Crowd Powered Media Delivery: Facilitating Ubiquitous Device-To-Device File Transfers

Colin Scott
Microsoft Research India

ABSTRACT

Device-to-device file transfers are pervasive in many emerging markets, but users typically only share content with close friends or informal media vendors. We seek to facilitate ubiquitous device-to-device file transfers beyond one's immediate social network. This extended abstract outlines the research challenges we need to address and our initial plans for addressing them.

1. INTRODUCTION

It is widely acknowledged that demand for Internet content exceeds the available supply of bandwidth in emerging markets. Internet companies have launched numerous ventures to address this shortage, ranging from fiber installations [9] to low Earth orbiting satellites [13], solar powered drones [8], and high altitude balloons [10].

These efforts largely ignore the capabilities of potential Internet users. While people in emerging markets await the arrival of affordable Internet access, they have found ways to make use of the high bandwidth connections that are already available to them: device-to-device file sharing over Bluetooth, USB sticks, and WiFi is pervasive, and informal economies have formed around device-to-device transfer of media content [17, 28, 37]. WiFi file transfer applications such as Zapya [16] and SHAREit [11]—the most popular means of device-to-device file transfer among smartphone users—claim to have more than 400 million registered users.

Despite its widespread popularity, device-to-device transfers are conducted in an ad hoc manner. Users find available content mostly through word of mouth, and share content primarily with friends, not people they do not know. In contrast to Internet content providers like Netflix, transfer applications do not provide recommendation systems for discovery of nearby content. Files are not checked for malware or integrity, and the spread of content is not tracked.

We argue that device-to-device transfers should be supplemented with *services*. Based on initial interviews with SHAREit users in India, we believe that there is widespread, latent desire for content that is not available from immediate friend groups: users would like to discover and transfer content within their geographic vicinity, and may even be willing to pay others to transport content to them. We seek to facilitate this 'crowd powered' delivery of media, by providing services such as a geo-tagged content directory and a remuneration system for transportation of content.

Realizing our goal requires us to address several research challenges. Are users willing to travel physical distances to share content, and if so, what is the most appropriate remuneration model? How should metadata be distributed to devices in offline and poorly connected areas? What is the most understandable interface for naming and authenticating content? Can we devise a sustainable business model? Can we arrange an agreeable relationship between copyright holders and consumers of (often pirated) content?

In this extended abstract we outline our agenda for facilitating crowd powered media delivery. We hope to refine our agenda with feedback from the ACM DEV community.

2. RELATED WORK

We start by discussing how our agenda relates to existing literature. We categorize related work according to subject.

Device-to-device transfers. The observation that physical data transfer can easily exceed the throughput of speed-of-light transfer is well known [12]. Physical data transfer has also been proposed for use in areas with limited or no network connectivity, in the context of delay tolerant networks (DTNs) [20, 21, 30] and sensor nets [29]. Subsequent work studied peer-to-peer data dissemination within ad-hoc mobile networks [18, 27, 31]. We restrict ourselves to file transfers (not interactive requests) across a single hop, and our focus is on lowering the human barriers to use.

More broadly, physical transportation (e.g., delivery of packages) via 'the crowd' has been studied in the past [32]. We contend that the economics of media delivery differ substantially from delivery of individual packages.

Resource constrained networking. Other work has focused on improving the usage of resource constrained networks, e.g., by keeping local content close to users [22], efficiently multiplexing or improving the reliability of uplinks [33, 36], providing appropriate peer-to-peer incentives [24], designing resilient networks [35], and providing content distribution hubs [26]. We also hope to make better use of bandwidth, by facilitating the already widespread usage of device-to-device file transfers.

Media sharing practices. Several studies have examined existing practices for obtaining and sharing media con-

ACM DEV '16 November 17-22, 2016, Nairobi, Kenya

ACM ISBN 978-1-4503-4649-8/16/11.

DOI: http://dx.doi.org/10.1145/3001913.3006640

tent [25, 28, 34, 37]. A key finding of these studies is that desire for media content (often obtained through device-to-device transfers) is widespread. We propose an intervention designed to amplify this desire.

Designing for resource constrained environments. Several studies have provided new interaction models designed for offline or resource constrained networks [19, 38]. These do not focus specifically on offline file sharing.

Products. Existing products implement device-to-device file transfers [11, 14–16], facilitate interactions with nearby strangers [2,4,7], and provide incentives for users to consume data [5, 6]. We are unaware of products that bridge these use-cases to facilitate crowd powered media delivery.

3. RESEARCH CHALLENGES

In this section we outline the research challenges we seek to address and our initial plans for addressing them. Except where noted, we focus our discussion on smartphone owners rather than feature phone owners, on the assumption that smartphone prices will drop to within the reach of price sensitive populations in the near future.

The core question we ask is the following:

> *Can we enable ubiquitous device-to-device media delivery beyond existing social networks?*

Device-to-device file transfers are already popular [28], but users currently only transfer files with friends or informal media vendors. The popularity of torrent [1] has demonstrated that many people are willing to share files (via wired network connections) with 'the crowd'. Here, we focus on enabling crowd powered device-to-device file sharing.

We have started our investigation by conducting an informal, preliminary survey of 41 CGNet Swara workshop attendees in Naya Raipur, India, as well as informal interviews with college students in Bangalore, India. The CGNet Swara workshop attendees are between 18 and 50 years old (most in their 30s), and live in rural areas in Chhattisgarh, Odisha, Jharkhand, Bihar and Madhya Pradesh. They are primarily farmers, with a few social activists, folk singers, and students among the attendees. Very few of the attendees make more than 10,000 rupees ($150 usd) per month. The Bangalore interviewees are all college students in their 20s who study in downtown Bangalore.

The results of our preliminary survey,[1] which we show in Table 1, are striking. Every workshop attendee said they would be comfortable sharing files with strangers, everyone we asked could name specific media content that they would like but cannot currently access, and all but three respondent would be willing to pay or be paid to deliver media across physical distances. So it seems that there is a latent desire for media content that is not immediately available, and a willingness to expend effort to obtain it.

So why are people not already transferring content with people outside their friend groups? There are several possible barriers to sharing more broadly, which we hope to lower through technological and monetary interventions.

Lack of information. The most obvious barrier is limited awareness of what files are available on other devices within a geographic vicinity. The potential benefit of obtaining desired media does not obviously outweigh the transaction cost of verbally asking people what content they have.

[1]Our survey results were gathered over four separate workshops, with a cumulative total of 41 attendees. Not all questions were asked in all workshops.

Question (shortened)	# Yes
Do you use SHAREit, Xender, Zapya?	15 / 41
Comfortable sharing files with strangers?	41 / 41
Comfortable sharing location with strangers?	36 / 41
Willing to pay to have media delivered?	41 / 41
Would you travel 5-7km for a mobile top-up?	38 / 41
Do you buy media from media vendors?	15 / 20
Any content you want, but can't find?	20 / 20
Would you listen to ads for mobile top-up?	18 / 20
Are you within cell range most of the time?	6 / 9

Table 1: CGNet Swara workshop survey responses. Surveys were conducted in four batches, and some questions were added between batches.

Here we propose a *location aware directory service*, where users can post (i) their location, (ii) their transfer method of choice, and (iii) what files (music, movies, pictures, apps) they are willing to share with others. Once a user discovers a file they want within their proximity, they can receive the file from that person. Transfers can take place without needing to coordinate or speak with the other person; a simple notification on the sender's device should suffice for coordination, since WiFi range extends as far as 80 meters.

In our initial prototype, we hope to avoid reimplementing file transfer functionality (already provided by SHAREit, Zapya) by allowing users to complete transfers using existing file transfer applications. In the longer term we plan to keep all transactions within a single application.

Designing such a directory service is non-trivial. We are considering two main use cases. To facilitate serendipitous content discovery between people already within WiFi range of each other (e.g., on a train, at a bus stop), we plan to provide a discovery protocol that automatically syncs metadata about available files whenever users turn on the directory service application.

For content discovery outside WiFi range, we need to consider whether users can receive cellular data. For urban users with cellular connectivity, it should be viable to maintain directory contents in a centralized database and periodically push (or pull) directory updates to all users within each geographic vicinity. These updates would consist of small computed deltas between geographic locations or available content, and should not incur significant data costs.

In rural areas, many users only infrequently come within cellular range. Here, the directory could store stable pieces of information, such as locations where users spend their time during known windows of time throughout the day, or information about what genres users prefer. Specific, more ephemeral directory contents could be transferred in the same way as the peer-to-peer content discovery protocol.

Transaction costs. Ideally, transactions should be mutually beneficial trades of content for both users. However, we would also like to encourage asymmetric transactions. Transactions might be asymmetric when one user has desirable content but the other does not; in these cases the sender has to incur some battery drainage without receiving anything in return. Another form of asymmetry would arise if one user has to travel a physical distance and coordinate a time and location to complete the transfer.

To enable asymmetric transactions we plan to provide a remuneration system. When asked about asymmetric transfers, the college students we interviewed mentioned that they would be willing to pay a small amount (usually not more than ˜$2 usd, which is roughly how much media vendors in peri-urban Bangalore currently charge [28]) to have someone bring them content they want. Such payments could be made through cash exchanges between users, mobile money applications, or a form of virtual currency that could eventually be cashed out.

Advertisers and Governments issuing public service announcements may also be willing to pay users for transferring certain types of content. We discuss this model more under 'legal challenges' below.

Trust. There are also extreme cases of transaction costs. To account for 'flaky' users who consistently fail to coordinate a time/location for transfer, we plan to provide a user reputation system. To protect users from malware obtained through file transfers, we have an opportunity to improve on existing file transfer applications: unlike transfer applications which run only on the device itself, a centralized database would have ample computational resources to check content hashes (uploaded by clients) against known malware blacklists.

UI issues. One remaining barrier may be the difficulty of naming content, e.g., for many users, URLs or content hashes will be inappropriate. For music and movies, we are considering providing a small icon of marketing material (album covers, movie posters) as the content's name.

Technical solutions to the challenges outlined above are within reach. So why have software developers not already provided these services? A few possible barriers remain:

Legal challenges. Media distribution technologies have a history of legal disputes with content providers. We hope to arrange an agreeable relationship between copyright holders, consumers, and intermediaries, and we believe that emerging markets present unique opportunities for arranging these business relationships.

After we submitted this extended abstract for review, YouTube announced a feature for transferring videos between devices in emerging markets [15]. YouTube's entrance into offline sharing represents strong validation of the feasibility of agreeable business relationships. We outline some possible business relationships below.

Existing Business Models. In the Internet's curated media business model, law-abiding media consumers pay twice: they pay their ISP for data connectivity, and they pay the content provider (e.g. Netflix) for the media content itself. The content provider pays for hosting costs. We depict these relationships in the left hand side of Figure 1.

It would be challenging to effectively implement this model for low-income populations in emerging markets such as India. Indian media consumers widely misunderstand piracy laws, and may already believe that they have 'paid' for the media by virtue of paying for cellular data [28]. Indeed, piracy is pervasive in India [23]. Convincing low-income Indians to change their behaviors and explicitly pay for media would be difficult, especially since consumers' demand becomes highly elastic once they are accustomed to free content [3].

Alternative Models. One crucial feature of our proposed directory service is that we remove the intrinsic costs of Internet connectivity. We argue that this enables new models for legality and business sustainability.

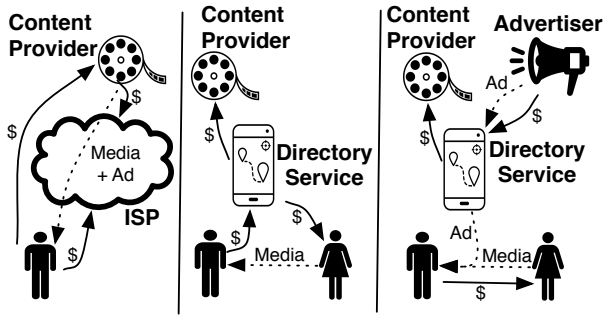

Figure 1: Alternate business models. The left side depicts the Internet's curated media model, where lawful consumers pay ISPs and content providers separately (bold lines) and content providers disseminate content over the network (dotted line). The middle depicts an offline curated model, where the directory service redirects a fraction of the price that would otherwise be spent on bandwidth towards content providers and peers. The right hand side depicts an advertisement-driven model, where copyright holders are entitled to a share of advertisement revenue, and peers pay each other directly for asymmetric transfers.

We depict one possible business model in the middle of Figure 1. This represents an offline curated model, where the directory service only indexes curated content, users pay the directory service upon completing a transfer with a peer,[2] the directory service pays a royalty to the content provider, and the directory service reimburses peers for asymmetric transfers. Because we are eliminating the cost of bandwidth, we hope that the consumer will be able to pay a fraction of what they would otherwise have had to pay for data.

We depict another possible model on the right hand side of Figure 1. The directory service injects advertisements into transferred content, and receives payments from advertisers per view. Peers pay each other directly for asymmetric transfers. In this model, content is not necessarily curated. Similar to YouTube's business model, we could give copyright holders the opportunity to either revoke content from the directory, or take a share in advertisement revenues.

Regardless of the specifics, the role of the directory service is to strengthen the position of both media consumers—by enabling transfers that would be too costly without information from the directory—and content providers—by providing a platform for tracking number of views and extracting revenue. Content providers currently do not derive much revenue from low-income populations, and have little recourse for copyright violations. Our platform would give them an means to serve low-quality versions of their content differentially priced towards a population that they are not currently able to sell to.

[2]To prevent users from subverting the revenue model by transferring content out-of-band, we would ensure that all transfers take place within a single monolithic application. We could additionally apply DRM technologies which are present in some phones.

Arranging these business relationships may take significant time. In the meantime, we plan to configure our directory service to only index a list of whitelisted content. In addition to known freeware content, India in particular has another source of legally permissible yet desirable content: Bollywood producers regularly release free music (videos) from their movies as an advertisement for the movie. Many of these music videos are viewed tens of millions of times after being released by the producer on YouTube.

If feasible, we plan to build this whitelist by checking the filenames of mp4 files, or checking URLs embedded in the mp4 metadata. We could have clients compute content watermarks and check these watermarks against a database of whitelisted content.

4. EVALUATION PLAN

We plan to target three initial populations: college students (especially at colleges without high speed Internet on campus) since they are price sensitive and generally have high technical literacy; cab drivers, who are also price sensitive, have great desire for 'timepass', and most interestingly, travel long physical distances and congregate together at places such as airport taxi stands; and CGNet Swara citizen listeners and reporters in rural areas. To encourage users to download our application, we are considering either simply paying users for the initial download, or bundling the application with desirable media content.

The key metrics we seek to understand are: the extent to which people share files compared to those who do not have access to our interventions; what kinds of new interactions or behaviors our interventions enable; whether directory updates incur significant data or latency costs; and whether the business model can be made sustainable.

Acknowledgements. We thank Bill Thies and Abhay Agarwal for helping shape these ideas, and Rachit Sharma for conducting the CGNet Swara surveys.

5. REFERENCES

[1] BitTorrent Website. http://www.bittorrent.com.
[2] BounceChat Website. http://info.bouncechat.com/.
[3] EconTalk: Chris Anderson on Free. http://www.econtalk.org/archives/2008/05/chris_anderson_1.html.
[4] FireChat Website. https://opengarden.com/firechat.
[5] Jana Website. https://www.jana.com/.
[6] Movivo Website. http://www.movivo.com/.
[7] Nearby Live Website. https://www.wnmlive.com/.
[8] Project Aquila. https://www.facebook.com/notes/mark-zuckerberg/the-technology-behind-aquila/10153916136506634/.
[9] Project Link. https://www.google.com/get/projectlink/.
[10] Project Loon. http://www.solveforx.com/loon/.
[11] SHAREit Website. http://shareit.lenovo.com/.
[12] Sneakernet Wikipedia Page. https://en.wikipedia.org/wiki/Sneakernet.
[13] SpaceX satellite development facility. https://en.wikipedia.org/wiki/SpaceX_satellite_development_facility.
[14] Xender Website. http://web.xender.com/.
[15] Youtube Go Announcement. https://youtube.googleblog.com/2016/09/youtube-go-youtube-reimagined-for-next.html.
[16] Zapya Website. http://www.izapya.com/.
[17] BBC. Cuban Internet Delivered Weekly by Hand. http://www.bbc.com/news/technology-33816655.
[18] A. S. Cacciapuoti, M. Caleffi, and L. Paura. Mobile P2P: Peer-to-Peer Systems Over Delay Tolerant Networks. CRC '10.
[19] J. Chen, S. Amershi, A. Dhananjay, and L. Subramanian. Comparing Web Interaction Models in Developing Regions. DEV '10.
[20] K. Fall. A Delay-Tolerant Network Architecture for Challenged Internets. SIGCOMM '03.
[21] S. Guo, M. H. Falaki, E. A. Oliver, S. Ur Rahman, A. Seth, M. A. Zaharia, and S. Keshav. Very Low-Cost Internet Access Using KioskNet. SIGCOMM CCR '08.
[22] D. L. Johnson, E. M. Belding, and C. Mudenda. Kwaabana: File Sharing for Rural Networks. DEV '13.
[23] J. Karaganis. *Media Piracy in Emerging Economies*. SSRC, 2011.
[24] U. W. Khan and U. Saif. BitTorrent for the Less Privileged. HotNets '11.
[25] N. Kumar and N. Rangaswamy. The Mobile Media Actor-Network in Urban India. CHI '13.
[26] J. LeBrun and C.-N. Chuah. Bluetooth Content Distribution Stations on Public Transit. Workshop on Decentralized Resource Sharing in Mobile Computing and Networking, '06.
[27] Y. Li, T. Wu, P. Hui, D. Jin, and S. Chen. Social-Aware D2D Communications: Qualitative Insights and Quantitative Analysis. IEEE Communications Magazine '14.
[28] J. O'Neill, K. Toyama, J. Chen, B. Tate, and A. Siddique. The Increasing Sophistication of Mobile Media Sharing in Lower-Middle-Class Bangalore. ICTD '16.
[29] U. Park and J. Heidemann. Data Muling with Mobile Phones for Sensornets. ENSS '11.
[30] A. Pentland, R. Fletcher, and A. Hasson. DakNet: Rethinking Connectivity in Developing Nations. IEEE Computer Magazine '04.
[31] T. Repantis and V. Kalogeraki. Mobile Peer-to-Peer Data Dissemination. MDM '05.
[32] A. Sadilek, J. Krumm, and E. Horvitz. Crowdphysics: Planned and Opportunistic Crowdsourcing for Physical Tasks. ICWSM '13.
[33] U. Saif, A. L. Chudhary, S. Butt, N. F. Butt, and G. Murtaza. A Peer-to-Peer Internet for the Developing World. ITID '09.
[34] N. Sambasivan and T. Smyth. The Human Infrastructure of ICTD. ICTD '10.
[35] P. Schmitt and E. Belding. Navigating Connectivity in Reduced Infrastructure Environments. Workshop on Computing within Limits, '13.
[36] P. Schmitt, R. Raghavendra, and E. Belding. Internet Media Upload Caching for Poorly-Connected Regions. DEV '15.
[37] T. N. Smyth, S. Kumar, I. Medhi, and K. Toyama. Where There's a Will There's a Way: Mobile Media Sharing in Urban India. CHI '10.
[38] S. P. Wyche and L. L. Murphy. Dead China-Make Phones Off the Grid: Investigating and Designing for Mobile Phone Use in Rural Africa. DIS '12.