao gồm N bộ cảm biến được bố trí một cách ngẫu nhiên trong vùng cần được thu thập dữ liệu. N_c trong N bộ cảm biến được chọn ngẫu nhiên để trở thành những nhóm trường với xác suất được chọn là N_c/N theo LEACH [4]. Mỗi nhóm sẽ có số lượng cảm biến trung bình là $(N/N_c - 1)$. Mỗi nhóm trường được giả định là có đủ khả năng lưu trữ và tạo ra các mẫu cảm biến dựa trên những dữ liệu thu được từ các bộ cảm biến còn lại không phải nhóm trường.

Công nghệ nén cảm biến

Công nghệ lấy mẫu nén (CS – Compressive sensing) cho phép khôi phục toàn bộ dữ liệu dựa trên một số lượng mẫu nhỏ hơn rất nhiều so với các phương pháp nén và lấy mẫu thông thường như Shannon /Nyquist. Điều kiện tiên quyết để sử dụng công nghệ này là tín hiệu phải "thưa - rỗng" trong miền thích hợp

Biểu diễn của tín hiệu:

Một tín hiệu, ví dụ $X = [x_1 x_2 \dots x_N]^T \in \mathbb{R}^N$, được định nghĩa là rỗng mức k nếu nó có biểu

diễn tín hiệu ở một miền nào đó thích hợp, ví dụ $\psi = [\psi_i, j] \in \mathbb{R}^{NxN}$ và $X = \psi\theta$ và θ có k thành phần khác 0 và các thành phần nhỏ còn lại có thể coi như bằng không.

Lấy mẫu tín hiệu và ma trận lấy mẫu

Các mẫu cảm biến được tạo ra dựa trên công thức $Y = \Phi X$, where $\Phi = [\phi_{i,j}] \in \mathbb{R}^{MxN}$ bao gồm các thành phần là các hê số Gaussian được tạo ra một cách ngẫu nhiên. Vector các mẫu cảm biến còn có thể được viết như sau: $Y = [y_1 y_2 \dots y_M]^T \in \mathbb{R}^M$.

Khôi phục tín hiệu

Với số lượng mẫu cảm biến nhất định $M = O(k \log N/k)$ có thể khôi phục được toàn bộ dữ liệu cảm biến như đã được đề cập ở [2].

$$\hat{\Theta} = \arg \min ||\Theta||_1, \text{ s.t. } Y = \Phi \Psi \Theta \quad (1)$$

Trên thực tế, những mẫu cảm biến khi thu thập được sẽ thường gắn với nhiễu như sau: $Y = \Phi X + e$, trong đó $\|e\|_2 = \epsilon$. Và dữ liệu sẽ được khôi phục theo thuật toán sau:

$$\hat{\Theta} = \arg \min ||\Theta||_1, \text{ s.t. } \|Y - \Phi \Psi \Theta\|_2 < \epsilon \quad (2)$$

THUẬT TOÁN CCS

Giả sử rằng trạm gốc BS cần M mẫu cảm biến để có thể khôi phục N giá trị cảm biến từ N bộ cảm biến đã biết. Mỗi một cụm/nhóm sẽ phải tao ra trung bình là M/N_c mẫu cảm biến. Những mẫu này sẽ được chuyển tiếp tới trạm gốc BS để khôi phục dữ liệu.

Trạm gốc BS nhận được hạt tạo hệ số ngẫu nhiên sẽ tạo ra ma trận khôi phục chéo chính (BDM). Theo [13], số lượng mẫu cảm biến cần thu thập để thỏa mãn điều kiện khôi phục dữ liệu với công nghệ CS là:

$$M = O(k\mu^2 \log^2(k) \log^4 N) \quad (4)$$

Trong đó, $\mu = \min\{\sqrt{N_c}, \mu\}$ và $1 \leq \mu \leq \sqrt{N}$. Sau đó, dựa trên tổng số mẫu cảm biến nhận được (tổng cộng là M mẫu) sẽ khôi phục toàn bộ dữ liệu từ N điểm trong mạng dựa trên công nghệ CS.

Thuật toán CCS bao gồm 3 pha như sau:

Pha thứ nhất – Phân cụm nhóm trong mạng

Trong pha này, N_c trong số N bộ cảm biến được chọn ngẫu nhiên để trở thành nhóm

trường, điều này sẽ giúp việc tiêu thụ năng lượng được cân bằng. Xác suất lựa chọn này là $p = N_c/N$.

Mỗi một bộ cảm biến sẽ tự tìm nhóm trường gần nhất để tạo nhóm và chỉ thuộc về một nhóm duy nhất.

Pha thứ hai – Tạo ra các mẫu cảm biến

Tất cả các bộ cảm biến sẽ gửi dữ liệu cảm biến của mình tới nhóm trường của nhóm mình. Toàn bộ dữ liệu ở mỗi nhóm sẽ được lưu lại ở bộ cảm biến nhóm trường, bao gồm cả dữ liệu của chính nó.

Mỗi nhóm trường sẽ tạo ra một ma trận Gaussian nhỏ rồi đem nhân với toàn bộ dữ liệu mà nó lưu trữ để tạo ra được một số lượng mẫu cảm biến.

Pha thứ ba – Thu thập mẫu cảm biến và khôi phục dữ liệu

Số lượng mẫu cảm biến tạo ra từ mỗi nhóm sẽ được gửi về trạm gốc. Dựa trên thuật toán khôi phục dữ liệu của CS theo công thức (1) và (2), toàn bộ dữ liệu từ N bộ cảm biến sẽ được khôi phục.

PHÂN TÍCH NĂNG LƯỢNG TIÊU THỤ

Chi phí truyền thông được xem xét ở đây bao gồm chi phí truyền dữ liệu giữa các bộ cảm biến và nhóm trường (được ký hiệu là $P_{intra-cluster}$) và chi phí truyền giữa các nhóm trường và trạm gốc (được ký hiệu là $P_{to BS}$). Tổng năng lượng tiêu thụ (được ký hiệu là P_{total}) được tính như sau:

$$P_{total} = (P_{intra-cluster} + P_{to BS}). \quad (9)$$

Phân tích $P_{intra-cluster}$

Phân bố của mạng cảm biến được giả sử là phân bố đều. Mạng này được phân chia làm N_c cụm/nhóm bằng nhau có số bộ cảm biến đều bằng N/N_c . Mỗi nhóm sẽ có các bộ cảm biến không phải nhóm trường là $(N/N_c - 1)$. Từ đây, ta có năng lượng tiêu thụ trong các nhóm là:

$$P_{intra-cluster} = N_c (N/N_c - 1) E[r^{\alpha}] \quad (10)$$

Trong đó r là biến ngẫu nhiên đại diện cho khoảng cách thực giữa các bộ cảm biến

thường và nhóm trường của chúng α là hệ số mũ suy giảm của tín hiệu truyền được chọn bằng 2 trong nghiên cứu này. Ta có thể dễ dàng tính $E[r^{\alpha}]$ như sau:

$$E[r^{\alpha}] = \iint (x^2 + y^2) \rho(x, y) dx dy \quad (11)$$

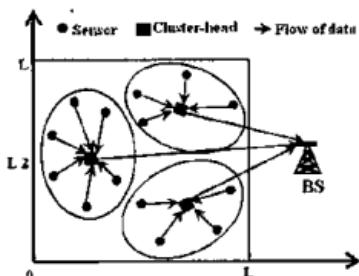
$$E[r^{\alpha}] = \iint r^2 \rho(r', \theta) r' dr' d\theta \quad (12)$$

Trong đó $\rho(r', \theta)$ là phân bố của các bộ cảm biến. Ta có thể giả định mỗi cụm/nhóm là một hình tròn có đường kính là $R = L/\sqrt{N_c}$ và hàm mật độ phân bố của các bộ cảm biến là đều ở các nhóm, do đó, $\rho(r', \theta) = 1/(L^2/N_c)$. Ta có:

$$E[r^2] = \frac{1}{L^2/N_c} \int_{\theta=0}^{2\pi} \int_{r'=0}^R r'^2 dr' d\theta = \frac{L^2}{2\pi N_c} \quad (13)$$

$$\text{Và } P_{intra-cluster} = \left(\frac{N_c}{N} - 1\right) \frac{L^2}{2\pi} \quad (14)$$

Phân tích $P_{to BS}$



Hình 1. Mạng cảm biến được phân nhóm và mẫu cảm biến được truyền trực tiếp tới trạm gốc BS

Để xác định được $P_{to BS}$, ta cần xác định được năng lượng tiêu thụ để truyền M mẫu cảm biến trực tiếp từ các nhóm trường tới trạm gốc BS như Hình 1. Các nhóm trường này cũng nằm trong phân bố ngẫu nhiên đều trong toàn mạng. Tổng năng lượng tiêu thụ trung bình được tính bằng:

$$P_{to BS} = M E[d^2] \quad (15)$$

Trong đó d là một biến ngẫu nhiên thể hiện khoảng cách bất kỳ giữa các nhóm trường và trạm gốc BS. Dựa trên phân bố xác suất đã cho, ta có thể tính $E[d^2]$ dễ dàng như sau:

$$E[d^2] = \int_0^L \int_L^L \left[(x - L)^2 + \left(y - \frac{L}{2}\right)^2 \right] f(x, y) dx dy \quad (16)$$

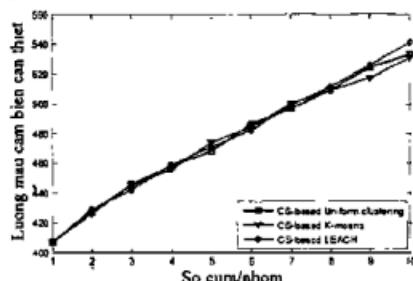
$$E[d^2] = \frac{1}{L} \left[\frac{(L-L)^2}{3} + \frac{L^2}{3} \right] + \frac{L^2}{12} \quad (17)$$

Trong đó $f(x, y) = \frac{1}{L^2}$ là phân bố đều của các nhóm trường cảm biến. Từ các công thức (15) và (17), ta có thể kết luận rằng P_{BS-BS} là độc lập với số lượng nhóm trong mạng. Sử dụng các công thức (9), (14), (15) và (17), tổng năng lượng tiêu thụ được tính như sau:

$$P_{total} = \left(\frac{N_c}{N} - 1\right) \frac{L^2}{2\pi} + \frac{N}{L} \left[\frac{(L-L_0)^2}{3} + \frac{L_0^2}{3} \right] + \frac{ML^2}{12} \quad (18)$$

KẾT QUẢ MÔ PHỎNG

Trong phần này, tín hiệu thua ngẫu nhiên được đưa vào mô phỏng và xem xét. Mang cảm biến bao gồm 2000 bộ cảm biến liên hệ không dây với nhau theo khoảng cách thực và được bố trí trên vùng cảm biến có kích cỡ 100×100 đơn vị độ dài. Hai thuật toán phân cụm phổ biến K-means và LEACH được sử dụng để so sánh với thuật toán CCS trong nghiên cứu này.



Hình 2. Số lượng mẫu cảm biến tăng khi số cụm tăng

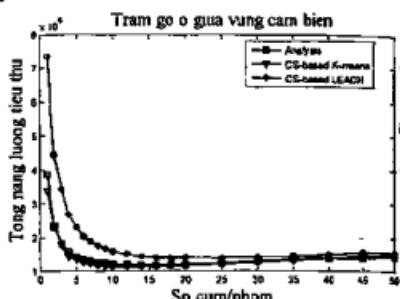
Hình 2 cho thấy, số lượng mẫu cảm biến yêu cầu phải tăng khi số lượng cụm/nhóm tăng lên để đáp ứng được độ chính xác của dữ liệu khôi phục ($Error-target = 0.1$). Điều này có thể giải thích như sau. Khi số cụm tăng, ma trận lấy mẫu trở nên thưa hơn dẫn tới số lượng lấy được không mang đầy đủ nội dung cần lấy mẫu. Do đó, số lượng mẫu yêu cầu phải tăng để đáp ứng cho việc khôi phục tin hiệu.

Tiếp sau là tổng năng lượng tiêu thụ của toàn mạng với những vị trí khác nhau của trạm gốc.

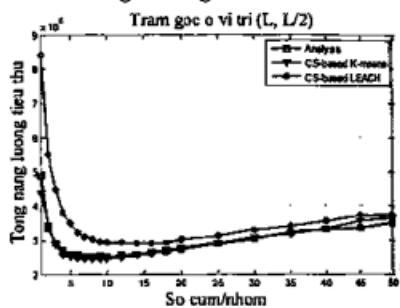
Hình 3 cho chúng ta thấy khi trạm gốc BS ở giữa vùng cảm biến, giá trị tối ưu cho số cụm/nhóm để có được tổng năng lượng tiêu thụ nhỏ nhất là $N_c^* = 14$.

Với vị trí của trạm gốc tại $1L$ ($L_i = L$) như Hình 4 thì giá trị tối ưu cho số cụm/nhóm để

có được tổng năng lượng tiêu thụ nhỏ nhất là $N_c^* = 9$.



Hình 3. Tổng năng lượng tiêu thụ khi trạm gốc ở giữa vùng cảm biến



Hình 4. Tổng năng lượng tiêu thụ của toàn mạng khi trạm gốc ở vị trí L

KẾT LUẬN VÀ HƯỚNG NGHIÊN CỨU TRONG TƯƠNG LAI

Công nghệ lấy mẫu nén áp dụng trong mạng cảm biến không dây có phân cụm đã chứng minh được hiệu quả. Thuật toán CCS không những làm giảm số lượng dữ liệu truyền trong mạng mà còn giúp tối ưu số lượng cụm trong mạng.

Trong những nghiên cứu tới, tác giả sẽ làm tăng hiệu quả việc áp dụng công nghệ nén cảm biến không chỉ trong mạng cảm biến không dây với các bộ cảm biến tĩnh mà còn sẽ triển khai với các bộ cảm biến động. Những bộ cảm biến được gắn lên phương tiện di chuyển được hoặc robot để tăng hiệu quả hoạt động của mạng không dây.

TÀI LIỆU THAM KHẢO

- I. Akyildiz, W. Su, Y. S. E. Cayirci (2002). "Wireless sensor networks a survey", *Comput Netw*, 38 (4), pp. 393–422.

- 2 D. L. Donoho (2006), "Compressed sensing", *Inf. Theory, IEEE Trans.*, 52, pp 1289–1306
3. M.T.Nguyen, N.Rahnavard (2013), "Cluster-based energy-efficient data collection in Wireless sensor networks utilizing compressive sensing", *Military Communications Conference, IEEE/MILCOM2013*, pp 1708–1713
4. M Handy, M. Haase, D Timmermann (2002), "Low energy adaptive clustering hierarchy with deterministic cluster-head selection", *Mobile and Wireless Communications Network, 2002. 4th International Work shop*, pp 368–372.
- 5 S P.Lloyd (1982), Least squares quantization in pcm, *IEEE Trans Inf. Theory*, 28, pp 29–137.
6. O.Younis, S Fahmy (2004), "Distributed clustering in ad-hoc sensor networks: ahybrid, energy-efficient approach", *INFOCOM*, volume1, p. 4
- 7 C Li, M. Ye, G Chen, J Wu (2005), "An energy-efficient unequal clustering mechanism for wireless sensor networks, in: Mobile Ad hoc and Sensor Systems Conference", *IEEE International Conferenceon*, 2005, pp 8pp –604
- 8 L. Xiang, J Luo, C Rosenberg (2013), "Compressed data aggregation: Energy-efficient and high-fidelity data collection", *Netw IEEE/ACM Trans.* 21(6), pp 1722–1735
9. R Xie, X Jia, "Transmission-efficient clustering method for wireless sensor networks using compressive sensing", *Parallel Distrib.Syst. IEEE Trans.* 25 (3) (2014) 806
10. T. S Rappaport, *Wireless Communications. Principles and Practice (2nd Edition)*, second, Prentice Hall, Upper Saddle River, New Jersey, USA, 2002.
11. T. Do, L Gan, N Nguyen, T Tran (2012), "Fast and efficient compressive sensing using Structurally random matrices", *Sig Process. IEEE Trans.* 60(1).
- 12 M. T Nguyen, K A. Teague and N Rahnavard, "CCS: Energy-Efficient Data Collection for Clustered Wireless Sensor Networks utilizing Block-wise Compressive Sensing", *ScienceDirect ELSEVIER (COMNET)*, Vol 106, 4 September 2016, pp 171–185

ABSTRACT

ENERGY-EFFICIENT DATA COLLECTION METHOD IN CLUSTERED WIRELESS SENSOR NETWORKS

Nguyễn Tuan Minh*

University of Technology - TNU

In this paper, Compressive sensing (CS) is introduced and applied to Wireless sensors networks (WSNs) for energy-efficient data collection methods. CS and WSNs support a lot of applications both military and civilian areas. Based on the fact that sensing data is often highly correlated, new algorithm is proposed to reduce significantly energy consumption for data collection in WSNs. Energy consumption for data transmitting in the networks is analyzed, formulated and simulated. Some optimal cases are suggested for the networks to consume to least energy.

Keywords: wireless sensor networks; compressive sensing; clustering, sensors, energy consumption; sparse signals