

TĂNG CƯỜNG BẢO MẬT TRONG HỆ THỐNG IOTS DỰA TRÊN CÔNG NGHỆ LẤY MẪU NÉN

Nguyễn Tuấn Minh¹, Đoàn Minh Cảnh^{2*}

¹Đại học Kỹ thuật Công nghiệp - ĐH Thái Nguyên

²Trường Cao đẳng nghề số 1-Bộ Quốc Phòng

TÓM TẮT

Hiện nay, công nghệ lấy mẫu nén (Compressed Sensing - CS) 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, các mẫu nén được thu từ các bộ cảm biến sẽ được thu lượm và khôi phục lại dữ liệu gốc một cách an toàn, tiêu thụ ít năng lượng và tiết kiệm chi phí truyền dẫn. Công nghệ vạn vật kết nối Internet (Internet of Things - IoTs) đang được triển khai mạnh mẽ trong nhiều lĩnh vực khác nhau tạo với những ứng dụng nổi bật như: nông nghiệp thông minh, nhà thông minh, thành phố thông minh. Nội dung bài báo sẽ tập trung vào các ứng dụng của IoTs cần được bảo mật thông tin và đề xuất phương án áp dụng CS để tăng cường bảo mật cho các hệ thống trên. Một số kết quả mô phỏng được trình bày để làm sáng tỏ phương pháp của bài báo.

Từ khóa: Mạng lưới vạn vật kết nối internet; công nghệ lấy mẫu nén; dữ liệu cảm biến; bảo mật dữ liệu

GIỚI THIỆU CHUNG

Internet of Things (IoT) đang cung cấp nhiều ứng dụng trong các lĩnh vực khác nhau. Mục tiêu chính là tạo một mạng dựa trên internet để kết nối mọi thứ bao gồm các thiết bị điện tử và nhu cầu của con người [1,2]. Các mạng có thể kết nối để điều khiển hệ thống từ xa. Các hệ thống trong nhà thông minh hoặc các lĩnh vực công nghiệp/quân sự có thể liên lạc với nhau cho các mục đích khác nhau. IoT hỗ trợ các mạng khác để đạt hiệu quả cao hơn.

Mạng cảm biến không dây truyền thông (WSN) thu thập dữ liệu từ khu vực đặt cảm biến được gửi đến trạm gốc (BS). BS có thể ở các vị trí cố định để thu thập dữ liệu cảm biến. Với sự tích hợp giữa các IoT và WSN, dữ liệu có thể được gửi qua internet để được lưu trữ ở mọi nơi cần thiết. BS có thể được thiết lập ở mọi nơi để có thể thu thập dữ liệu [3].

IoT tạo điều kiện thuận lợi cho nhà thông minh với nhiều loại ứng dụng [4, 5]. Tất cả các thiết bị điện tử được kết nối vào internet. Chủ sở hữu có thể kiểm soát từng thiết bị như bật / tắt hoặc thiết lập lịch cho tất cả các thiết bị được kết nối. Ngoài ra, các loại cảm biến khác nhau được lắp đặt bên trong ngôi nhà

cho mục đích giám sát. Điều kiện không khí, hệ thống báo cháy hoạt động dựa trên thông tin được cung cấp bởi các cảm biến.

Công nghệ lấy mẫu nén (Compressive sensing - CS) [6] đã được ứng dụng trong mạng cảm biến để làm giảm năng lượng tiêu thụ [7,8]. 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.

Với sự gia tăng nhanh chóng trong việc sử dụng ứng dụng IoT, một số vấn đề bảo mật và riêng tư được quan sát thấy. Khi gần như mọi thứ sẽ được kết nối với nhau, vấn đề này sẽ trở nên rõ ràng hơn, và tiếp xúc thường xuyên sẽ tiết lộ lỗ hổng bảo mật và điểm yếu. Những hiểm họa an toàn đối với các dịch vụ trong IoT là do nguyên nhân hạn chế về năng lực tính toán, năng lượng và băng thông kết nối. Các loại mối đe dọa khác nhau đến mô hình IoT được mô tả gồm: tấn công từ chối dịch vụ (DoS), loại tấn công này làm cho máy tính hoặc tài nguyên mạng không khả dụng cho người sử dụng như dự kiến [9,10]. Do khả năng bộ nhớ thấp và nguồn lực tính toán hạn chế, phần lớn nguồn tài nguyên của các thiết bị trong IoTs dễ bị tấn công đe dọa; Các cuộc tấn công vật lý, loại tấn công này can thiệp

* Email: caodang1bqp@gmail.com

vào các thành phần phần cứng. Do tính chất không được giám sát và phân phối của IoT, hầu hết các thiết bị thường hoạt động trong môi trường ngoài trời, rất nhạy cảm với các cuộc tấn công vật lý. Những cuộc tấn công này có thể khai thác được những dữ liệu mật, khóa... từ thiết bị.

Trong nghiên cứu này, tác giả đề xuất một phương pháp mới với mục đích chính là nén dữ liệu và tăng cường bảo mật cho dữ liệu dựa trên công nghệ lấy mẫu nén. Các phương pháp này không chỉ tiết kiệm năng lượng truyền khi dữ liệu truyền giảm đáng kể mà còn bảo mật được dữ liệu truyền. Những đóng góp chính của nghiên cứu này được liệt kê như sau:

1. Áp dụng thuật toán truyền dữ liệu từ IoTs được thiết lập dựa trên phép biến đổi Wavelet [12].
 2. Đề xuất thuật toán truyền dữ liệu mới dựa trên công nghệ lấy mẫu nén xử lý dữ liệu từ IoTs.
 3. Cung cấp các kết quả mô phỏng xử lý dữ liệu để làm rõ hiệu quả các thuật toán.
- Phần còn lại của bài báo cáo sẽ được trình bày như sau: công nghệ nén cảm biến sẽ được trình bày ở phần II. Phần III và Phần IV sẽ trình bày về hai thuật toán xử lý dữ liệu từ IoTs cùng các kết quả mô phỏng tương ứng. Phần cuối cùng (V) sẽ là kết luận và những gợi ý cho những nghiên cứu tương lai.

CÔNG NGHỆ LẤY MẪU NÉN

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.

Tín hiệu cảm biến

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 $\psi = [\psi_i, j] \in \mathbb{R}^{N \times N}$ 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 $\Phi = [\phi_{i,j}] \in \mathbb{R}^{M \times N}$ bao gồm các thành phần là các hệ 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 [6].

$$\hat{\Theta} = \argmin \|\Theta\|_1, \text{ và } Y = \Phi \Psi \Theta \quad (1)$$

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:

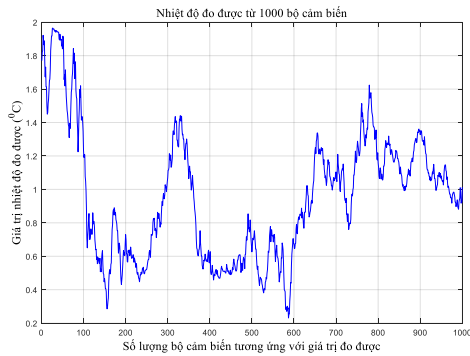
$$\hat{\Theta} = \argmin \|\Theta\|_1, \text{ và } \|Y - \Phi \Psi \Theta\|_2 < \epsilon \quad (2)$$

THUẬT TOÁN XỬ LÝ DỮ LIỆU DỰA TRÊN BIẾN ĐỔI WAVELET

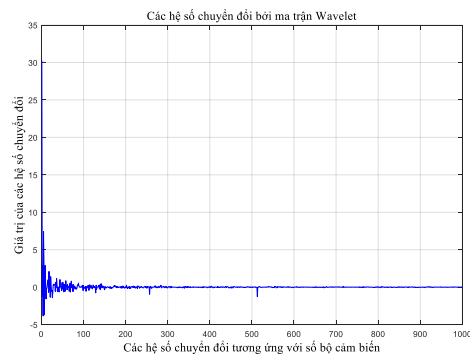
Trong phần này, chúng tôi áp dụng lý thuyết biến đổi Wavelet để xử lý dữ liệu thu được từ các hệ thống IoTs để nén dữ liệu cảm biến và truyền đi một số lượng nhất định mẫu [12]. Hình 1 thể hiện 1000 giá trị cảm biến thu từ 1000 bộ cảm biến trong hệ thống IoTs.

Dữ liệu thu thập được từ IoTs sẽ được nhân với các hệ số Wavelet thay ma trận Wavelet với kích cỡ tùy thuộc vào kích cỡ của dữ liệu để nhận được các hệ số đã biến đổi. Các hệ số này tự phân loại hệ số lớn và bé như trong Hình 2. Thuật toán này sẽ chỉ gửi các hệ số lớn đến trạm gốc (BS) hoặc trung tâm xử lý dữ liệu. Do vậy, theo như Hình 2, chỉ một lượng mẫu rất nhỏ cần phải gửi đi nên sẽ tiết kiệm được nhiều năng lượng truyền.

BS thu thập các hệ số lớn và khôi phục lại tất cả dữ liệu ban đầu. Hệ số nhỏ được coi bằng không (‘0’) ở phía đầu thu. BS sẽ nhân trở lại các hệ số với ma trận Wavelet để thu được toàn bộ dữ liệu ban đầu.

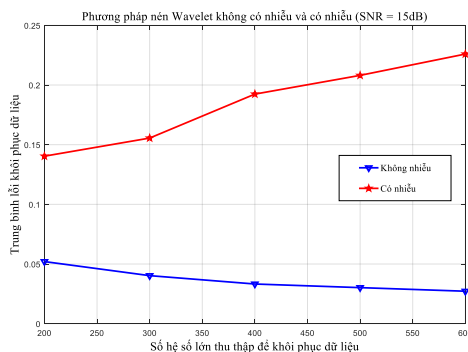


Hình 1. Dữ liệu cảm biến nhiệt độ thu từ 1000 bộ cảm biến trong hệ thống IoTs



Hình 2. Dữ liệu cảm biến sau biến đổi Wavelet sẽ trở thành các hệ số lớn (từ 0 tới 300 như hình vẽ) và còn lại là các hệ số bé có thể coi bằng không ('0')

Đặc tính bảo mật ở đây là phần hệ số lớn được chuyển đi hoàn toàn an toàn cho dữ liệu gốc và giảm chi phí truyền dẫn. Người truy cập khác không thể khôi phục dữ liệu dựa trên các hệ số lớn nếu không biết thông tin về biến đổi Wavelet.



Hình 3. Hệ số lớn tăng và chất lượng khôi phục dữ liệu trong các môi trường có nhiễu và không có nhiễu

Tuy nhiên, phương pháp này không cản trở được nhiều xâm nhập vào dữ liệu, như Hình

3. Do vậy, phương pháp này nên chỉ sử dụng ở môi trường truyền ngắn, ít bị ảnh hưởng bởi nhiễu.

THUẬT TOÁN XỬ LÝ DỮ LIỆU DỰA TRÊN CÔNG NGHỆ LẤY MẪU NÉN (CS)

Thuật toán này với khả năng xử lý được nhiều nhờ công nghệ lấy mẫu nén như đã trình bày ở phần II.

Thuật toán này bao gồm 3 pha như sau:

Pha thứ nhất – Thu dữ liệu từ IoTs

Trong pha này, dữ liệu được tập hợp lại từ hệ thống IoTs chờ xử lý. Dữ liệu có thể là hình ảnh, videos, nhiệt độ, độ ẩm, ... sẽ được phân chia theo từng cụm, từng khoảng thời gian tùy vào ứng dụng để chờ xử lý

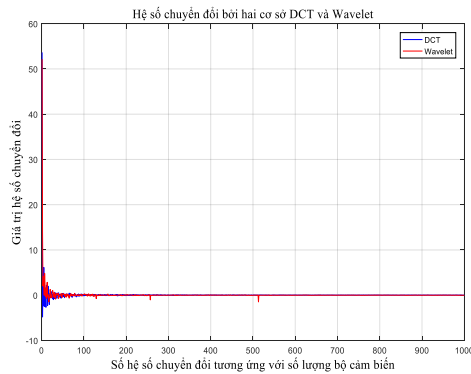
Pha thứ hai – Tạo ra các mẫu cảm biến

Ở pha này, bộ xử lý trong hệ thống IoTs sẽ tạo ra một ma trận ngẫu nhiên Gaussian rồi đem nhân với toàn bộ dữ liệu mà nó lưu trữ để tạo ra được một số lượng mẫu cảm biến nhất định. Kích cỡ của ma trận lấy mẫu là $M \times N$ sẽ quyết định số mẫu cảm biến truyền đi là M . Thông thường chỉ cần M mẫu cảm biến sẽ đủ để khôi phục N giá trị cảm biến, trong đó, $M \ll N$.

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

Số lượng mẫu cảm biến tạo ra 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 bộ cảm biến ở trong các hệ thống IoTs sẽ được khôi phục.

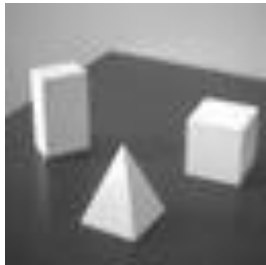
Với thuật toán trên, hai cơ sở Wavelet và DCT được lựa chọn để làm rộng dữ liệu đảm bảo áp dụng được công nghệ nén cảm biến. Hình 4 đưa ra so sánh khả năng phân tích và phân loại dữ liệu của Wavelet và DCT. Hai phương pháp biến đổi về cơ bản là giống nhau, đảm bảo số lượng mẫu lớn và năng lượng của tín hiệu tập trung ở phần đầu như hình vẽ. Số lượng hệ số lớn tập trung khoảng ở 100 giá trị đầu trong số 1000 hệ số được chuyển đổi.



Hình 4. Sử dụng hai cơ sở là Wavelet và DCT để làm rõ dữ liệu trong quá trình khôi phục dữ liệu với công nghệ nén cảm biến

Với số lượng mẫu cảm biến nhất định được gửi về BS để khôi phục toàn bộ dữ liệu. Ở đây dữ liệu từ IoTs được chọn có thể là dữ liệu cảm biến nhiệt độ, độ ẩm đọc từ các bộ cảm biến, dữ liệu ngẫu nhiên, dữ liệu ảnh, v.v.

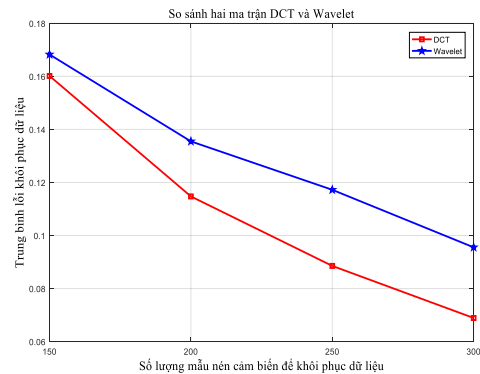
Tất cả dữ liệu có thể nén hay nói cách khác có độ tương quan cao đều có thể áp dụng được công nghệ lấy mẫu nén. Trong phần này, tác giả đã chọn hình ảnh để thực hiện mô phỏng như Hình 5.



Hình 5. Hình ảnh được chọn ra 1000 mẫu để thực hiện mô phỏng nén và khôi phục dữ liệu sử dụng công nghệ lấy mẫu nén

Sau khi thực hiện khôi phục dữ liệu sử dụng công nghệ lấy mẫu nén. Tác giả đã đánh giá để làm rõ chất lượng khôi phục dữ liệu dựa trên số lượng mẫu sử dụng. Chú ý ở đây là, số lượng mẫu càng lớn sẽ cho chất lượng khôi phục dữ liệu càng cao. Điều đó có nghĩa là lỗi khôi phục sẽ nhỏ dần và được thể hiện ở Hình 6. Để nâng cao chất lượng dữ liệu khôi phục ở phía thu, cách thức đơn giản là tăng số mẫu cảm biến. Tuy nhiên, số lượng mẫu cảm biến cần truyền vẫn là rất nhỏ so với tổng số lượng giá trị dữ liệu. Trên hình vẽ, số lượng mẫu

cảm biến tối đa là 300 so với 1000 mẫu dữ liệu cảm biến.



Hình 6. Chất lượng khôi phục ảnh với tổng số dữ liệu ảnh là 1000 giá trị vô hướng số mẫu nén tăng trong khi lỗi khôi phục giảm dần

Hơn nữa, tác giả có so sánh việc hiệu quả sử dụng giữa Wavelet và DCT trong việc khôi phục dữ liệu như Hình 6. Kết quả trong việc khôi phục ảnh với cả DCT tốt hơn Wavelet, tuy việc khôi phục là hiệu quả.

Tính bảo mật được đóng góp ở nghiên cứu này là với số mẫu bảo mật có thể bị mất mát, tuy nhiên, dữ liệu vẫn được khôi phục đầy đủ. Mặt khác, hackers nhận được một số mẫu cảm biến cũng không thể khôi phục được dữ liệu của hệ thống IoTs hiện tại vì không thể biết ma trận lấy mẫu cũng như ma trận là ngẫu nhiên dữ liệu được sử dụng trong các hệ thống IoTs.

KẾT LUẬN VÀ HƯỚNG NGHIÊN CỨU TRONG TƯƠNG LAI

Công nghệ lấy mẫu nén áp dụng trong mạng IoT đã chứng minh được hiệu quả. Với số mẫu chuyển đi rất nhỏ so với toàn bộ khối lượng dữ liệu ban đầu, kết quả khôi phục dữ liệu đáp ứng được yêu cầu sử dụng của hệ thống và người dùng. Hơn nữa, các thuật toán đã đáp ứng được nhu cầu bảo mật trong hệ thống IoTs.

Trong những nghiên cứu tới, tác giả sẽ làm tăng hiệu quả việc áp dụng công nghệ nén cảm biến không chỉ với các đối tượng dữ liệu nêu trên mà còn triển khai đa dạng hơn. Ngoài ra, mã hóa cũng đang được nghiên cứu để áp dụng công nghệ nén cảm biến để tăng cường khả năng bảo mật cho hệ thống IoTs.

TÀI LIỆU THAM KHẢO

1. L. Atzori, A. Iera, and G. Morabito, "The internet of things: A survey", *Comput. Netw.*, vol. 54, no. 15, pp. 2787–2805, Oct. 2010.
2. L. Li, "Study on Security Architecture in the Internet of Things," *International Conference on Measurement, Information and Control (MIC)*, 2012, vol. 1, May 18-20, pp. 374-377
3. 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)*. Volume 106, 4 September 2016, Pages 171–185.
4. D. Bandyopadhyay and J. Sen, "Internet of things: Applications and challenges in technology and standardization", *Wireless Personal Communications*, vol. 58, no. 1, pp. 49–69, 2011.
5. O. Vermesan and P. Friess, "*Internet of Things: Converging Technologies for Smart Environments and Integrated Ecosystems*", River Publishers, 2013.
6. D.L. Donoho, Compressed sensing, *Inf. Theory*, *IEEE Trans.* 52(2006) 1289– 1306.
7. M.T. Nguyen, N. Rahnavard, Cluster-based energy-efficient data collection in Wireless sensor networks utilizing compressive sensing, in: *Military Communications Conference, IEEE/ MILCOM2013*, pp.1708–1713.
8. M. T. Nguyen, H. M. La, and K. A. Teague, "Collaborative and compressed mobile sensing for data collection in distributed robotic networks," *IEEE Transactions on Control of Network Systems*, pp. 1–1, 2017.
9. G. Yang, J. Xu, W. Chen, Z.-H. Qi, and H.-Y. Wang, "Security characteristic and technology in the internet of things," *Nanjing Youdian Daxue Xuebao(Ziran Kexue Ban)/ Journal of Nanjing University of Posts and Telecommunications(Natural Nanjing University of Posts and Telecommunications)*, vol. 30, no. 4, 2010.
10. L. Li, "Study on Security Architecture in the Internet of Things," *International Conference on Measurement, Information and Control (MIC)*, 2012, vol. 1, May 18-20, pp. 374-377.
11. Q. Gou, L. Yan, Y. Liu and Y. Li, "Construction and Strategies in IoT Security System," *IEEE International Conference on Green Computing and Communications and IEEE Internet of Things and IEEE Cyber*, Aug 20-23, 2013, 1129-1132.
12. Minh Nguyen, Huyen Nguyen, Keith Teague, "Wavelet-based Energy Efficient Data Collection Algorithm in Wireless Sensor Networks," *ICSES Transactions on Computer Networks and Communications (ITCNC)*, vol. 4, no. 2, pp. 3-10, Jun. 2018.

ABSTRACT

COMPRESSED SENSING BASED SECURITY IMPROVEMENT
IN THE INTERNET OF THINGS

Nguyen Tuan Minh¹, Doan Minh Canh^{2*}

¹University of Technology - TNU

²The Vocational College No1 – Ministry of National Defense

Recently, Compressed sensing (CS) supports many applications in different fields, in both civil and military applications. Based on the high correlation characteristic of sensing data, CS measurements collected from sensors are forwarded and are reconstructed perfectly and energy-efficiently at a base-station. Internet of Things (IoTs) is being exploited with different promising applications, such as smart agriculture, smart home, smart city, etc. This paper focuses on IoTs applications involved security methods and proposes a new method utilizing CS to improve security issues for such networks. Simulation results are provided to clarify the methods.

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

Ngày nhận bài: 05/6/2018; Ngày hoàn thiện: 14/9/2018; Ngày duyệt đăng: 30/11/2018

* Email: caodang1bqp@gmail.com