IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Volume-3, Issue- 10 E-ISSN: 2347-2693

A Secure and Effective Content Transmission in Content Based Networks Using Bloom Cast

S. Hemanjali[1*] and I.M.V. Krishna[2]

[1*,2] Dept. of Computer Science and Engineering,
Prasad V Potluri Siddhartha Institute of Technology, Vijayawada, India

www.ijcseonline.org

Received: Sep/26/2015 Revised: Oct/26/2015 Accepted: Oct /26/2015 Published: Oct /31/ 2015

***Abstract*—** P2P network stands amidst one of the best and attractive network tool. This approach made the allocation of contents through internet accessible. For unstructured P2P networks Bloom Cast deliver as an efficient technique for text sending scheme. In order to conclude a guaranteed recall at a communication cost of the network, Bloom Cast makes copies of the capacity in the network constantly at a random across the P2P networks. To abutment random node sampling and network size appraisal Bloom Cast appropriate its hybrid network which is an aggregate of a lightweight DHT (distributed hash table) and an unstructured P2P overlay. There are prospects of malicious codes and false agreement in unstructured P2P networks. At times it achieves false identities resulting in false activity with other character. The normal method here uses the conception of DHT and influence management which arrange efficient file transmission. In the expected system, we propose a novel strategy, called Bloom Cast, an efficient and effective full-text transportation scheme, in satisfied networks. The query demand independent archetype strategy, we propose a novel strategy, called Bloom Cast, to support efficient and adequate text transmission. Bloom Cast are mathematically that the recall can be guaranteed at a communication cost of O (square root N), where N is the size of the network.

***Keywords*—**P2P network, Bloom Cast, Bloom Filter, Content Sharing, Distributed Hash Table.

1. INTRODUCTION

Peer to peer (p2p) file allocation networks are very attractive these days. Millions of people use these networks to share excessive amounts of digital fulfilled. Most of this fulfilled is copyrighted material that is extremely distributed, like movies and music. This fact makes the p2p file sharing systems very controversial and very attractive to millions of people. The fact that 60% of all Internet traffic subsist of p2p file allocation in the end of 2004 [14], and that this percentage is still developing, shows its popularity P2p file sharing is however not only about copyright infringement. The theory trailing p2p networks shows abundant advantages correlated to conventional client-server systems to conceive many legal opportunities for p2p systems in the near eventual. One beneficial property of p2p systems is that they automatically scale with the number of users, in comparison to a central server with compulsion that grow with the number of users. Examples of legal operation of p2p networks are the Skype network [22] for world-wide telephony and the circulation of free software through Bit torrent [3]. The difference between p2p networks and consolidate solutions is essentially that in p2p networks content and knowledge are dispersed over the network users. This announces many exploration questions. The Parallel and Distributed Systems group is researching these topics by conniving various solutions and testing them in real p2p file sharing networks. Currently, new appearance is added to an actual Bit torrent client to analyze new possibilities of file allocation.

The background of one of these appearance is consider in this report, namely Bloom filters. A Bloom filter is a data construction that can be used for bandwidth adequate communication of data sets. It also empower network users to send sets of contented while protecting their privacy , since a Bloom filter can only be read by users that have analogous content.

2. RELATED WORK

A. Content oriented networks

A content oriented network is a network architecture in which network control is occupying on a content ID rather of an IP address. Content authority broadcast their content intelligence to all nodes in the network through flooding, and each node sets a routing table based on the content advice.

In this paper, process of broadcasting content advice to all nodes in the network through inundation is

attribute to as ”advertising”. A routing table entry is the content ID in a content oriented network. Each node checks its clobber table and forwards a query to detect the content (Fig. 1). A content oriented network is classified as an architecture that focuses only on content disclosure or one that focuses on both content disclosure and arrangement. i3 (Internet Indirection Infrastructure) [9], DONA (Data-Oriented Network Architecture) [10], and TRIAD [11] are architectures that focus on content disclosure. In particular, TRIAD is said to be the dependent of a pioneering study. CCN (Content Centric Network) [12][13] is an architecture focusing on both content disclosure and provision. In CCN, each node has a table called an impending interest table (PIT) that stores information announce where the content query has passed and content foundation is achieve by using the PIT. To realize a content oriented network, it is

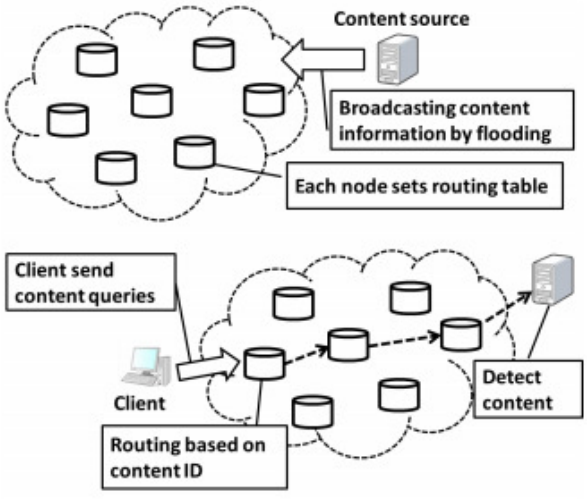

Figure 1: Overview of Content Oriented Networks

B. Routing scheme using Bloom filters (BFs)

A large amount of comfortable exists in a network but the memory of each node is defined. The significant increase in the number of routing table entrance is one of the problems that occur in content oriented networks. Routing schemes using BFs, i.e., data structures capable of abbreviate data have therefore captivate a lot of attention [6]. A BF is a space-efficient data structure that is used to test whether an element is an associate of a BF. An empty BF is a bit array of any number of bits, all set to 0. The process to add a component is as follows, first, the elements are fed to each of the hash functions to get the array location. The bits at this location are set to one. To test whether a component is in the BF, it is fed to each of the hash activity to check the position in the BF. Some problems with BFs are that the positions in a BF may have by chance been set to 1 during the injection of other elements. This result in a false positive and makes it impassable to remove an element from the BF because false negatives are not acceptable. In a current study, the amount of memory usage is about seventy five percent contraction by using BFs [8].

C. Bloom filters

A Bloom filter is a data structure that can be used to perform a set of elements. One can first add material of a set to the structure. Later on, the architecture can be interrogated for the membership of elements. So the elements themselves are not stored in the Bloom filter, only their membership. Bloom filters were named after Burton H. Bloom, who imported them in [2]. In this paper he correlated the space/time trade-offs of contrasting types of hash-tables. He construct that a new type of hash-table, which is now known as a Bloom filter desired less time to dismiss elements that are not in the table and less arena to store these elements. The deficiency of this Bloom filter is that a small error anticipation is imported when verification if elements are in the filter. A Bloom filter subsists of a bit array of m bits, which are all originally set to 0. Adding the elements of a set S = fx1; x2; : : : ; xn g of n details is done as follows. For each detail xi that is added, k contrasting hash functions h1; : : : ; hk each with a dimension f1; : : : ; mg are used to calculate k contrasting hash values h1(xi); : : : ; hk(xi). We conclude that these hash functions map each component to a accidental number inflexible over their range. Then, the bits hj(xi) are set to 1, for j = 1; 2; : : : ; k.

The bits in the Bloom filter can be set to 1 numerous times, but only the first time this has aftermath. The addition of the elements is show in Figure 2. When we want to check if a convinced element is in our Bloom filter, a similar approach is used as during the extension of elements. The same k hash action is calculated over the element y. Then we test if the bits hi (y) for i = 1; : : : ; k are equal to 1. If one or more of these bits are still 0, the aspect is absolutely not in the set. If all bits are 1, the aspect was apparently in the set, admitting there is a small probability that the approved bits were set to 1 due to the addition of different aspect. Then we have a false positive. There is a trade-off between the contingency of false likeness and the size of the Bloom filter. The false-positive probability can be determined from m and k in the following way .The probability p of one of the m bits still being zero after the addition of n aspect is

$$p = \left(1 - \frac{1}{m}\right)^{km} \approx e^{-kn/m} \quad (2.1)$$

The probability of a false positive f is then equal to the probability that all the k bits that we test are equal to 1, which is equal to

$$f = (1 - p)^k = \left(1 - \left(1 - \frac{1}{m}\right)\right)^k \approx (1 - e^{-kn/m})^k \quad (2.2)$$

By taking the cognate of Equation 2.2, from plain calculations it follows that for a given m and n, the value of k that decrease the false-positive anticipation f is equal to

$$k = \frac{m}{n} \ln 2, \quad (2.3)$$

Which gives a false-positive probability of?

$$f = \left(\frac{1}{2^{\ln 2}}\right)^{m/n} \approx 0.62^{m/n} \quad (2.4)$$

And the anticipation p of one of the bits still being 0 equal to

$$p = e^{-kn/m} = \frac{1}{2} \quad (2.5)$$

The choicest values for k calculated with Equation 2.3 are not natural numbers. So the optimum has to be bowed to a common number to get a useful k value.

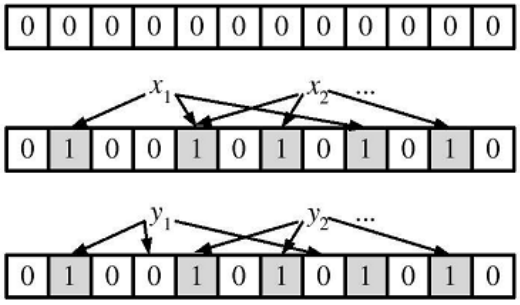

Figure 2: The bit array of a Bloom filter

In figure 2 The bit array of a Bloom filter, Top to bottom: an empty Bloom filter, adding component and checking the membership of elements (taken from [4]) most efficient to make k the essential number smaller than the excellent value, so that less calculation is desired to add and test aspect in the Bloom filter. In Table 1 are some examples of error probabilities and the sizes of the analogous Bloom filters.

bits/element	m/n	2	8	16	24
number of hash functions	k	1.39	5.55	11.1	16.6
false-positive probability	f	0.393	0.0216	$4.59 \cdot 10^{-4}$	$9.84 \cdot 10^{-6}$

Table1: The false-positive anticipation for Bloom filters with different sizes and an excellent number of hash functions. For large k, using k hash functions appeal much gauge. In [19] M. Mitzen-Macher considers the use of linear merger of two hash functions h1 and h2 to create k new hash-functions. The error-contingency f does not increase by this.

3. IMPLEMENTATION

Creating an Unstructured P2P Network:

Peer-to-peer (P2P) gauge or networking is a distributed application building that divides the burden among the peers. Peers are active and more honored participants in the application. They are said to form a P2P network of nodes. These P2P applications become attractive due to some files allocation systems such as Napster. This approach paved a way to new architecture and philosophies in many areas of human cooperation. Peer -to-peer networking has no restriction against technology. It covers only social course where peer-to-peer is aggressive. In such situation peer-to-peer processes are currently emerging throughout society. Peer-to-peer system appliance an abstruse glaze network which is built at Application Layer on the top of the physical network topology.

These overlays are independent from the physical network geology and are used for indicator and peer discovery. The contents are common through the Internet Protocol (IP) network. Anonymous peer-to-peer systems are interruption in the network, and appliance extra routing layers to ambiguous the existence of the source or harbor of queries.

Content Searching Using Bloom Cast

In this branch we are going to have a accurate discussion about the concepts of contented searching and bloom cast.

3.1. Content Searching

In unstructured peer to peer, information about one peer is unknown to the other. (i.e.) to empower communication between the peers, the peers in the network should know some advice about the alternative peer in the network. The proposed system uses the assigned hash table

where each and every peer has the isolated hash table. The information gathered in the hash table is established on Reputation authority (tracking users past activity). It helps to achieve the file searching operation comfortably. In a contented search function, the input is a set of access representing a user's concern and the output is a set of resources consist of these keywords. In the contented search context, resources perform text documents or metadata of familiar resources. Some of this capability is software applications, computer platforms, or data volumes. Content search is useful when a user does not know the exact capability names of concern; this case is accepted in P2P-based searches as well as in web searches. Flooding is the basic approach of seeking in unstructured P2P networks; however, large figure of unnecessary traffic is seen in blind inundation based search mechanism. This greatly changes the achievement of P2P systems. The farther study shows that a large amount of this undesirable traffic is divinable and can be bypass while curious in P2P networks. The bloom hash table is used to store the basics which help in effective searching of basics with desired competence. Information about the path-name is also administered by the bloom hash table. This design empower resource discovery without awareness of where the corresponding data items are stored.

3.2. Bloom Cast

Bloom Cast is a novel archetype strategy to support active and adequate full-text betterment. Different from the WP scheme, accidental node examine of a lightweight DHT is employ by the Bloom Cast. Here we generate the excellent number of replicas of the contented in the appropriate workspace. The size of the networks is not build upon on any factor since it is an unstructured P2P network. The size of the network is represent here as N. By further depict the optimal number of Bloom Filters rather of the raw documents, Bloom Cast accomplish guaranteed recall rate which results in devaluation of communication cost for archetype. Based on the Bloom Filter membership evidence we can easily architecture a query appraisal language to abutment full-text multi keyword search. Bloom Cast hybrid P2P network has three types of nodes: they are structured peers, natural peers, and bootstrap peers. A Bloom Cast peer stores an assortment of documents and continues a local storage also known as repository. A bootstrap node maintains a limited list of Bloom Cast nodes it accepts are directly in the system. There are many ways to appliance the bootstrap mechanism in the earlier P2P designs.

3.3. Bloom Filters

Bloom Filters to encode the conveyed lists while recursively converge the matching document set. A Bloom Filter is an adequate data architecture method that is used to test whether the element apply to that set or not. False positive comeback results are also achievable, but false negatives are not achievable; i.e. a query arrival either it is „inside the set" or „not inside the set". Elements can only be added to the set and cannot be detached. When more components are added to the set then the contingency of false portrait increases. Bloom Casting is a secure source definite multicast approach, which deportation the membership authority and per group forwarding state from the multicast routers to the source. It uses in -packet Bloom filter (BF) to cipher the forwarding tree. Bloom Casting separates multicast group management and most cast forwarding. It sends a Bloom Cast Join (BC JOIN) message towards the source AS. The message consists of an originally empty collector Bloom filter. While the message excursion upstream towards the source, each AS records promoting information in the authority packet by embed the corresponding link mask into a collector. After this, it performs a bit alteration on the representative. The figure for Bloom Filter and their memory cache is arranged here to show the interconnections between source and definitive multicast agreement. In Bloom Cast the passage routers do not keep any group -definite state. But in traditional IP multicast access the forwarding information is equipped in routers on the distribution tree.

4. CONCLUSION

By check the results of the bloom hash table, we construct that this is significantly faster than an orderly hash table using the same amount of recollection, hence it can backing better throughput for router operation that use hash tables. We here introduce an adequate and secured full-text retrieval design in an unstructured P2P networks using Bloom Cast method. Bloom Cast method guarantees the recall with high probability. Thus it is considered more adequate. The overall communication cost of a full-text search is diminished below an academic bound. Thus it is one among the sense that Bloom Cast is efficient and adequate among other design. Moreover the communication cost for replication is also diminished because we depict

Bloom Filters rather of the raw certificate across the network. During the deportation of files there is achievable of distribution of ailment, worms and Trojan horses and malicious peers to affected this the self-certification (MD5 with RSA) is used, it administer authentication and authorization. It easily finds the malicious become visible and aborts the transaction. Therefore the expected method provides the adequate and secure communication between the peers. Further the content examine method helps the user to find the rate of the hacked (modified) comfortable by fixing the chance level using the discipline data sets that are pre-defined by the user.

REFERENCES

[1] aSmallW orld. http://www .asmallworld.net.

[2] Burton H. Bloom. Space/time trade-offs in hash coding with allowable errors. Comm.of the ACM, 13(7):422–426, **1970**.

[3] Title=Bittorentprotocol,http://www.bittorrent.com/protocol/html Bram Cohen.

[4] A. Broder and M. Mitzenmacher. Network applications of bloom lters: A survey .In Proc. of the 40th Annual Allerton Conference on Communication, Control and Computing, pages 636 – 646, Illinios, USA, **2002**.

[5] Bernard Chazelle, Joe Kilian, Ronitt Rubinfeld, and A yellet T al. The bloomier filter an efficient data structure for static support lookup tables. In SODA '04: Proc. of the Fifteenth Annual ACM-SIAM Symp. on Discrete Algorithms, pages 30–39, Philadel-phia, P A, USA, **2004**.

[6] Saar Cohen and Y ossi Matias. Spectral bloom filters. In SIGMOD '03: Proc. of the 2003 ACM SIGMOD Int. conference on Management of data, pages 241–252, New York, NY , USA, **2003**. ACM Press.

[7] S. Dharmapurikar, P . Krishnamurthy, T . Sproull, and J. Lockwood. Deep packetinspection using parallel bloom filters. Micro, IEEE, 24(1):52 – 61, **2003**.

[8] Henrich Eisenhardt, Muller. Spektrale bloom-filter fur peer-to-peer information re-trieval. In Proceedings Informatik 2004, pages 44–48, **2004**.

[9] Li Fan, Pei Cao, Jussara Almeida, and Andrei Z. Broder. Summary cache: a scal-able wide-area web cache sharing protocol. IEEE/ACM Trans. Netw., 8(3):281–293, **2000**.

[10] Mohd. Amjad, "Wireless Network Security: Susceptibilities, Extortion and kiosk", International Journal of Computer Sciences and Engineering, Volume-01, Issue-03, Page No (10-14), Nov -**2013**, E-ISSN: 2347-2693

[11] Maciej Ceglowski Joshua Schachter. Loaf, a simple emailextension for contactrecognition using bloomlter, http://loaf.cantbedone.org, **2004**.

[12] R.A. Ferreira, M.K. Ramanathan, A. Awan, A. Grama, and S. Jagannathan, "Search with Probabilistic Guarantees in Unstructured Peer-to-Peer Networks," Proc. IEEE Fifth Int"l Conf. Peer to Peer Computing (P2P "05), pp. 165-172, **2005**.

[13] E. Cohen and S. Shenker, "Replication Strategies in Unstructured Peer-to-Peer Networks," Proc. ACM SIGCOMM "02. pp. 177-190, **2002**.

[14] P. Reynolds and A. Vahdat, "Efficient Peer-to-Peer Keyword Searching," Proc. ACM/IFIP/USENIX 2003 Int"l Conf. Middleware (Middleware "03), pp. 21-40, **2003**.

[15] H. Song, S. Dharmapurikar, J. Turner, and J. Lockwood, "Fast Hash Table Lookup Using Extended Bloom Filter: An Aid to Network Processing," Proc. ACM SIGCOMM, **2005**.

[16] C. Tang and S. Dwarkadas, "Hybrid Global-Local Indexing for Effcient Peer-to-Peer Information Retrieval," Proc. First Conf. Symp. Networked Systems Design and Implementation (NSDI "04),p. 16, **2004**.

[17] D. Li, J. Cao, X. Lu, and K. Chen, "Efficient Range Query Processing in Peer-to-Peer Systems," IEEE Trans. Knowledge and Data Eng., vol. 21, no. 1, pp. 78-91, Jan. **2008**.

[18] G.S. Manku, "Routing Networks for Distributed Hash Tables," Proc. ACM 22nd Ann. Symp. Principles of Distributed Computing (PODC "03), pp. 133-142, **2003**.

[19] H. Shen, Y. Shu, and B. Yu, "Efficient Semantic-Based Content Search in P2P Network," IEEE Trans. Knowledge and Data Eng.,vol. 16, no. 7, pp. 813-826, July **2004**.

[20] A. Broder and M. Mitzenmacher, "Network Applications of Bloom Filters: A Survey," Internet Math., vol. 1, no. 4, pp. 485-509, **2004**.

[21] V. King and J. Saia, "Choosing a Random Peer," Proc. ACM 23rdAnn. Symp. Principles of Distributed Computing (PODC "04), pp. 125-130, **2004**.

[22] I. Stoica, R. Morris, D. Karger, M.F. Kaashoek, and H.Balakrishnan, "Chord: A Scalable Peer -to-Peer Lookup Service for Internet Applications," Proc. ACM SIGCOMM "01, pp. 149 -160, **2001**.

AUTHORS PROFILE

Ms. Hemanjali Sanapala received her B.Tech degree in Information Technology from Andhra University Vizag, Andhra Pradesh. And she is currently pursuing M.Tech Degree in Computer Science & Engineering in Prasad V. Potluri Siddhartha Institute of Technology (Autonomous) Vijayawada, Andhra Pradesh, India and is affiliated to Jawaharlal Nehru Technological University Kakinada, Andhra Pradesh.

Mr. I. M. V. Krishna been working as Assistant Professor in the Department of Computer Science and Engineering Prasad V. Potluri Siddhartha
Institute of Technology (Autonomous) Vijayawada, Andhra Pradesh and is affiliated to JNTU-K, Kakinada,. He obtained M.Tech, and he had more than 7 Years 4 months of teaching experience. His areas of interests are Computer Networks, Distributed Systems.