

Low-Latency Parallel Transport in Anonymous Peer-to-Peer Overlays

Igor Margasiński¹ and Michał Pióro^{1,2}

¹ Institute of Telecommunications, Warsaw University of Technology, Poland

² Department of Electro- and Information Technology, Lund University, Sweden
{I.Margasinski,M.Pioro}@tele.pw.edu.pl

Abstract. The paper presents a design and discusses configuration aspects of an overlay transport protocol based on an idea of the peer-to-peer direct and anonymous distribution overlay (P2PRIV). We estimate a secure configuration of the protocol and examine a correlation between the P2PRIV's anonymous path lengths and latency. An increase of the path lengths speaks strongly in favor of the parallel solution's anonymity, as in classical cascade networks. In the paper we evaluate the new protocol in a scope of a trade-off between anonymity and traffic performance and show that the presented solution allows effectively increasing anonymity with relatively low impact on anonymous transport latency.

Keywords: Communication system traffic, communication system security, privacy, peer-to-peer overlays, overlay networks management.

1 Introduction

Nowadays anonymous networks impose a trade-off between anonymity and traffic performance. A high level of anonymity can be achieved primarily at a high traffic expense. In particular, the latency constitutes a crucial factor for performance of anonymous networks, as the basic common mechanism used to achieve network anonymization is a traffic forwarding by a set of middleman nodes. Long anonymous path lengths usually speak strongly in favor of anonymity. However, the traffic performance is reduced and a download time elongated by increasing path lengths of the content's transport. Each of forwarding nodes introduces a delay imposed by specific anonymization techniques applied (such as batching and multiple asymmetric encryptions in *Mix-nets* [3]) and bandwidth limitations of routing links. Furthermore, when we consider the environment of peer-to-peer overlays, insufficient throughputs links become an onerous issue while the content relaying by personal or *soho* computers deepen latency of the anonymous transport.

On March 2008, a new anonymous peer-to-peer architecture was proposed by the authors in [12]. In the new solution called P2PRIV (peer-to-peer direct and anonymous distribution overlay) only control messages are sent over anonymous paths, called "cloning cascades" (CC). The P2PRIV uses the well known anonymous techniques for anonymization of a specific management communications adjusted to

provide further anonymous and direct parallel transport of the shared information content. In this paper we continue the research and present design details of the P2PRIV protocol, and analyze an impact of the protocol configuration, i.e. CC path lengths, on the solution’s anonymity and latency. We study the trade-off between anonymity and performance and show that the new solution tolerates an increase of anonymous path lengths with a significant advance in anonymity and with a significantly lower impact on traffic performance.

2 Related Work

A basic common mechanism used to achieve P2P anonymization is the traffic forwarding by a set of middleman network nodes. These peers anonymize a traffic in accordance with various anonymous techniques ranging from heuristics with encryption methods (applied in *Freenet* [4], *Gnunet* [1]) to highly anonymous *mixing* [3] (e.g. in *Free Haven* [7], *Tarzan* [9], *MorphMix* [14]). Mix-net is a network of many intermediate nodes called *Mixes*. Anonymous messages are routed throughout Mixes which aggregate, permute randomly, and encrypts received messages with public keys of successive nodes. The identity of both a sender and a receiver is never disclosed to any single proxy Mix, and due to specific mixing of asymmetrically encrypted packets an attack based on traffic analysis is unlikely to succeed. Still, highly secure mixing networks impose high latency. Various low-latency – yet less secure – techniques, such as so called *virtual link encryption* (*PipeNet* [5]), *fixed shared routes/cascades* (*JAP* [2]), and *onion-routing* multiple public key encryption (applied in *Onion Routing* [10], [13], and widely used *TOR* [8], [16], [11]), were introduced. Another, simple low-latency technique as involving only a symmetric cryptography, was introduced for *CROWDS* system [15]. *CROWDS* anonymizes Web traffic and utilizes an idea of *random walk* algorithm by traffic forwarding via a random group of nodes before its delivery. *CROWDS* member, who wishes to send an anonymous message, selects a random node of the *CROWDS* network (the so called “Jondo”) and then sends the node the message. The message contains a destination address, but the source address is neglected. Then, the selected node flips an asymmetric coin to decide whether to forward the message to the next random node or send it directly do the destination (random walk). All random nodes repeat the same activity and finally the message is sent to its destination (based on the included address). The decision whether to forward the message to a next proxy node or to the message’s receiver is random. However, commonly the selection of a proxy is more probable than directing the message to the destination. This probabilistic forwarding assures anonymity, because none of network nodes can ascertain the message’s origin. The coin asymmetry is described by a probability p_f . The proxy node forwards the encrypted message to the next random proxy node with the probability p_f and sends it to the destination with a probability $1 - p_f$. Then the mean forwarding path length of network random walk equals

$$P = \sum_{i=2}^{\infty} i p_f^{i-2} (1 - p_f) = \frac{p_f - 2}{p_f - 1} . \quad (1)$$

The P2PRIV system omits the canon of information content forwarding. In P2PRIV only a token is sent over random walk path, called “cloning cascade” (Step 1). The traffic generated by token relaying can be additionally anonymized by Mix-net, as numerous and short messages can be effectively exchanged by the Mix cascades. Then, based on information included in the token and after a random delay, the selected (“cloned”) nodes download content data specified by an initiator directly (Step 2). The novel idea of the P2PRIV consists in the parallel content transport separating the anonymization process from the transport function (Figure 1).

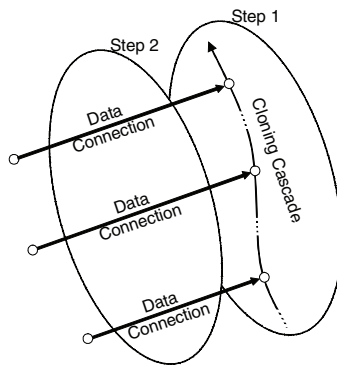


Fig. 1. P2PRIV parallel transport of content

3 Parallel Overlay Transport

Let us describe the protocol design from the point of view of a “cloned peer”. Notice that the P2PRIV is a pure P2P overlay, hence all P2PRIV peers are symmetric and represent the same functionality. Figure 2 shows a finite state machine diagram of a single peer. The basic states of the peer are: “LISTEN”, “CLONED”, and “DATA CONNECT”. LISTEN represents an idle operation of a peer waiting for requests (marked on the diagram as “recv:”) from other peers or from the user (marked with “user:” symbol). CLONED state is triggered when the peer joins a Cloning Cascade (compare with Step 1 of P2PRIV operation). DATA CONNECT corresponds to the Data Connection step in which P2PRIV node downloads a specified content (Step 2 of P2PRIV operation). The diagram contains also: “COIN FLIP” and “LOOKUP” states. COIN FLIP applies to a decision process based on a binary random selection and LOOKUP initiates a DHT lookup procedure. The P2PRIV communications is based on five messages: “CLONE”, “FIND”, “FOUND”, “GET” and “PUT”. The following pseudo-code includes a description of their detail roles (Figure 3, 4, 5, and 6). and is organized as a one infinite loop with three main sections. The first section (starting from the Line 3, Figure 3) refers to a CC joining process and a further content download. When a peer receives a CLONE message (token) containing a file id it will check whether the message comes from other peer or from a local user. Then it starts a “Clone and Download” subroutine presented in Figure 4. The next section (starting from the Line 8, Figure 3) refers to a reaction of a peer to a content look-up

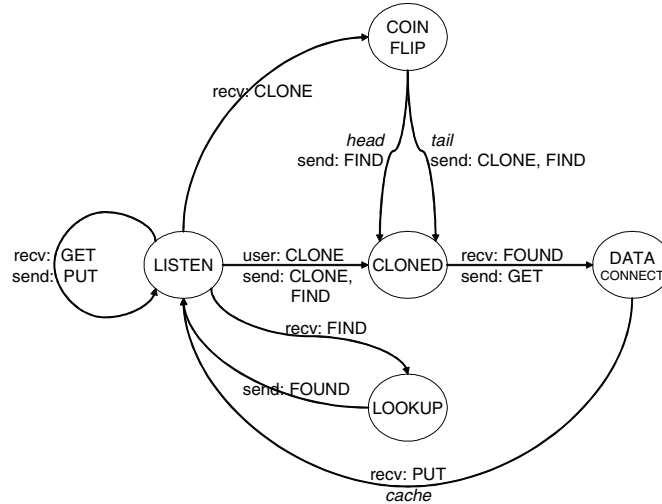


Fig. 2. State diagram of P2PRIV peer

request. If a FIND message is received from other P2PRIV node, then a “Find and Reply” subroutine will begin (see Figure 5), where a P2PRIV node looks for a specified content on behalf of other anonymous P2PRIV node. The last section (starting from Line 11, Figure 3) refers to reaction to an upload request from other node. In this case the node was pointed out by DHT interface as storage of the specified file. If it occurs and a GET message is received then an “Upload” subroutines will begin (Figure 6).

The Clone and Download subroutine describes a process of a CC joining and fulfilling tasks of a clone: a location of a requested content and its download. If the request is originated anonymously (via the Mix-net) from other node, the P2PRIV peer will flip an asymmetric coin to decide whether to forward the token anonymously to the next random peer with probability p_f (Line 5, Figure 4). Otherwise, if the request originates from a user, the forwarding process will surely proceed (Line 8, Figure 4).

```

(1)  while  $\infty$  do
(2)    User_Request := false
(3)    if Mix_Receive ({“CLONE”, File_Id}, Irrelevant_Return_Addr)
(4)    or
(5)    User_Request := User_Api_Read ({“CLONE”, File_Id})
(6)    then Clone_and_Download (File_Id, User_Request)
(7)
(8)    if Mix_Receive ({“FIND”, File_Id}, Some_Proxy_Addr)
(9)    then Find_And_Reply (File_Id, Some_Proxy_Addr)
(10)
(11)   if Tcp_Receive ({“GET”, File_Id}, Clone_or_Initiator_Addr)
(12)   then Upload (File_Id, Clone_or_Initiator_Addr)
  
```

Fig. 3. Pseudo-code description of a P2PRIV single node operation

```

(1) subr. Clone_and_Download (Download_File_Id, Is_Request_from_User)
(2)   CC_Forward:=false
(3)   if (Is_Request_from_User)
(4)     CC_Forward:=true
(5)   else if (Coin_Flip (pf))
(6)     CC_Forward:=true
(7)   if (CC_Forward)
(8)     Mix_Send ({"CLONE", Download_File_Id}, First_Random_Addr)
(9)
(10)  Mix_Send ({"FIND", Download_File_Id}, Second_Random_Addr)
(11)
(12)  while not
(13)    Mix_Receive
(14)      ({"FOUND", Download_File_Id, "@", File_Owner_Addr},
(15)       Second_Random_Addr)
(16)  do Wait
(17)  after (randomTime) Tcp_Send ({"GET", Download_File_Id},
(18)   File_Owner_Addr)
(19)
(20)  while not Tcp_Receive ({"PUT", File}, File_Owner_Addr)
(21)  do Wait
(22)  Cache (File)
(23)  if Is_Request_From_User
(24)    Alert ("Requested file has been anonymously downloaded!")

```

Fig. 4. Pseudo-code description of the Clone and Download subroutine

Next, the P2PRIV node sends a FIND message to a randomly chosen peer via means of a Mix-net. The selected node will perform a DHT look-up procedure and reply with an address of the peer which stores the file. This reply is sending on a Chaumian untraceable return address allowing a preservation of anonymity of this request. After receiving the reply message (FOUND) with an address of a storage and a file id (Line 13, Figure 4), the P2PRIV peer sends a randomly delayed request for the file (GET) directly to the storage node and waits for PUT reply message (Line 17, Figure 4).

```

(1) subroutine Find_and_Reply (Lookup_File_Id, Reply_Dest_Addr)
(2)   if File_Owner_Addr:=Lookup (Key (Lookup_File_Id))
(3)     Mix_Send
(4)       ({"FOUND", Lookup_File_Id, "@", File_Owner_Addr},
(5)        reply_Dest_Addr)

```

Fig. 5. Pseudo-code description of the Find and Reply subroutine

This request can be performed directly as the destination node cannot detect whether this request originates from the real initiation or from one of clones. Finally, the requested file is downloaded and cached (Lines 20-22, Figure 4). If the P2PRIV node is the real initiator, then the user will be informed about the successful download (Line 24, Figure 4). As it was described earlier (compare with Clone and Download

```

(1) subroutine upload (upload_File_Id, upload_Dest_Addr)
(2)   Tcp_Send ({"PUT", Upload_File_Id}, Upload_Dest_Addr)

```

Fig. 6. Pseudo-code description of the Upload subroutine

subroutine), the P2PRIV node can send the FIND message to the other P2PRIV node. The “Find and Reply” subroutine describes a reaction to this request. Here the P2PRIV node becomes an anonymizing proxy, which initiates the DHT look-up procedure for a specified file. Additionally, this proxy is reached by means of Mix-net. In the line 2 (Figure 5) a DHT look-up procedure is executed and a result is sent back to an initiator on its untraceable address provided by means of Mix-net.

The last subroutine “Upload” (Figure 6) contains a short reaction to GET request (see Clone and Download subroutine where GET requests are initiated).

4 Path Lengths

In this section we will analyze the impact of p_f configuration on the system’s anonymity and estimate secure lengths of CC. The analysis will cover small overlays ($N = 50$ network nodes), where it seems to be more difficult to hide a real initiator, and large networks ($N = 1000$ nodes) – more realistic for open P2Ps. We will analyze static, adaptive, passive, and active attacks using anonymity model of P2PRIV presented in [12]. We will stick to three variants of network collaboration levels:

- $C = 10\%$: of colluding nodes, scenario usually considered in the state of the art;
- $C = 5\%$: more realistic collaboration level for large and public access overlays (where, apart from Sybil attack, it is more difficult for the adversary to prevail the level of honest nodes); and
- $C = 20\%$: scenario for small overlays (among a small population of honest nodes it is more easy for the adversary to introduce the significant range of colluding nodes).

Let us analyze an uncertainty of finding a real initiator of the content download. The passive-static adversary can distinguish two sets of peers $\{S_1, S_2\}$ among all N nodes and assign their members probabilities of being the initiator $\{p_1, p_2\}$. S_1 consists of peers which are directly connected to colluding nodes C and S_2 are remaining honest nodes. From [12] the sizes of S_1 and S_2 sets equal:

$$S_1 = \frac{C}{N}n = \frac{C(p_f - 2)(N - C)}{(p_f - 1)N^2}, \quad S_2 = N - C - S_1, \quad (2)$$

and the members of those sets are recognized as the initiator with probabilities:

$$p_1 = \frac{(p_f - 1)N}{(p_f - 2)(N - C)}, \quad p_2 = \frac{(p_f - 1)N}{(p_f - 1)N^2 - (p_f - 2)C}. \quad (3)$$

According to the information theory [18] and ([6], [17]) entropy of P2PRIV for this scenario will be

$$H_{psP2PRIV} = -\sum_{i=1}^N p_i \log_2(p_i) = -S_1 p_1 \log_2(p_1) - S_2 p_2 \log_2(p_2) , \quad (4)$$

$$H_{psP2PRIV} = \begin{cases} 0 & p_1 = 1 \vee p_2 = 1 \\ \frac{C}{N} \log_2(p_1^{-1}) & p_2 = 0 \\ \left(1 - \frac{C}{N}\right) \log_2(p_2^{-1}) & p_1 = 0 \\ \frac{C}{N} \log_2(p_1^{-1}) + \left(1 - \frac{C}{N}\right) \log_2(p_2^{-1}) & p_1 \in (0,1) \wedge p_2 \in (0,1) . \end{cases} \quad (5)$$

Figure 7 shows the entropy H of small and large P2PRIV overlays as a function of the parameter p_f .

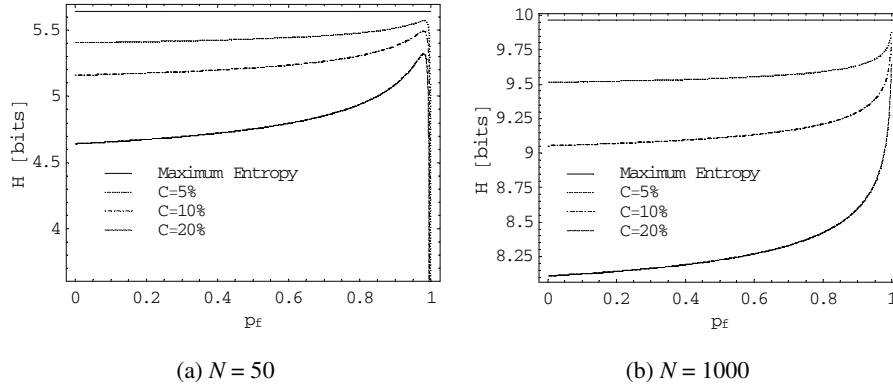


Fig. 7. Entropy of P2PRIV, passive-static attacks

In both cases a p_f increase is in favor of the entropy of P2PRIV system. However, in small overlays (Figure 7a, $N = 50$ nodes) high p_f values (close to 1) eventually degrades the entropy. For $p_f = 0.9$ a cascade mean length reaches value of 6 nodes. This is a relatively large value for small networks. It is significant that in small networks it is easier for the adversary to become a member of CC and degrade P2PRIV entropy (from about 5 bits to even 0). Therefore, when we consider small overlays, p_f configuration should not be close to 1.

The analyzed static attack shows realistic capabilities of the adversary and exemplifies a critical point of view on the expansion of forwarding paths. However, a more pessimistic attack – adaptive observation – is possible. In adaptive attacks it is assumed that the adversary has yet colluding nodes among network nodes, which actively anonymize specified request (e.g., CC in P2PRIV). This scenario shows the

effectiveness of the system's anonymity protection among network nodes actively involved in hiding the initiator. If the colluding node belongs to a CC involved in the download of an observed content then based on model from [12] the number of nodes that communicate directly with colluding nodes (S_{1pa}) and the remaining nodes (S_{2pa}) are

$$S_{1pa} = \frac{(N-C)((p_f-2)C + (p_f-1)N)}{(p_f-1)N^2}, \quad S_{2pa} = N - C - S_{1pa}, \quad (6)$$

with assigned probability of being the initiator p_1 for S_{1pa} and

$$p_{2pa} = \frac{(p_f-1)N(N + (p_f-2)C)}{(p_f-2)(C-N)(C(p_f-2) + N(N - p_f N + p_f - 1))} \quad (7)$$

for S_{2pa} . Then entropy in this attack scenario is

$$H_{paP2PRIV} = \begin{cases} 0 & p_1 = 1 \vee p_{2pa} = 1 \\ \vartheta \log_2(p_1^{-1}) & p_{2pa} = 0 \\ \zeta \log_2(p_{2pa}^{-1}) & p_1 = 0 \\ \vartheta \log_2(p_1^{-1}) + \zeta \log_2(p_{2pa}^{-1}) & p_1 \in (0,1) \wedge p_{2pa} \in (0,1) \end{cases}, \text{ where} \quad (8)$$

$$\vartheta = \frac{(p_f-2)C + (p_f-1)N}{(p_f-2)N}, \quad \zeta = \frac{(p_f-2)C + N}{(p_f-2)N}.$$

Figure 8 shows the entropy H of small and large P2PRIV overlays.

Similarly to the previous scenario of passive-static attacks we can observe that large values of p_f increase the entropy. In small networks, values close to maximum finally negatively impact the system entropy. Contrary to the static attacks, low p_f value significantly impacts the system entropy and for large networks gives the

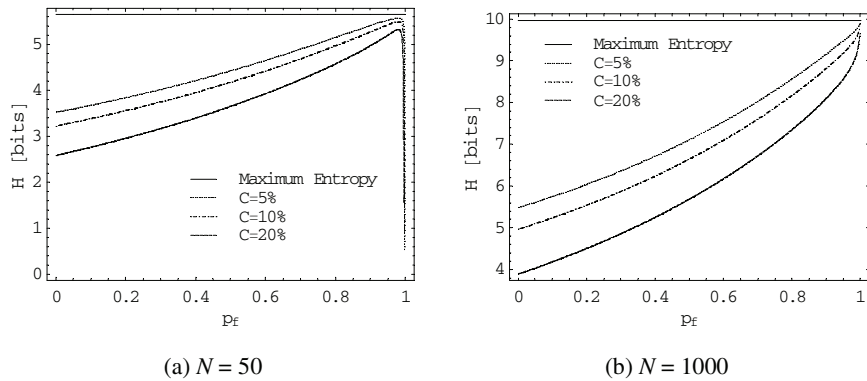


Fig. 8. Entropy of P2PRIV, passive-adaptive attacks

adversary about 5 bits (from a total number of 10 bits) of information about the origin of a specified request.

We will analyze active attacks (static and adaptive) enabling attackers to change protocol operation to disclose more information next. Active attacks, adjusted to the proposed P2P architecture, were proposed in [12]. These attacks allow the adversary to intercept the CC as this action cannot be detected quickly in the new, parallel architecture. Using results from [12] we can stress that the active adversary can distinguish two sets of peer among all overlay network nodes:

$$S_{1as} = \frac{C(N-C)(N^3 + p_f^2(C-N)^3 + p_f N(2C^3 - 3NC + N^2))}{(N + p_f(C-N))N^4}, S_{2as} = N - C - S_{1as} \quad (9)$$

with probabilities of being initiator, accordingly for static attacks:

$$p_{1a} = \frac{N^3(p_f(C-N) + N)}{(N-C)(p_f^2(C-N)^3 + N^3 + p_f N(2C^3 - 3NC + N^2))}, \quad (10)$$

$$p_{2as} = \frac{N^3(N + p_f(C-N))}{p_f^2 C(C-N)^3 + N^3(C-N^2) + p_f N(2C^3 - 3C^2 N - (N-1)CN^2 + N^4)} \cdot \quad (11)$$

Taking into account adaptive possibilities of the adversary:

$$S_{1aa} = \frac{C}{N} \left(\frac{(N-C)(N^3 + p_f^2(C-N)^3 + p_f N(2C^3 - 3NC + N^2))}{(N + p_f(C-N))N^3} - 1 \right) + 1, S_{2aa} = N - C - S_{1aa} \quad (12)$$

with probability p_{1a} and

$$p_{2aa} = \zeta \left[N - C + \frac{C}{N} \left(\frac{1 + (C-N)(N^3 + p_f N(2C^2 - 3NC + N^2) + p_f^2(C-N)^3)}{N^3(N + p_f(C-N))} \right) \right]^{-1}. \quad (13)$$

Finally the entropies of P2PRIV for active attacks are

$$H_{asP2PRIV} = \begin{cases} 0 & p_{1a} = 1 \vee p_{2as} = 1 \\ \frac{C}{N} \log_2(p_{1a}^{-1}) & p_{2as} = 0 \\ \left(1 - \frac{C}{N}\right) \log_2(p_{2as}^{-1}) & p_{1a} = 0 \\ \frac{C}{N} \log_2(p_{1a}^{-1}) + \left(1 - \frac{C}{N}\right) \log_2(p_{2as}^{-1}) & p_{1a} \in (0,1) \wedge p_{2as} \in (0,1) \end{cases} \quad (14)$$

$$H_{aaP2PRIV} = \begin{cases} 0 & p_{1a} = 1 \vee p_{2aa} = 1 \\ \psi \log_2(p_{1a}^{-1}) & p_{2aa} = 0 \\ \zeta \log_2(p_{2aa}^{-1}) & p_{1a} = 0 \\ \psi \log_2(p_{1a}^{-1}) + \zeta \log_2(p_{2aa}^{-1}) & p_{1a} \in (0,1) \wedge p_{2aa} \in (0,1) \end{cases}, \text{ where}$$

$$\psi = \frac{N^3(N+C) - p_f N(N^3 - 2C^3 + 3NC^2 - 2N^2C) + p_f^2 C(C-N)^3}{(N^3 + p_f N(2C^2 - 3CN + N^2) + p_f^2 (C-N)^3)N}$$

$$\zeta = \frac{p_f N(2N^3 - 2C^3 + 5C^2N - 5CN^2) - CN^3 - p_f^2 (C-N)^4}{(N^3 + p_f N(2C^2 - 3CN + N^2) + p_f^2 (C-N)^3)N}.$$

Figure 9 shows the entropies for P2PRIV static and adaptive attacks against active adversary attempting to intercept the token's random walk.

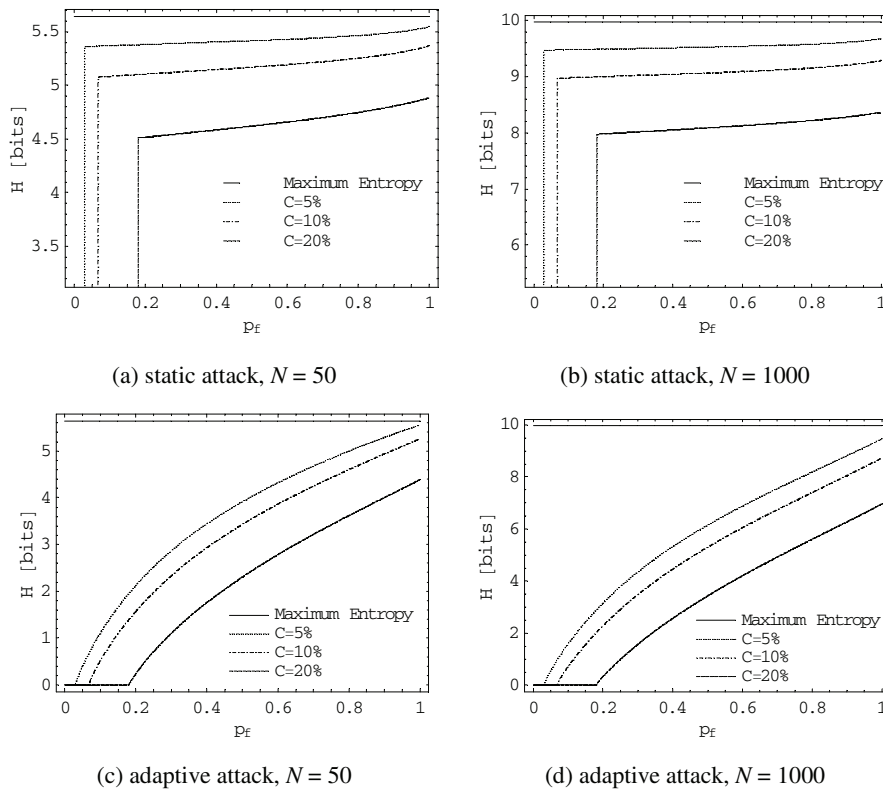


Fig. 9. Entropy of P2PRIV, active attacks

Taking into account far reaching possibilities of active adversary, which is able to imperceptibly break the cloning cascade, we have observed that P2PRIV can still assure proper level of anonymity. In configuration of high p_f values anonymity of P2PRIV is still close to maximum. However, active-adaptive scenarios show that secure configuration of the system should not cover p_f values lower than 0.5 ($P = 3$). For low p_f the P2PRIV entropy against active-adaptive adversary quickly reaches 0 value, which means that the initiator of the request becomes completely exposed.

Based on realistic assumptions of an adversary capabilities corresponding to the environment of public peer-to-peer overlays, we have observed that the new protocol

assures high entropy. Passive attacks reveal that the solution is robust against the specific character of a static attack and that the effective penetration of a proper cloning cascade is difficult to carry out for the adversary. The analysis of active attacks strengthened our pronouncement, obtained by the analysis of passive-adaptive attacks, that p_f configuration should not cover low values (below 0.5) as in this configuration entropy is close to a minimum. Additionally, we have observed that in small overlay networks configuration of p_f close to 1 also negatively impacts the entropy. The overlay is dedicated to public and wide usage; however it should also be able to work under temporal conditions of small number of users.

Serjantov *et al.* [17] used the entropy measure to calculate the effective anonymity set size of anonymous systems. Let H be the entropy of the system, then

$$S_A = 2^H . \tag{15}$$

This quantification allows stressing anonymity of a particular system as an equivalent of the perfect system with 2^H users. Table 1 contains a summary of the results obtained throughout analyzed scenarios. For the purpose of results' legibility, we have calculated the effective anonymity set size of the new protocol for characteristic and permissible values of average path lengths P as a function of p_f parameter (compare with equation 1). Here, we can observe how numerous would be the perfectly homogeneous "crowd" of peers surrounding and hiding the initiator of the request for networks of 50 and 1000 nodes.

Table 1. Effective anonymity set size for P2PRIV

	p_f	0.5	0.66	0.8	0.88	0.5	0.66	0.8	0.88
	P	3	4	6	10	3	4	6	10
Attack	C	$N = 50$				$N = 1000$			
Passive-Static		43	44	45	46	750	760	770	790
Passive-Adaptive	5%	21	28	36	42	140	230	380	550
Active-Static		43	43	44	45	740	750	760	790
Active-Adaptive		15	24	32	38	71	160	290	430
Passive-Static		37	38	40	41	550	570	590	620
Passive-Adaptive	10%	18	24	31	38	100	170	290	430
Active-Static		36	37	38	39	540	550	570	590
Active-Adaptive		11	17	24	30	40	92	170	250
Passive-Static		27	28	31	34	300	320	340	380
Passive-Adaptive	20%	13	17	24	30	51	93	160	260
Active-Static		25	26	27	28	270	280	300	310
Active-Adaptive		5	8.5	12	16	11	26	49	74

5 Latency

Based on the previous discussion and results showed in Table 1 we will evaluate latency of P2PRIV with configurations starting from $p_f = 1/2$ (3 parallel links) and CROWDS with corresponding configurations $p_f = 2/3$ (3 cascade links) and higher.

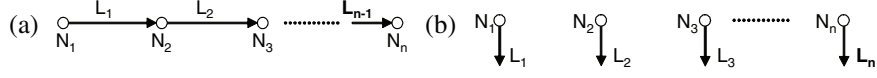


Fig. 10. Number of nodes (P) and links (L) for cascade (a) and parallel (b) transport

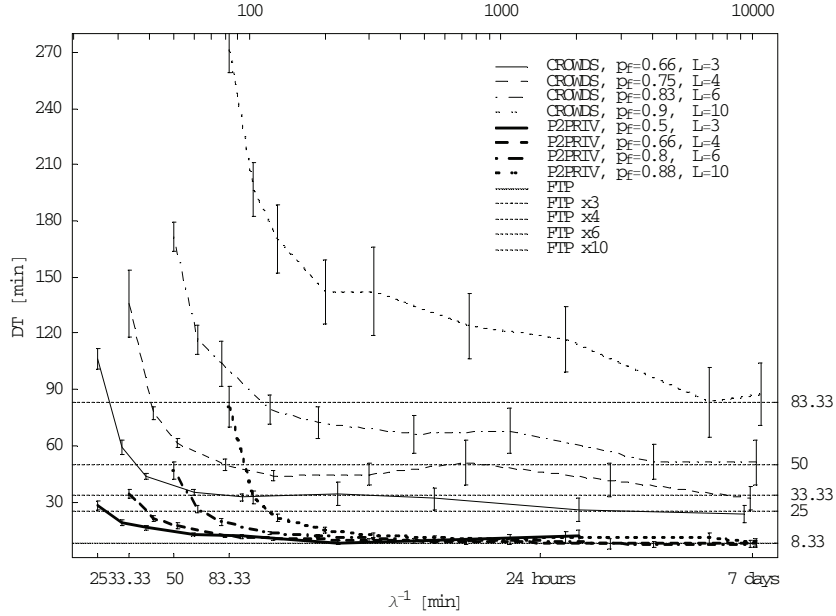


Fig. 11. Latency for CROWDS and P2PRIV

Notice that CROWDS includes one less link (L) for the same path length (P) because of its cascade transport architecture (Figure 10).

Let us evaluate the latency of the proposed protocol using anonymous traffic simulation model described in [12]. In the peer-to-peer anonymous traffic’s simulation environment each peer of the simulated network retrieves the same algorithm suitable to a simulated protocol (for example “Jondo” of CROWDS) and estimates a download time DT quoted as an average time required for a content transport after a submission of the request by the user (initiator). The simulator traces tasks for each symmetric peer independently. We use Poisson distribution to model a request arrival process, where mean request arrival rate λ represent an intensity of requests for a content per single peer. The simulation results are compared to the classical anonymous network CROWDS and analytically calculated DT optimum value μ_{\min}^{-1} , referred to as FTP.

$$\mu_{\min} = \frac{B}{V} = 0,002[s^{-1}], \tag{16}$$

where average link throughput between peers is $B = 512$ kb/s and average file size of the shared content $V = 32$ MB. Thus, FTP (equaled to $8\frac{1}{2}$ minutes) represents a theoretical download time between two directly connected nodes (ratio of average content size to average link throughput).

Figure 11 illustrates latency of P2PRIV and CROWDS. The content transport of P2PRIV is at least three times faster (for $L = 3$ configurations) than CROWDS DT. The observed increase of DT for the close to maximum request arrival rate is lower for P2PRIV. CROWDS introduces multiplications of DT for path lengths increases, while P2PRIV retains the DT level close to optimum regardless of CC lengths for medium- and low-loaded networks.

6 Conclusions

The paper introduces an overlay transport protocol based on a concept of a peer-to-peer direct and anonymous distribution overlay (P2PRIV). The idea of the P2PRIV consists in a parallel content transport instead of the widespread cascade transmission between chaining nodes. This feature allows for a separation of the anonymization process from the content transport function. In the paper we presented a design of the P2PRIV protocol for direct and anonymous P2P download by using a finite state machine diagram and a pseudo-code description. We analyzed the latency of the new protocol and a classical low-latency anonymous network, and compared the results to optimum values. We studied an impact of the protocol configuration, i.e., “cloning cascade” path lengths on the solution’s anonymity and latency, and considered an impact of the new solution on a shape of the present-day’s trade-off between anonymity and traffic performance. We found that the new solution allows both for increasing of anonymous path lengths and for assuring higher anonymity with significantly lower impact on transport’s latency in contrast to the classical low-latency anonymous network (CROWDS). Each successive node from a forwarding path of CROWDS increases anonymity, however, it also considerably increases a latency of the content transport. As one can expect, the average latency imposed by the classical network is at least a multiplication of average number of links between a sender and a receiver, and an average download time of content’s portion between two neighbor nodes. P2PRIV dissolve this issue. The impact of anonymization path lengths on the latency of the new parallel transport is significantly lower and in a case of medium- to low-loaded networks it is negligible. We believe that the results bring new possibilities of overcoming borders of existing trade-off between anonymity and traffic performance of anonymous networks.

The proposed solution is dedicated to the transport of a large content. However, it is motivating to analyze its usefulness in a wider range of applications. The vital question is if the new system can assure generic anonymous communications. In this case, it should be studied how robust P2PRIV is against long-term intersection attacks. Having in mind a single access to particular content by a single user, characteristic for a sharing of static, large content, we did not analyze this issue. However, in systems of general purpose, the problem of multiple accesses to the same resource should be considered. In peer-to-peer overlays it is difficult to ensure a long-term availability of individual peers, so the cascade is deemed to change over time (with the exception of

the initiator). One direction of solving this issue for the generic purpose anonymous system, based on the P2PRIV architecture, is to allow the initiator to maintain the composition of the cloning cascade for a long period of time.

Our future work will also include analysis of an integration of the P2PRIV with publication of content functions. For a pervasive use of the presented system it is necessary to include in it the anonymous publication service, which together with the described P2PRIV anonymous transport can provide a comprehensive anonymous communications overlay. The publication process can be treated as a separated function of the overlay, however, to achieve a coherent prototype, the publication scheme should be based on a concurrent with P2PRIV set of security primitives. The other motivating goal is selecting such a known anonymous publication system that can share as many security primitives embedded in the P2PRIV as possible.

References

1. Bennett, K., Grothoff, C.: GAP – practical anonymous networking. In: Dingledine, R. (ed.) PET 2003. LNCS, vol. 2760. Springer, Heidelberg (2003)
2. Berthold, O., Federrath, H., Köpsell, S.: Web MIXes: A system for anonymous and unobservable Internet access. In: Federrath, H. (ed.) Designing Privacy Enhancing Technologies. LNCS, vol. 2009. Springer, Heidelberg (2001)
3. Chaum, D.: Untraceable electronic mail, return addresses, and digital pseudonyms. *Communications of the ACM* 4(2) (February 1981)
4. Clarke, I., Sandberg, O., Wiley, B., Hong, T.W.: Freenet: A distributed anonymous information storage and retrieval system. In: Proceedings of Designing Privacy Enhancing Technologies: Workshop on Design Issues in Anonymity and Unobservability, pp. 46–66 (July 2000)
5. Dai, W.: Popenet 1.1. Usenet post (1996)
6. Diaz, C., Seys, S., Claessens, J., Preneel, B.: Towards measuring anonymity. In: Dingledine, R., Syverson, P.F. (eds.) PET 2002. LNCS, vol. 2482. Springer, Heidelberg (2003)
7. Dingledine, R., Freedman, M.J., Molnar, D.: The free haven project: Distributed anonymous storage service. In: Federrath, H. (ed.) Designing Privacy Enhancing Technologies. LNCS, vol. 2009. Springer, Heidelberg (2001)
8. Dingledine, R., Mathewson, N., Syverson, P.: Tor: The second generation onion router. In: Proceedings of the 13th USENIX Security Symposium (August 2004)
9. Freedman, M.J., Morris, R.: Tarzan: A peer-to-peer anonymizing network layer. In: 9th ACM Conference on Computer and Communications Security, Washington, DC (November 2002)
10. Goldschlag, D., Reed, M., Syverson, P.: Hiding Routing Information. In: Anderson, R. (ed.) IH 1996. LNCS, vol. 1174, pp. 137–150. Springer, Heidelberg (1996)
11. Loesing, K., Sandmann, W., Wilms, C., Wirtz, G.: Performance Measurements and Statistics of Tor Hidden Services. In: Proceedings of the 2008 International Symposium on Applications and the Internet (SAINT), Turku, Finland (July 2008)
12. Margasiński, I., Pióro, M.: A Concept of an Anonymous Direct P2P Distribution Overlay System. In: Proceedings of the 22nd IEEE International Conference on Advanced Information Networking and Applications (AINA 2008), March 2008, pp. 590–597 (2008) ISSN 1550-445X, ISBN 978-0-7695-3095-6
13. Reed, M., Syverson, P., Goldschlag, D.: Anonymous Connections and Onion Routing. *IEEE Journal on Selected Areas in Communications* 16(4), 482–494 (1998)

14. Rennhard, M., Plattner, B.: Introducing MorphMix: Peer-to-Peer based Anonymous Internet Usage with Collusion Detection. In: Proceedings of the Workshop on Privacy in the Electronic Society (WPES 2002), Washington, DC, USA (November 2002)
15. Reiter, M.K., Rubin, A.D.: Crowds: Anonymity for web transactions. *ACM Transactions on Information and System Security* 1(1) (June 1998)
16. Snader, R., Borisov, N.: A Tune-up for Tor: Improving Security and Performance in the Tor Network. In: Proceedings of the Network and Distributed Security Symposium (NDSS 2008) (February 2008)
17. Serjantov, A., Danezis, G.: Towards an information theoretic metric for anonymity. In: Dingledine, R., Syverson, P.F. (eds.) PET 2002. LNCS, vol. 2482. Springer, Heidelberg (2003)
18. Shannon, E.: A Mathematical Theory Of Communication. *The Bell System Technical Journal* 27, 379–423, 623–656 (1948)