Mục lục

Ns2, perl igen 0.15 qca, dinh tuyen tren qca , (gnuplot, dia, Kile)

[**I.** **Giới thiệu** 3](#_Toc371628323)

[**II.** **Sơ lược** 6](#_Toc371628324)

[1. Ký hiệu 6](#_Toc371628325)

[2. Dynamic social network 6](#_Toc371628326)

[3. Hàm mục tiêu 6](#_Toc371628327)

[4. Định nghĩa vấn đề 7](#_Toc371628328)

[**III.** **Mô tả thuật toán và lý thuyết liên quan** 7](#_Toc371628329)

**Các thuật toán thích nghi để phát hiện cấu trúc cộng đồng trong mạng xã hội năng động**

Nam P. Nguyen, Thang N. Dinh, Ying Xuan, My T. Thai

Department of Computer and Information Science and Engineering, University of Florida

Email: {fnanguyen, tdinh, yxuan, mythaig}@cise.uﬂ.edu

***Abstracts(tóm tắt)* – Mạng xã hội có tính chất rất đặc biệt: cấu trúc cộng đồng. Sự hiểu biết về cấu trúc cộng đồng mạng là một lợi thế lớn. Nó không chỉ cung cấp những thông tin hữu ích trong phát triển thêm về chiến lược nhận thức xã hội (social-aware) cho các vấn đề mạng xã hội mà còn hứa hẹn một loạt ứng dụng kích hoạt bằng mạng điện thoại, chẳng hạn như định tuyến trong mạng di động Ad Hoc (MANETs) và luồn sâu vào các chính sách ngăn chặn của mạng di động. Thật tiếc, sự hiểu biết về cấu trúc này là một thách thức lớn, đặc biệt là trong các mạng xã hội năng động nơi mà các hoạt động và tương tác xã hội đang phát triển một cách nhanh chóng. Chúng ta có thể xác định cấu trúc cộng đồng mạng nhanh chóng và hiểu quả không? Chúng ta có thể cập nhật cấu trúc mạng dựa trên những thông tin đã được biết trước thay vì tính toán lại từ đầu không? Trong bài báo này, chúng tôi trình bày về Quick Community Adaptation(QCA) – Sự thích ứng cộng đồng nhanh chóng, một sự thích ứng dựa trên định nghĩa và chỉ ra cấu trúc cộng đồng của mạng xã hội trực tuyến năng động. Cách tiếp cận của chúng tôi không chỉ nhanh chóng mà còn cập nhật hiệu quả cộng đồng mạng, thông qua một loạt các thay đổi, bằng cách chỉ sử dụng các cấu trúc xác định từ các ảnh chụp (snapshots) mạng trước đó, nhưng lại có khả năng đưa ra sự thay đổi của cấu trúc cộng đồng theo thời gian. Để làm rõ tính hiệu quả của thuật toán, chúng tôi kiểm tra tổng quát QCA trên mạng xã hội thực tế bao gồm email, arXiv và Facebook. Cuối cùng chúng tôi chứng minh khả năng áp dụng của thuật toán qua một ứng dụng thực tế trong MANETs. Kết quả so sánh cho thấy rằng: chiến lược định tuyến social-aware sử dụng QCA như một phát hiện cộng đồng mạng tốt hơn phương pháp có sẵn hiện nay.**

# **Giới thiệu**

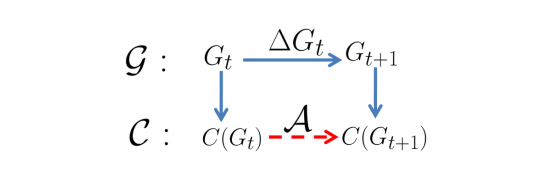
Nhiều mạng xã hội thể hiện đặc điểm của sự bao hàm cấu trúc cộng đồng[1][2], có nghĩa là : chúng tự phân chia thành các nhóm đỉnh với các kết nối dày hơn bên trong mỗi nhóm và ít kết nối hơn giữa các nhóm giao nhau, nơi mà đỉnh và các kết nối đại diện cho mạng người dùng(network users) và tương tác xã hội của họ. Các thành viên trong mỗi cộng đồng mạng xã hội thường chia sẻ những điểm chung như ảnh, phim, âm nhạc hoặc các chủ đề tranh luận và do đó, họ có xu huớng giao tiếp thường xuyên với nhau hơn là giao tiếp với các thành viên bên ngoài cộng đồng của họ. Việc phát hiện các cộng đồng trong mạng lưới là việc tập hợp các đỉnh của mạng lại thành các nhóm bằng cách các nút (nodes) trong mỗi nhóm được kết nối nhiều bên trong và ít bên ngoài.

Cần lưu ý để phân biệt giữa phát hiện cộng đồng mạng và phân nhóm đồ thị. Hai vấn đề này cùng có một mục tiêu là phân vùng các nút mạng (network nodes) thành các nhóm; tuy nhiên, số lượng các nhóm được định nghĩa trước hoặc đưa ra một phần của đầu vào phân nhóm đồ thị trong khi các cộng đồng thường chưa biết trong việc phát hiện các cộng đồng. Phát hiện các cộng đồng trong mạng lưới cung cấp cho chúng ta sự hiểu biết hữu ích về cấu trúc bên trong cũng như các nguyên tắc tổ chức của nó. Hơn nữa, hiểu biết về cấu trúc của cộng đồng mạng có thể cung cấp cho chúng ta những điểm hữu ích hơn về một số phần chưa được khám phá của mạng, do đó giúp ngăn ngừa các mối nguy hiểm tiềm năng như virus hoặc sự lan truyền mầm bệnh. Nghiên cứu dựa trên phát hiện cộng đồng trong mạng tĩnh (static networks) cs thể được tìm thấy trong một cuộc khảo sát [3], cũng như công việc của [4][5][6][7] và sự tham khảo trong đó.

Tuy nhiên, các mạng xã hội ở thế giới thực không phải lúc nào cũng tĩnh (static). Trong thực tế, phần lớn các mạng xã hội (như Facebook, Bebo và Twitter) phát triển, mở rộng về kích thước và không gian như sự gia tăng người dùng của họ, do đó tự giúp đỡ bản thân họ trong lĩnh vực mạng lưới năng động. Một mạng lưới năng động là một loại tiến hóa đặc biệt của mạng lưới phức tạp trong đó thường xuyên thay đổi theo thời gian. Trong khả năng phán đoán của một mạng xã hội trực tuyến, ví dụ như Facebook, Twitter hoặc Flickr, các thay đổi thay đổi thường được giới thiệu bởi những người dùng gia nhập vào hoặc thu hồi từ một hoặc nhiều nhóm hoặc các cộng đồng, qua bạn bè và bạn bè kết nối cùng nhau hoặc qua người mới kết bạn với những người khác. Bất kì những khả năng này có vẻ như ảnh hưởng nhỏ tới cấu trúc địa phương (local structure) của một mạng lưới, sự năng dộng của mạng lưới trong thời gian dài, mặt khác, có thể dẫn tới một sự thay đổi đáng kể trong cấu trúc cộng đồng mạng, vì vậy đặt ra nhu cầu xác định lại. Tuy nhiên, sự thay đổi nhanh chóng và không thể xác định trước của mạng xã hội năng động khiến nó trở thành một vấn để vô cùng phức tạp nhưng đầy thách thức.

Mặc dù, người ta có thể chạy bất kì phương thức phát hiện cộng đồng tĩnh (static community) nào, [4][5][6][8], để tìm thấy những cấu trúc cộng đồng mới bất cứ khi nào mạng lưới được cập nhật, tuy nhiên có thể gặp phải một số nhược điểm không thể bỏ qua: (1) thời gian hoạt động dài của phương thức tĩnh cụ thể trên mạng lưới lớn, (2) the trap of local optima và (3) hầu như cùng phản ứng lại những thay đổi nhỏ với một phần của mạng lưới (the almost same reaction to a small change to some local part of the network). Cách tốt hơn, hiệu quả và ít tốn thời gian hơn để thực hiện công việc này là cập nhật thích nghi (adaptively update) cộng đồng mạng từ những cấu trúc đã biết trước, điều này giúp tránh tranh cãi của việc tính toán lại từ đầu. Phương pháp giải quyết thích nghi này là phần trọng tâm chính của nghiên cứu trong bài báo này. Trong Hình 1, chúng tôi tổng quát hóa ý tưởng sự thích nghi cấu trúc mạng lưới cộng đồng năng động.

Phát hiện cấu trúc cộng đồng của một mạng xã hội năng động là một lợi ích đáng kể. Để đưa ra ý nghĩa của nó, hãy để ý đến định tuyến vấn đề trong mạng lưới (communication network) nơi các nút (nodes) và các liên kết (links) trình bày về giao tiếp người và di động (mobile) tương ứng.



Hình 1: Mạng tiến triển từ thời gian t tới t+1 dưới sự thay đổi ΔGt. Thuật toán thích nghi A nhanh chóng tìm được cấu trúc cộng đồng mới C (Gt + 1) dựa trên cấu trúc trước đó C (Gt) cùng với sự thay đổi ΔGt.

Do tính di động của các nút (nodes) và tính không ổn định của các liên kết (links) của mạng lưới, việc thiết kế mô hình định tuyến hiệu quả là cực kì thách thức. Tuy nhiên, từ khi con người có xu hướng hình thành các nhóm giao tiếp tự nhiên, tồn tại các nhóm nút kết nối bên trong dày hơn bên ngoài nằm dưới MANET (Mobile Ad-hoc Network) như một sự phản chiếu, vì thế, hình thành cấu trúc cộng đồng trên MANET này. Một thuật toán định tuyến hiệu quả, ngay khi nó phát hiện ra cấu trúc cộng đồng mạng, có thể định tuyến hoặc gửi những tin nhắn ngay lập tức tới các nút cộng đồng tương đương (hoặc các liên quan) như là đích đến. Làm theo các này, chúng ta có thể tránh những tin nhắn thừa (unnecessary mesages) chuyển tiếp thông qua các nút trong các cộng đồng mạng khác nhau, vì vậy có thể giảm số bản sao tin nhắn cũng như giảm chi phí thông tin, điều này là cần thiết trong MANETs.

Các đóng góp của bài báo này:

* Chúng tôi đề xuất QCA – thuật toán nhanh và thích nghi cho việc xác định một cách hiệu quả cấu trúc cộng đồng trong mạng xã hội năng động. Cách tiếp cận này có tính toán đến phát hiện cấu trúc mạng trước đây và chỉ quá trình thay đổi mạng lưới, do đó làm giảm đáng kể chi phí và thời gian xử lý.
* Chúng tôi nghiên cứu tính năng động của mạng xã hội và chứng minh kết quả lý thuyết liên quan, đó là những tính năng cơ bản của phương pháp.
* Chúng tôi đánh giá bao quát các thuật toán trên các mạng xã hội năng động thực tế khác nhau bao gồm: Enron email network, ArXiv citation network (mạng trích dẫn ArXiv) và Facebook. Kết quả thực nghiệm cho thấy phương pháp của chúng tôi không chỉ được cạnh tranh mà còn phát hiện cấu trúc cộng đồng chất lượng cao, một cách kịp thời.
* Như một ứng dụng, chúng tôi dùng QCA như cốt lõi của xác định cộng đồng trong định tuyến chiến lược MANETs. Kết quả mô phỏng cho thấy rằng QCA làm tốt hơn các phương pháp có sẵn hiện tại và xác nhận việc áp dụng phương pháp đề xuất trong điện toán di động (mobile computing).

Phần còn lại của bài báo được tổ chứ như sau: Phần *II*: Giới thiệu sơ bộ và định nghĩa vấn đề. Phần *III*: Đưa ra đầy đủ mô tả các thuật toán và lý thuyết liên quan. Phần *IV*: Cho thấy kết quả thử nghiệm của cách tiếp cận trên các bộ dữ liệu khác nhau trong thực tế. Phần *V*: Giới thiệu một ứng dụng thực tế trong cách tiếp cận MANETs. Phần *VI*: Thảo luận các công việc liên quan và cuối cùng là tổng kết công việc trong phần *VII*.

# **Sơ lược**

## Ký hiệu

* Cho *G = (V, E)* là đồ thị có hướng không trọng số (unweighted) với *N* nút và *M* liên kết trong mạng lưới.
* Cho *C = {C1, C2, …, Ck}* biểu thị các cộng đồng phân chia trong đó *Ci ∈ C* là một cộng đồng của *G*.
* Với mỗi đỉnh *u* biểu thị bằng *du*, *C(u)* và *NC(u)* lần lượt là cộng đồng có chứa *u* và các cộng đồng lân cận.
* Bất kì *S ⊆ V*, đặt *ms, ds,*  lần lượt là số liên kết trong *S*, tổng số đỉnh của *S*, và số kết nối từ *u* tới *S.*

## Dynamic social network

* Cho *Gs* = (*Vs*, *Es*) là thời gian phụ thuộc vào ảnhchụp mạng (network snapshot recorded) tại thời điểm *s* .
* Δ*Vs*, Δ*Es* là các bộ đỉnh và liên kết được giới thiệu hoặc loại bỏ tại thời điểm *s* và đặt Δ*Gs*= (Δ*Vs*, Δ*Es*) biểu thị sự thay đổi của toàn bộ mạng lưới.
* Ảnh chụp mạng tiếp theo *Gs+1* là hiện tại cùng với những thay đổi, chẳng hạn: *Gs+1 = Gs* ∪ *ΔGs.*
* Mạng lưới năng động *G* là một chuỗi các ảnh chụp mạng theo thời gian

*G = (G0, G1,...,Gs)*

## Hàm mục tiêu

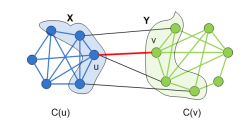
* Để định lượng cấu trúc cộng đồng mạng, chúng ta đưa vào biện pháp Q được định nghĩa như sau:
* Δ*Vs*, Δ*Es* là các bộ đỉnh và liên kết được giới thiệu hoặc loại bỏ tại thời điểm *s* và đặt Δ*Gs*= (Δ*Vs*, Δ*Es*) biểu thị sự thay đổi của toàn bộ mạng lưới.
* Ảnh chụp mạng tiếp theo *Gs+1* là hiện tại cùng với những thay đổi, chẳng hạn: *Gs+1 = Gs* ∪ *ΔGs.*
* Mạng lưới năng động G là một chuỗi các ảnh chụp mạng theo thời gian G = (G0, G1,...,Gs).

## Định nghĩa vấn đề

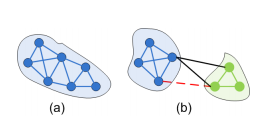
* G = (G0, G1,...,Gs) G0 là mạng ban đầu (original network) và G1, G2,...,Gs là các ảnh chụp thu được qua Δ*G1,* Δ*G2,...,* Δ*Gs*  .
* Cần đưa ra một thuật toán thích nghi để phát hiện và xác định hiệu quả cấu trúc cộng đồng mạng tại bất kì thời điểm .
* Thông qua việc sử dụng thông tin từ các ảnh chụp trước đó.
* Đồng thời tìm ra sự tiến hóa của cấu trúc cộng đồng mạng.

# **Mô tả thuật toán và lý thuyết liên quan**

Đầu tiên, chúng ta bàn về sự sự thay đổi topo mạng (network topology) ảnh hưởng tới cấu trúc cộng đồng của nó như thế nào. Giả sử *G = (V, E)* là mạng hiện tại và *C={C1, C2,…,C3}* là cấu trúc cộng đồng mạng tương ứng của nó. Chúng ta sử dụng giới hạn liên kết cộng đồng nội bộ (intra-community links) để biểu thị các cạnh mà 2 đầu mút (endpoints) của chúng thuộc cùng một cộng đồng và các cạnh có đầu mút thuộc các cộng đồng khác nhau. Với mỗi cộng đồng *C* của *G*, số kết nối của *C* với các động đồng khác là ít hơn số kết nối bên trong *C* với nhau, tức là, các nút trong *C* kết nối với nhau bên trong nhiều (densely) hơn bên ngoài. Việc thêm các kết nối nội bộ hoặc loại bỏ các kết nối giữ các cộng đồng của *G* làm cho cộng đồng trở lên vững mạnh và làm cho cấu trúc của *G* rõ ràng hơn. Ngược lại, việc loại bỏ các kết nối nội bộ và thêm vào các kết nối liên cộng đồng sẽ làm mất đi cấu trúc của *G.* Tuy nhiên, khi hai cộng đồng có đứt quãng (distraction), việc thêm vào hoặc loại đi các kết nối làm chúng hút nhau (attractive) và do đó, đề ra khả năng là họ sẽ được kết hợp để tạo ra một cộng đồng mới. Quá trình cập nhật cộng đồng là vô cùng khó khăn, bất kỳ thay đổi không đáng kể trong cấu trúc liên kết mạng có thể dẫn tới sự cố bất ngờ (sự thay đổi không mong đợi/ unexpected transformation) của cấu trúc(kết cấu) cộng đồng. Chúng ta sẽ thảo luận chi tiết về các hành vi có thể của kết cấu cộng động mạng năng động ở mục III-A.

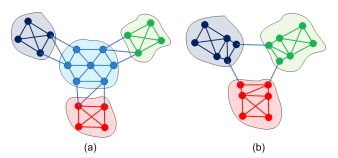


Hình 2: Khi biên (u, v) tham gia C(u) và C(v) được giới thiệu. Kiểm tra sự thay đổi thành viên được thực hiện trên bộ X và Y



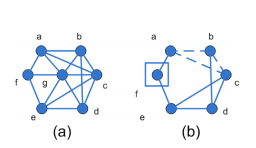
Hình 3: (a) Cộng đồng ban đầu

(b) Sau khi một cạnh (đường đứt đoạn đỏ) bị bỏ đi, cộng đồng được chia thành 2 cộng đồng nhỏ hơn



Hình 4: (a) Mạng ban đầu với bốn cộng đồng

(b) Sau nút có mức độ cao nhất được lấy ra, các nút còn lại tham gia vào các mô-đun khác, tạo ra mạng lưới với ba cộng đồng



Hình 5: (a) Cộng đồng ban đầu

(b) Khi nút trung tâm g được lấy ra, a sẽ tìm thấy b, c, d, e sau đó là f.

Để phản ánh những thay đổi với một mạng xã hội, biểu đồ cơ bản của nó được cập nhật một cách liên tục bằng cách thêm, bớt một nút hoặc tập các nút, hoặc thêm bớt cạnh hoặc tập các cạnh. Trong thực tế, việc khai báo hoặc loại bỏ một tập các nút (hoặc tập các cạnh) có thể được phân tích như một chuỗi các nút (cạnh) thêm vào (hoặc bớt/đi), trong đó, một nút duy nhất (hoặc một cạnh duy nhất) được khai báo (hoặc loại bỏ) tại cùng một thời điểm. Khả năng quan sát này giúp chúng ta nghiên cứu về sự thay đổi mạng nhưu một tập các sự kiện đơn giản (simple events), sự kiện đó có thể là một trong các: *newNode*, *removeNode*, *newEdge*, *removeEdge* chi tiết như dưới đây:

* *newNode (V+u)*: Một nút *u* mới với các cạnh có liên quan của nó được khai báo. *u* có thể đi với 0 hoặc nhiều hơn một cạnh.
* *removeNode (V-u)*: Một nút *u* và các cạnh liền kề bị loại bỏ khỏi mạng lưới.
* *newEdge (E+e)*: Một cạnh *e* mới nối 2 nút đã tồn tại được khai báo.
* *removeEdge (E-e)*: Một cạnh *e* đã có trong mạng bị loại bỏ.

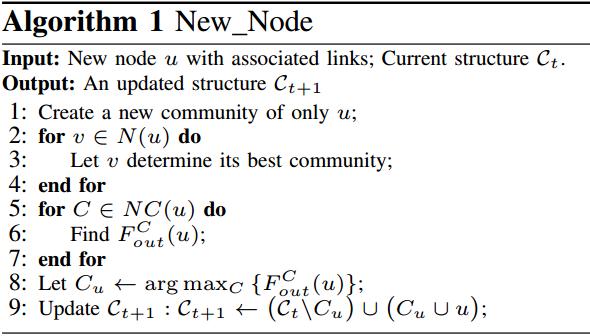
A. Thuật toán

Sự tiếp cận của chúng ta đòi hỏi một kết cấu cộng đồng ban đầu C0, được cho là cấu trúc cơ bản, để xử lý xa hơn. Khi mô hình đầu vào bị hạn chế như mạng vô hướng không trọng số, kết cấu cộng đồng ban đầu này có thể đạt được bằng cách biểu diễn bất kì phương pháp phát hiện cộng đồng mạng có sẵn [4][5][8]. Để đạt được kết cấu cở bản, chúng ta chọn một kết cấu cộng đồng mạng với mô-đun cao, trong khoảng thời gian hợp lý [3].

*1) New node*: Chúng ta cùng xem xét trường hợp đầu tiên khi nút *u* và các kết nối liên quan được khai báo. Lưu ý là nút *u* này có thể không có cạnh liền kề hoặc có thể đi với một hoặc nhiều cạnh kề hoặc nhiều cộng đồng. Nếu *u* không có cạnh liền kề, chúng ta tạo một cộng đồng mới chỉ có *u* và và để các cộng đồng khác cũng như mô-đun tổng thể Q còn nguyên vẹn. Một trường hợp thú vị xảy ra, khi u đi với các cạnh kết nối một hoặc nhiều cộng đồng đã tồn tại. Trong trạng thái sau này, chúng ta cần xác định cộng đồng nào mà u nên tham gia để tối ưu hóa tính mô-đun. Có một vài phương thức cục bộ được giới thiệu cho cho công việc này, ví dụ như các thuật toán của [4][8]. Phương thức của chúng ta được lấy cảm hứng từ một phương pháp tiếp cận vật lý (physical approach) được đề xuất ở [10], trong đó mỗi nút bị ảnh hưởng bởi 2 yếu tố:  (giữ *u* ở lại trong cộng đồng *C*) và (buộc *u* tham gia vào cộng đồng *C*) được định nghĩa như sau:

trong đó *doutS* trái nghĩa với *dS.*

Một nút *u* có thể xác định thành viên cộng đồng (community membership) của nó bằng cách tính toán những ảnh hưởng và cho phép bản thân nó tham gia vào cộng đồng có *Fout(u)* cao nhất (nếu *Fout(u) >* ) hoặc mặt khác, đặt *u* trong cộng đồng hiện tại. Theo Định lí 1, chúng ta kết nối một liên kết giữa ảnh hưởng và mục tiêu chức năng, tức là việc một nút mới tham gia vào cộng đồng với ảnh hưởng bên ngoài cao nhất (highest outer force) sẽ tối ưu hóa các mô-đun. Đây là ý tưởng trung tâm để xử lý trường hợp đầu tiên khi một nút mới và các liên kết liên quan của nó được khai báo. Qui trình chi tiết được trình bày trong Hình 6



Hình 6: Thuật toán 1 New\_Node

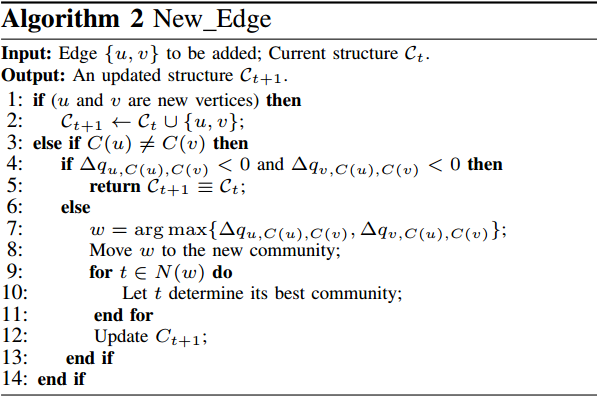
***Định lí******1****: Giả sử, C là một cộng đồng cung cấp Fout(u) tối đa, khi một nút u mới với bậc p được khai báo tới G, sau đó việc u tham gia C sẽ mang lại đóng góp tính mô-đun (modularity) tối đa.*

*2) New edge:* Trong trường hợp có một cạnh *e = (u, v)* nối hai đỉnh *u, v* đã được khai báo, chúng ta chia thành 2 trường hợp nhỏ hơn: *e* là một liên kết nội bộ cộng đồng (intra-community) (hoàn toàn trong một cộng đồng *C*) hoặc một liên kết liên cộng đồng (inter-community) (kết nối hai cộng đồng *C(u)* và *C(v)*). Theo Bổ đề 1, nếu *e* bên trong cộng đồng *C* thì sự có mặt của nó sẽ tăng cường cấu trúc bên trong *C*. Hơn nữa, theo Định lí 2, chúng ta biết rằng, việc thêm cạnh *e* không thể chia cộng đồng C hiện tại ra các mô-đun nhỏ hơn. Vì vậy, trong trường hợp này, chúng ta cho phép kết cấu mạng này giữ nguyên.

Một tình huống thú vị xảy ra khi cạnh *e* là một liên kết kết nối *C(u)* và *C(v)* vì sự có mặt của *e* có thể làm cho *u* (hoặc *v*) rời khỏi các mô-đun hiện tại của nó và tham gia vào một cộng đồng mới. Hơn nữa, nếu *u* (hoặc *v*) quyết định thay đổi tư cách thành viên (membership status) của nó thì cuối cùng nó có thể thông báo cộng đồng mới cho tất cả các các nút lân cận và một số có thể thay đổi thành viên như một hệ quả. Theo Bổ đề 2, chúng ta biết rằng, nếu *u* (hoặc *v*) có thể thay đổi phân cụm của nó thì *C(v)* (hoặc *C(u)*) là một cộng đồng mới phù hợp nhất cho nó. Nhưng làm thế nào để chúng ta có thể quyết định nhanh chóng và hiệu quả, liệu *u* (hoặc *v*) có thể thay đổi thành viên hay không hoặc tạo thành một cấu trúc cộng đồng tốt hơn khi *e* được thêm vào? Để kết thúc, chúng ta cung cấp một tiêu chuẩn để kiểm tra cho sự thay đổi thành viên của *u* và *v* trong Định lí 3. Ở đây, cả Δqv, C, D và Δqv, C, D  không đáp ứng được các tiêu chí, chúng ta có thể bảo vệ một cách an toàn cấu trúc cộng đồng mạng hiện tại và tiếp tục đi (keep going) (Hệ quả 1). Mặt khác, chúng ta chuyển u (hoặc v) tới cộng đồng mới của nó và để hang xóm (neighbors) xác định mô-đun tốt nhất để tham gia vào, sử dụng tìm kiếm cục bộ và swapping (đổi chỗ) để tối ưu hóa mô-đun thu được. Hình 2(a) mô tả thủ tục cho trường hợp này. Chi tiết thuật toán được mô tả trong Hình 7.

**Bổ đề 1**: Với bất kì *C ∈ C* , nếu *dC ≤ M – 1* thì thêm một cạnh bên trong *C* sẽ tăng cường tính mô-đun (modularity contribution).

**Định lí 2**: Nếu *C* là một cộng đồng của *G* ở thời điểm hiện tại, thì việc thêm bất kì kết nối trong nội bộ cộng đồng của *C* sẽ không chia ra các mô-đun nhỏ hơn.



Hình 7: Thuật toán 2 New\_edge

**Bổ đề 2**: Khi một cạnh mới (u, v) kết nối cộng đồng C(u) và C(v) được khai báo, thì C(u) (hoặc C(v)) là đề cử tốt nhất cho u(hoặc v).

**Định lí 3**: Giả sử rằng một cạnh mới (u, v) được thêm vào G. Để C ≡ C(u) và D ≡ C(v). Nếu cả Δqv, C, D ≡ Δqv, C, D