

LỰA CHỌN SIÊU NÚT TỐI ƯU CHO MẠNG P2P QUY MÔ LỚN

Vũ Thị Thúy Hà*, Vũ Văn San*, Nguyễn Hồng Đức*

*Học Viện Công Nghệ Bưu chính Viễn thông

Tóm tắt: Với sự phát triển nhanh chóng của mạng ngang hàng P2P, một số ứng dụng mới như P2PSIP đã nổi lên như một xu hướng mới trong lĩnh vực truyền thông đa phương tiện qua mạng internet. P2PSIP có khả năng khắc phục những nhược điểm của hệ thống SIP thông thường. Trong hệ thống P2PSIP cần một số các nút hoạt động như proxies và gateways gọi là siêu nút (SN) và khi mạng có kích thước lớn thì chi phí của việc lựa chọn SN tăng rất nhanh với độ phức tạp bản tin trao đổi là $O(N^2)$.

Bài báo đề xuất giải thuật bầu chọn siêu nút SNS (Super Node Selection) có tính tới các yếu tố trễ, độ ổn định và chi phí để duy trì độ ổn định mạng. Qua phân tích và kết quả mô phỏng giải thuật bầu chọn siêu nút SNS khi triển khai trên mạng ngang hàng Chord phân cấp mở rộng (Chord_SL) cải thiện hiệu năng so với các nghiên cứu trước đây.

Từ khóa: Mạng ngang hàng, siêu nút, nút thông thường, bảng băm phân tán, giao thức khởi tạo phiên, siêu-siêu nút, mạng Chord phân cấp mở rộng, lựa chọn siêu nút

I. ĐẶT VẤN ĐỀ

Khoảng vài năm trở lại đây, thế giới đã chứng kiến sự bùng nổ của Internet băng thông rộng, cùng với nó là sự phát triển mạnh mẽ của các ứng dụng ngang hàng. Với nhiều ưu điểm hứa hẹn như tính hiệu quả, linh hoạt và khả năng mở rộng cao. Đặc biệt các mạng ngang hàng P2P dựa trên bảng băm phân tán DHT đã và đang thu hút được nhiều sự quan tâm từ cộng đồng nghiên cứu. Tuy nhiên, mạng ngang hàng dựa trên DHT truyền thống chỉ cung cấp cấu trúc một chiều và không sử dụng đặc tính phân nhóm vốn có của một số ứng dụng (ví dụ: dịch vụ truyền hình hội nghị, thoại hội nghị qua mạng P2P, P2PSIP...). Mô hình truyền thống không thể giải quyết được các vấn đề liên quan tới tỷ lệ ra nhập và rời mạng cao và tính không đồng nhất của các nút trong mạng P2P.

Để cải thiện hiệu năng định tuyến và thích ứng với mạng không ổn định và tính không đồng nhất của các nút, việc phát triển một mô hình phân cấp P2P là cần

thiết. Nghiên cứu [1] đề xuất mô hình Chord_SL phân cấp mở rộng, việc cấu trúc và xây dựng mạng phân cấp dựa trên vị trí của các nút tham gia đã giải quyết được vấn đề không đồng nhất hiệu năng giữa mạng chồng phủ và mạng nền tảng. Cấu trúc mạng được chia làm hai lớp: Lớp liên miền (superlayer)

quản lý K cụm nội miền và các lớp nội miền (local layer) có n nút $n = N / K$.

Mô hình phân cấp Chord_SL đã cải thiện hơn so với các nghiên cứu phân cấp của các nghiên cứu trước. Tuy nhiên trong nghiên cứu [1] vẫn chưa đưa ra giải thuật lựa chọn SN mà SN được gán cố định. Vì vậy hiệu năng mạng giảm khi SN bị lỗi hoặc SN rời mạng.

Do các nút tham gia vào mạng ngang hàng là không đồng nhất vì vậy để xây dựng mạng phân cấp ổn định và hiệu quả, giải thuật bầu chọn các nút có năng lực làm siêu nút có ảnh hưởng rất lớn tới hiệu năng của hệ thống. Qua nghiên cứu và khảo sát nhiều nghiên cứu đã đưa ra giải thuật bầu chọn siêu nút trong mạng ngang hàng phân cấp [2-10]. Việc bầu chọn siêu nút dựa vào khoảng cách để giảm trễ được các nghiên cứu [2], [6], [7] đề xuất. Những nghiên cứu này tập trung giảm trễ truyền thông giữa các nút bằng cách khám phá sự lân cận của mạng hơn là khám phá năng lực của nút để xây dựng mạng phân cấp siêu nút hiệu quả.

SG-1 là một giao thức lựa chọn siêu nút nổi tiếng xem xét năng lực của nút, nhưng thiếu cơ chế ra quyết định phù hợp làm cho giải thuật hội tụ chậm và tiêu tốn mào đầu điều khiển trong quá trình lựa chọn siêu nút [10].

Nghiên cứu [3], [5] đã đề xuất một kỹ thuật học máy tự động để cải thiện giải thuật SG1. Tuy nhiên việc lựa chọn siêu nút phải dung hòa giữa giảm trễ truyền thông, hiệu quả tìm kiếm và lựa chọn những siêu nút có năng lực đủ mạnh dựa vào một số các đặc tính ví dụ như: Thời gian sống của nút, khả năng xử lý, băng thông....

Phần nội dung tiếp bài báo đề xuất giải thuật bầu chọn siêu nút được gọi là SNS (Super Node Selection) có tính tới các yếu tố tối ưu như: trễ, độ ổn định của mạng phân cấp và chi phí để duy trì độ ổn định mạng. Qua phân tích và kết quả mô phỏng giải thuật SNS khi triển khai trong mạng ngang hàng phân cấp cải thiện hiệu năng so với các nghiên cứu trước đây [4], [10].

II. ĐỀ XUẤT GIẢI THUẬT SNS

A. Xây dựng hàm mục tiêu

Trong mạng phân cấp, các nút trong vòng nội miền sau một chu kỳ sẽ được chạy thuật toán để đánh giá khả năng của nút. Một nút với hiệu năng vượt trội sẽ được chọn để trở thành một SN, trong khi nút với hiệu năng thấp sẽ hoạt động như một nút thông thường ON. Các tham số đánh giá khả năng của nút bao gồm: Băng thông, khả năng xử lý, thời gian trực tuyến. Trong

LỰA CHỌN SIÊU NÚT TỐI ƯU CHO MẠNG P2P QUY MÔ LỚN

nghiên cứu đề xuất hàm giá đa biến để đánh giá năng lực của nút :

$$u(x_1, x_2, \dots, x_n) = \sqrt[n]{\prod_{i=1}^n f_i(x_i)} \quad (1)$$

Trong đó:

x_1, x_2, \dots, x_n tương ứng là băng thông, khả năng xử lý, thời gian trực tuyến, ...

$f_i(x_i)$ là hàm chi phí tương ứng với các biến x_1, x_2, \dots, x_n .

Trong phương trình nêu trên tất cả các tham số có ảnh hưởng tới hiệu năng hoạt động của các nút ở mức độ khác nhau. Bằng cách kết hợp tất cả các yếu tố nêu trên, hàm mục tiêu G của giải thuật đề xuất được xác định như sau:

$$G = \sqrt[3]{\frac{t_{on(p)} \cdot P_{(p)} \cdot B_{(p)}}{t_{on(SN)} \cdot P_{(SN)} \cdot B_{(SN)}}} \quad (2)$$

Trong đó:

$t_{on(p)}, P_{(p)}, B_{(p)}$ lần lượt là thời gian hoạt động trung bình của nút, khả năng xử lý CPU (MIPS Million Instruction Per Second), băng thông của nút p ;

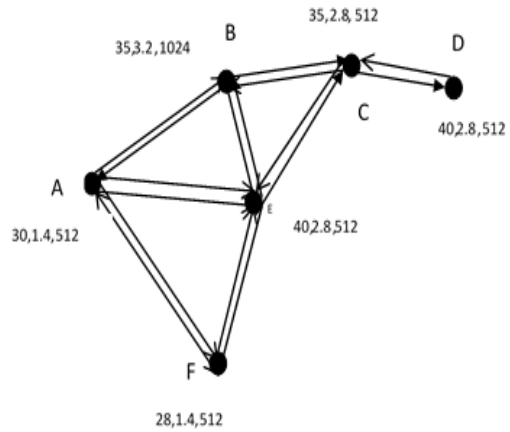
$t_{on(SN)}, P_{(SN)}, B_{(SN)}$ lần lượt là các giá trị yêu cầu tối thiểu của các tham số đối với một nút SN, giá trị tham số được chọn tùy theo mục tiêu của từng dịch vụ triển khai.

B. Thuật toán bầu chọn SNS

Thuật toán bầu chọn SN sẽ được cập nhật cùng với quá trình chạy ổn định stabilization() của giải thuật Chord trong lớp nội miền. Mỗi nút quảng bá các tham số hiệu năng tới các nút hàng xóm trong nội miền. Từng nút sẽ đưa ra bình chọn nút tốt nhất theo kết quả tính toán từ thông tin nó nhận được. Mỗi nút chịu trách nhiệm cho các tính toán bầu chọn của chính mình. Nút nào được bình chọn nhiều nhất sẽ được lựa chọn SN. Khi SN được lựa chọn, nó cần quảng bá thông tin của nó bao gồm cả địa chỉ IP và định danh ... đến tất cả các nút có liên quan trong cùng lớp nội miền.

Mô hình hóa một phiên trực tuyến trên một mạng ngang hàng như một đồ thị có hướng $G = \{V, E\}$, trong đó V là tập các đỉnh đại diện cho các nút và E là tập các cạnh đại diện cho các liên kết lớp chồng phủ. Quá trình lựa chọn SN được thể hiện trong hình 1. Giả sử có 6 nút trong lớp nội miền và các tham số hiệu năng của tất cả các nút được liệt kê trong hình.

Các tham số $t_{on(SN)}, P_{(SN)}, B_{(SN)}$ lần lượt là : 30 phút, 1.0MIPS và 512Kbps tương ứng.



Hình 1. Hiệu năng của các nút tham gia lớp nội miền Chord_SL

Năng lực của tất cả các nút được tính toán theo hàm mục tiêu G. Nếu tính toán kết quả của nút j lớn hơn so với i thì $G_{ij} = 1$, ngược lại, $G_{ij} = 0$, nếu kết quả tính toán của hai nút j và i là như nhau $G_{ij} = G_{ji} = 1$.

Kết quả được sắp xếp trong một ma trận, các nút đưa ra bình chọn được liệt kê như là hàng của ma trận, các nút đã nhận được bình chọn được liệt kê như là các cột. Vì vậy, tổng của mỗi cột là tổng số phiếu bình chọn của tất cả các nút đã nhận được. Ví dụ trong hình 2 nút B nhận được nhiều bình chọn nhất sẽ được chọn là SN của nhóm.

	A	B	C	D	E	F
A	1	1	1	1	1	0
B	0	1	0	0	0	0
C	0	1	1	1	0	0
D	0	1	0	1	1	0
E	0	1	0	1	1	0
F	1	1	1	1	1	1
S	2	6	3	5	4	1

Hình 2. Ma trận bầu chọn SN

Thuật toán 1. Mã giả thuật toán bầu chọn SN
// Tạo một vòng tròn Chord mới.

```
n.create()
predecessor = nil;
successor = n;
// join a Chord ring containing node n'.
n.join(n')
predecessor = nil;
```

```

successor = n'.find_successor(n);
VOTE(n,k){
    a = n; i = k; j = k;
    while (i < N){
        a = a.successor; t=j+1;
        G1 = n.t_on * n.P * n.B;

        G2 = a.t_on * a.P * a.B;
        if(G1==G2){ G[j][t]=1; G[t][j]=1;}
        else{
            if(G1>G2){G[j][t]=1;G[t][j]=0;}
            else{G[j][t]=0;G[t][j]=1;}
        }
        i++; j++;
    }
}
FIND_SN(n){
    for i=1 to m G[i][i]=1;
    k=1; t=n;
    while (k<N){
        VOTE(t,k);
        t=t.successor; k++;
    }
    for i=1 to N S[i]=0;
    for i=1 to N {
        for j=1 to N S[i]=S[i] + G[j][i];
    }
    SN = n; Smax = S[1]; j=1;
    for i=2 to N {
        if (S[i]>Smax){
            Smax = S[i];
            while(j<i){SN=SN.successor; j++;}
            j=i;
        }
    }
    return SN;
}

```

III. PHÂN TÍCH HIỆU NĂNG

Để phân tích hiệu năng của thuật toán lựa chọn siêu nút đề xuất, nghiên cứu dựa vào chi phí bản tin dùng lựa chọn siêu nút và duy trì độ ổn định của mạng chồng phủ siêu nút. Chi phí cho các bản tin trao đổi trong quá trình bầu chọn SN được tính:

a) Mỗi nút phát quảng bá các tham số của nó tới các nút hàng xóm;

b) Các nút có năng lực thấp hơn (bảng thông, khả năng xử lý,...) sẽ bầu chọn cho các nút có năng lực cao hơn, trong quá trình này các nút có năng lực yếu sẽ gửi bản tin bầu chọn cho các nút có năng lực cao hơn;

c) Sau khi SN được lựa chọn sẽ phát quảng bá các bản tin thông báo các tham số của nút đến các nút trong miền mà nó quản lý. Chi phí bản tin để lựa chọn SN, nếu triển khai thuật toán bầu chọn trên mạng Chord không phân cấp với số nút trong mạng N, chi phí để lựa chọn SN sẽ được tính:

$$A_N^2 + C_N^2 + N - 1 = N(N - 1) + \frac{N(N-1)}{2} + (N - 1) \quad (3)$$

$$= \frac{(3N+2)(N-1)}{2} \approx O(N^2)$$

Trong đó:

A_N^2 bản tin để phát quảng bá các tham số của nút tới các nút hàng xóm.

C_N^2 bản tin được dùng để các nút có năng lực yếu bầu cho các nút có năng lực cao hơn.

$(N - 1)$ bản tin được truyền từ nút được chọn làm SN để thông báo các tham số của nó tới các nút trong mạng. Nếu triển khai trên mô hình phân cấp Chord_SL [1], với số nút là N và mạng được chia làm K cụm. Chi phí để lựa chọn SN của thuật toán đề xuất là :

$$(A_{N/K}^2 + C_{N/K}^2 + \frac{N}{K} - 1) \times K = \frac{(3N+2K)(N-K)}{2K} \approx O(\frac{N^2}{K}) \quad (4)$$

Trong đó:

$A_{N/K}^2$ bản tin để phát quảng bá các tham số của nút tới các nút hàng xóm.

$C_{N/K}^2$ bản tin được dùng để các nút có năng lực yếu bầu cho các nút có năng lực cao hơn.

$(\frac{N}{K} - 1)$ bản tin được truyền từ nút được chọn làm SN để thông báo các tham số của nó tới các nút trong mạng.

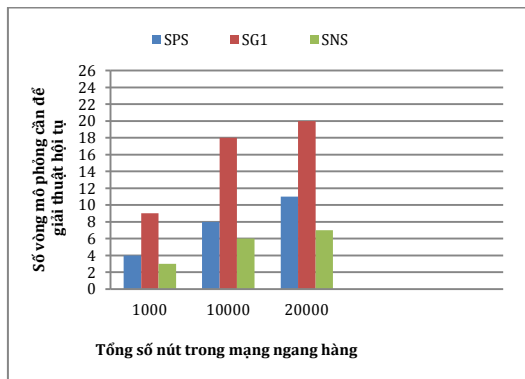
Từ công thức (3) và (4) rõ ràng thuật toán bầu chọn siêu nút SNS khi triển khai qua mô hình phân cấp có chi phí giảm so với mô hình không phân cấp truyền thống.

Nghiên cứu sử dụng phần mềm PeerSim [11] để thực hiện mô phỏng so sánh hiệu năng giải thuật đề xuất và giải thuật SPS [4] và SG1[10]. Phân tích mô phỏng tập trung đánh giá khả năng mở rộng của SNS với các nghiên cứu trước đây. Hay nói cách khác chúng tôi kiểm tra sự thay đổi thời gian hội tụ

của các giải thuật khi kích thước mạng tăng từ 1000 đến 20000 nút.

Kết quả mô phỏng hình 3 cho thấy với kích thước mạng tăng thì số vòng mô phỏng của giải thuật SPS và SG1 tăng hơn so với SNS. Lý do là các nút trong mô hình nghiên cứu của tác giả được cấu trúc phân cấp dựa trên vị trí của các nút tham gia, vì vậy siêu

nút SN là các nút không những gần về định danh mà còn lân cận với các nút trong nhóm. Ngoài ra mỗi nhóm còn duy trì một ma trận chứa dữ liệu được bầu chọn siêu nút từ các nút trong nhóm.



Hình 3. Khả năng mở rộng của thuật toán

IV. KẾT LUẬN

Mạng ngang hàng phân cấp có hiệu năng cải thiện hơn so với các mạng ngang hàng truyền thống [1]. Tuy nhiên do đặc tính của mạng ngang hàng là mạng không ổn định và các nút tham gia vào mạng không đồng nhất (băng thông, khả năng xử lý, thời gian trực tuyến ...). Vì vậy việc lựa chọn SN trong các mạng phân cấp là rất khó khăn. Bài báo đề xuất giải thuật bầu chọn siêu nút SNS trong mạng ngang hàng phân cấp. SN được bầu chọn là nút có năng lực đảm bảo các yêu cầu về băng thông, khả năng xử lý, nút có độ ổn định cao.

Giải thuật bầu chọn khi triển khai trên mạng Chord_SL phân cấp cải thiện hiệu năng hơn khi triển khai trên mạng ngang hàng truyền thống. Qua phân

tích mô phỏng giải thuật bầu chọn siêu nút thích ứng với các mạng có quy mô lớn, khi tăng kích thước mạng thời gian hội tụ của giải thuật giảm nhiều so với các nghiên cứu [4], [10].

TÀI LIỆU THAM KHẢO

[1] Vũ Thị Thúy Hà, Lê Hữu Lập, Lê Nhật Thăng “Xây dựng mô hình Chord-DHT phân cấp tối ưu hỗ trợ dịch vụ trên nền P2P”, Tạp chí Khoa học và Công nghệ, Viện Hàn lâm Khoa học và Công nghệ Việt Nam, tập 52, số 6C, trang 94-105, 2014.
 [2] Chung, W.H., 2016. A Super-Peer Selection Strategy for Peer-to-Peer Systems.
 [3] Amirazodi, N., Saghiri, A.M. and Meybodi, M., 2016. An adaptive algorithm for super-peer selection considering peer’s

capacity in mobile peer-to-peer networks based on learning automata. Peer-to-Peer Networking and Applications, pp.1-16.
 [4] Liu, M., Harjula, E., Ylianttila, M.: An Efficient Selection Algorithm for Building a SuperPeer Overlay, Journal of Internet Services and Applications, 4(4), (2013)
 [5] Gholami, S., Meybodi, M.R. and Saghiri, A.M., 2014. A learning automata-based version of SG-1 protocol for super-peer selection in peer-to-peer networks. In Recent Advances in Information and Communication Technology (pp. 189-201). Springer International Publishing.
 [6] Jesi GP, Montresor A, Babaoglu O (2007) Proximity-aware superpeer overlay topologies. IEEE Trans Network Serv Manag 4(2):74–83.
 [7] Yu J, Li M (2008) CBT: a proximity-aware peer clustering system in large scale BitTorrent-like Peer-to-Peer networks. Comput Comm 31(3):591–602.
 [8] Jelasity M, Montresor A, Babaoglu O (2009) T-Man: gossip-based fast overlay topology construction. Comput Netw Elsevier 53(13):2321–2339.
 [9] Min S, Holliday J, Cho DS (2006) Optimal Super-peer Selection for Largescale P2P System. Proc Hybrid Inform Tech:588–593
 [10] Montresor, Alberto. "A robust protocol for building superpeer overlay topologies." Peer-to-Peer Computing, 2004. Proceedings. Proceedings. Fourth International Conference on. IEEE, 2004.
 [11] Jelasity, Márk, et al. "PeerSim P2P simulator." 2011-05-08]. <http://peersim.sourceforge.net> (2009).
 [12] Lei Shi, Jing Zhou, Qi Huang, "A Chord-based super-node selection algorithm for load balancing in hybrid P2P networks", Mechatronic Sciences, Electric Engineering and Computer (MEC), Proceedings 2013 International Conference on, 20-22 Dec. 2013
 [13] Merz, Peter; Priebe, Matthias; Wolf, Steffen, " Super-peer selection in peer-to-peer network using network coordinates," Proc.-Int. Conf. Internet Web Appl. Serv., ICIW, Jun. 2008, pp. 385-390.
 [14] I. Stoica, R. Morris, D. Karger, F. Kaashoek and H. Balakrishnan, "Chord: a scalable peer-to-peer lookup service for internet applications," in Proceedings of ACM SIGCOMM'01, San Diego, CA, Aug. 2001.
 [15] Vũ Thị Thúy Hà, Lê Hữu Lập, Lê Nhật Thăng “Cải thiện hiệu năng giao thức định tuyến Chord trong mạng ngang hàng”, Tạp chí nghiên cứu Khoa học và Công nghệ Quân sự, Viện khoa học Công nghệ và Quân sự, số 23, trang 40-46, 2013.
 [16] Imtiaz, Waqas Ahmed, Shimul Shil, and A. K. Rahamn. "Three layer hierarchical model for chord." arXiv preprint arXiv:1303.1751 (2013).

OPTIMAL SUPERNODE SELECTION FOR LARGE - SCALE P2P NETWORKS

Abstract: With the rapid development of P2P, a number of new applications such as P2PSIP has emerged as a new trend in the field of multimedia communications over the Internet.

P2PSIP capable of overcoming the disadvantages of conventional SIP system. In those scenarios, there need some nodes acting as proxies and gateways are called Supernodes (SNs) and when the network is large, the cost of selecting SN increases very rapidly with message complexity of traditional selection algorithms was $O(N^2)$.

This paper proposes a new SNS selection algorithm, which takes into account the optimization factors of delay, the stability of the network, and the cost of maintaining the network stability.

Through analysis and simulation results, the SNS super-node selection algorithm deployed in the Chord_SL network improved performance compared to previous studies.

Keywords: Peer to peer, session initiation protocol, super node, ordinary node, distributed hash table, ultrasuper-peer, chord super – large, super node selection.



Vũ Thị Thúy Hà, nhận bằng Thạc sỹ CNTT năm 2001 tại Đại học Quốc gia Hà Nội, NCS Học viện Công nghệ BCV. Hiện là Giảng viên khoa Viễn thông 1. Lĩnh vực quan tâm: Phân tích đánh giá hiệu năng mạng, mô hình hóa và mô phỏng, mạng chồng phủ ngang hàng, nén và xử lý dữ liệu truyền thông đa phương tiện .



Vũ Văn San tốt nghiệp tiến sĩ chuyên ngành điện tử - viễn thông năm 2000 tại Học viện Công nghệ Bưu chính Viễn thông. Hiện đang làm việc tại Học viện Công nghệ Bưu chính Viễn thông. Lĩnh vực nghiên cứu chính: Hệ thống thông tin quang, Truyền dẫn và xử lý tín hiệu số, Hệ thống thông tin và truyền thông.



Nguyễn Hồng Đức: Sinh viên năm 4 Khoa Viễn thông 1 Học viện Công nghệ Bưu chính Viễn thông. Lĩnh vực quan tâm: Phân tích đánh giá hiệu năng mạng, mô hình hóa và mô phỏng mạng truyền thông.