

膨大な数の極小データの 効率的な配送基盤技術の研究開発

平成26年10月7日(火)

株式会社 KDDI研究所

国立大学法人 九州工業大学

国立大学法人 佐賀大学

株式会社 ネットワーク応用技術研究所

アウトライン

- 課題の位置づけ
- 研究開発成果
 - 課題(1)ビッグデータのネットワーク配送基盤技術
 - DPIを取り入れたパケット識別
 - パケット集約時のパケット構成方法
 - 課題(2)ビッグデータ用ネットワーク配送基盤の異常検出技術
 - 長周期通信(数分から数十分)の異常検知を10分で実現

課題の位置づけ

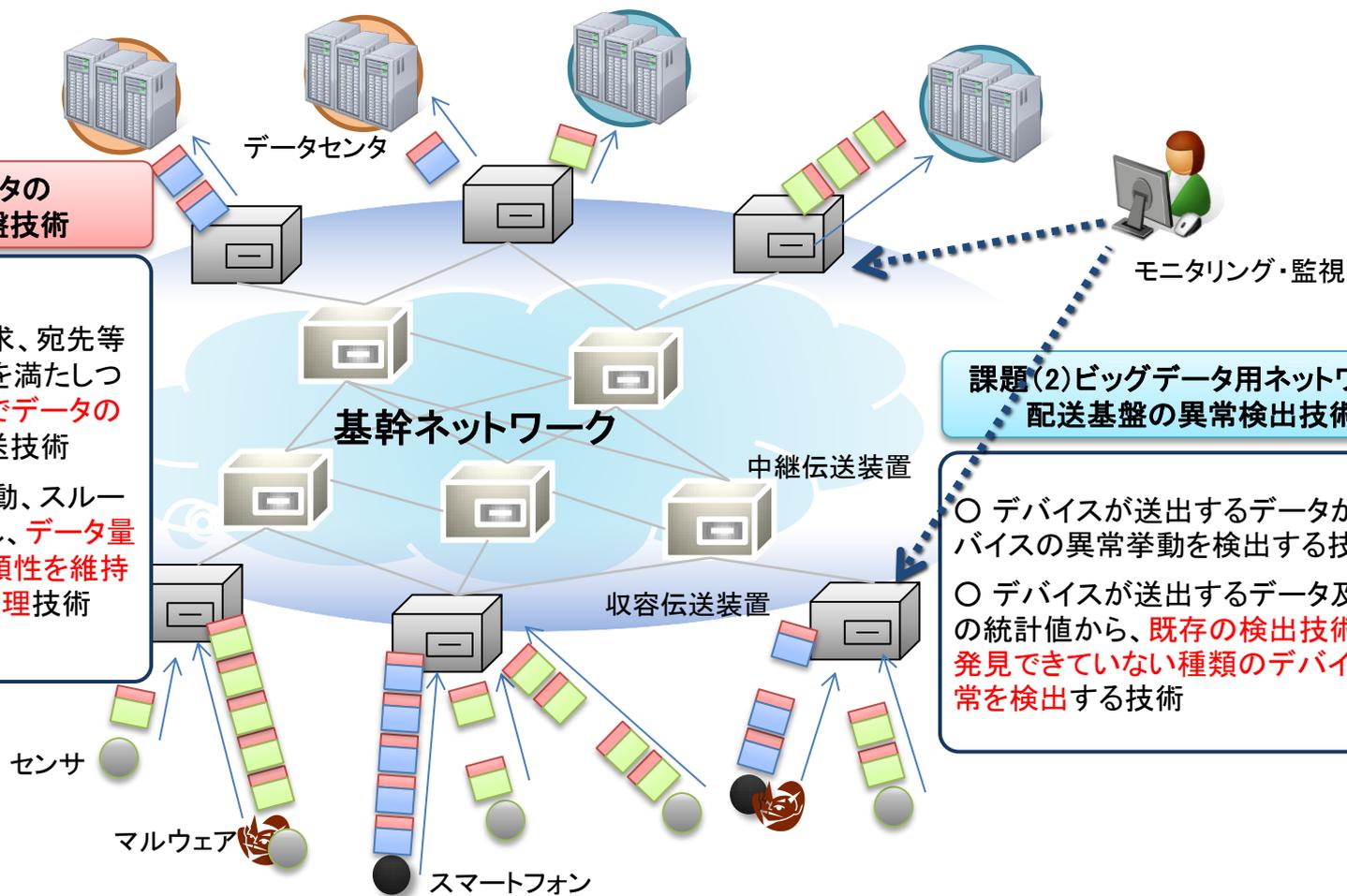
スマートフォンや小型化・低価格化が進むセンサ等の多様なデバイスによって生成される膨大な数の極小データを、信頼性を担保しながらサービス毎の要求品質を考慮して効率よくネットワークに收容する基盤技術等を確立。

課題(1) ビッグデータのネットワーク配送基盤技術

- データの種類や品質要求、宛先等の情報に基づき通信品質を満たしつつ、**基幹ネットワーク全体でデータの收容能力を増大**させる配送技術
- 通信特性情報(遅延変動、スループット等)を効率的に収集し、**データ量が増えても通信品質や信頼性を維持するスケーラブルな配送管理技術**

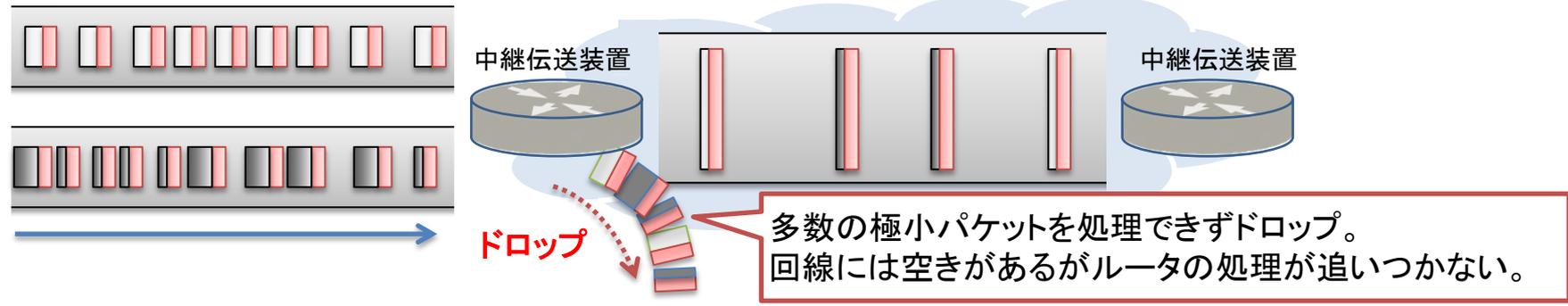
課題(2) ビッグデータ用ネットワーク配送基盤の異常検出技術

- デバイスが送出するデータからデバイスの異常挙動を検出する技術
- デバイスが送出するデータ及びその統計値から、**既存の検出技術では発見できていない種類のデバイス異常を検出**する技術

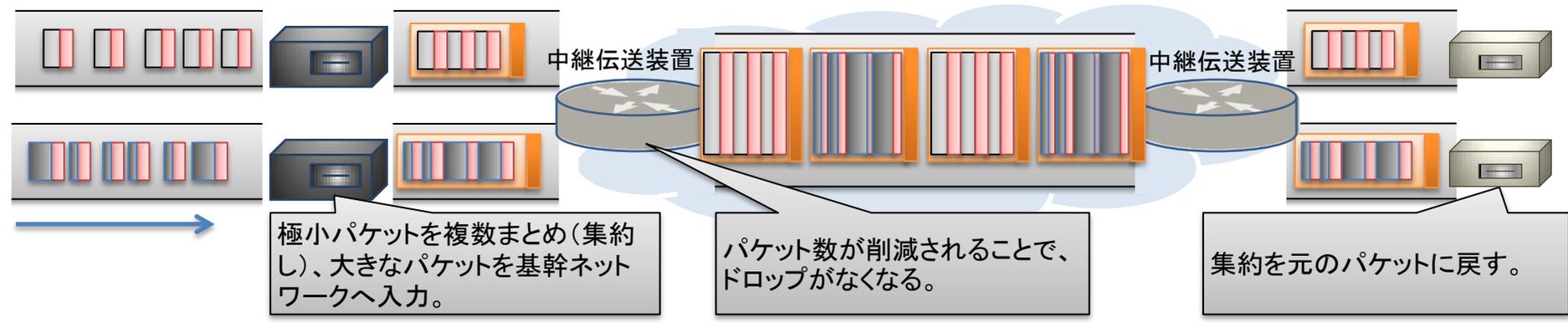


課題(1)のポイントと開発技術の概要

○ 既存技術



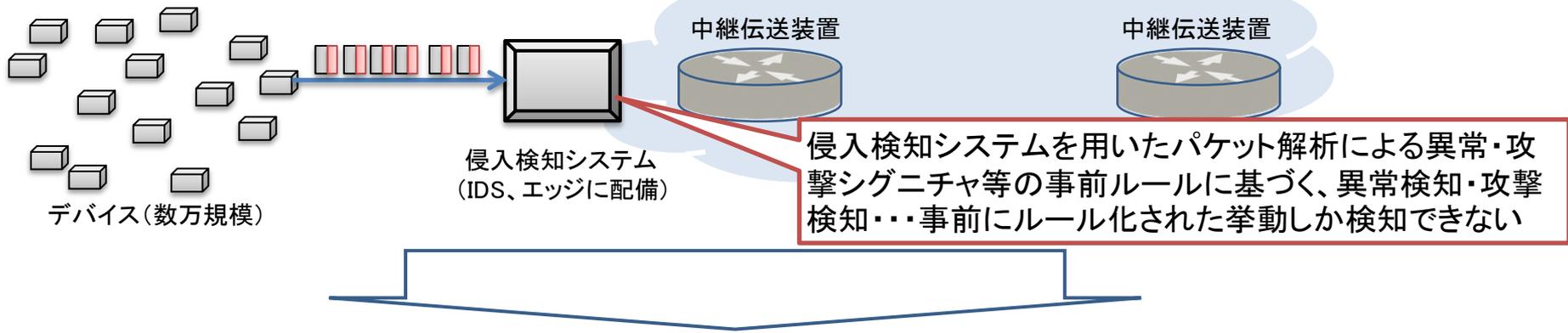
○ 本研究開発で確立する「ビッグデータ配送基盤技術」



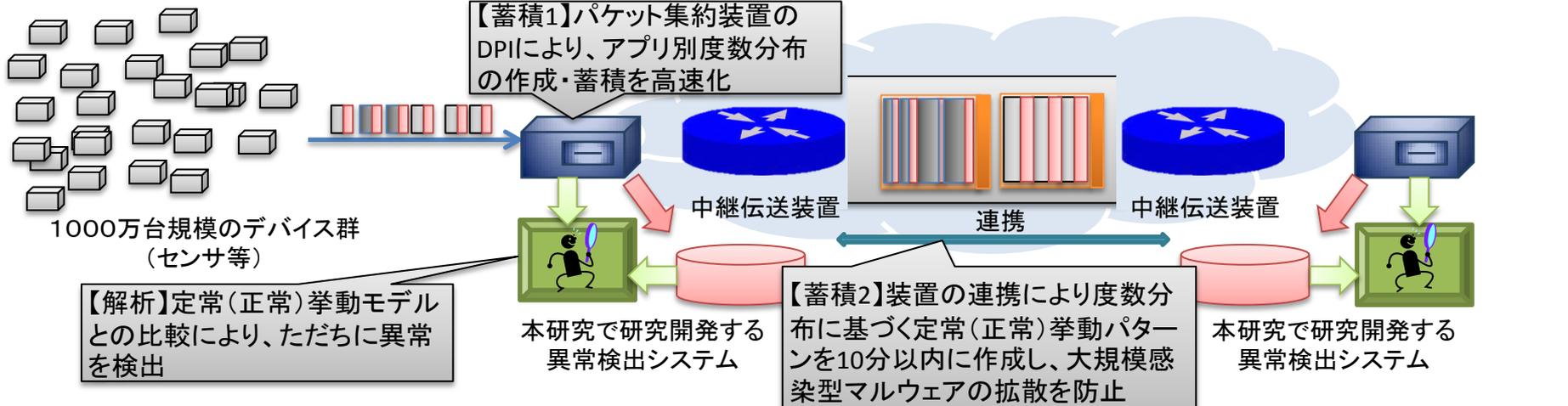
スマートフォンやセンサ等の多様な無線通信デバイスによって生成されるデータの種類や品質要求、宛先等の情報に基づき通信品質を満たしつつ、基幹ネットワーク全体でデータの收容能力を増大させる配送技術
→基幹ネットワークで用いられる既存の中継伝送装置当りのフロー收容能力を5倍以上に向上

課題(2)のポイントと開発技術の概要

○ 既存技術



○ 本研究開発で確立する「ビッグデータ配送基盤の異常検出技術」

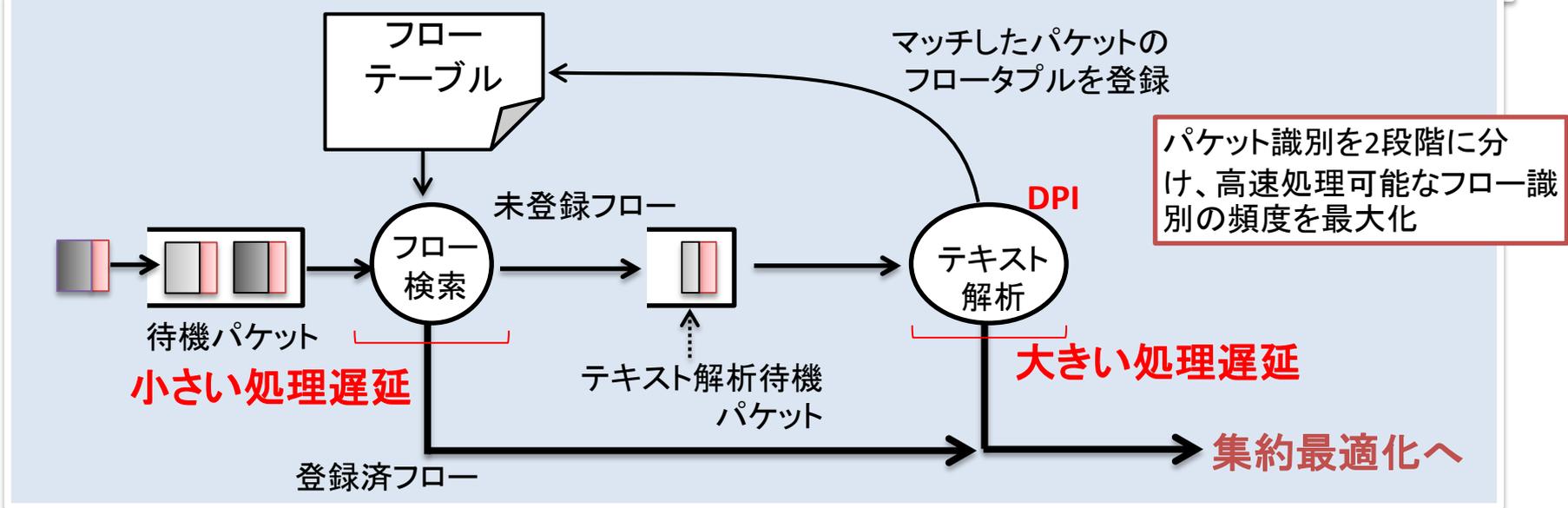
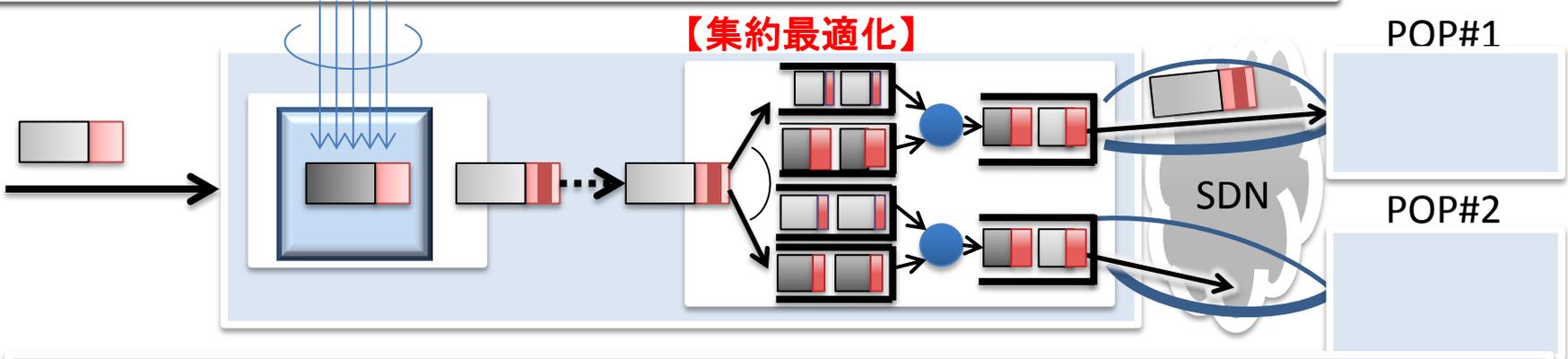


デバイスが送出するデータを蓄積し、デバイスの挙動又はデバイスが送出するデータのヘッダ情報、サービス種別及び通信特性情報などにより挙動を分析することで、ただちに挙動異常を検出
 →ネットワークに接続された3000万台のデバイスの挙動異常を検出

課題(1)ビッグデータのネットワーク配送基盤技術

【高速DPI】

```
If ((ip.dst == 10.1/16) && http.content_type contains "video/mp4")  
{POP ID =X; Content ID =Y}
```



課題(1)ビッグデータのネットワーク配送基盤技術

【課題解決の方向】

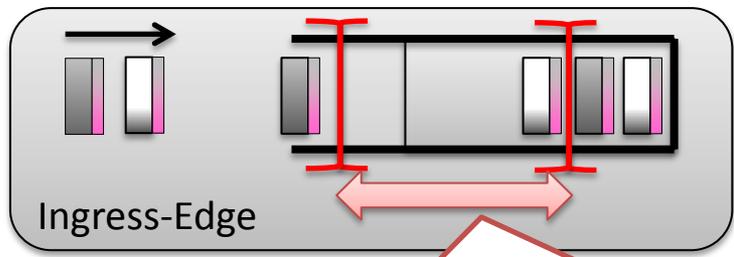
- 通信品質を維持可能な集約待機時間を調整

【課題解決の手法】

- 上記の連携により調整した集約待機時間を基に到着パケットを クラス分けし、優先度が高いパケットから順にパケットを集約

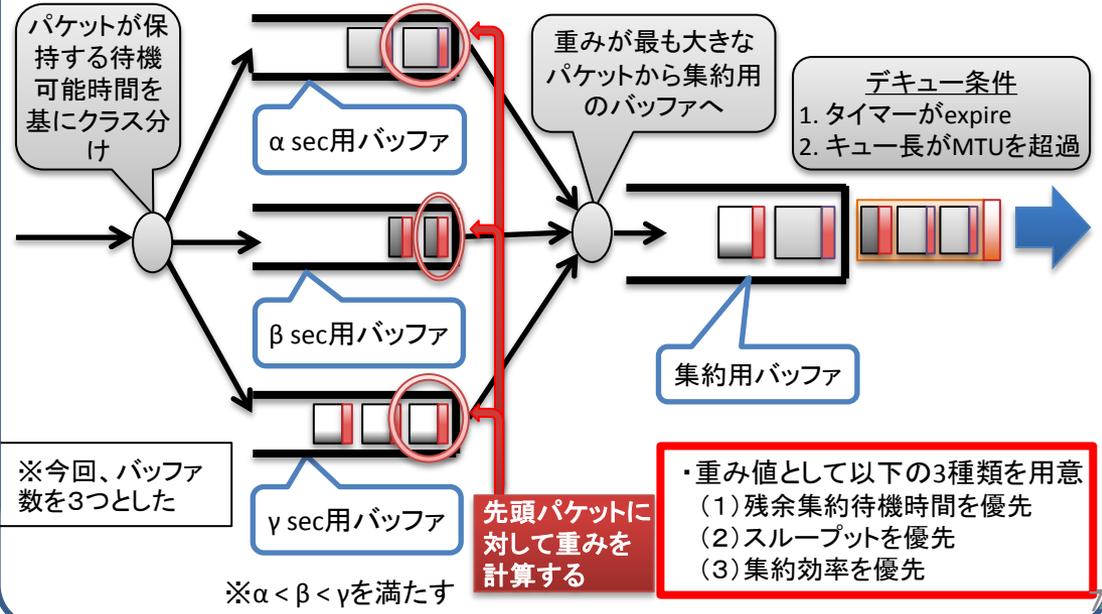
⇒到着トラヒックの過負荷時や偏在時にも適切にパケットを集約

(1) 集約待機時間調整手法に関する研究



- コアネットワークの変化に応じて集約待機時間を調整
- 遅延時間の変化
 - 遅延時間増加 → 集約待機時間短縮
 - 流入パケットレートの変化
 - パケットレート増加 → 集約待機時間拡大

(2) 集約パケット構成方式に関する研究開発



※今回、バッファ数を3つとした

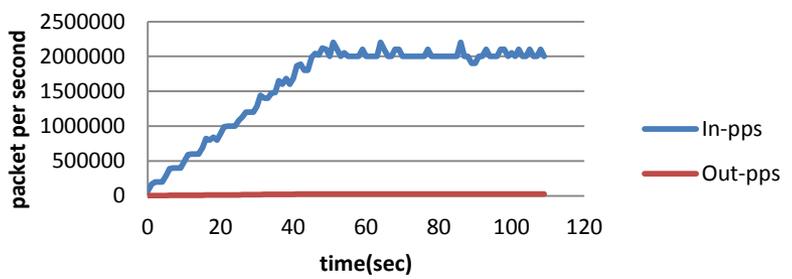
※ $\alpha < \beta < \gamma$ を満たす

- ・重み値として以下の3種類を用意
 - (1) 残余集約待機時間を優先
 - (2) スループットを優先
 - (3) 集約効率を優先

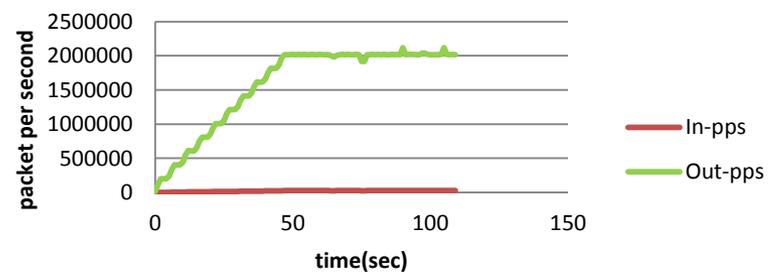
課題(1)ビッグデータのネットワーク配送基盤技術



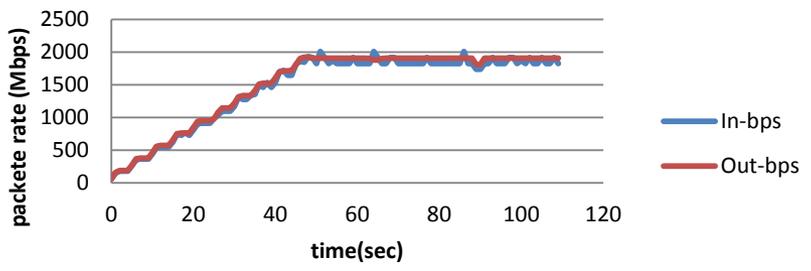
パケットレート @ ingress



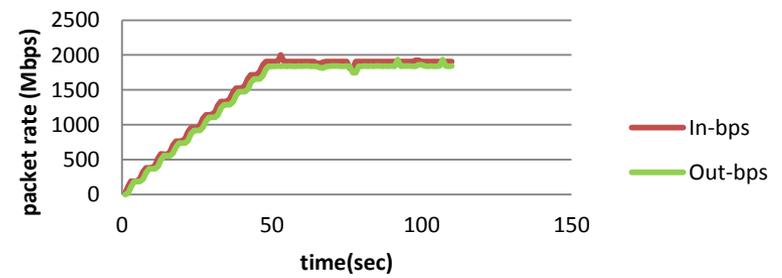
パケットレート @ egress



ビットレート @ ingress

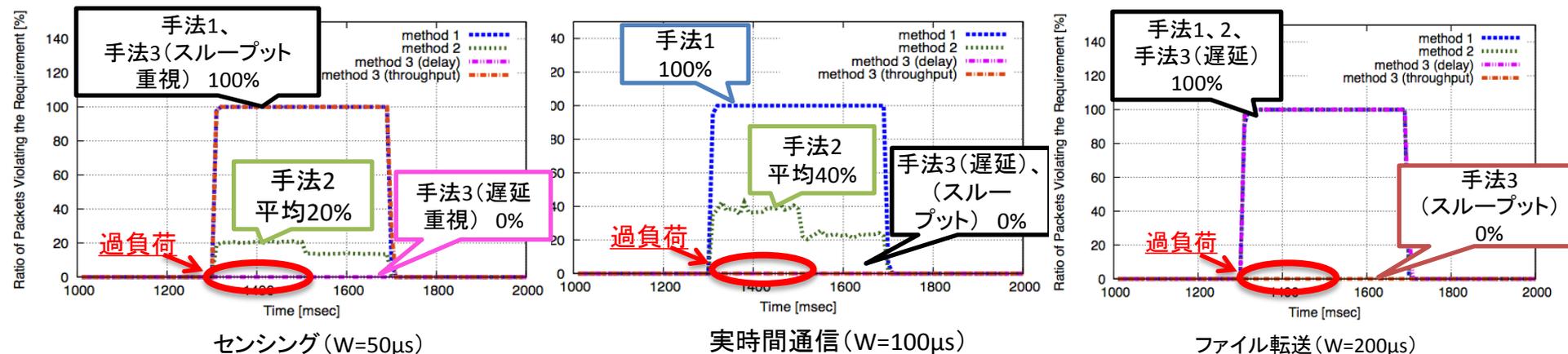


ビットレート @ egress



課題(1)ビッグデータのネットワーク配送基盤技術

- (1)入カトラヒックが過負荷の場合、(2)トラヒック量に偏りがある場合を評価
- 手法1(バッファ数1)、手法2(バッファ数3、各バッファで集約)と比較



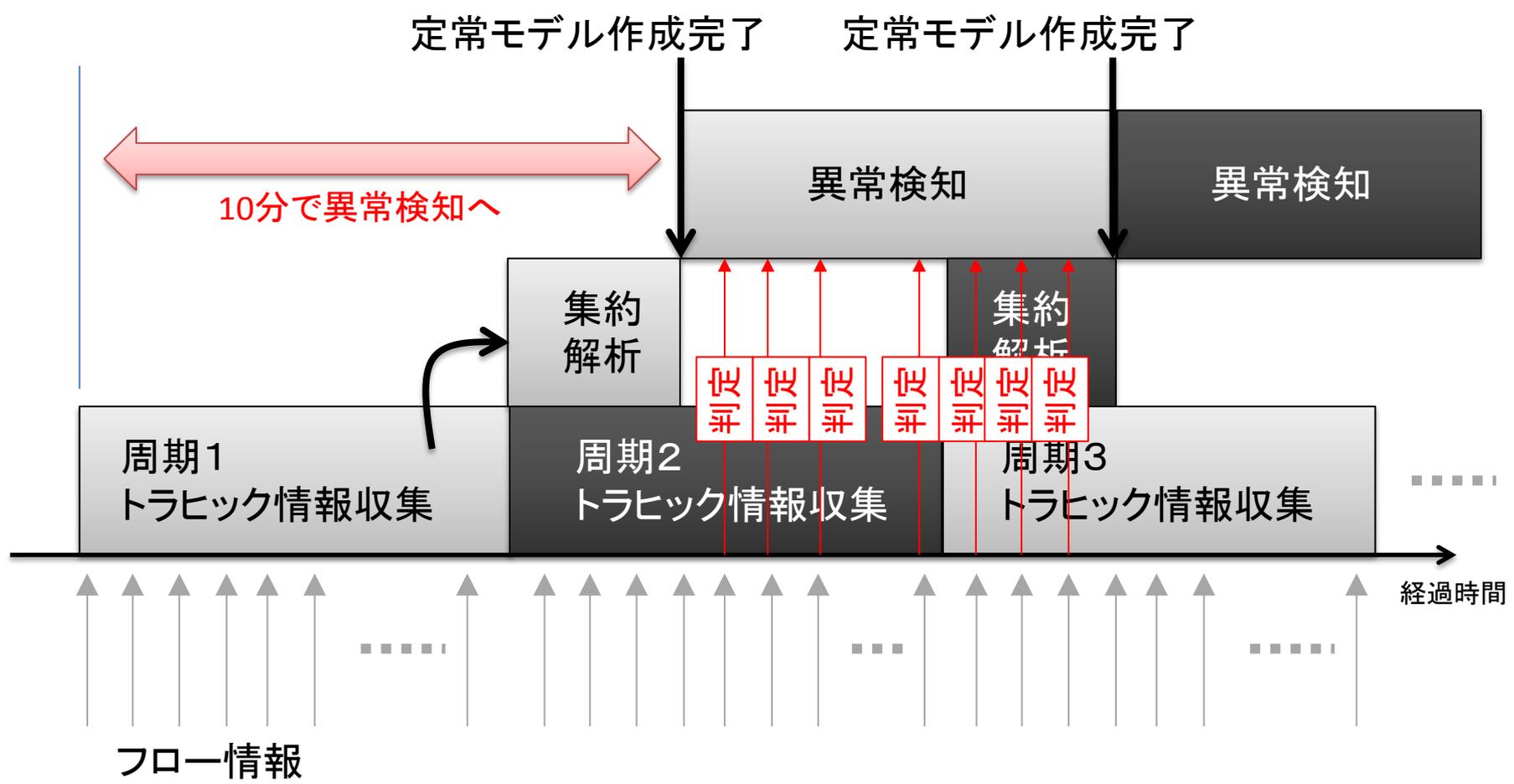
【成果】

トラヒックが過負荷時の結果

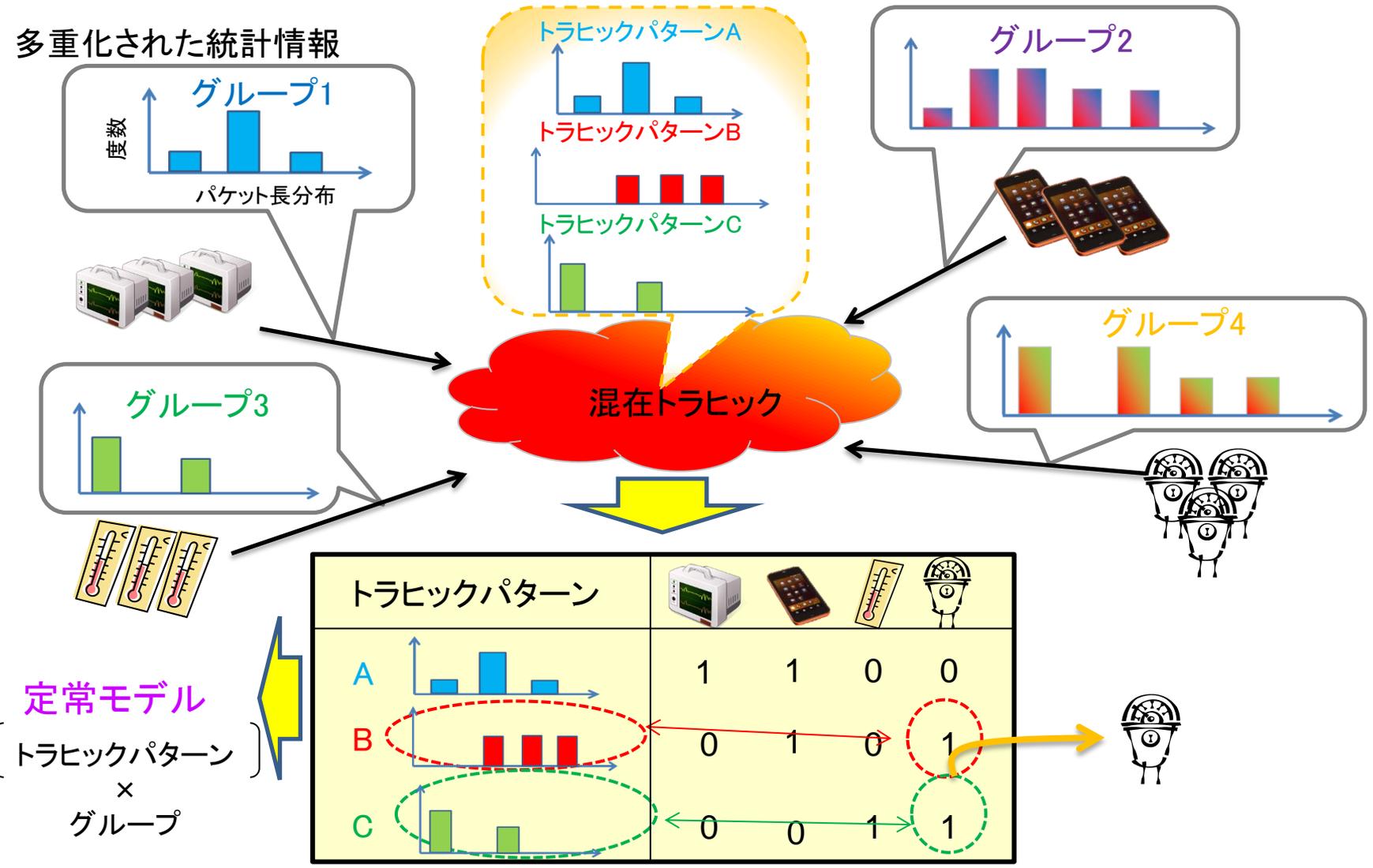
- 通信品質を維持しつつ、パケット数を削減するための集約待機時間に基づく集約パケット構成手法を考案し、実装した
- トラヒックの過負荷時、過小時、偏り時にも対応できることを確認
 - 過負荷時 ⇒ 複数バッファによるクラス分け
 - 過小時 ⇒ 複数バッファから集約パケットを構成するためのバッファを用意
 - 必要バッファ数 ⇒ 要求が類似するアプリ群ごとにバッファを用意することで削減可能

課題(2):ビッグデータ用ネットワーク配送基盤 の異常検出技術

収集～集約～解析、の流れ

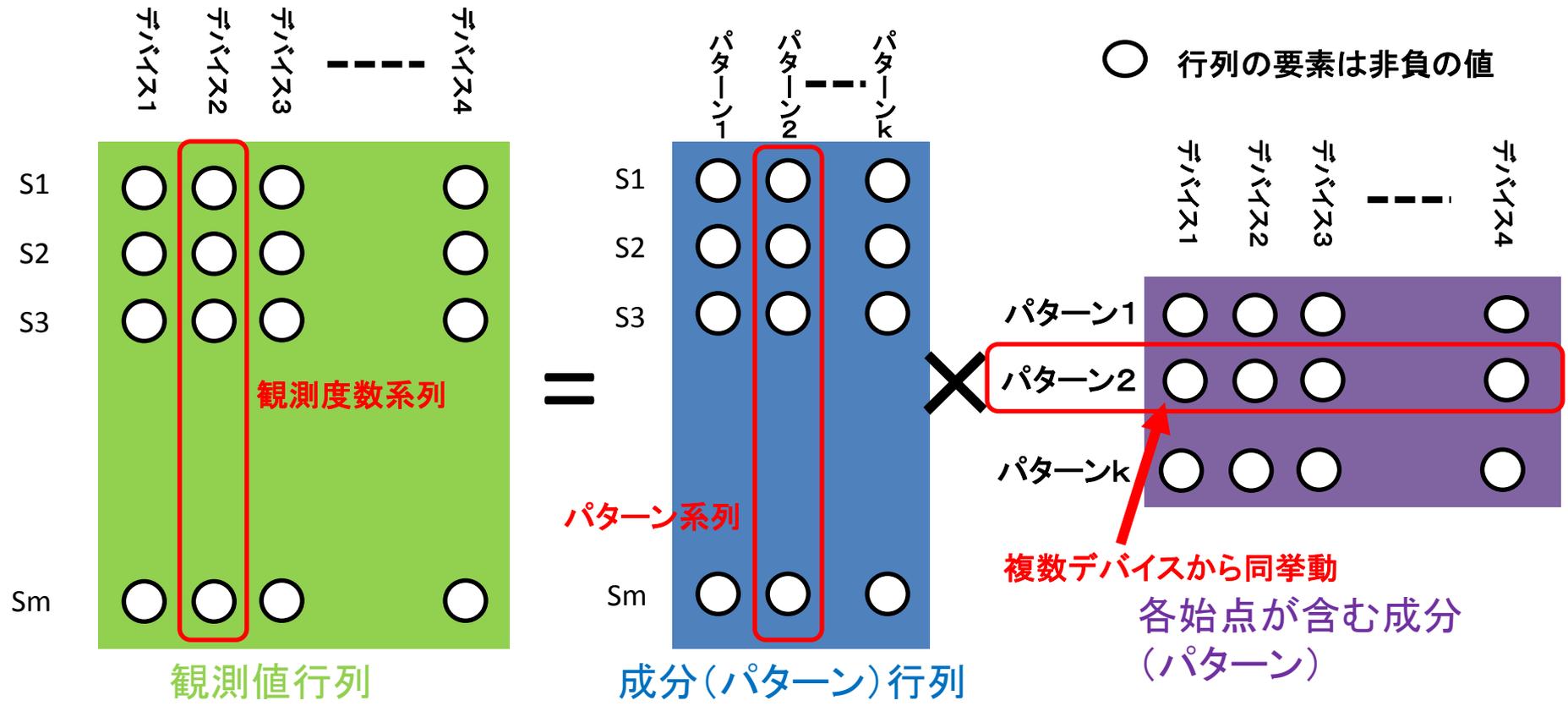


課題(2):ビッグデータ用ネットワーク配送基盤 の異常検出技術



課題(2):ビッグデータ用ネットワーク配送基盤の異常検出技術

NMF (Non-Negative Matrix Factorization = 非負値行列分解)



V を **W** と **H** に分解 (W, Hを生成) → 成分(パターン)を抽出
 基底行列 係数行列

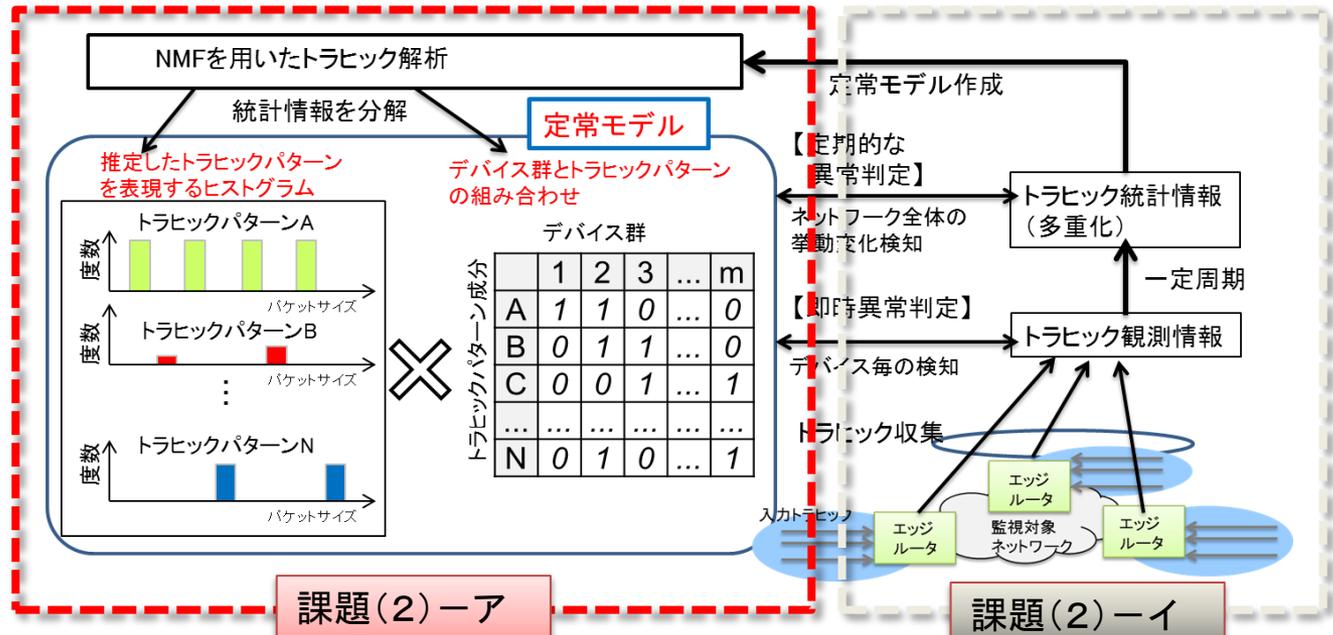
課題(2)-ア: パケット集約トラヒックの異常検出技術の成果(1/2)

【課題解決の方向】

- 複数デバイスから送出されるトラヒックを一括して解析することで、高速な定常モデル作成

【課題解決の手法】

- 複数デバイスからのトラヒックに、NMF（非負値行列因子分解）を適用することで、トラヒックのクラスタリングと、ヒストグラム生成を同時に実施
- NMFを高速に実行するために、GPUにより実装



研究開発全体:目標

1. ビッグデータのネットワーク配送基盤技術
基幹ネットワークで用いられる既存の**中継伝送装置当りのフロー収容能力を5倍以上に向上**
 - フロー: 数 bps ~ 数十 bps のデータを送信するデバイスの通信
2. ビッグデータ用ネットワーク配送基盤の異常検出技術
3000万台デバイスがデータを送出する環境において、挙動異常を検出するのに必要な**データ蓄積の時間を10分間以内**
 - 平均15分間に1回の間隔でデータが送信されると仮定

実施体制

研究責任者：中島 康之 (KDDI研究所)

(1) ビッグデータのネットワーク配送基盤技術

研究リーダー：北辻 佳憲 (KDDI研究所)

研究者：池永 全志、中村 豊、大西 圭、福田 豊、塚本 和也

野林 大起 (九州工業大学)

金 大善、梅木 智光 (KDDI研究所)

(2) ビッグデータ用ネットワーク配送基盤の異常検出技術

研究リーダー：中村 勝一 (ネットワーク応用技術研究所)

研究者：堀 良彰、フォン ヤオカイ (佐賀大学)

井上 達、永田 晃、富永 浩之、元田 恵一、森山 敦文、

Thin Thin Nwe、稲田 和彦、高石 良一、大石 一夫

(ネットワーク応用技術研究所)