情報処理学会研究報告
IPSJ SIG Technical Report

Vol.2009-DPS-140 No.10
2009/9/11

高レアリティブロック発生を抑制するP2Pファイル共有手法

松本 敬[†1] 遠藤 伶[†1]
高木健士[†1] 重野寛[†2]

P2P ファイル共有において，レアリティが高く入手の難しいブロックが発生し，ブロック収集効率が下がるブロックのレアリティ問題が存在する．ブロックとは，共有するファイルを予め決められたサイズに分割した断片のことである．そこで，本稿ではブロック収集効率を上げるためにブロックのレアリティを考慮した P2P ファイル共有手法 CAS の提案を行う．CAS ではレアリティ問題の原因であるブロックの初期分散速度とピア離脱の 2 点に対処することで，ブロック収集効率を上げる．さらに，シミュレーション評価を行い CAS のブロック収集効率における有用性及び悪意あるピアがネットワークに存在する場合の対応性を示す．

P2P file sharing method controlling high rarities of blocks

Kei MATSUMOTO,[†1] Rei ENDO,[†1] Kenji TAKAGI[†1]
and Hiroshi SHIGENO[†2]

This paper discusses a rarity problem of blocks in P2P file sharing that degenerates the blocks collection efficiency of peer by the blocks that have high rarity and thus difficult to obtain it. Blocks are the divided fragments of shared files as which size is decided previously. The aim of this proposal is to improve the blocks collection efficiency of peer in P2P. We propose CAS that considers the rarities of blocks. In the proposal, CAS deals both the initial dispersion speed of blocks and the peer departure to improve the blocks collection efficiency. In addition, we evaluated the proposal by simulation and showed the availability of CAS for blocks collection efficiency of peer and the correspondence analysis of CAS when malicious peers exist in P2P Network.

†1 慶應義塾大学大学院理工学研究科
Graduate School of Science and Technology, Keio University
†2 慶應義塾大学理工学部
Faculty of Science and Technology, Keio University

1. はじめに

P2P 技術を用いた代表的なサービスとしてファイル共有が挙げられる．P2P ファイル共有では，対等な立場でネットワークに接続するピアによりファイルを分割した断片（ブロック）が交換される．ピア間で交換を行い，全ブロックを集めて元ファイルを復元させることでファイル共有を実現している．P2P ファイル共有ネットワークに参加する各ピアはアップロード，すなわち他ピアへのブロック送信とダウンロード，すなわち他ピアからブロック受信を行う．また P2P 技術を使用しているため，リソース（ファイルのアップロード帯域）をネットワークに参加しているピアが提供している．これにより，ピアの増加に対応して柔軟にネットワークリソースを向上させられるためスケーラビリティがあるのも特長である．

そして，これらのファイル共有サービスで問題となっているのが，ブロックのレアリティ問題である．これは，ネットワーク内に存在するブロック数が種類によりばらつき，ネットワーク内に数が少なく入手困難なブロックが発生する問題である．この入手困難なブロックの発生により，ブロック収集効率が落ち，ファイル復元に時間がかかってしまう．

ブロック数が種類によってばらつく原因の 1 つが，ブロックの初期分散速度である．共有する元ファイルを所持するピア（シード）から，最初にブロックがアップロードされる際に，ネットワークへのブロックの広がり方が遅いと，ネットワーク全体にブロックが行き渡るまで時間がかかる．また，ネットワークからのピア離脱もブロック数が種類によってばらつく原因の 1 つである．ピアが離脱する際に，そのピアの所持していたブロックがネットワークから消失してしまう．そのため，そのブロックのネットワーク内に存在する数が少なくなってしまう．このようにネットワーク内に数が少ないブロックのことを，本稿ではレアリティが高いブロックと呼ぶ．このレアリティが高いブロックの存在により，ピアが未所持のブロックを入手できない状況が起こり，全てのブロックを揃えるのに時間がかかってしまう．そのため，P2P ファイル共有ではブロックのレアリティ問題を考慮する必要がある．

また，ファイル共有の抱える別の問題として，ブロックのダウンロードだけを行いアップロードを行わないピア（悪意あるピア）の free-riding 問題[1] が存在する．アップロードを行わないピアは，ネットワークにリソースを提供しない．そのため，このようなピアが存在するとファイル共有の性能が落ちてしまう．代表的な P2P ファイル共有ソフトである BitTorrent[2] では，free-riding 問題の対策として TFT(Tit-for-Tat)[3),4)] を用いている．

© 2009 Information Processing Society of Japan

本稿ではブロックのレアリティ問題を考慮した P2P ファイル共有手法 CAS (Carrot and Stick) を提案する．CAS では，各ピアの入手ブロック数に応じて必須アップロード数を変化させる．必須アップロード数とは，新規ブロックをダウンロードするために他ピアに対して必要なアップロード数条件のことを指す．CAS では，ブロック所持率が少ない間は必須アップロード数を 0 にする．これにより，既存手法ではアップロード帯域により制限され，使用できずに無駄になっていたダウンロード帯域を有効活用できる．そして，ブロックの初期分散速度が速くなる．ピア離脱に対しても，ピア離脱前の必須アップロード数を多く設定し，ダウンロード制限を行うことで対処する．その結果，ブロックがネットワーク内を循環し易くなりレアリティ問題が抑制される．CAS によりブロックのレアリティが考慮され，ブロック収集効率の改善が可能となる．さらにシミュレーション評価を行い，CAS のブロック収集効率における有用性，及び free-riding 問題で挙げた悪意あるピアへの対応性を示す．

2. ブロックのレアリティ問題

ブロックのレアリティ問題は，ピアがブロック交換を行っていく過程で，ネットワーク内に存在するブロック数が種類によってばらつくことで起こる．ばらつきが起こる原因は様々だが，ばらついた結果ネットワーク内にレアリティの高いブロックが発生する．ネットワーク内に存在する数が少ないレアリティの高いブロックは，通常のブロックに比べ入手が困難である．そのため，レアリティの高いブロックが発生するとブロック収集効率が低下する．

以下で，ブロックのレアリティ問題発生の主な原因であるブロックの初期分散速度とネットワークからのピア離脱の 2 点について説明する．

2.1 ブロックの初期分散速度

ブロックの初期分散速度とは，共有する元ファイルを所持するシードから，最初にブロックがアップロードされる際のネットワークへ広がる速度である．この速度が速い場合は，ネットワークの広範囲へとブロックが広がる．その結果，多くのピアが何らかのブロックを所持できネットワーク全体でブロック交換が行われるようになる．それに対し，ブロックの初期分散速度が遅い場合は，ネットワークへブロックが広がり難く，大半のピアが暫くの間ブロックを所持していない状態となる．そして，シードに近い一部のピア間でしかブロック交換が行われず，ネットワーク内のブロック数は増え難くなる．結果として，初期分散速度が遅いとレアリティの高いブロックが発生しやすくなる．

2.2 ピア離脱

P2P ファイル共有では，ネットワークからのピア離脱が発生する．ピア離脱が発生する主な要因は，ファイルを復元したピアによるものである．ブロック交換を行うことで全種類のブロックを入手したピアは，ファイル復元が可能となる．ファイルの復元に成功したピアは，ネットワーク内にいる理由がなくなりネットワークから離脱する．ピアが離脱すると，そのピアが所持していたブロックがネットワークから消失してしまう．その結果，各ブロックのネットワーク内に存在する数が減少し，レアリティが高くなる．

3. 関連研究

代表的な P2P ファイル共有ソフト BitTorrent で使用されている既存手法の TFT（Tit-for-Tat），及び BitTorrent 上で悪意ある行動をとる BitThief について説明する．

TFT の特徴は，ブロックをアップロードしないとダウンロードできなくなることである．この特徴により，ブロックのアップロードを行わないピアの free-riding 問題に対処できる．

具体的には，各ピアがブロック交換を行う際に以下の式に従うことで実現している[5]．

$$UP_{A\to B} - DOWN_A \leq n \quad (1)$$

1 式はピア A とピア B がブロック交換を行う時にピア A 視点で見た場合の式である．1 式の $UP_{A\to B}$ はピア A がピア B へアップロードしたブロック数で，$DOWN_A$ はピア A がピア B からダウンロードしたブロック数である．この式に従いピア A とピア B は左辺が n 以内ならばブロック交換を続ける．しかし，左辺が n 個を超えた場合，ピア B がピア A へのブロックのアップロードを止めるためブロック交換は中止される．つまり，n 個までピア A はピア B から多くのブロックをダウンロードできる．n 個までのダウンロードを許容するのは，ある程度の許容幅がないとブロックを所持していないネットワークへの新規参加ピアがブロックをダウンロードできなくなるからである．n 個より多くのブロックをダウンロードしたい場合は，相手へブロックをアップロードし続ける必要がある．そのため，他ピアへブロックをアップロードする動機付けが行われており，free-riding 問題対策となる．図 1 にピア A とピア B がブロック交換している様子を示す．

しかし，この手法ではブロックのダウンロード可能数がアップロード数に依存する問題がある．つまり，ADSL 等のダウンロード帯域に比べアップロード帯域が小さい環境で TFT を適用すると，ピアがダウンロード帯域を十分に活かせず，入手できるブロック数が減る．本提案の目標であるブロック収集効率の観点からすると，この手法は効率が悪い．

© 2009 Information Processing Society of Japan

Vol.2009-DPS-140 No.10
2009/9/11

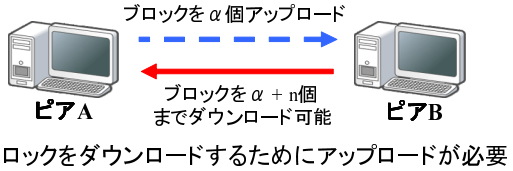


図 1 Tit-for-Tat

BitThief[6] では，他ピアへブロックのアップロード無しにダウンロードできるように BitTorrent を改良している．つまり，free-riding 問題対策手法である TFT 上で free-riding を行う手法である．BitThief の特徴は，TFT で許容しているブロック n 個までのダウンロードを最大限に活かす点にある．具体的には，常にアップロードするブロックを持たない新規参加ピアに偽装し続けることで実現している．

BitThief は，一般的には悲劇的な結果になることが知られている．BitThief を使用しているピアは，P2P ファイル共有において悪意あるピアと同等の動きをする．これは，ブロックのアップロードを行わないピアは，ネットワークにリソースを提供しないため，ネットワーク全体としての性能が落とすからである．

4. 提　　案

本稿では，P2P ファイル共有においてブロック収集効率が下がってしまうブロックのレアリティ問題に着目し，ブロックのレアリティを考慮した P2P ファイル配信手法 Carrot and Stick（CAS）を提案する．CAS ではブロックのレアリティ問題の原因であるブロックの初期分散速度とネットワークからのピア離脱の 2 点に対処し，ブロック収集効率を改善する．

図 2 に CAS におけるピアの挙動を示す．CAS では所持ブロック数の少ないピアは，ダウンロード帯域を有効活用しブロックの収集を行う．そして所持ブロック数が多いピアは，必須アップロード数に達するまでブロックのアップロードを行う．

4.1 ダウンロード帯域の有効活用

P2P ファイル共有では，ブロックの初期分散速度の速さがブロックのレアリティ問題の原因となっている．既存手法の TFT では，使用可能なダウンロード帯域が制限されブロックの収集効率が悪くなっている．そこで，初期分散速度を速くすることでブロックのレアリティ問題を抑制する．その方法として，ピアのダウンロード帯域を有効活用する．

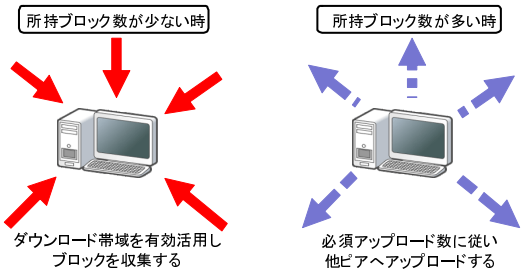


図 2 CAS

CAS では，ネットワークへの新規参加ピア等のブロック所持数が少ないピアに対して，ブロックの初期分散速度を速くするために，所持ブロック数がある一定数（4.3 節参照）に達するまで必須アップロード数を 0 とし，自由なブロックのダウンロードを許す．

4.2 ピア離脱の制約

ピア離脱が発生すると，離脱したピアの所持ブロックがネットワークから消失する．その結果，ブロックのレアリティが上がりファイル収集効率が悪くなる．そこで，ピア離脱寸前に必須アップロード数を多く設定し，ピア離脱を制約する．具体的には，所持ブロック数が多いピアに対する必須アップロード数を多く設定する．

4.3 必須アップロード数の決定

CAS では必須アップロード数の値を変動させることでピアを制御している．ネットワークへ参加したばかりのピアに対しては必須アップロード数を 0 とし，所持ブロック数が多く離脱寸前のピアに対しては必須アップロード数を極力大きくするのが理想的である．

そこで，CAS では必須アップロード数の値を決めるために，値が急激に上昇する特性を持つ指数関数を用いる．しかし，指数関数をそのまま使用すると，ネットワーク新規参加ピアにも小さい値ながら必須アップロード数が発生するため以下のような式に変形した．

$$(\frac{N^{100}}{2} - N^{x})/N^{100} \times \text{全ブロック数} \times 2 \qquad (2)$$

N とは指数関数の底を，x は所持ブロック率をそれぞれ表している．この式を用いることで，特定の所持ブロック率（図 3 の g）になるまでアップロード数は 0 となり，ダウンロード帯域を有効活用できる．また，N の値を変更させることで指数関数の傾きが変わるために，N の値により必須アップロードを 0 から切り替えるタイミングを調整できる．

図 3 に CAS における所持ブロック数と必須アップロード数についての関係を示す．横軸

© 2009 Information Processing Society of Japan

Vol.2009-DPS-140 No.10
2009/9/11

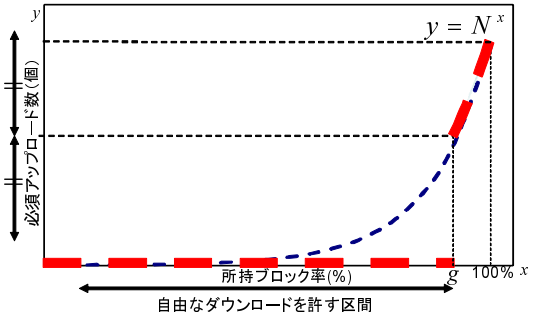


図 3 CAS の特性

はファイル復元に必要なブロック数を 100% とした際に，ピアが現在所持しているブロック数の比率を表す所持ブロック率，縦軸は必須アップロード数である．

4.4 動作アルゴリズム詳細

未所持ブロックを探すピアを *Receiver*，そのブロックを所持するピアを *Sender* とする．

(1) ブロック交換の開始

ブロック交換は，*Receiver* の未所持ブロックを持つ隣人ピア *Sender* に対して，ブロック要求と *Receiver* の所持ブロック数，過去アップロード合計数情報を送信し開始する．過去アップロード合計数とは，今までに *Receiver* が他ピアに対してアップロードを行ったブロックの合計数である．

(2) 必須アップロード数の計算

ブロック要求を受け取った *Sender* は，*Receiver* の所持ブロック数情報から *Receiver* の現在の必須アップロード数を 2 式を使用し計算する．

(3) アップロードの判断

Sender は，*Receiver* の必須アップロード数と過去アップロード合計数を比較する．もし，必須アップロード数よりも過去アップロード合計数が多い場合は，*Receiver* は新規ブロックを入手出来る状態と判断する．そして，*Sender* は *Receiver* に対して要求ブロックのアップロードを行う．逆に，必須アップロード数よりも過去アップロード合計数が少ない場合は，*Receiver* は新規ブロックを入手できない状態と判断する．そして，*Sender* は *Receiver* に対して要求ブロックのアップロードを行わない．

5. シミュレーション評価

CAS のブロック収集効率の有用性及び悪意あるピアへの対応性を示すため，P2P ファイル共有ソフト BitTorrent の既存手法 TFT との比較評価を行った．

5.1 シミュレーション条件

表 1 にシミュレーションで使用した条件を示す．ダウンロード数やアップロード数は, 通常ピアは ADSL，高性能ピアは FTTH から性能比を算出した[7]．

本シミュレーションの単位時間は Round である．各ピアは Round 単位でブロックのアップロードとダウンロードを行う．また，既存手法 1 式の n の値は一般的な 2[8] とした．

評価項目は，ファイル復元時間，ブロックの普及率，悪意あるピアがネットワーク内に存在する場合のファイル復元時間とした．シミュレーション回数は各 50 回行った．

表 1 シミュレーション条件

総ピア数	1000
シード数	1
ファイル分割数	10000 ブロック
隣人更新周期	3Round
最大隣人数	10
高性能ピア最大アップロード数	15 ブロック
高性能ピア最大ダウンロード数	15 ブロック
通常ピア最大アップロード数	3 ブロック
通常ピア最大ダウンロード数	10 ブロック
シード最大アップロード数	9 ブロック
高性能ピアと通常ピアの比率	1:2

5.1.1 シミュレーションシナリオ

- ブロックを収集しファイル復元を目的としたピアをネットワークへ参加させる．
- 全ブロックを所持するシードが，Round 毎 3 ピアにブロックをアップロードする．
- 各ピアが隣人リスト内のどの隣人へブロックを渡すかはランダムに決定する．
- 隣人へのアップロードは Round 単位で行い, 既存方式と CAS の違いを下にまとめた．
  - 既存方式
    隣人からダウンロードしたブロック数を使用し，1 式に従いアップロードを行う．
  - CAS
    隣人リストからランダムに決定した隣人へ, ランダムにアップロードを行う．

© 2009 Information Processing Society of Japan

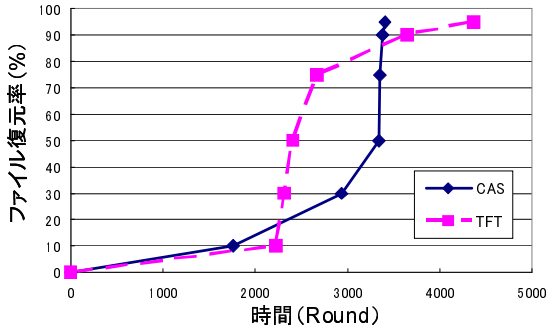


図 4 ファイル復元率の比較

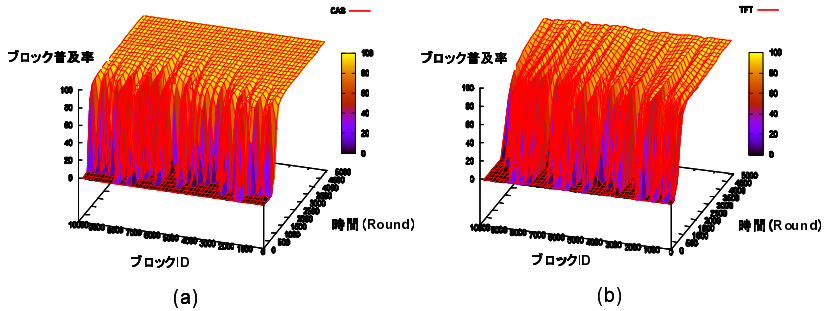


図 5 ブロックの普及率推移．(a) は CAS，(b) は既存手法．

- 隣人からのダウンロードは Round 単位で行い, 既存方式と CAS の違いを下にまとめた．
  - 既存方式
    隣人へアップロードしたブロック数を使用し，1 式に従いダウンロードを行う．
  - CAS
    2 式に従い, ダウンロード可能な状態ならダウンロードを行う．
- 全種類のブロックを入手しファイル復元したピアはネットワークから離脱する．
- ネットワーク参加ピアの 95% がファイル復元した時点でシミュレーションを終了する．

5.2 シミュレーション結果

CAS と既存手法におけるファイル復元率，ブロックの普及率推移，及びネットワーク内に悪意あるピアが存在する場合の CAS の対応性について考察する．

5.2.1 ファイル復元率

図 4 に CAS と既存手法のファイル復元率の比較を示す．図において横軸は時間（単位は Round），縦軸はピアのファイル復元率（単位は %）を示している．ファイル復元率とは，ブロックを全種類入手し，ファイル復元に成功したピア数を表す．具体的には，ファイル復元に成功したピア数を総ピア数で割って求めた．

既存手法は，早い段階で多くのピアがファイル復元を成功させる．しかし，ファイル復元できずにネットワークに取り残されたピアは，ファイル復元に非常に時間がかかってしまう．

それに対し CAS は，ピア離脱の制約によりファイル復元の速さは緩やかである．しかし，ネットワークに取り残されるピアが発生し難い．これは，高レアリティブロック発生の抑制により，既存手法に比べてピアが欲しいブロックを見つけ易くファイル復元が容易だからである．その結果，ネットワーク内の全ピアがファイル復元に要する時間を短縮できる．

5.2.2 ブロック普及率

図 5 に CAS と既存手法のブロック普及率推移比較を示す．(a) は CAS，(b) は既存手法である．ブロック普及率は，各ブロックがネットワークにどのくらい普及しているかを表す．具体的には，各ブロックの所持ピア数を総ピア数で割って求めた．

CAS と既存手法の違いは，ブロックの普及速度と普及の安定性である．ブロックの普及速度は，各ブロックの普及率が 100% になるまでの速度を表す．CAS の方が既存手法よりブロックの普及速度が速い．これは，ブロックの初期分散速度を考慮しているからである．普及の安定性は，普及率 100% のブロックが多いほど安定である．CAS は全ブロックが普及率 100% で安定性が高いのに対し，既存手法は普及率が 100% でないブロックが多く安定性が低い．このブロック普及率のばらつきが，高レアリティブロック発生の要因である．

5.2.3 悪意あるピアが存在する場合

ここで悪意あるピアとは，アップロードを行わずダウンロードだけ行うピアとする．

図 6 に CAS の悪意あるピアによるファイル復元時間の影響を示し，図 7 に通常ピアと悪意あるピアのファイル復元ピア数推移を示す．図 7 の (a) は悪意あるピアが全体の 10%，(b) は 50% 存在する場合である．各図の横軸は時間（単位は Round），図 6 の縦軸はピアのファイル復元率（単位は %），図 7 の縦軸はファイル復元ピア数を示している．

図 6 を見ると，悪意あるピアが多いほどファイル復元時間が長くなるのが分かる．また悪

© 2009 Information Processing Society of Japan

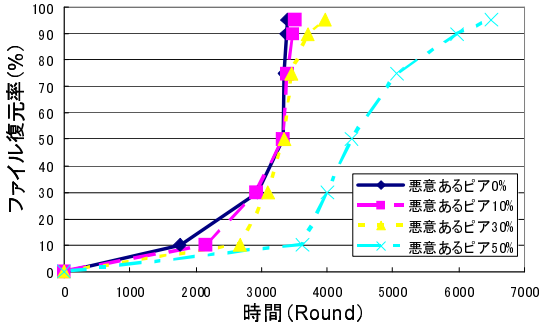


図 6 悪意あるピアによるファイル復元時間の影響

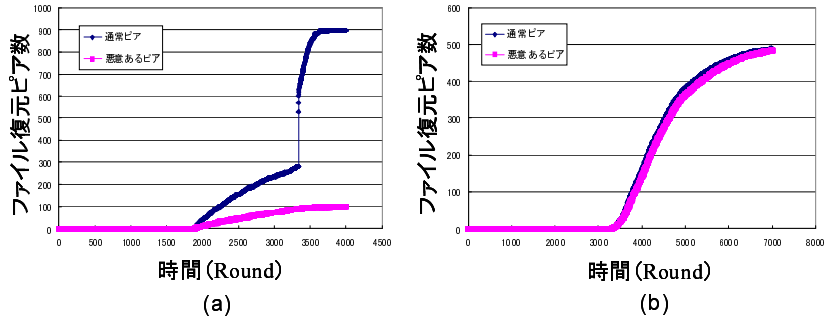


図 7 通常ピアと悪意あるピアのファイル復元ピア数推移．(a) は悪意あるピアが 10%，(b) は 50% 存在する場合．

意あるピアの比率が大きくなるほどファイル復元にかかる時間への影響も大きくなる．これは，悪意あるピアがリソースを提供しないためブロック交換効率が下がるためである．次に図 7 を見ると，悪意あるピアが 10%，50% 存在する場合の両方とも，通常ピアと悪意あるピアの全ピアがファイル復元にかかる時間に大きな差は見られない．悪意あるピアが 50% の場合では，ほぼ同じである．これは，CAS ではブロックをダウンロードするための動機付けを行わないからである．これらから，悪意ある行動をとってもファイル復元時間が長くなるだけで，悪意ある行動をとるメリットは存在しない．

6. おわりに

本稿では，ブロックのレアリティ問題を考慮した P2P ファイル共有手法 CAS を提案した．ブロックのレアリティ問題とは，ネットワークに存在するブロック数が種類によりばらつき，入手困難なブロックが発生する問題である．入手困難なブロックの発生によりブロック収集効率が落ちる．そのため，ブロックのレアリティ問題を考慮した手法が必要である．

ブロックのレアリティ問題の原因はブロックの初期分散速度とピア離脱である．CAS では，初期分散速度の問題を改善するために，既存手法では無駄となっていたダウンロード帯域を有効活用している．そして，ピア離脱の問題を改善するために必須アップロード数を設定し，ピア離脱を制限した．これらにより，高レアリティブロック発生の抑制を実現する．

CAS をシミュレーションにより評価した結果，既存手法に比べ，ネットワーク内に存在する全ピアのファイル復元に要する時間が短縮するのを確認した．また，高レアリティブロック発生の抑制を確認し，悪意ある行動をとってもメリットがないことを確認した．

以上より CAS は P2P ファイル共有において効率的なブロック収集手法として有用である．

謝辞

本研究の一部はグローバル COE プログラム「アクセス空間支援基盤技術の高度国際連携」により行われました。

参 考 文 献

1) L.Ratnasamy and L. Liu. Free Riding:A new challenge to peer-to-peer file sharing systems. In HICSS, 2003.
2) Bram Cohen. Incentives build robustness in bittorrent. In P2PEcon, June, 2003.
3) A.Legout, G.Urvoy-Keller, and P.Michiardi. Rarest first and choke algorithms are enough. In IMC, 2006.
4) G.Neglia, G.L.Presti, H.Zhang and D.Towsley. A network formation game approach to study BitTorrent tit-for-tat. In NET-COOP, 2007.
5) S.Jun and M.Ahamad. Incentives in BitTorrent induce free riding. In P2PEcon, 2005.
6) T.Locher, P.Moor, S.Schmid, and R.Wattenhofer. Free riding in BitTorrent is cheap. In HotNets, November 2006.
7) 大橋省吾, 大根田進 and 町田達彦. FTTH 時代を担うファースト 1 マイルソリューション. SWCC グループ, 2005.
8) C. Gkantsidis and P. R. Rodriguez. Network Coding for Large Scale Content Distribution. In IEEE/Infocom, 2005.

© 2009 Information Processing Society of Japan