ABSTRACT

In this paper we enhance the Digital Content Mediator (DCM) approach, a legitimate online service that uses financial incentives as an effective weapon to fight against online piracy. We provide needed network security support for the DCM service.

The DCM mediator is a trusted notary to ensure fair and legitimate deals between digital content selling peers and buying peers. (1) In our design, the mediator sees no raw bits of digital contents. This saves storage and communication resource for the central mediator. (2) For the seller and buyer in a DCM-legitimized transaction, one wants payment and the other wants the content. The DCM protocol ensures that neither of them can stop the protocol in the middle to steal its service without serving the other party. (3) A digital content may have many legitimate copies from large amount of sellers. In a large-scale random network like the Internet, transaction fairness is defined as the condition that a buyer wants to buy the copy from the seller with shortest downloading delay (i.e., largest seller-to-buyer pairwise bandwidth) given the same amount of financial charge. DCM employs flow network security countermeasures to ensure that a seller keeps its bandwidth promises. Our experiments on the Internet confirm the effectiveness of our design.

1. INTRODUCTION

Although a recent innovation, peer-to-peer (P2P) digital content sharing systems are now one of the most popular Internet applications and have become a major source of Internet traffic. As the recent U.S. Supreme Court ruling (MGM vs. Grokster) indicates, it is important that these systems be used in a legitimate way. Unfortunately, initial designs for P2P systems embody approaches to security and information legitimacy that have lead to significant legal concerns. For example, the original Napster had been used as a major piracy resource and was ordered to shut down.

Despite these concerns, peer-to-peer systems have several technical advantages over centralized systems, for example, (1) scalability that automatically expands to meet the demand/load; (2) robustness against individual node failures; (3) fully-distributed content storage model without hotspots; (4) variety of data repositories which mean it’s difficult to "lose" data due to replicated copy; and (5) multiple sources that can deliver component parts of a content element to a single destination.

With these concerns in mind the digital content mediator (DCM) was created [18] and extended to exploit the advantages of peer-to-peer networks for legitimate digital content distribution purposes. Currently, the DCM design has following features:

1. Scalability and robustness: DCM is a proper design for large-scale networks (e.g., the Internet) with fully-distributed content storage. Failures at some individual content storage peers do not affect the transaction system. In particular, as the mediator itself is a single authorized and trusted notary, its network site is vulnerable to network-based denial-of-service attacks (although the mediator’s operating and database systems can be protected by fault-tolerant cluster computing). In our design, the mediator sees no raw bits of digital contents. This not only saves storage and maintenance cost in the mediator’s local systems, and also avoids denial-of-service attacks against the mediator.

2. Transaction fairness and flow network security: Many peers may possess legitimate copies of the same digital content. DCM provides incentives for the peers to sell the stored contents to make legitimate profits, rather than to share the contents in illegitimate manners. Fairness in content delivery must be ensured. Otherwise the unfairly treated storage peers may return to illegitimate distribution. In Definition 1, we define transaction fairness as the condition that a buyer wants to buy a copy from the seller with largest seller-to-buyer pairwise bandwidth (i.e., shortest delay) upon paying the same amount of financial charge. This way, every bps bandwidth is equally treated in all pairwise seller-to-buyer connections. DCM employs flow network security countermeasures to ensure transaction fairness.

3. Atomic service exchange: For the seller and buyer in a DCM-legitimized transaction, one wants payment and the other wants content. Like other transaction systems, the DCM protocol ensures that either neither of them is served or both of them are served. Neither of them can stop the DCM protocol in the middle to steal its service without serving the other party.

The paper is organized as follows. In Section 2 we present the problem statement in details. Section 3 describes related work. The DCM protocol design details are specified in Section 4. We show our implementation and evaluation results in Section 5. And finally Section 6 concludes this paper.

2. PROBLEM STATEMENT

A. Network assumptions

P2P networks are open networks, and peers can join and leave the network of their own free will. In this paper, we also assume that each of them can join and leave with any pseudonym (e.g., email address, self-chosen identity), because this is the Internet
status quo. The effectiveness of our security countermeasure is independent of a peer’s choice of its ID.

We assume that it is possible to know the estimated network distance (measured using hop count, end-to-end latency or some other network metrics) between any pair of peers. In the current Internet, this assumption can be realized by existing schemes like Vivaldi [1].

In this paper, the network distance is defined by the real-time average available bandwidth between any pair of peers. A downloader tries to minimize its content downloading delay and get the needed content as early as possible. For example, if by paying the same amount of money a buyer can acquire the needed content from a seller $A$ in 3 hours and another seller $B$ in 10 hours, the buyer will always pay $A$ rather than $B$ assuming all other factors are equal. The challenge is how to enforce this rule.

B. Roles in a DCM transaction

A high-level overview of the DCM system architecture is shown in Figure 1.

![Figure 1: DCM High-level System Architecture](image)

### Table 1: Roles in a DCM transaction

<table>
<thead>
<tr>
<th>Role</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M$</td>
<td>The DCM mediator</td>
</tr>
<tr>
<td>$O$</td>
<td>The copyright owner of a content (e.g., 20th Century Fox)</td>
</tr>
<tr>
<td>$B$</td>
<td>Buyer (Destination) who wants to buy a digital content</td>
</tr>
<tr>
<td>$S_i$</td>
<td>The $i$-th seller (Source) of the digital content</td>
</tr>
</tbody>
</table>

The DCM follows a mediator approach that is very different from common digital portals. (1) To realize the system only one mediator $M$ is needed (Table 1) although $M$ may be implemented in a distributed manner to achieve better load balancing and network resilience. Even in such a distributed implementation $M$ only plays a single mediator role. (2) For each buyer $B$ who wants to buy a digital content, at least one or more seller(s) $S_i$ are capable of providing the needed digital content (e.g., 9.5 gigabytes of data for DVD9). During a transaction, we assume $M$, $O$, $B$, $\{S_i\}_{i=1}^n$ are fixed roles with unique and verifiable network addresses.

C. Mediator’s role

The mediator $M$ is a trusted notary. It processes digital payments and digital legitimacy proof materials. Compared to the size of the raw digital content, the materials handled by $M$ are short in length. To ensure the practicality of DCM in a large-scale network, $M$ does not handle raw bits of the digital contents in transaction, and indeed it is a key innovation of the DCM approach that $M$ never needs to ‘see’ the transferred content or know of its subject matter.

**DDoS attack against the mediator** Distributed DoS attack against an online site is always possible. Fortunately, many countermeasures against non-compliant DDoS attacks have been proposed and developed [2—6]. These countermeasures should not reduce protocol-compliant traffic towards the protected site (otherwise such a countermeasure itself is denying the service). Therefore, in this paper we only consider attacks compliant to DCM design. Malicious sellers and buyers would try to incur DCM-allowed traffic toward the mediator $M$ to shut down its service. Raw digital contents sent to the mediator are considered as an attack and the responsible sender has to pay the price immediately.

In particular, if we allow $M$ to process the huge digital content itself, then the malicious attackers can try all means to direct the content bits toward $M$ and deplete its network resource. This is one reason why $M$ only processes short-length legitimacy and cryptographic proofs in the DCM approach.

D. Legitimacy and digital content source

The content owner $O$ owns the content’s copyright. The mediator $M$ must enforce legitimate transactions between any buyer $B$ and any seller $S_i$. To invoke the peer node’s incentive to stop piracy, DCM must ensure that:

- The seller $S_i$ who stores the digital content, must be able to make legitimate profit from the transaction, and $B$ gets the requested content with legitimacy proof. Otherwise, as $M$ is no longer indispensable, the two parties $B$ and $S_i$ would simply ignore $M$ and exchange digital content privately in an illegitimate manner (per status quo).

- All nodes that have legitimately acquired the same digital content are equally treated as content sellers. Otherwise, those mistreated nodes will soon translate the biased treatment into piracy.

DCM design assumes that hardcopies of legitimate digital content media (e.g., DVD) can be bought by out-of-band means. A legitimate digital credential can be pre-generated and appended to every such legitimate hardcopy (e.g., in the form of a verifiable license file stored on DVD). The legitimate digital credentials can be efficiently verified. *Anybody who has bought such a legitimate hardcopy can convert it into softcopy and be an online seller once registered with DCM. Anybody who has legitimately bought a softcopy from a preceding seller via DCM can equally be an online seller.* Therefore, after a time lapse, there are potentially many legitimate sellers of the same digital content. The mechanisms by which this ‘seeding’ process occurs are outside of the scope of this paper.
E. Selfish sellers and buyers

Peers seeking profit are selfish. A seller \( S_i \) or a buyer \( B \) will not offer any extra service to the other side once it gets what it wants (payment or content, respectively).

A seller \( S_i \) wants payment. Selfishness implies that \( S_i \) will take \( B \)'s payment as soon as possible, and stop delivering digital content to \( B \). Moreover, \( S_i \) will try all means to sell its own legitimate softcopy to as many buyers as possible by depriving the other competing sellers’ chances.

A buyer \( B \) wants the content. Buyer’s selfishness means that \( B \) will try to acquire a copy of \( S_i \)'s digital content without completing the payment transaction.

F. Transaction fairness

**Definition 1: Anti-piracy fairness factor \( \Psi \).** Assuming other network and social conditions are identical, those nodes with larger bandwidth can (illegally) disseminate more (copyrighted) contents compared to those nodes with smaller bandwidth, thus cause more damage to copyrights. In a large-scale random network of \( N \) nodes, if an anti-piracy countermeasure \( X \) successfully makes a set of \( K \) nodes \( \{x_1, x_2, ..., x_K\} \) follow legitimate means rather than doing piracy, \( X \)'s fairness factor \( \Psi \) is a probabilistic value between 0 and 1, and a larger \( \Psi \) means less piracy.

As previously described, after an initial setup and many legitimate DCM transactions, there may be many legitimate sellers for the same digital content. DCM should enforce fairness amongst many sellers for each purchase transaction. Otherwise, the unfairly treated nodes may return to illegitimate distribution. In this paper, transaction fairness means that DCM will implement a larger \( \Psi \) by granting transaction to the "best" seller \( S_j \) so that the buyer \( B \) enjoys the minimal downloading delay. In other words, every bit-per-second bandwidth is equally respected in all peer-to-peer connections. However, this is an open challenge in DCM design, due to the selfishness attacks described below:

(Fairness attack against a buyer) A selfish seller would try to wrongfully claim that it has a larger average available bandwidth when the network distance measuring scheme (e.g., via Vivaldi [1], Spruce [14]) probes its site. This can be done by temporarily shutting down all other network sessions during the probe, then reviving these bandwidth-consuming sessions later. As a result, the seller can attract more buyers and sell more copies of its digital content. But the chances of other competing sellers are deprived, and the unfortunate buyers have to suffer much longer downloading delay than what is expected.

G. Summary

In summary, in this paper we would provide cryptographic and network security support to properly execute a legitimate P2P transaction even in the presence of malicious and selfish sellers and buyers, who are capable of launching protocol-compliant DDoS attack against the mediator \( M \) and/or breaking the transaction fairness requirement to make unfair profits.

3. RELATED WORK

**Digital Content Mediator and Digital Right Management** Telcordia has previously published work on the development of the Digital Content Mediator infrastructure [18] with particular reference to it’s application in the traditional Telecom operator environment. This environment is particularly relevant to the DCM concept because content never flows through any of the DCM components that come under the responsibility of the operator. Thus, the DCM approach lends itself to the development of a common carrier concept for content analogous to the way in which the voice streams of the regular telephone call have legal protection.

It is evident to the interested reader that there have been many, highly successful, approaches to the problem of Digital Rights Management (DRM), notably by Microsoft [19] and Apple [20] but these do not explicitly support peer-to-peer distribution and have not been widely applied in this space. The DCM approach is compatible with any existing DRM technique provided that the content itself is stored in a constant form. That is to say that, once header and footer information is removed, if the same sequence of bytes represents the content on each users devices then the DCM can be used effectively to track it’s distribution.

**Non-intrusive bandwidth estimation** Bandwidth estimation has been extensively studied, and various techniques have been proposed in the last few years [7][8][9][13]. Among them, the packet-pair based techniques perform much less intrusive than the others. The basic Packet Pair algorithm [10] relies on the fact that if two packets sent back-to-back are queued next to each other at the narrow link, they will exit the link with "dispersion" \( T \) given by:

\[
T = \frac{L}{B},
\]

where \( L \) is the size of the second packet, and \( B \) is the bandwidth of the narrow link.

If the two packets have the same size, their transmission delays are the same. This means that after the narrow link, a dispersion of \( T \) will be maintained between the packets even if faster links are traversed downstream of the narrow link. This is shown in Figure 2.

![Figure 2: Bottle link capacity B = L / T.](image)

The Packet Pair algorithm assumes that the packets will queue next to each other at the narrow link. However, previous researchers [8]have also noted that capacity estimates resulting from packet pair dispersion can be inaccurate due to compression or expansion of dispersion induced by cross traffic. In order to provide accurate capacity estimates, another recently proposed approach, CapProbe [13], combines dispersion and delay measures to filter out packet pair samples “distorted” by cross-traffic.
More specifically, whenever an incorrect value of capacity is estimated, the sum of the delays of the packet pair packets, which we call the delay sum, includes cross-traffic induced queuing delay. We refer to this delay sum, which does not include any cross-traffic queuing delay, as the minimum delay sum. The dispersion of such a packet pair sample is not distorted by cross-traffic and will reflect the correct capacity.

Thus, assuming that at least one packet pair sample goes through without cross-traffic interference, we get a sample that measures the correct capacity and the delay sum of whose packets is equal to the minimum delay sum. This sample can easily be identified since its delay sum will be the minimum among delay sums of all packet pair samples.

It is easy to see that CapProbe is extremely efficient in the use of resources. In fact, CapProbe needs to perform only one addition and one comparison operation per packet pair sample. Compared to some of the previous schemes [7][8][9], this represents an order of magnitude reduction in workload. As a result, the CapProbe algorithm is ideal in this study to provide fast and accurate capacity estimation of the selected end-to-end Internet paths.

4. DESIGN

A. Design assumptions

We assume that end-to-end privacy and digital signature implementations are available to all peers. We further assume that data origin authentication based on digital signature is supported such that each message sending is non-repudiable.

We also assume that a network distance measuring scheme (e.g., Vivaldi [11]) periodically (e.g., every 10 minutes) measures the network topology and the mediator M is aware of the measurement result. Here pairwise distance is defined as the real-time available bandwidth on the bottleneck link between a pair of peer nodes. We assume the mediator M is supported by financial institutions to approve a legitimate transaction and fulfill the payment service.

B. Cryptographic support

For efficiency purposes, the needed digital content D in a transaction is divided into pieces (we will use 10Mbytes/80Mbits per piece in this paper). The buyer can switch seller/source from a piece to another. Let’s use following notions:

<table>
<thead>
<tr>
<th>Table 2: Notions used in a DCM transaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>D</td>
</tr>
<tr>
<td>tag</td>
</tr>
<tr>
<td>Dij</td>
</tr>
<tr>
<td>authi</td>
</tr>
<tr>
<td>enc</td>
</tr>
<tr>
<td>h</td>
</tr>
<tr>
<td>{m}</td>
</tr>
</tbody>
</table>

Recall by our assumption, every message in the following pur-
(Selfish service provisioning in Section 2.E) We assume that the content owner \(O\) will offer the mediator \(M\) the encryption keys \(k_i\) and \(auth_j = henc(k_i, D_j)\), the authenticators corresponding to the keys and the owner’s contents. The previously described cryptographic design enforces that a seller must get paid after it uploads the encrypted content data to the buyer, thus a selfish seller cannot get paid before it delivers. On the other hand, before the buyer sends the payment to \(M\), it merely sees encrypted pseudo-random bits. The buyer must obtain the one-time decryption key to translate the encrypted bits into useful digital contents. Neither the seller nor the buyer can acquire what it wants before giving the other side the needed service. Both sides cannot gain advantage by stopping in the middle of the protocol.

In summary, the cryptographic support prevents protocol-compliant DDoS attack against the mediator. It also ensures that either the selfish buyer or the selfish seller must offer the thing needed by the other side before it acquires what is desired.

C. Flow network security support

Flow network security attacks are feasible at step (f), where the seller \(S\) tries to spend less bandwidth (thus larger uploading delay) than it promised to serve the buyer \(B\). We adopt a network-based countermeasure to cope with the transaction fairness attacks (Section 2.F). More specifically, we employed two accurate and timely network measurement techniques (i.e. AdHoc Probe and Spruce) in the DCM system in order to detect selfish peers and enforce fairness in the network. We briefly recapitulate AdHoc Probe and Spruce as follows.

AdHoc Probe is a recently proposed technique, which is based on CapProbe algorithm [13], designed for effective path capacity estimation in wired and wireless networks. Similar to CapProbe, AdHoc Probe probes the network using back-to-back packet pairs, and it uses the dispersion of one “good” sample, which is obtained by applying the minimum delay sum filter, to estimate the link capacity. However, AdHoc Probe is a one-way technique, and it does not implement the convergence test and probing adaptation of the original CapProbe algorithm in order to further simplify the algorithm. As a result, it is ideal to be integrated into other data transmission protocols and became passive estimation techniques [16][17].

Spruce is a network measurement tool designed for available bandwidth estimation. Different from bottleneck link capacity, the available bandwidth is the residual bandwidth of the tightest link on a path, i.e. it is the maximum achievable data throughput while coexisting with other network flows. Spruce probes the network using packet pairs. However, instead of using back-to-back packet pairs, Spruce sends two packets with an interval, which is equal to the packet length divided by the link capacity. After filtering out extremely congested samples, Spruce estimates the available bandwidth by averaging the result of each sample. This scheme has been proved to be fast and accurate [14], and the simplicity of this algorithm makes it applicable to be easily integrated into other data transmission protocols.

Using AdHoc Probe estimates, we provide Spruce the information of the link capacity, and Spruce is thus able to determine the interval between the sampling packet pairs. Moreover, the bottleneck link capacity (\(C\)) estimate can be used as the upper bound of the data throughput (i.e. assuming there is no traffic on the bottleneck link), and the available bandwidth (\(B\)) estimate can be used as the currently maximum achievable data throughput. As a result, selfish behavior is then possible to be identified if the seller over-claims the bandwidth (i.e. \(B>C\)), or if the seller intentionally slows the data transmission while the available bandwidth is still rich (i.e. data throughput \(< B\)). We present evaluation of these network measurement tools in the next section.

5. EVALUATION

In this section, we present Internet experiment results evaluating feasibilities of using network measurement tools (i.e. AdHoc Probe and Spruce) to detect selfish peers in DCM networks. Four Internet hosts and three Internet paths were selected for the experiments. The selected Internet hosts are listed in Table 3, and the topology and path properties (i.e. link capacity and round trip time) of the selected paths are illustrated in Figure 3. For each selected path, we repeated the experiments of AdHoc Probe and Spruce estimation for 20 runs. Table 4 shows the averaged results.

Table 3: Description of the participating hosts in the experiments

<table>
<thead>
<tr>
<th>Abbrev.</th>
<th>Full Name</th>
<th>Host IP Address</th>
</tr>
</thead>
<tbody>
<tr>
<td>UCLA</td>
<td>University of California at Los Angeles</td>
<td>131.179.136.159</td>
</tr>
<tr>
<td>PITT</td>
<td>University of Pittsburgh</td>
<td>130.49.221.41</td>
</tr>
<tr>
<td>UW</td>
<td>University of Washington</td>
<td>128.208.4.197</td>
</tr>
<tr>
<td>UCSD</td>
<td>University of California at San Diego</td>
<td>132.239.17.226</td>
</tr>
</tbody>
</table>

Figure 3: Topology and path properties of selected Internet paths.

The results confirm that AdHoc Probe is able to accurately measure bottleneck link capacity of an Internet path. Specifically, the accuracy of AdHoc Probe estimates is higher than 95% in all experiments. It should be noted that, though AdHoc Probe spent around 40 seconds for capacity estimation, the measurement speed can be further improved by changing the probing...
rate. However, we do not address this issue in this paper since it is out of the scope of this study.

### Table 4: Experiment results of link capacity and available bandwidth estimates

<table>
<thead>
<tr>
<th>Source</th>
<th>Sink</th>
<th>Capacity (Mbps)</th>
<th>RTT (ms)</th>
<th>AdHoc Probe Estimate (Mbps)</th>
<th>Time (seconds)</th>
<th>Spruce Estimate (Mbps)</th>
<th>Time (seconds)</th>
</tr>
</thead>
<tbody>
<tr>
<td>UCLA</td>
<td>100</td>
<td>98</td>
<td>95.6</td>
<td>43</td>
<td>86.9</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>UW</td>
<td>100</td>
<td>33</td>
<td>97.2</td>
<td>42</td>
<td>91.4</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>UCSD</td>
<td>100</td>
<td>4.5</td>
<td>98.3</td>
<td>42</td>
<td>91.6</td>
<td>11</td>
<td></td>
</tr>
</tbody>
</table>

Additionally, Table 4 shows that the available bandwidth of the selected Internet paths is around 90% of the corresponding link capacity, i.e., the link utilization is around 10%. This is due to the fact that the tight link of the estimated Internet path is on the first/last hop, which is lightly loaded (i.e., less than 10%) in the experiments.

Therefore, through Internet experiments, we have verified that current network measurement tools, such as AdHoc Probe and Spruce, are able to accurately and fast measure link capacity and available bandwidth of a peer-to-peer Internet path. Moreover, the simplicity of these techniques makes them ideal to be integrated into other data transmission protocols to achieve passive network measurement. For instance, the AdHoc Probe concepts have been integrated into TFRC and TCP, and the resulting protocols are called TFRC Probe and TCP Probe.

Given the link capacity and available bandwidth estimates, it soon becomes possible to detect selfish peers in DCM networks. For instance, a selfish seller can be easily identified when the available bandwidth estimates are much lower than the claimed bandwidth (i.e., the seller over-claims the bandwidth), or the cases that the achieved data throughput is much lower than the available bandwidth (i.e., the seller intentionally slows the data transmission).

### 6. CONCLUSIONS

In this paper, we extend the Digital Content Mediator (DCM) [18], a legitimate online service that uses financial incentives as an effective weapon to fight against online piracy. We provide needed network security support for the DCM service.

The DCM design is suitable in a scalable network with distributed storage that is robust against various failures and denial-of-service attacks. The DCM design implements atomic service exchange such that selfish sellers and buyers cannot steal service without serving the other side. With the extension presented in this paper, the DCM design ensures transaction fairness, which depends on how much bandwidth consumed by seller-to-buyer content delivery. This way, every bit-per-second (bps) bandwidth is equally treated in all pairwise seller-to-buyer connections, and the buyer enjoys the least content delivery delay. We devise flow network security supports to ensure transaction fairness. Our experiments on the Internet confirm the effectiveness of our design.