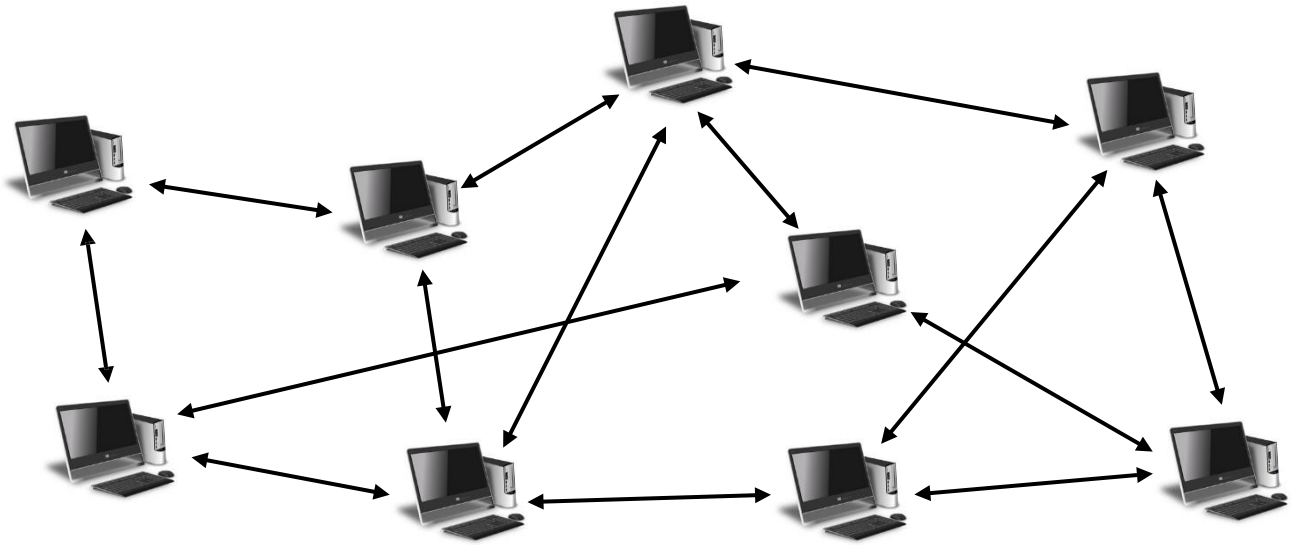
# IPFSにおけるSybil攻撃のオンデマンド検知

NS研究会　３月４日

立命館大学　情報理工学部
姫野貴一・上山憲昭

# IPFS(InterPlanetary File System)

- P2Pネットワーク上で，完全自律分散でデータを分散して保存するシステム　→　Web3が目指す分散型のデータ共有を実現
- 本質は「権威的サーバのない自律分散型の名前解決メカニズム」
  - すべてのノードが対等な立場で振る舞い、データやコンテンツのやり取りを行う
- DHT（distributed hash table）を用いてコンテンツの取得を行う



# IPFS(InterPlanetary File System)

## IPFSの用語

- CID（Content Identifier）
  - 各コンテンツを識別するID
  - データのハッシュ値に基づいて生成
  - 同じ内容　→　同じCID
  - 内容が1ビットでも変わるとCIDも変化　→　改ざん検知が可能

- PID（Peer ID）
  - 各ピアを識別するID
  - 公開鍵をハッシュ化して生成
  - ノードごとに一意

→PID、CIDはDHTルーティングの距離計算に利用される

# IPFS(InterPlanetary File System)

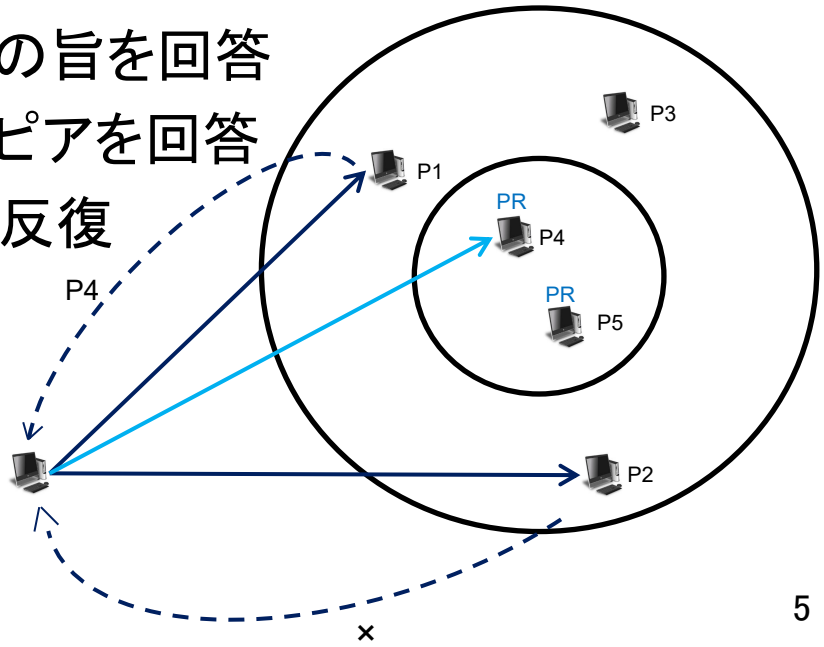
## IPFSの用語

- DHT（Distributed Hash Table）
  - 分散型の検索テーブル
  - CIDとPIDのXOR距離に基づき，よりCIDに近いPIDを持つノードへ逐次転送を行う
  - 中央サーバなしでProviderノードを発見可能

- ProviderRecord
  - 「どのPIDがどのCIDを提供可能か」を示す情報
    - Peer ID
    - IPアドレス
    - ポート番号などの接続情報
  - CIDにXOR距離の近いPIDを持つピアに保存される
  - 一定時間で期限切れ（TTLあり）

# IPFSのコンテンツ取得の流れ

- IPFSはDHTを用いてPUTとGETの対象ピアを選択
- 各ピアはルーティングテーブル(PIDとIPアドレスの組)を保持
- DHT探索処理(DHTウォーク)
  - ノードは自身のルーティングテーブルからCIDとのXOR距離が近いα個のピアを選出、並列に問い合わせ
  - 問い合わせを受けたピアの動作:
    - 自身がCIDのデータを保持 ⇒ その旨を回答
    - そうでない場合 ⇒ 自身より近いピアを回答
  - ProviderRecord(PR)を見つけるまで反復



# IPFSの優位性・課題

優位性

- 分散性：中央管理者を必要とせず、耐検閲性・耐改ざん性に優れる
- 拡張性：既存のWebやアプリケーションに統合でき、Web3の基盤として活用可能
- 障害耐性

課題

- Sybil攻撃への耐性がない

# 研究の背景

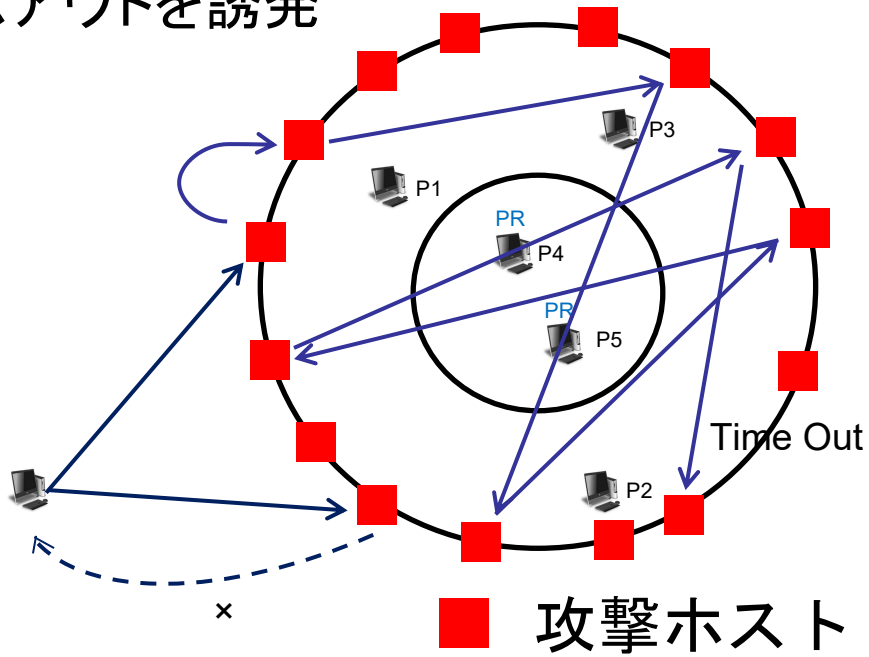
- 一般的なSybil攻撃
  - 攻撃者が多数の偽ID（Sybilノード）を生成し，ネットワークに参加
  - 単一の実体で多数のノードを装い，分散システムの多数決や探索を不正に操作

- **代表的な具体例**
  - ブロックチェーン（例：PoS系ネットワーク）
    - 攻撃者が多数のノードIDを作成してネットワークに参加
    - ピア選択や投票に影響を与え，合意形成を偏らせる
  - 無線アドホックネットワーク
    - 攻撃者が1台の端末から多数の偽ノードIDを生成
    - ターゲット付近にSybilノードを密集配置
    - ルーティングでは目標に近いノード、候補が多く見つかる方向が優先される性質を悪用
      → **ルーティングを誘導する攻撃**

# 研究の背景

- IPFSにおけるシビル攻撃
  - 攻撃者がターゲットCIDにXOR距離の近いPIDを大量生成し、該当ノードをIPFSネットワークに参加させる
  - ルーティングはCIDとPIDのXOR距離に基づく
    →CID近傍にSybilノードを配置すると命中しやすい
  - 一度攻撃ノードに到達すると，攻撃ノード間で要求が転送され続ける
    または要求を無視することでタイムアウトを誘発

→ Sybilノードにより**探索が妨害**される



# Peer ID（PID）の生成と可変性

- PIDの生成
  - PIDは各ピアを識別する一意なID
  - ノード生成時に作成される公開鍵から導出される
  - 公開鍵をハッシュ化して生成

- PIDの可変性
  - 鍵ペアを再生成するとPIDも変化する
  - ノードは任意のタイミングで鍵ペアを作り直せる
  - 鍵生成はローカルで完結し、外部承認を必要としない
  - 生成コストが低く短時間で多数生成可能

# 既存研究の紹介

## 既存研究のアプローチ

- 定期検知：一定の時間間隔ごとに全体を走査して検知
  →不要な検査が多く，検知コストが増大する傾向にある

- Push時検知：コンテンツをネットワークに公開する際に検知
  →コンテンツを公開後の後出しSybilを検出できない

- 遅延ベース検知：ユーザのコンテンツ要求から取得完了まの時間遅延を指標に検知
  →検知遅延および誤検知の問題が存在

# 本研究の目的

- 本研究の目的
  - 既存方式には，コストまたは検知性能の課題がある
  - 本研究では，要求発生時のみ計測を行う
    オンデマンド検知方式を導入
  - 検知コストの削減と検知精度の両立を目的として，
    性能およびコスト効果を評価する

# アプローチ（1/3）

- 本研究のアプローチ（オンデマンド検知）
  - ユーザ要求をトリガー：コンテンツ要求が発生した時のみ検知を実施
    - 人気コンテンツ：要求回数が多いため検査頻度が上がり、検知率を高められる
    - 不人気コンテンツ：不人気コンテンツはリクエストが少ないため、検査が最小限でコストを抑えられる

# アプローチ(2/3)

- 分布のエントロピーに基づく検知
  - KLダイバージェンスを利用：要求対象CID周辺のPeer ID（PID）分布を観測
    - 正常時の分布と比較
    - エントロピー差から攻撃の有無を判定

  - $D_{KL}(P||Q) = \sum_{i=1}^{n} p_i log_2 \frac{p_i}{q_i}$
    - P=$\{P_i\}$：観測分布
    - Q=$\{q_i\}$：基準分布

# アプローチ(3/3)

- 攻撃検知後の対応
  - KLダイバージェンスによりSybil攻撃を検知
  - 攻撃検知時にProvider Recordの追加配布を実行
  - 再配布にはSDS(SR-DHT-STORE)を適用
    - 近距離から遠距離まで段階的に分散配置
    - 距離が遠い層ほど配布数を削減
    - 攻撃範囲外にもProvider Recordを配置し，Sybilノードを回避

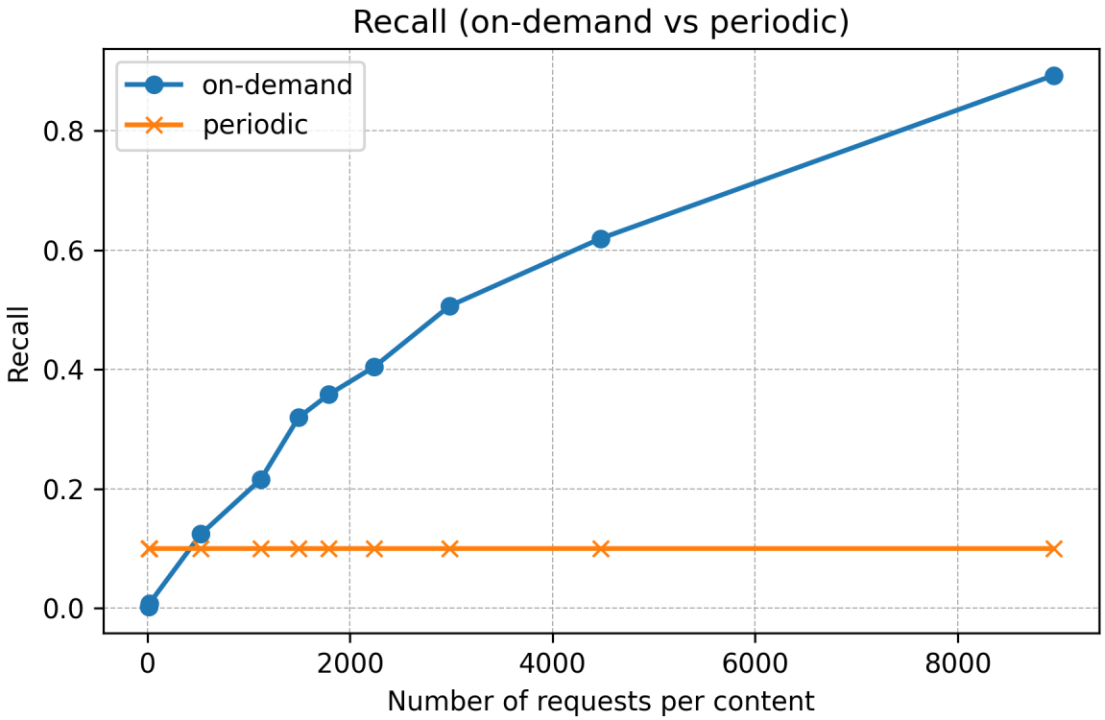
# 評価概要

- 評価概要
  - コンテンツ取得率を指標とした性能評価
  - 要求数に対する取得率の変化を分析

- ProviderRecordの配布制約条件
  - Provider Record の過剰配布を防ぐため，1ピアが追加で配布できる PR の数を最大5件に制限する
  - 追加配布を一度行ったピアは，180分間は新たな PR の再配布を行わない

# 性能評価（1/4）

- オンデマンド検知方式の特性により，人気コンテンツに対して多くの攻撃が検知される傾向が確認された。
- このように，影響の大きい人気コンテンツに重点的に検知を行う設計は，実運用を考慮した場合に合理的であるといえる。



# 性能評価(2/4)

- 1コンテンツあたりシビルノード最大100ノードで攻撃した時の評価
- シビルノードは20minutesごとに増減させてる
- 定期検知方式は30分間隔で実行
- 定期検知もオンデマンド検知も同じ条件で評価

| 定期検知 | |
|---|---|
| 真陽性(TP) | 1830 |
| 偽陰性(FN) | 16380 |
| 偽陽性(FP) | 21 |
| 真陰性(TN) | 176 |

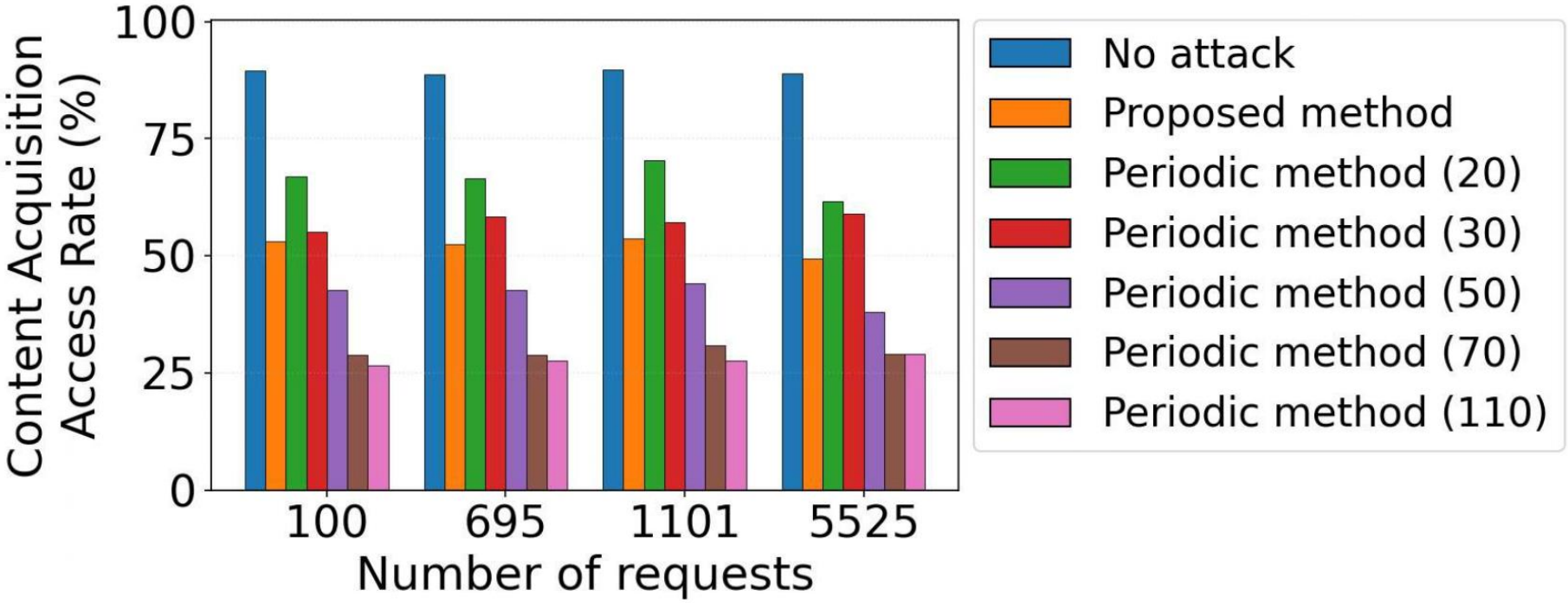
| オンデマンド検知 | |
|---|---|
| 真陽性(TP) | 6285 |
| 偽陰性(FN) | 11925 |
| 偽陽性(FP) | 60 |
| 真陰性(TN) | 137 |

| 定期検知 | |
|---|---|
| 検知率 | 0.1005 |
| 精度 | 0.9887 |
| 偽陽性率 | 0.1066 |

| オンデマンド検知 | |
|---|---|
| 検知率 | 0.3451 |
| 精度 | 0.9905 |
| 偽陽性率 | 0.3046 |

# 性能評価（3/4）

- 提案方式は，定期検知法を30分間隔で実行した場合とほぼ同等のコンテンツ取得成功率を達成できている。



# 性能評価（4/4）

- 全体結果の比較
- 攻撃なし環境：
  コンテンツ取得率：88.24%　平均ホップ数：2.17

- 既存研究手法（30分間隔）：（調査回数　330,000）
  攻撃対象コンテンツの平均取得成功率：57.33%

- 既存研究手法（110分間隔）：（調査回数　90,000）
  攻撃対象コンテンツの平均取得成功率：27.71%

- 提案手法：（調査回数　100,000）
  攻撃対象コンテンツの平均取得成功率：52.15%

# まとめ

- 提案手法により、検知コストを抑制しつつ、コンテンツ取得成功率および平均ホップ数を高い水準で維持することを実現
  - 既存の検知アルゴリズムでは防御コストが高くスケーラビリティに課題

- 今後の展望
  - 選任アルゴリズムなどの分散合意手法を導入することで，検知コストおよび Provider Record の再配布数をさらに削減することを目指す。

# 参考文献


- Active Sybil Attack and Efficient Defense Strategy in IPFS DHT (Nettoら, 2025, arXiv)
- Web3 Sybil Avoidance Using Network Latency（Stokkinkら, 2023, *Computer Networks*）
- Mapping the Interplanetary Filesystem（Henningsenら, 2020, arXiv）
- IPFS – Content Addressed, Versioned, P2P File System（Benet, 2014, arXiv）
- Sybil Attack Strikes Again: Denying Content Access in IPFS (Cholez & Ignat, arXiv)