

Adaptive Probabilistic Algorithm for Peer Networks

K.G.S. Venkatesan¹, K.P.Kaliyamurthie²

¹Assistant Professor, Department of Computer Science Engineering, Bharath University, Chennai, India

²Professor & Head, Department of Computer Science Engineering, Bharath University, Chennai, India

ABSTRACT: In unstructured peer-to-peer networks, every node doesn't have world data regarding the entire topology and also the location of different nodes. Search algorithms to find the queried resources and to route the message to the target node. Flooding and RW square measure 2 blind search algorithms by that question messages square measure sent to neighbors with none information regarding the potential locations of the queried resources. Both algorithms are not suitable to route a message to target in effective manner. The proposed algorithm is dynamic search (DS), which is a generalization of flooding and RW. Dynamic Search uses a knowledge-based search mechanism that is Adaptive Probabilistic Search (APS). Each node could relay query messages more intelligently to reach the target node.

I.INTRODUCTION

Dynamic property of unstructured the whole topology and the location of queried resources.[7] Search algorithms to route the message to the target node. It can be classified into two categories: breadth first search (BFS)-based methods, and depth first search (DFS)-based methods.

These 2 kinds of search algorithms tend to be inefficient, either generating an excessive amount of load on the system or not meeting users' necessities. stochastic process (RW) may be a conservative search algorithmic rule, that belongs to DFS-based strategies. By RW, the question supply simply sends one question message (walker)to at least one of its neighbors. If this neighbor doesn't own the queried resource, it keeps on causation the walk at least one of its neighbors, apart from the one the question message comes from, and thus, the search value is reduced. Moreover, the success rate of every question by RW is additionally low because of constant coverage issue [1]. Increasing the quantity of walkers would possibly facilitate improve the search time and success rate, however the result is restricted because of the link degree and redundant path.[13]

II. RELATED WORKS

Flooding and RW square measure 2 typical samples of blind search algorithms by that question messages square measure sent to neighbors with one data concerning the attainable locations of the queried resources or any preference for the directions to send. another blind search algorithms embody changed BFS (MBFS), directed BFS, increasing ring [2], and random periodical flooding (RPF) Knowledge-based search algorithms cash in of the data learned from previous search results and route question messages with totally different weights supported the data [5].

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 7, July 2015

III. DYNAMIC SEARCH ALGORITHM

3.1 Operation of Dynamic Search Algorithm

DS is designed as a generalization of flooding, MBFS, and RW. There are two phases in DS [9].

3.1.1 Phase 1. When $h < n$

DS acts as flooding or MBFS. The number of neighbors that a query source sends the query messages to depends on the predefined transmission

```

Algorithm: The pseudo-code of dynamic search DS
Input: query source  $s$ , queried resource  $f$ , transmission probability  $p$ 
Output: the location information of  $f$ 
DS( $s, f, p$ )
/* the operation of  $s$  */
 $h \leftarrow 0$ 
if ( $h \leq n$ )
     $h \leftarrow h + 1$ 
     $s$  choose  $p$  portion of its neighbors
     $m_i$  carrying  $h$  visits these chosen neighbors
elseif ( $h > n$ )
     $h \leftarrow h + 1$ 
     $m_i$  carrying  $h$  visits one neighbor of  $s$ 
/* the operation of  $r$  */
foreach ( $r$ )
    if ( $r$  has the location information of  $f$ )
         $r$  returns the information to  $s$ 
         $m_i$  stops
    elseif ( $h > TTL$ )
         $m_i$  stops
    elseif ( $h \leq n$ )
         $h \leftarrow h + 1$ 
         $r$  choose  $p$  portion of its neighbors
         $m_i$  carrying  $h$  visits these chosen neighbors
    elseif ( $h > n$ )
         $h \leftarrow h + 1$ 
         $m_i$  carrying  $h$  visits one neighbor of  $r$ 

```

Probability p , query source is d . It send the query messages to $d < p$ neighbors.[4] When p is equal to 1, DS resembles flooding. Otherwise, it operates as MBFS with the transmission probability p . [14]

3.1.2 Phase 2. When $h > n$

The search strategy switches to RW. Each node that receives the query message would send the query message to one of its neighbors the number of nodes visited by DS at hop $h = n$ is the coverage c_n .

3.2 Knowledge-Based Dynamic Search

Knowledge-based search algorithms, including APS, biased RW, RI, local indices, and intelligent search, are applicable to combine with our DS algorithm. [15].

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 7, July 2015

IV. PERFORMANCE EVALUATION

DS based on some performance metrics, including the success rate, search time, query hits, query messages, query efficiency, and search efficiency [4].

4.1 Network Model

Summarize Newman's work about the random graph. Let $G_0(x)$ be the generating function for the distribution of the vertex degree k . $G_0(x)$ can be represented as

$$G_0(x) = \sum_{k=1}^m p_k x^k \quad (1)$$

Based on the generating function, the average degree of a randomly chosen vertex is given by

$$z_1 = \langle k \rangle = \sum_{k=1}^m k p_k = G_0'(1). \quad (2)$$

The average number of second neighbors is

$$z_2 = \left[\frac{d}{dx} G_0(G_1(x)) \right]_{x=1} = G_0'(1) G_1' \quad (3)$$

Where $G_1(x)$ is given by

$$G_1(x) = \frac{G_0'(x)}{G_0'(1)} \quad (4)$$

Due to the difficulties to correctly measure and sample the operational P2P networks

4.1.1 Power-Law Graphs

The power-law random graph with the degree exponent τ p_k is proportional to $k^{-\tau}$. That is, $p_k \propto k^{-\tau}$

(5)

the power-law distribution is obtained:

$$G_0'(1) \cong \frac{1}{\tau - 2} (1 - m^{2-\tau}) \quad (6)$$

$$\text{And } G_1'(1) \cong \frac{1}{G_0'(1)} \cdot \frac{m^{3-\tau}}{3 - \tau}, \quad (7)$$

assuming $2 < \tau < 3$.

4.1.2 Bimodal Topologies The bimodal network topology few ultra-peers are connected to a large number of nodes, and the rest have few neighbors [3]. peer belongs to the other part with few neighbors is thus $p_{\text{few}} = 1 - p_{\text{ultra}}$.

4.2 Performance Analysis

4.2.1 Success Rate (SR)

The success rate (SR) is the probability that a query is successful. SR can be calculated as

$$SR = 1 - (1 - R)^c \quad (8)$$

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 7, July 2015

where R is the replication ratio, and C is the coverage.

4.2.2 Search Time (ST)

The search time (ST) as the time it takes to guarantee the query success with success rate requirement SR_{req} . For MBFS search algorithms, this situation occurs when

$$\begin{aligned}
 & p \cdot G'_0(1) + p^2 \cdot G'_0(1) \cdot G'_0(1) + p^3 \cdot G'_0(1) \cdot (G'_0(1))^2 \\
 & + \dots + p^{ST_{MBFS}} \cdot G'_0(1) \cdot (G'_0(1))^{ST_{MBFS}-1} \\
 & = \log_{(1-R)}(1 - SR_{req})
 \end{aligned} \tag{9}$$

Thus, ST for MBFS is

$$ST_{MBFS} = \log_{p \cdot G'_1(1)} \left(\frac{p \cdot G'_1(1) \cdot \log_{(1-R)}(1 - SR_{req})}{p \cdot G'_0(1)} + 1 \right) \tag{10}$$

ST for RW is obviously $\log_{(1-R)}(1 - SR_{req})$

. When k is larger than 1, assume that

$$G'_0(1) \cdot (G'_1(1))^{t-1} \leq k \leq G'_0(1) \cdot (G'_1(1))^t, \tag{11}$$

ST of RW can be expressed as

$$\begin{aligned}
 & G'_0(1) + G'_0(1) \cdot G'_1(1) + \dots + G'_0(1) \cdot (G'_1(1))^{t-1} \\
 & + k \cdot (ST_{RW} - t) = \log_{(1-R)}(1 - SR_{req})
 \end{aligned} \tag{12}$$

ST for RW is

$$ST_{RW} = \frac{\log_{(1-R)}(1 - SR_{req}) - \sum_{i=0}^{t-1} G'_0(1) \cdot (G'_1(1))^i}{k} + t \tag{13}$$

Query message is smaller than or equal to the decision threshold n, ST of DS is equal to (10). When h is larger than n, the calculation can be expressed as

$$\begin{aligned}
 & p \cdot G'_0(1) + p^2 \cdot G'_0(1) \cdot G'_0(1) + p^3 \cdot G'_0(1) \cdot (G'_1(1))^2 \\
 & + \dots + p^n \cdot G'_0(1) \cdot (G'_1(1))^{n-1} + p^n \cdot G'_0(1) \\
 & \cdot (G'_1(1))^{n-1} \cdot (ST_{DS} - n) = \log_{(1-R)}(1 - SR_{req})
 \end{aligned} \tag{14}$$

Therefore, ST for DS is

$$\begin{aligned}
 & ST_{DS} \cdot n + \frac{\log_{(1-R)}(1 - SR_{req})}{p^n \cdot G'_0(1) \cdot (G'_1(1))^{n-1}} \\
 & - \frac{(p \cdot G'_1(1))^n - 1}{(p \cdot G'_1(1) - 1) \cdot (G'_1(1))^{n-1}} \\
 & \cong n + \frac{\log_{(1-R)}(1 - SR_{req})}{p^n \cdot G'_0(1) \cdot (G'_1(1))^{n-1}} - 1
 \end{aligned} \tag{15}$$

We compare ST's for DS and RW with one walker. The improvement ratio is

$$\frac{ST_{RW} - ST_{DS}}{ST_{RW}} \cong 1 - \frac{G'_1(1)}{G'_0(1)} \cdot \frac{1}{(p \cdot G'_1(1))^n} \tag{16}$$

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 7, July 2015

4.2.3 Query Hits (QH)

The replication ratio R in the network, and the coverage is C [7]. The number of query hits is R . First analyze the coverage C_h at the h^{th} hop. the coverage C_h at the h^{th} hop can be written as

$$C_h = \begin{cases} \sum_{i=1}^n P_i(V_h), & \text{for } h = 1, \\ \sum_{i=1}^n \prod_{j=1}^{h-1} [1 - P_j(V_j)] \cdot P_i(V_h), & \text{for } h \geq 2, \end{cases} \quad (17)$$

The visiting probability $P_i(V_h)$ of flooding or MBFS is

$$P_i(V_h) = \begin{cases} p \cdot p_i \cdot G_0'(1), & \text{for } h = 1, \\ 1 - [1 - p \cdot p_i \cdot G_1'(1)]^{C_{h-1}}, & \text{for } h \geq 2 \end{cases} \quad (18)$$

p_i can be written as

$$p_i = \frac{m / i^{1\tau}}{\sum_{i=1}^n m / i^{1\tau}}, \quad (19)$$

First calculate the probability that a vertex i is the candidate of RW, i.e.,

$$P_i(R_h) = \begin{cases} p_i \cdot G_0'(1), & \text{for } h = 1, \\ 1 - [1 - p_i \cdot G_1'(1)]^{C_{h-1}}, & \text{for } h \geq 2, \end{cases} \quad (20)$$

Then, the average number of candidates of RW at hop h is

$$r_h = \sum_{i=1}^n P_i(R_h). \quad (21)$$

Hence, the probability that vertex i is visited at hop h for RW is

$$P_i(V_h) = P_i(R_h) \cdot [1 - (1 - 1/r_h)^k] \quad (22)$$

where k is the number of walkers. The visiting probability

$P_i(V_h)$ of DS is given by

$$P_i(V_h) = \begin{cases} p \cdot p_i \cdot G_0'(1), & \text{for } h = 1, \\ 1 - [1 - p \cdot p_i \cdot G_1'(1)]^{C_{h-1}}, & \text{for } 2 \leq h \leq n, \\ P_i(R_h) \cdot [1 - (1 - 1/r_h)^{C_n}], & \text{for } h > n. \end{cases} \quad (23)$$

4.2.4 Query Messages (QM)

The query message e_h generated at hop h is given by

$$e_h = \begin{cases} p \cdot G_0'(1), & \text{for } h = 1, \\ p \cdot G_1'(1) \cdot C_{h-1}, & \text{for } h \geq 2. \end{cases} \quad (24)$$

The calculation of query messages for DS depends on h and n . The query messages e_h generated at hop h for DS can be written as

$$e_h = \begin{cases} p \cdot G_0'(1), & \text{for } h = 1, \\ p \cdot G_1'(1) \cdot c_{h-1}, & \text{for } 2 \leq h \leq n, \\ c_n, & \text{for } h > n \end{cases} \quad (25)$$

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 7, July 2015

4.2.5 Query Efficiency (QE)

The number of query hits (QE) and the number of query messages (QM) are the well-known performance metrics [7] for the evaluations of search algorithms

QE is defined as

$$QE = \frac{\sum_{h=1}^{TTL} QH(h)}{QM} \cdot \frac{1}{R}, \quad (26)$$

where QH(h) is the query hits at the hth hop, QM is the total number of query messages generated during the query, and R is the replication ratio of the queried object.

$$QE_2 = \frac{\sum_{h=1}^{TTL} QH(h) / h}{QM} \cdot \frac{1}{R} \quad (27)$$

4.2.6 Search Efficiency (SE)

The search efficiency (SE) is proposed as a unified performance metric for search algorithms [6]. the success rate SR, SE is defined as

$$SE = \frac{\sum_{h=1}^{TTL} QH(h) / h}{QM} \cdot \frac{SR}{R} \quad (28)$$

Where QH(h)/h is the query hits in the hth hop weighted by the hop count, QM is the total number of query messages generated during the query [12]. The hth hop and the replication rate R. Therefore, (28) can be written as

$$SE = \frac{\sum_{h=1}^{TTL} C_h \cdot R / h}{\sum_{h=1}^{TTL} e_h} \cdot \frac{1 - (1 - R)^{\sum_{h=1}^{TTL} C_h}}{R} \quad (29)$$

Where C_h is the coverage at the hth hop, e_h is the query messages generated at the hth hop, and R is the replication ratio. [16]

V. PERFORMANCE OF DYNAMIC SEARCH

5.1 Effects of Parameters n and p of DS

First, N is set as 10,000. Power-law topology is adopted and the exponent τ is set as 2.1, which is analog to the real world situation. The case n = 1 is analog to RW with K equal to the number of first neighbors, which is roughly 3.55 in this case. The case n = 7 is equal to the flooding. As this figure shows, DS with n = 7 sends the query messages aggressively in the first three hops and gets good Search Efficiency [10].

5.2 Search Time

In this case, N is set as 10,000, R is set as 0.01, and TTL is set as 7. Similar results can be obtained when the parameters are set as other values. The walkers K for RW are set as 1 and 32.[17]

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 7, July 2015

VI. CONCLUSION

We have proposed the DS algorithm, which is a generalization of the flooding, MBFS, and RW. DS overcomes the disadvantages in which each search algorithm performs well. Flooding or MBFS for the short term search and RW for the long-term search [11]. We analyze the performance of DS based on some metrics including the success rate, search time and search efficiency. Its search performances could be further improved.

V. ACKNOWLEDGMENT

The author would like to thank the Vice Chancellor, Dean-Engineering, Director, Secretary, Correspondent, Principal, HOD of Computer Science & Engineering **Dr. K.P. Kaliyamurthie**, Bharath University, Chennai for their motivation and constant encouragement. The author would like to specially thank **Dr. A. Kumaravel, Dean CSE** for his guidance and for critical review of this manuscript and for his valuable input and fruitful discussions in completing the work and the Faculty Members of Department of Computer Science & Engineering. Also, he takes privilege in extending gratitude to his parents and family members who rendered their support throughout this Research work.

REFERENCES

- [1] D. Stutzbach, R. Rejaie, N. Duffield, "Sampling Techniques for Large, Dynamic Graphs," Proc. Ninth IEEE Global Internet Symp. (Global Internet '06), Apr. 2006.
- [2] K.G.S.Venkatesan and M.Elamurugaselvam, "Using the conceptual cohesion of classes for fault prediction in object-oriented system", International journal of Advanced & Innovative Research, Vol. 2, Issue 4, pp. 75 – 80, April 2013.
- [3] Kamatchi P., Selvaraj S., Kandaswamy M., "Synthesis, magnetic and electrochemical studies of binuclear copper(II) complexes derived from unsymmetrical polydentate ligands", Polyhedron, ISSN : 0277-5387, 24(8) (2005) PP.900-908.
- [4] A.H. Rasti, D. Stutzbach, and R. Rejaie, "On the Long-Term Evolution of the Two-Tier Gnutella Overlay," Proc. Ninth IEEE Global Internet Symp. (Global Internet '06), Apr. 2006.
- [5] K.G.S.Venkatesan, "Automatic detection & control of Malware spread in decentralized peer to peer network", International journal of Innovative Research in Computer & Communication Engg., Vol. 1, Issue 1, September – 2013.
- [6] D. Milojicic, V. Kalogeraki, R. Lukose, K. Nagaraja, J. Pruyne, B. Richard, S. Rollins, and Z. Xu, "Peer-to-Peer Computing," Technical Report HPL-2002-57, HP, 2002.
- [7] K. Sripanidkulchai, Queries and Its Implications on Scalability of the Popularity of Gnutella, white paper, Carnegie Mellon Univ., Feb. 2001.
- [8] Babu T.A., Joseph N.M., Sharmila V., "Academic dishonesty among undergraduates from private medical schools in India. Are we on the right track?", Medical Teacher, ISSN : 0142-159X, 33(9) (2011) PP.759-761.
- [9] S.Sathish Raja and K.G.S. Venkatesan, "Email spam zombies scrutinizer in email sending network infrastructures", International journal of Scientific & Engineering Research, Vol. 4, Issue 4, PP. 366 – 373, April 2013.
- [10] M. Jovanovic, F. Annexstein, and K. Berman, "Scalability Issues in Large Peer-to-Peer Networks: A Case Study of Gnutella," technical report, Laboratory for Networks and Applied Graph Theory, Univ. of Cincinnati, 2001.
- [11] Suresh V., Jaikumar S., Arunachalam G., "Anti diabetic activity of ethanol extract of stem bark of *Nyctanthes arbor-tristis* linn", Research Journal of Pharmaceutical, Biological and Chemical Sciences, ISSN : 0975-8585, 1(4) (2010) PP.311-317.
- [12] K.G.S.Venkatesan and M.Elamurugaselvam, "Design based object oriented Metrics to measure coupling & cohesion", International journal of Advanced & Innovative Research, Vol. 2, Issue 5, pp. 778 – 785, 2013.
- [13] Singamsetty P., Panchumarthy S., "Automatic fuzzy parameter selection in dynamic fuzzy voter for safety critical systems", International Journal of Fuzzy System Applications, ISSN : 1562-2479 , 2(2) (2012) PP. 68-90..
- [14] K.P.Kaliyamurthie, D.Parameswari and R.Udayakumar "QOS Aware Privacy Preserving Location Monitoring in Wireless Sensor Network", Indian Journal of Science and Technology, Vol.6(5S) , PP: 4648-4652, May 2013, ISSN : 0974-6846.
- [15] Kiran B., Kareem S.A., Illamani V., Chitralkha S., "Case of Phthiriasis palpebrarum with blepheroconjunctivitis", Indian Journal of Medical Microbiology, ISSN : 0255-0857, 30(3) (2012) PP. 354-356..
- [16] Kumaravel. A. 2013. "Cryptography Automata", Indian Journal of Science & Technology, 6 (5s) : 4561 – 4566.
- [17] K.P.Kaliyamurthie, R. Udayakumar and D. Parameswari, "Highly Secured Online Voting System over Network". Indian Journal of Science & Technology ,Volume 6(6S), PP:4831-4836, June 2013, ISSN : 0974-6846.
- [18].Dr.A.Muthu Kumaravel, KNOWLEDGE BASED WEB SERVICE, International Journal of Innovative Research in Computer and Communication Engineering, ISSN(Online): 2320-9801,pp 5881-5888, Vol. 2, Issue 9, September 2014



ISSN(Online): 2319-8753
ISSN (Print): 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 7, July 2015

- [19].Dr.A.Muthu Kumaravel, Data Representation in web portals, International Journal of Innovative Research in Computer and Communication Engineering, ISSN(Online): 2320-9801, pp 5693-5699, Vol. 2, Issue 9, September 2014
- [20].Dr.Kathir.Viswalingam, Mr.G.Ayyappan, A Victimization Optical Back Propagation Technique in Content Based Mostly Spam Filtering, International Journal of Innovative Research in Computer and Communication Engineering, ISSN(Online): 2320-9801, pp 7279-7283, Vol. 2, Issue 12, December 2014
- [21].KannanSubramanian, FACE RECOGNITION USING EIGENFACE AND SUPPORT VECTOR MACHINE, International Journal of Innovative Research in Computer and Communication Engineering, ISSN(Online): 2320-9801, pp 4974-4980, Vol. 2, Issue 7, July 2014.
- [22].Vinothlakshmi.S, To Provide Security & Integrity for Storage Services in Cloud Computing, International Journal of Innovative Research in Computer and Communication Engineering, ISSN(Online): 2320-9801, pp 2381-2385, Volume 1, Issue 10, December 2013