ISSN 2278-3091

International Journal of Advanced Trends in Computer Science and Engineering, Vol.2 , No.2, Pages : 67-73 (2013)
Special Issue of NCRTECE 2013 - Held during 8-9 February, 2013 in SMK Fomra Institute of Technology, OMR, Thaiyur, Kelambakkam, Chennai

P2P ASSISTED MULTIMEDIA SHARING IN USER GENERATED CONTENTS

[1] Ponnaruvi P., [2] Dheivanai S. ., [3] Ramesh Kumar M. .,
[1] M.E- Student Vel Tech Multitech Dr.Rangarajan Dr.Sakunthala Engg College
E-mail:goldfallz@gmail.com, , Phone No: 9443311074
[2] M.E- Student Vel Tech Multitech Dr.Rangarajan Dr.Sakunthala Engg College
E-mail:dheivani54@gmail.com, , Phone No: 9790926154
[3] Asst Professor Vel Tech Multitech Dr.Rangarajan Dr.Sakunthala Engg College
E-mail:maestro.ramesh@gmail.com, Phone No: 9994483033

Abstract

Online forums are being eminent social network for conveying and discussing ideas regarding various topics. These forums are currently deployed as client-server model. As, that model is incapable of managing wide range of largely increasing multimedia contents efficiently. Alternatively, P2P assisted multimedia sharing is proposed. M-Board that can meet with increasing user demands for sharing multimedia contents efficiently with more server bandwidth is designed. By maintaining a list of systems those are remaining in online beyond the predefined threshold are all designated as stable nodes and other systems are randomly assigned as child nodes. Especially refreshing schema is proposed for maintaining recently refreshed list of available stable nodes. Here, by employing the authentication scheme in the P2P based multimedia sharing security issues are resolved. Additionally Instead of splitting and storing the contents at nodes at the time of increased demand the system with more capacity will be found and there multimedia files are replicated.

Keywords: Multimedia Board, Peer to Peer, Distributed Hash Table.

I. INTRODUCTION

1.1 Definition of P2P

Peer-to-peer (abbreviated to P2P) computer network is one in which each computer in the network can act as a client or server for the other computers in the network, allowing shared access to various resources such as files, peripherals, and sensors without the need for a central server. P2P networks can be set up within the home, a business, or over the Internet. Each network type requires all computers in the network to use the same or a compatible program to connect to each other and access files and other resources found on the other computer.

P2P is a distributed application architecture that partitions tasks or workloads among peers. Peers are equally privileged participants in the application. The owner of each computer on a P2P network would set aside a portion of its resources - such as processing power, disk storage, or network bandwidth - to be made directly available to other network participant without the need for central coordination by servers or stable hosts.

With this model, peers are both suppliers and consumers of resources, in contrast to the traditional client server model where only the server supply (send), and clients consume (receive). Emerging collaborative P2P systems are going beyond the era of peers doing similar things while sharing resources, and are looking for diverse peers that can bring in unique resources and capabilities to a virtual community thereby empowering it to engage in greater tasks beyond that can be accomplished by individual peers, yet are beneficial to all the peers.

1.2 Architecture of P2P systems

Peer-to-peer systems often implement an abstract overlay network, built at Application Layer, on top of the native or physical network topology. Such overlays are used for indexing and peer discovery and make the P2P system independent from the physical network topology.

A pure P2P network does not have the notion of clients or servers but only equal *peer* nodes that

ISSN 2278-3091

International Journal of Advanced Trends in Computer Science and Engineering, Vol.2 , No.2, Pages : 67-73 (2013)
Special Issue of NCRTECE 2013 - Held during 8-9 February, 2013 in SMK Fomra Institute of Technology, OMR, Thaiyur, Kelambakkam, Chennai

simultaneously function as both "clients" and "servers" to the other nodes on the network. This model of network arrangement differs from the client–server model where communication is usually to and from a central server. A typical example of a file transfer that does not use the P2P model is the File Transfer Protocol (FTP) service in which the client and server programs are distinct: the clients initiate the transfer, and the servers satisfy these requests.

The P2P overlay network consists of all the participating peers as network nodes. There are links between any two nodes that know each other: i.e. if a participating peer knows the location of another peer in the P2P network, then there is a directed edge from the former node to the latter in the overlay network. Based on how the nodes in the overlay network are linked to each other, we can classify the P2P networks as structured or unstructured.

In **structured** *P2P* networks, peers are organized following specific criteria and algorithms, which lead to overlays with specific topologies and properties. They typically use distributed hash table (DHT) based indexing, such as in the Chord system (MIT). [4]. Structured P2P systems are appropriate for large-scale implementations due to high scalability and some guarantees on performance (typically approximating O(log N), where N is the number of nodes in the P2P system).

Unstructured *P2P* networks do not impose any structure on the overlay networks. Peers in these networks connect in an ad-hoc fashion based on some loose set of rules. Ideally, unstructured P2P systems would have absolutely no centralized elements/nodes, but in practice there are several types of unstructured systems with various degrees of centralization. Three categories can easily be seen:

- In pure peer-to-peer systems the entire network consists solely of equipotent peers. There is only one routing layer, as there are no preferred nodes with any special infrastructure function.
- In centralized peer-to-peer systems, a central server is used for indexing functions and to bootstrap the entire system. Although this has similarities with a structured architecture, the connections between peers are not determined by any algorithm.
- Hybrid peer-to-peer systems allow such infrastructure nodes to exist, often called supernodes.

1.3Structured systems

Structured P2P networks employ a globally consistent protocol to ensure that any node can efficiently route a search to some peer that has the desired file/resource, even if the resource is extremely rare. Such a guarantee necessitates a more structured pattern of overlay links. The most common type of structured P2P networks implement a distributed hash table (DHT), in which a variant of consistent hashing is used to assign ownership of each file to a particular peer, in a way analogous to a traditional hash table's assignment of each key to a particular array slot. Though the term DHT is commonly used to refer to the structured overlay, in practice, DHT is a data structure implemented on top of a structured overlay.

1.4 Distributed hash tables

Distributed hash tables (DHTs) are a class of decentralized distributed systems that provide a lookup service similar to a hash table: (*key*, *value*) pairs are stored in the DHT, and any participating node can efficiently retrieve the value associated with a given key. Responsibility for maintaining the mapping from keys to values is distributed among the nodes, in such a way that a change in the set of participants causes a minimal amount of disruption. This allows DHTs to scale to extremely large numbers of nodes and to handle continual node arrivals, departures, and failures.

1.6 Indexing and resource discovery

Older peer-to-peer networks duplicate resources across each node in the network configured to carry that type of information. This allows local searching, but requires much traffic.

Modern networks use central coordinating servers and directed search requests. Central servers are typically used for listing potential peers (Tor), coordinating their activities (Folding @home), and searching (Napster, e Mule). Decentralized searching was first done by flooding search requests out across peers. More efficient directed search strategies, including supernodes and distributed hash tables, are now used.

ISSN 2278-3091

International Journal of Advanced Trends in Computer Science and Engineering, Vol.2 , No.2, Pages : 67-73 (2013)
Special Issue of NCRTECE 2013 - Held during 8-9 February, 2013 in SMK Fomra Institute of Technology, OMR, Thaiyur, Kelambakkam, Chennai

1.7 Peer to Peer like systems

In modern definitions of peer-to-peer technology, the term implies the general architectural concepts outlined in this article. However, the basic concept of peer-to-peer computing was envisioned in earlier software systems and networking discussions, reaching back to principles stated in the first Request for Comments, RFC 1.

A distributed messaging system that is often likened as an early peer-to-peer architecture is the USENET network news system that is in principle a client–server model from the user or client perspective, when they read or post news articles. However, news servers communicate with one another as peers to propagate Usenet news articles over the entire group of network servers. The same consideration applies to SMTP email in the sense that the core email relaying network of Mail transfer agents has a peer-to-peer character, while the periphery of e-mail clients and their direct connections is strictly a client–server relationship.

II. RELATED WORKS

Peer-to-Peer Overlay Networks aims at providing a deep view into the peer to peer network structure and comparing it with classical client server architecture, It defines the Peer to Peer network as like factually an overlay network for distributed Object store search and storing, Some common operation procedures in P2P are also presented along with its classifications in various generations, Second, P2P performance metric is discussed in order to understand and differentiated the practical P2P architectures and protocols. Third, the current P2P architecture is been classified and compared in detail. Then, object search protocols are carefully discussed with corresponding P2P architecture.

They also reviewed the current P2P modelling outputs and listed the general modelling approaches. They present multiple modeling approaches such as 1ST Generation of P2P, 2nd Generation of P2P, 3rd Generation of P2P and also the types involved in each generation of P2P.

Finally, several novel applications based on P2P techniques and future research directions of P2P overlay networks are briefly investigated. On the whole, this paper provides a clear and complete survey of P2P networks according to clues of its development and evolution process.

Multimedia Content analysis techniques are analyzed and their proposed methods are also discussed in Multimedia Content Analysis, Management and Retrieval Trends and Challenges automatically providing annotation to images appears to be the most valuable, yet the least utilized application. Almost all applications today work with data elements through verbal metadata.

It is the least utilized because automatic annotation is far from being even 50% accurate when the number of semantic classes is large (say exceeding 100).

Of course, the large search engines scrape text near a picture and use it for metadata which is not very accurate either.

An application of these techniques to affective content classification clearly requires the prior specification of the affective content categories (e.g. “happy”, “sad”, “exciting”) that are to be searched for in data, which then needs to be followed by training these categories using a suitable training data set.

The content diversity problem discussed above further propagates into the problem of feature-based affect representation. Although the existing solutions for the cognitive MCA have not yet reached the desired level of maturity, the need has already emerged for starting a parallel research effort in this field that targets the extraction of the affective content. The need for affective MCA theory and algorithms stems from the inability of the cognitive MCA principles to adequately address some of the grand challenges in the field, such as personalized music/video recommendation.

III. P2P ASSISTED FORUM IMPLEMENTATION SYSTEM ARCHITECTURE AND

3.1 P2P Construction

The P2P model has two structures: unstructured and structured P2P (DHT). Unstructured P2Ps are mostly gossip and flooding-based, which incur large amounts of communication overhead in

ISSN 2278-3091

International Journal of Advanced Trends in Computer Science and Engineering, Vol.2 , No.2, Pages : 67-73 (2013)
Special Issue of NCRTECE 2013 - Held during 8-9 February, 2013 in SMK Fomra Institute of Technology, OMR, Thaiyur, Kelambakkam, Chennai

the network. A DHT needs to maintain its topology in churn where node joins and departures lead to high maintenance overhead and decreased lookup efficiency.

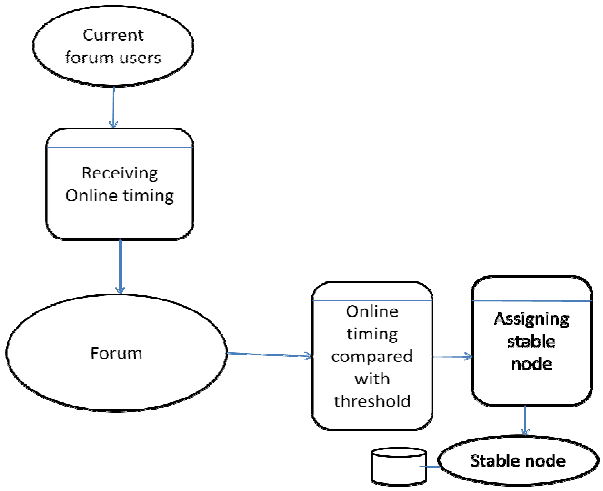

Fig 3.1: P2P Construction Level 1

MBoard intelligently forms stable nodes into a DHT to assist content discovery by aggregating content indices and matchmaking content requesters to providers. Specifically, MBoard builds a two-tier structure, with the DHT in the upper tier and other nodes connecting to the stable nodes in the lower tier.

A DHT needs to maintain its topology in churn, where node joins and departures lead to high maintenance overhead and decreased lookup efficiency. The observation implies that nodes are very dynamic in forums. Therefore, DHTs are not an optimal choice. However, the high communication overhead of unstructured P2P makes it a poor choice as well. Through which, we know that there are a fair number of stable nodes, which remain active in the forum most of the day.

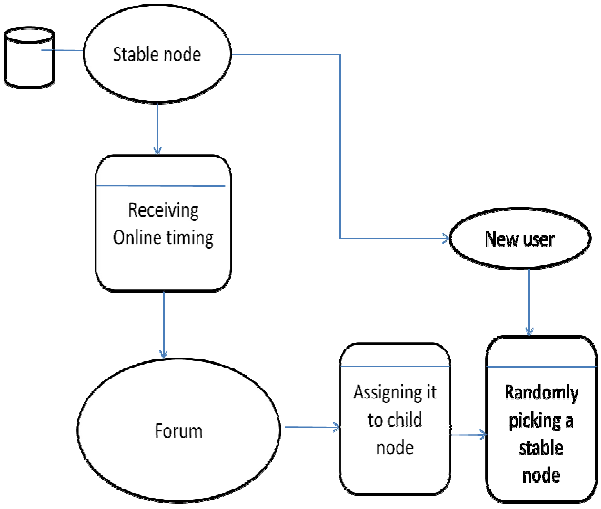

Fig 3.2: P2P Construction Level 2

3.2 Forum Maintenance:

A provider registered to a broker will not be in service forever due to a number of reasons: (1) it goes offline; (2) it stops providing uploading service; or (3) it deletes the cached videos. Therefore, the brokers need to update the index information in time in order to ensure that the chosen providers are in service.

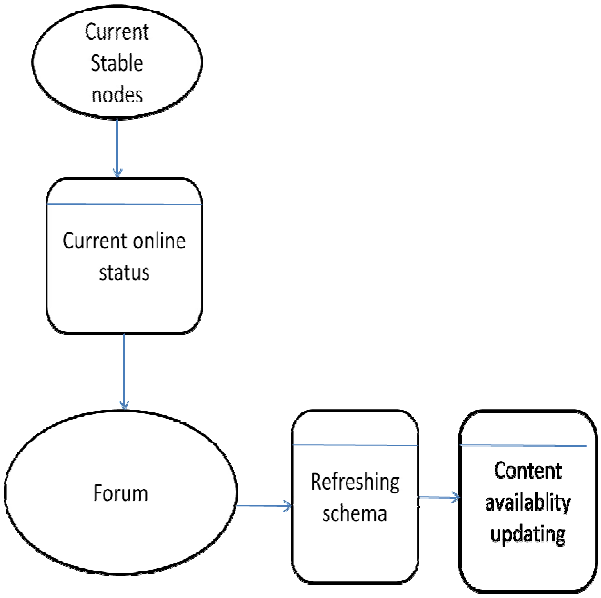

Fig 3.3: Forum Maintenance

ISSN 2278-3091

International Journal of Advanced Trends in Computer Science and Engineering, Vol.2 , No.2, Pages : 67-73 (2013)
Special Issue of NCRTECE 2013 - Held during 8-9 February, 2013 in SMK Fomra Institute of Technology, OMR, Thaiyur, Kelambakkam, Chennai

In order to minimize the overhead, MBoard lets brokers automatically discard the registered indices which were reported a certain time period ago. We call this time period the *refreshing time*, denoted as *tr*. MBoard sets *tr* equal to the continuous online time of the majority of the nodes.

3.3 Multimedia Content Retrieval:

When a node is downloading and viewing media content, it can upload the content simultaneously. In order to efficiently share media content, MBoard uses segmented media content to avoid the possibility of downloading failure and enable users to share existing media segments while downloading others.

MBoard specifies the segment size as 15 MBytes, the largest size of most media content in YouTube and the bit rate of the videos on YouTube, so that users do not need to split their videos in most cases and the rare long videos are automatically cut into segments by the MBoard client.

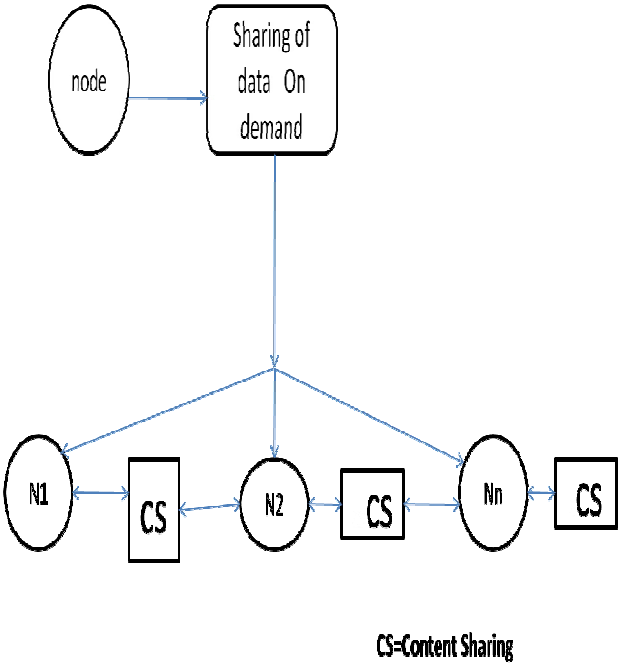

Fig 3.4: Multimedia content sharing

To retrieve a media segment, a requester asks its parent to send a request Lookup (key). The request will be forwarded to the broker of the segment that holds the registered index of the providers of the segment. The broker looks for the providers of the requested segment and returns a set of the freshest registered providers to the requester. The broker returns a number of providers rather than a single provider in order to increase the probability that at least one provider is available.

3.4Architectural representation

As both the structured and unstructured P2P does not suit for our purpose here, a two tier structure which is partially utilizing the structured P2P is designed. In that structure stable nodes are all placed in the upper tier and all other systems newly joining the systems are placed at the lower tier.

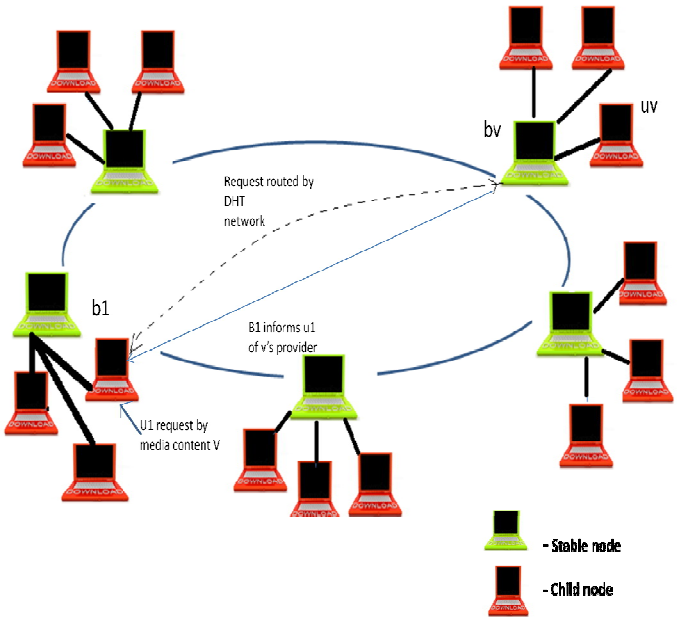

Fig 3.5: System Architecture

A DHT needs to maintain its topology in churn, where node joins and departures lead to high maintenance overhead and decreased lookup efficiency. On the whole, observations imply that nodes are very dynamic in forums. Therefore, DHTs are not an optimal choice. However, the high communication overhead of unstructured P2P makes it a poor choice as well. Through that, we know that

there are a fair number of stable nodes, which remain active in the forum most of the day.

Hence, MBoard intelligently forms a certain number of stable nodes into a DHT to assist content discovery by aggregating content indices and matchmaking content requesters to providers. Specifically, MBoard builds a two-tier structure, with the DHT in the upper tier and other nodes connecting to the selected stable nodes in the lower tier. The nodes connected to a stable node are called child nodes of the stable node. Since the selected stable nodes perform media content indexing, they are called brokers.

A DHT uses a consistent hash function to hash the identifier of nodes (e.g., IP addresses) and data objects (e.g., file names) to keys. It has two functions: Insert(key,object) and Lookup(key) to store the object with the key to its owner node and retrieve the object with the key.

In a DHT, each node maintains a routing table for log n neighbours. In order for a new node to join in the DHT overlay, it must know at least one other node already within the DHT. In MBoard, we use the Pastry DHT, but any kind of DHT can be used. For the best performance, the number of brokers N should not be large in order to avoid long routing latency.

The server maintains a list of stable nodes that are not selected as brokers and a list of brokers that currently serve in the DHT. The principle of stable node selection is that the longer a node is online daily, the higher probability it has of staying in the DHT. When a node u's daily online time exceeds a predefined threshold, it reports to the server. Then, the server adds node u to its stable node list.

The nodes in the stable node list are ordered according to their capacities. The highest capacity stable node becomes a broker. Specifically, the server assigns a bootstrap node from its broker list to the highest capacity stable node, and the node joins the DHT using the DHT node join protocol. Each time a stable node leaves the network and the node executes the DHT departure protocol and notifies the server.

IV.PERFORMANCE EVALUATIONS

4.1 The Effectiveness and Efficiency of MBoard

The CDF of the percentage of users for a forum versus playback delay with and without MBoard, i.e., the traditional server-client model. We see that with MBoard, more than 95 percent of the nodes achieve a very low delay before starting to download video data and 99 percent of all nodes have a delay less than 20 seconds.

On the contrary, without peer assistance, only 60 percent of all nodes have a delay less than 20 seconds. This is caused by the limited server upload bandwidth. When a large number of requests are sent to the server, most of them have to wait in the queue for processing due to the bandwidth limit of the server. Since MBoard allows nodes to request videos from peers, it achieves a much lower overall delay and also the CDF of user's video playback interruptions. Without MBoard, only 20 percent of nodes have no interruptions, and 80 percent of nodes have at least one interruption. With MBoard, 85 percent have no interruptions and only 11 percent of nodes have at least one interruption.

The results are consistent to those in MBoard the number of interruptions is substantially lower than without. The result implies that MBoard can enhance the users' playback smoothness of the server-client model due to its P2P model.

4.2 The Effect of Stable Nodes

In this test, we want to show the effectiveness of using stable nodes by comparing stable node settings with all nodes in the DHT. Fig. 10 shows the CDF of the percentage of users versus video playback delay when all nodes are on the DHT and when only stable nodes are on the DHT, respectively.

When only stable nodes are on the DHT, 30 percent of nodes have no more 0.6 second delay, 50 percent of nodes have no more 0.7 second delay, and 93 percent of nodes have no more 0.9 second delay. While when all nodes are on the DHT, 30 percent have no more 3 second delay, 50 percent of nodes have no more than 3.5 second delay, and 93 percent of nodes have no more 5 second delay. In both cases, around 96 percent of users have delays no more than 10 seconds.

Thus, more users ask the server for video segments. Later, as more and more peers have requested video segments, users can retrieve video segments from their peers, leading to a higher and constant P2P contribution percentage. We also see that the percentage when all nodes are included in the DHT structure is lower than when only stable nodes

are included. This is because the increased churn and larger playback delay result in higher data transmission failures when unstable nodes function as brokers in the DHT. Upon failure, the requesters ask the server for the video segments.

4.3 The Effect of Forum Popularity

For this test, we aim to show the effectiveness of MBoard at different levels of forum popularity. The forum popularity is the number of thread accesses during a certain period of time. We calculated the popularities of the 21 forums and ordered the forums in an ascending order of the popularity. We chose the last, two-thirds and one-third popularities in the list as the highest popularity, medium popularity, and low popularity, respectively, and tested the MBoard forum with different popularities.

Our survey shows that CDF of nodes' video playback delays for different forum popularities. We see that the low popularity and medium popularity tests have no users with a delay greater than 28 seconds, and over 98 percent of users have virtually no delay. Furthermore, while the high popularity test has less than 1 percent of users with the highest delay of 57 seconds, almost 99 percent have a delay of less than 28 seconds.

V.CONCLUSION

Most forums presently employ the server-client model, where the server replies requested content to the clients. The trace data collected from DISBoards show the rapid daily growth of user generated media content and users in forums, which becomes a hurdle for forums in meeting user demand due to limited server bandwidth. Through the analysis of trace data from DISBoards, it is found that their large group of users and user activity patterns meet the basic environmental requirements of employing a P2P model. Also, the existence of stable nodes, thread characteristics and media content patterns provide us with a direction to optimize the design of a P2P-based media sharing system .Hence, by creating a MBoard the P2P structure is been implemented and additional features are added to still increase the performance of the forum to provide multimedia content uploading and downloading efficiently.

REFERENCES

[1] "Internet Usage Statistics," http://www.internetworldstats.com/ stats.htm, 2011.
[2] "4chan," http://www.4chan.org/, 2011.
[3] "Gaia Online," http://www.gaiaonline.com/, 2011.
[4] "Ultimate Guitar," http://www.ultimate-guitar.com/, 2011.
[5] "The DIS Discussion Forums," http://www.disboards.com/,2011.
[6] "PPLive," http://www.pplive.com, 2011.
[7] "Google Launches YouTube Video Ads," http://www.readwriteweb.com, 2011.
[8] X. Ochoa and E. Duval, "Quantitative Analysis of User-Generated Content on the Web," Proc. First Int'l Workshop Understanding Web Evolution (WebEvolve '08), 2008.
[9] Y. Niu, Y. min Wang, H. Chen, M. Ma, and F. Hsu, "A Quantitative Study of Forum Spamming Using Context based Analysis," Proc. Network and Distributed System Security (NDSS) Symp., 2007.
[10] Y. Zhu, "Measurement and Analysis of an Online Content Voting Network: A Case Study of Digg," Proc. 19th Int'l Conf. World Wide Web (WWW' 10), 2010.
[11] "Digg website," http://www.digg.com, 2011.
[12] H. Yu, D. Zheng, B.Y. Zhao, and W. Zheng, "Understanding User Behavior in Large-Scale Video-on-Demand Systems," Proc. EuroSys Conf., 2006.
[13] X. Cheng, C. Dale, and J. Liu, "Statistics and Social Network of YouTube Videos," Proc. 16th Int'l Workshop Quality of Service (IWQoS), 2008.
[14] S. Annapureddy, "Providing Video-on-Demand Using Peer-to-Peer Networks," Proc. Internet Protocol TeleVision (IPTV) Workshop, pp. 238-247, 2006.
[15] C. Huang, J. Li, and K.W. Ross, "Can Internet Video-on-Demand Be Profitable?," Proc. SIGCOMM, 2007.