Exploiting cluster multicast for P2P streaming application in cellular system

Mohammad Zulhasnine
Systems and Computer Engineering
Carleton University, Ottawa, Canada
Email: mzulhasn@sce.carleton.ca

Changcheng Huang
Systems and Computer Engineering
Carleton University, Ottawa, Canada
Email: huang@sce.carleton.ca

Anand Srinivasan
EION Inc., Ottawa, Canada
Email: anand@eion.com

Abstract—**Streaming over peer-to-peer (P2P) network is popular, however causes needless traffic traversal through multiple links due to the mismatch between the physical and the overlay network. Cellular channels are limited in number and expensive. Because of the magnitude of contents per unit time and the nature of playing same contents throughout the entire system, collaborative streaming approach is the key to an efficient P2P streaming system. In this paper, we propose a collaborative streaming system where some cellular peers download contents from the Internet peers, and then share the contents with the remaining cellular peers by employing device-to-device (D2D) multicast application in order to avoid bottleneck at the eNodeB (eNB), and to reduce streaming cost. We present the broadcasters/agents and their optimal assisted peers selection problem as stable admission assignment and formulate the problem as integer linear programming (ILP) problem. We also present a distributed algorithm to select agents among the cellular peers with suitable number of assisted peers for each agent to tackle the retransmission problem. The cellular peers change role as a broadcaster or a multicast receiver to ensure fairness. We also perform extensive simulation to show the efficacy of our design and to verify our claims.**

I. INTRODUCTION

As more appealing applications and services are offered, we are being less wired on the contrary. Forecast conducted by Cisco [1] suggests that mobile video traffic would exceed Exabyte per month by 2012-2013. Applications such as live video streaming, Internet radio, and video conference have proliferated by riding over the P2P network. Enabling these P2P features to cellular users is still in limbo due to limitations caused by heterogeneity, mobility and time-varying capacities of the wireless channel. Besides, cellular links are expensive and inadequate than that of the Internet link. And the paradoxical reality is that peers scatter randomly which leads to needless traffic traversal through multiple links within a provider's network. The work in [2] emphasized on the scalable topology control protocol to discover neighboring wireless peers and save valuable wireless bandwidth. However the proposed model requires super peers to maintain indexes of shared file and peers' location. One way to save the wireless bandwidth is to broadcast/multicast contents from the eNB or to employ cache at the access point. However, this scenario requires Internet Service Provider (ISP) cooperation which is most unlikely in the context of the P2P network. The streaming system, named COSMOS [3], has the motivation of reducing streaming costs through collaborative sharing. Few cellular peers pull streams from the base stations and then share the contents with remaining users using free broadcast channel such as WiFi or Bluetooth. COSMOS employs dynamic broadcast instead of fixed broadcast where peers determine the broadcasting scope depending on its local density, and thereby reduce flooding and channel redundancy. However the schemes do not mention how the Internet peers and the close-proximity wireless peers are organized. Also there is no feedback mechanism to recover any missing segments of the content.

We just open the Pandora's box by stating that feedback mechanism is essential to ensure quality while multicasting over the wireless channel. Multicasting over the wireless network may exploit the inherent broadcast nature of the shared wireless channel. Ironically the same physical property poses a challenge on incorporating feedback mechanism while ensuring quality. Feedback mechanism may overwhelm the sender specially in a dense network. When traditional Automatic Repeat re-Quest (ARQ) method is implemented in multicasting over the wireless channel, high bandwidth and extreme coordination are required to process acknowledgements (ACKs) by each receiver. This leads to feedback implosion problem. An alternative solution to this problem is to implement negative ACK (NACK) where receiver(s) only send(s) NACK feedback to the sender upon receiving an erroneous frame. Even with NACK based protocol, significant bandwidth consumes and overhead accrues if the sender needs to identify receivers that did not receive a correct frame.

For streaming applications, users need to download a certain range of segments timely. Segment arriving after its scheduled playback time is useless. Meeting time constraint is therefore critical for the streaming service. In wireless multicast method, the sender has to transmit at the lowest rate sustainable to the receiver at the worst condition environment. This way valuable capacity is wasted because of the prolong channel occupancy. The authors in [4] proposed a dynamic rate-adaptation mechanism where the sender adapts multicast rate based on the Quality of Experience (QoE) feedbacks by the receivers. This way the sender transmits at a rate higher than the lowest instant sustainable rate of the receiver with the worst channel condition. Recently, D2D communication as an underlay to cellular networks has been introduced as a technol-

978-1-4673-5939-9/13/$31.00

ogy component to the Long Term Evolution Advanced (LTE-Adv) [5]. D2D communication as an underlaying LTE network empowers user-driven rich multimedia applications and also has proven to be network efficient offloading eNB traffic. In this paper, we propose a topology-aware P2P streaming system that integrates the cellular peers to the Internet peers. The key contributions to the work are:

- We infer the broadcaster/agent selection problem as stable admission assignment and formulate the problem as ILP. The proposed system avoids bottleneck at the base station, and saves wireless bandwidth as much as possible but still solves the feedback implosion problem and achieve effective multicast rate.
- We also propose and implement a novel distributed algorithm to select the broadcasters where peers need information of only one-hop away neighbors. We show an application of D2D communication by designing an streaming system where some cellular peers download from the eNB, and then share with neighboring peers. We also provide the details of content dissemination technique.

II. AGENT AND ASSISTED PEER SELECTION AS STABLE ADMISSION PROBLEM

Our proposed P2P streaming system comprises of both Internet and cellular peers. The streaming source is located in the Internet. We need to design a scheme to disseminate streaming contents among the peers efficiently. To save the expensive wireless bandwidth, and to avoid congestion at the base station, we need to select broadcasters among the cellular peers that download contents from the Internet peers, and then share with other neighboring cellular peers through multicasting mechanism. The peers that broadcast, we name them agents, and the peers that receive contents from the agents are assisted peers. We consider the following issues to design an efficient streaming system:

- Minimizing the number of agents would save wireless bandwidth by delivering contents to the assisted peers with minimized number of transmission.
- There should be enough number of agents so that none of them has to feed more than a specific number of assisted peers due to limitations of retransmission
- An agent should prefer a closer one among two assisted peers and vice versa. This would increase the multicasting rate.

Cluster based broadcasting approach has been around for long time. The cluster heads (agents in our case) selection criteria proposed in the past are based on node's connectivity [6], or node's mobility [7]. However we need to find agents that serves as many assisted peers as possible however not exceeding a certain limit to tackle the retransmission problem. In addition to that agents should choose closer peers as assistance to maximize the multicasting rate. The problem is rather close to many-to-one matching or stable admission problem [8] which finds a stable assignment between universities and students under a strict order of individual preference. The stable assignment is stable when there is no university and student pair who prefer each other as an outcome of the assignments. In our problem, universities are analogous to agents and students are analogous to assisted peers. A university on one side can offer admission to as many students as it may afford, and a student can be admitted by one university. Likewise one agent may assist a limited number of peers through broadcasting, and an assisted peer may only get assistance from one agent. Preparing preference is simple and straightforward. Optimum performance is achieved from stable admission algorithm by making use of the individual preference. To increase the multicast rate within cluster, an agent prefers a closer peer over a distant peer and vice versa. In addition to solving admission problem, we need to minimize the number of agents/universities as well. Also the roles of peers have not determined beforehand. Therefore we formulate the stable admission or the agent selection problem as a ILP framework.

In the steady state, cellular peers under an arbitrary eNB station is represented by a directed graph $G = (\mathcal{N}, \mathcal{L})$. Peers are labeled through 1 to N. The presence of a link, represented by an order pair (u, v) of distinct nodes, means that peer u and v are in direct communication range. The subset $\mathcal{A} \subseteq \mathcal{N}$ represents a set of agents. $\mathcal{N}_u$ is a set of open neighborhood of peer u. One agent can handle maximum S assisted peers due to limitations of retransmission. For all $u \in \mathcal{N}$, d_u^{eNB} represents the distance of peer u from the eNB, whereas d_{uv} represents the distance between two peers u and v. We define binary variable x_u which satisfies $x_u = 1$ if $u \in \mathcal{A}$, and '0' otherwise. Let y_{uv} represents a binary variable such that $y_{uv} = 1$ if peer u assists peer v, and '0' otherwise. The symbol $>_u$ indicates the preference orderings of peer u. For example, $j >_u v$ indicates that agent u prefers peer j over peer v as an assisted peer. Similarly $i >_v u$ denotes that an assisted peer v prefers agent i over peer u as an agent. Agents and assisted peers always prefer closer peer to maximize the multicasting rate. Therefore if $d_{uj} \leq d_{uv}$ then $j >_u v$ else $v >_u j$. The following ILP model describes the problem of selecting appropriate agents:

$$\text{Minimize} \quad \sum_{u=1}^{N} x_u d_u^{eNB} \tag{1}$$

Subject to:

$$\sum_{\forall u \in \mathcal{N}_v} y_{uv} + x_v = 1, \quad \forall v \in \mathcal{N} \tag{2}$$

$$\sum_{\forall v \in \mathcal{N}_u} y_{uv} \leq S, \quad \forall u \in \mathcal{N} \tag{3}$$

$$x_u \geq y_{uv}, \quad \forall u \in \mathcal{N}, \forall v \in \mathcal{N}_u \tag{4}$$

$$\begin{aligned} S y_{uv} + S x_v + S \sum_{i >_v u} y_{iv} \\ + \sum_{j >_u v} y_{uj} \geq S x_u, \quad \forall (u,v) \in \mathcal{L} \end{aligned} \tag{5}$$

The objective function in Eq. 1 minimizes the number of agents which is analogous to minimizing the number of broadcast transmission. The constraints in Eq. 2 ensure that each peer is either serving as an agent or getting assistance from only one agent. The constraints in Eq. 3 guarantee that an agent assists at most S peers. The constraints in Eq. 4 meet the requirements that an assisted peer is only served by an agent. The constraints in Eq. 5 ensure that suffice number of agents are selected and the selected pair is a match. The formulated ILP problem ensures that there are enough number of agents to avoid feedback implosion problem and therefore differs from the problem of finding minimum dominating set.

III. THE PROPOSED P2P STREAMING SYSTEM

The ILP-based agent selection problem, presented in section (II), requires the knowledge of the global topology. The optimal solution becomes unsolvable or at least intractable in polynomial time due to the increase size of the solution space. A distributed clustering algorithm is required not only due to the enormous size of the solution space, but more importantly the P2P system is self-organizing and decentralized in nature. In this section, we first describe distributed agent selection problem that captures the objective and the constraints presented in the ILP formulation, and then provide the detail description of the content dissemination technique. Our main focus is on the cellular part of the streaming network, and the air interface of the cellular network is the LTE.

A. Agent selection and Cluster formation

In our design, agents which have typically good connection to the eNB, first download contents from the eNB through the Internet Gateway and then share the contents with peers more closer to them in terms of mutual distance. P2P streaming applications are entirely user-driven, ISP/mobile-operator's cooperation is not feasible. The eNB does not participate in the streaming process and is completely oblivious to the User Equipments' (UEs) applications. It only allocates D2D resources whenever there is any request from the UEs. Cellular users may employ network controlled beacons to discover peers around its neighborhood [9]. Alternatively, the users may utilize the topology-aware C-Chord [10] system to communicate with neighboring device directly instead of connecting through the eNB. We now briefly present the agent selection algorithm. Peer only communicates to one-hop neighbors to determine its role as an agent or as an assisted peer in a distributive manner. Peers who do not have neighbors within maximum D2D distance, download contents through the eNB.

- At the beginning of the agent selection, peers exchange information of their eNB distance with their one-hop neighbors. If the distance information is unavailable, peers exchange exponential moving average of their signal-to-noise ratio (SNR) corresponding to the eNB with their one-hop neighbors.
- Each peer compares its own eNB distance to that of their neighbors. Peer with smallest distance when compared to that of its neighbors, declares it as an agent and waits for assistance requests. If the metric is SNR, peer with highest SNR compared to its neighbors declares itself as an agent.
- Each subordinate contacts the nearest agent for assistance. If the request for assistance is rejected by one agent, it contacts the next agent in order of mutual distance. In case a subordinate does not have any agent, however have potential neighbors to serve; it declares itself as an agent.
- Each agent accepts requests from at most S subordinate first with the preference to select closest one first. Due to peer dynamics, if an agent does not have enough assisted peers; it contacts the closest agent in order of mutual distance, and invites it to become an assisted peer.

Algorithm (1) describes how a cellular peer u determines its role as an agent or an assisted peer. The value of the integer variable s_u is the final outcome of the algorithm that defines the role of the peer u. Figure 1 illustrates the outcome of the distributed solution to the problem of selecting agents at a particular instant of time. For clarity, only a part of the network is shown. The agents (marked by the square) download streaming contents from the eNB, and then disseminate contents to their one-hop assisted peers

Algorithm 1: The functionality of peer u to determine the broadcasting role in a distributed manner.

```
Symbol definition:
$S$ :        maximum number of assisted peers
$T$ :        an integer larger than $S$
$s_u$ :      status of peer $u$; undetermined (-1), agent (0 to $S$), assisted ($T > S$)
$\mathcal{N}_u$ :  set consisting one-hop neighbor of peer $u$
$d_u^{eNB}$ : distance of peer $u$ from the eNB
$req_u$ :    number of request for assistance to $u$
Initial assignment:
Send $d_u^{eNB}$ to all $v \in \mathcal{N}_u$
Receive $d_v^{eNB}$ from all $v \in \mathcal{N}_u$
$s_u = -1$ (status undetermined);
$req_u = 0$;
begin
    while $s_u == -1$ do
        if $d_u^{eNB} \leq \min\left(d_v^{eNB};\ \forall v \in \mathcal{N}_u\ \&\ s_v < 0\right)$ then
            $s_u = 0$; (set itself as agent)
            while wait for assistance request do
                if Receive request for assistance then
                    $req_u = req_u + 1$;
                end
            end
            while $s_u < S\ \&\ req_u > 0$ do
                Accept request from the closest one;
                $s_u = s_u + 1$ ;
                $req_u = req_u - 1$;
            end
        else
            Find agents for assistance;
            if $found\ \&\ s_{agent} < S$ then
                Send request to the nearest agent for assistance
                if $accept$ then
                    $s_u = T$; (status is assisted)
                end
            end
        end
        if $s_v >= S;\ \forall v \in \mathcal{N}_u$ then
            $s_u = -2$; force termination; no potential neighbor left
        end
    end
end
```

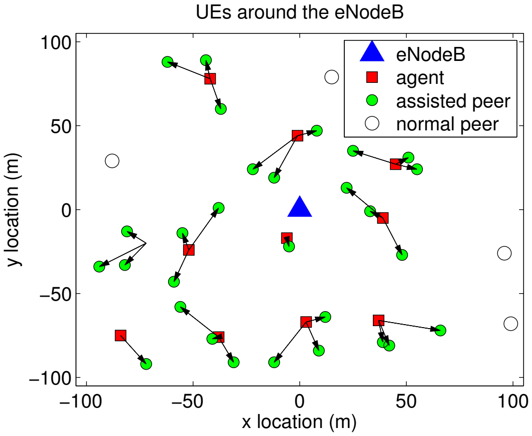

Fig. 1. Cellular peers under an eNB from clusters. Agents download contents from the Internet peers, and then share with multicast receivers employing D2D multicast. The max D2D distance $d_{max} = 30$ meters. The maximum number of assisted peers an agent assist, $S = 3$. For clarity, only a part of the network is shown.

(marked by filled circle) exploiting the broadcast nature of the wireless medium. The maximum number of assisted peers (S), an agent assists, is set to '3'. For an example, two agents i, j were selected with maximum number of '3' assisted peers per agent even though all of them are close. The agents employ D2D multicast mechanism; there is no collision as the eNB allocates channels appropriately. Therefore, agents i, j concurrently disseminate contents to their assisted peers. Peers only request D2D resources for multicast application; the eNB does not require devising P2P streaming system.

Algorithm complexity: We measure the performance by communication complexity. Assuming for an arbitrary eNB, $\Delta = \max\{|\mathcal{N}(u)| \forall u \in \mathcal{N}$ denotes the maximum degree among N cellular peers. Then the total number of messages to disseminate neighboring information among peers at the beginning of the algorithm is $O(N\Delta)$. If a peer determines its state to be an agent, it needs to send at most S communication messages of 'acceptance' to the assisted peers. Therefore total $O(AS)$ messages are required for A number of agents. Any peer seeking assistance, requires $\lceil \frac{\Delta}{S} \rceil$ messages in the worst case scenario to send request for assistance. So the total number of messages in this case is $O\left((N - A)\lceil \frac{\Delta}{S} \rceil\right)$. The number of messages requires is scalable, and the price, paid to it, is negligible when compared to the amount of cellular bandwidth saving to play the bandwidth hungry media.

Extension to the algorithm: To ensure fairness of shared streaming cost, each cellular peer maintains an additional information on shared amount. F_u, an integer variable, represents the shared amount of cellular peer u. The value of F_u is increased one unit each time cellular peer u broadcasts streaming content around its neighbors. When the value of F_u reaches a limit F^{lim}, it may choose to refrain itself to be an agent. When most of the neighbors reaches the limit F^{lim}, F_u is reset to zero for the peers around that neighborhood.

B. The content dissemination technique

In the proposed P2P streaming system, a source in the Internet progressively generates new chunks with consecutive time-stamp. We now discuss how to synchronize the content download and broadcast session. Let us assume agents download M chunks within α seconds. The broadcasting session completes within β seconds. Then for δ seconds, cellular peers download missing chunks from the neighboring peers, or from the Internet. During this δ seconds, peers also exchange information to determine the agents on the next cycle. Here we describe how to disseminate M chunks within $\alpha + \beta + \delta$ seconds beginning from a time-stamp t.

1) From t to $t + \alpha$ seconds, selected agents download M chunks from the Internet peers. An agent might miss few chunks from the M chunks, even if the agent selection algorithm always favors cellular peers with target SNR as the candidates. If a cellular peer does not have a neighbor, it simply download M chunks from the Internet within time-stamps $t, t + \alpha + \beta + \delta$.
2) From $t + \alpha$ to $t + \alpha + \beta$ seconds, the agents share M chunks (excluding any missing chunks) with neighboring peers employing D2D multicast applications. The eNB is not aware of the application, only remains in control of mode switching between cellular to D2D role and vice versa. Selecting multicast transmission rate is critical to avoid packet errors. We illustrate the multicast mechanism at the end of this discussion.
3) Peers in a streaming system maintain lists of multiple senders. From $t + \alpha + \beta$ to $t + \alpha + \beta + \delta$ time, multicast receivers (assisted peers) prepare a list of missing chunks, and send request for missing chunks to the agents. If time allows, peers continuously download missing chunks otherwise play media at a degrade rate. During this period, each cellular peer runs the distributed agent selection algorithm for the next cycle.

If the streaming rate is R kbps, then each cellular peer needs $R \times (\alpha + \beta + \delta)$ kb content in $\alpha + \beta + \delta$ seconds. Each broadcaster has to download at the rate of $R \times (\alpha + \beta + \delta)/\alpha$ kbps, and then transmits at the rate $R \times (\alpha + \beta + \delta)/\beta$ kbps employing multicast scheme.

D2D multicast reliability mechanism: In our agent selection algorithms, we allow agents to support a limited number of assisted peers. By limiting the multicast group into smaller size we avoid limiting multicast transmission rate to a lowest instant sustainable rate of the weakest peers around the neighborhood. Limiting the number of assisted peers for each agent also make retransmission possible. The agent adapts multicast rate based on the QoE feedbacks by the assisted peers [4] to avoid huge error or grab a higher rate if achievable. As the agents prefer closest peers as assisted peers, the adapted multicast transmission rate is also high. For realtime traffics like User Data Protocol (UDP)-based streaming, in-time delivery is more crucial than providing hard core guarantee of the correct packet arrival. It is more acceptable to allow few erroneous packets than waiting for retransmission as long as the target level of user satisfaction is achieved. Our P2P architecture offers multiple-peer selection capability. In case any assisted peer missed few pieces of video chunks, it simply finds an alternate Internet peer to download the missing chunks typically at a lower rate without affecting other peers in the cluster.

IV. PERFORMANCE EVALUATION

We perform extensive simulations to evaluate the performance of our proposed collaborative P2P streaming system. We developed a simulator using $C{+}{+}$ programming language that realistically modeled the LTE system [11]. We run event-driven program on top this simulator.

Topology: We implement the streaming system on several different networks that include peers from the Internet and also from the cellular networks. Peers join the system with the Poisson arrival rate (λ). Peers stay in the system unless the streaming ends or the peers fail. The physical links between Internet peers are generated by using a stochastic loss model. The available bandwidth is set randomly in the range $[0.75R_0, 1.25R_0]$, where R_0 is the base bandwidth. We also set the maximum bandwidth of the cellular peers to $0.1R_0$, 10 times smaller than the wired network. We set $R_0 = 300$ $kbps$, which is realistic as sources are tapped while uploading. The simulator in [11] captures the time varying capacity of the wireless channel. There are 7-cells in 3 sectors hexagonal layout.

Workload: Each peer participates to download a 97.5 MB video, that is divided into 390 chunks of size 256 KB.

Parameter setting: We choose $\alpha = 5$ seconds which is sufficiently larger than round-trip-time (RTT). α does not have any impact on the system performance as long as it remains larger than RTT. The values of β, and δ are 4, and 5 seconds respectively. The run time of the simulation is 6 minutes excluding the initial setup time.

First we compare the proposed distributed algorithm of agent selection with the centralized solution. Then we measure the cellular bandwidth saving and congestion avoidance performance of the proposed P2P streaming system.

A. Performance of the distributed agent selection algorithm

To evaluate the performance of the proposed distributed agent selection algorithm, we monitor the contents received at the cellular peers, and we repeat the procedure for varying number of the cellular peers. Our distributed algorithm is solvable for any size of the network. Finding optimal solution is NP-hard, and such a global formulation is not realizable in the context of the P2P network. We find the optimal solution by solving ILP formulation with the free optimization software LP_Solve [12] for comparison person. Due to the mismatch between the physical and overlay network, the aggregate content received at the cellular peers is the lowest for the traditional P2P streaming system for all network size. See Fig. 2. All aggregate contents are normalized to the highest aggregate content of all systems. The aggregate content received with distributed solution is comparable with that of the optimal solution. Hence, our proposed distributed agent selection algorithm performs close to the optimal solution.

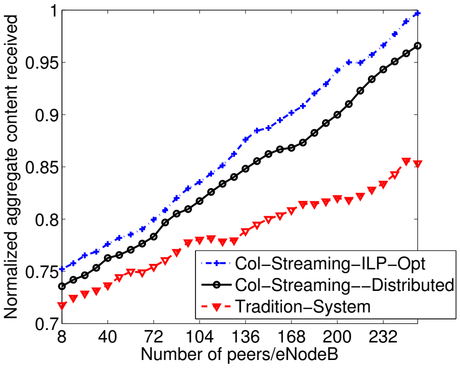

Fig. 2. Performance measurement of the distributed agent selection algorithm with varying number of cellular peers in terms of throughput at the cellular receivers averaged over time.

B. Bandwidth saving and eNB traffic measurement

To investigate the eNB congestion and wireless bandwidth saving, we monitor the traffic through a particular eNB, and also calculate the aggregate throughput at the cellular peers (under the same eNB) averaged over time. As the patterns remain the same, we only present the simulation result for a particular eNB. Figure 3 shows the eNB traffic and aggregate throughput of the receivers averaged over time for the proposed and traditional P2P streaming system. In all cases, streaming contents are normalized to the highest traffic for the comparison purposes. In traditional P2P streaming system, aggregate throughput at the receivers (marked as 'Trad-Scheme-Rec-TP') is always less than the amount traffic through the eNB (marked as 'Trad-Scheme-eNB-Traffic'). This is due to the fact that in traditional P2P system, all the cellular peers download contents through the eNB, and do not share content between them. Being oblivious to network architecture, cellular peers in traditional system, often make peering decision with other cellular peers through the eNB and accrues two-way wireless bandwidth. Whereas in our proposed P2P system streaming, the eNB traffic (marked as 'Prop-Scheme-eNB-Traffic') is less than the aggregate throughput at the receivers (marked as 'Prop-Scheme-Rec-TP'). With increase number of cellular peers, more and more peers join the D2D multicast application and offload the traffic through the eNB. For an example, with 128 cellular peers at the eNB, the eNB traffic for the proposed P2P streaming system is much less than that of tradition P2P streaming system. Hence, the proposed P2P streaming system offloads eNB traffic by taking the advantage of D2D multicasting technique.

C. Streaming cost and fairness issue

The streaming cost is defined by the amount of contents downloaded through the eNB by the cellular peers. The cost is per user averaged over time, and then normalized to the highest cost of all systems for the comparison purpose. We

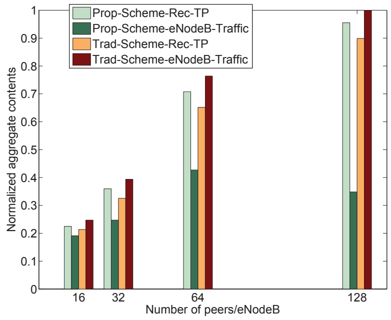

Fig. 3. The normalized eNB traffic and aggregate throughput at the cellular receivers averaged over time.

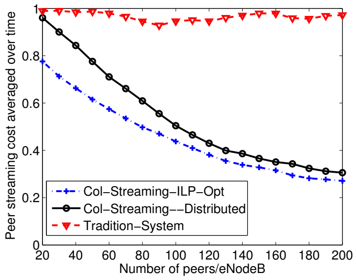

Fig. 4. Per peer streaming cost averaged over time with varying number of cellular peers.

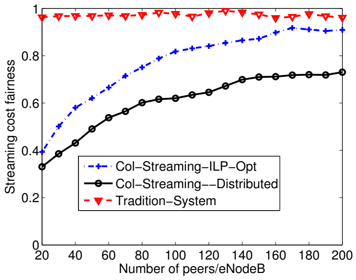

Fig. 5. Comparison of streaming cost fairness for different P2P systems.

do count the streaming cost for D2D multicast; as the eNB allocates shared resources for the D2D multicast application, and cellular mode operation by some interference avoidance technique [11]. Figure 4 depicts the per peer streaming cost averaged over time with varying number of cellular peers. The streaming cost is the highest for the traditional P2P streaming system for all network size. The per peer streaming cost is high for both distributed and optimal algorithm, however decreases with the increase of the network size. This is due to the fact that with the increase size of the network, a single broadcaster is able to share streaming with more cellular peers. The per peer streaming cost is always comparable to that of optimal solution for all network size. Hence the proposed streaming system effectively reduces the streaming cost.

We measure the cost fairness using Jain's fairness [13] index between 0 and 1. The larger the value of index, the more the fairness is; and the cost is distributed among the cellular peers. Figure 5 illustrates the fairness of the streaming cost for different P2P systems. Although traditional P2P system shows the highest cost fairness, that aspect is meaningless as peers do not share streams in such a system. The variation is due to the difference in number missing chunks among the peers. Streaming cost fairness is lower for both distributed and optimal solution, however increases with network size. As more and more peers share the streaming cost with increase number of network size, the fairness of streaming cost increases. Streaming cost fairness for the distributed algorithm closely follows that of the optimal solution.

V. CONCLUSION

In this work, we have designed P2P streaming multimedia distribution system for cellular system where peers with close proximity form clusters and communicate directly. We have proposed and implemented a distributed algorithm to select appropriate number of broadcasters with limited number of assisted peers to ensure the quality of the streaming content. We have also presented the details of the content dissemination technique that schedules the direct download from the Internet, and the broadcasting session. In the simulation, we have shown that the proposed distributed agent selection algorithm has performed close to the optimal solution. The proposed system has offloaded eNB traffic, and saved cellular channels by taking the advantage of D2D multicasting technique. The proposed streaming system has effectively reduced more and more streaming cost, and showed better and better cost fairness with the increase of the network size. Although in our simulation, we have scheduled few peers leaving the system and few others join the system at random time, we have not measured the degree of fault tolerance. Measuring the degree of fault tolerance even with concurrent join/failure is our ongoing task.

REFERENCES

[1] Cisco, "Visual networking index: Forecast and methodology: Global mobile data traffic forecast update, 2011-2016," *White Paper*, 2012.

[2] C.-M. Huang, T.-H. Hsu, and M.-F. Hsu, "Network-aware P2P file sharing over the wireless mobile networks," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 1, pp. 204–210, Jan. 2007.

[3] M.-F. Leung and S. H. G. Chan, "Broadcast-based peer-to-peer collaborative video streaming among mobiles," *IEEE Trans. Broadcast.*, vol. 53, no. 1, pp. 350–361, march 2007.

[4] K. Piamrat, A. Ksentini, J. M. Bonnin, and C. Viho, "Q-DRAM: QoE-based dynamic rate adaptation mechanism for multicast in wireless networks," in *IEEE GLOBECOM*, 2009, pp. 1–6.

[5] K. Doppler, M. Rinne, C. Wijting, C. Ribeiro, and K. Hugl, "Device-to-device communication as an underlay to LTE-advanced networks," *IEEE Commun. Mag.*, vol. 47, no. 12, pp. 42–49, Dec. 2009.

[6] G. Chen, F. G. Nocetti, J. S. Gonzalez, and I. Stojmenovic, "Connectivity based k-hop clustering in wireless networks," in *Proc. Annual Hawaii International Conference on System Sciences*, 2002, p. 2450.

[7] P. Basu, N. Khan, and T. D. C. Little, "A mobility based metric for clustering in mobile ad hoc networks," in *International Conference Distributed Computing Systems Workshop*, apr 2001, pp. 413–418.

[8] M. Baïou and M. Balinski, "Many-to-many matching: stable polyandrous polygamy (or polygamous polyandry)," *Discrete Appl. Math.*, vol. 101, no. 1-3, pp. 1–12, April 2000.

[9] K. Doppler, C. B. Ribeiro, and J. Kneckt, "Advances in D2D communications: Energy efficient service and device discovery radio," in *International Conference on Wireless Communication, Vehicular Technology, Information Theory and Aerospace Electronic Systems Technology (Wireless VITAE)*, march 2011, pp. 1–6.

[10] M. Zulhasnine, C. Huang, and A. Srinivasan, "Topology-aware integration of cellular users into the P2P system," in *IEEE Vehicular Technology Conference Fall (VTC) Fall*, 2011, pp. 1–5.

[11] M. Zulhasnine, C. Huang, and A. Srinivasan, "Efficient resource allocation for device-to-device communication underlaying LTE network," in *IEEE International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, 2010.

[12] "Free linear (integer) programming solver—lp_solve (version 5.5)," 2010. [Online]. Available: http://lpsolve.sourceforge.net/

[13] R. Jain, D.-M. Chiu, and W. Hawe, "A quantitative measure of fairness and discrimination for resource allocation in shared computer systems," Tech. Rep. DEC-TR-301, 1984.