

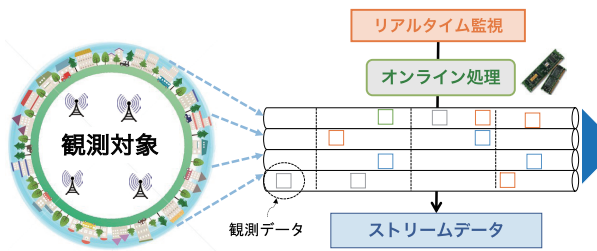
【代表的な研究テーマ】

□ **ストリームデータの要約**

□ **データ駆動の科学発見**

キーワード：リアルタイム分析、異常検知、省メモリアルゴリズム、ストリームデータ基盤

- クラウドサービスの普及とIoTの発展に伴い、人や社会に関わる多様な事象のデータ利活用が広がっています。観測系から常に生成され続けるストリーム型のデータでは、観測系の変化や異常をいち早く検出するリアルタイム解析への応用が期待されています。
- 本研究では、このようにストリーム型ビッグデータのインメモリ管理を実現するソリューションとしてデータ要約 (summary) と呼ばれる情報圧縮技法を扱っています。
- サマリ**と呼ばれる**超軽量データ構造**を構築することで、任意の関係クエリに対し、高速に応答することが可能となります。

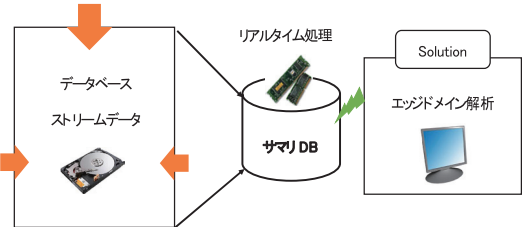


サマリの構築法

- 元のデータベースそのまま
- 可逆圧縮 (きちんと管理)
 - 空間計算量 & 時間計算量は良くても $O(n)$ (n はデータ量)
- 非可逆圧縮 (ゆるく管理)
 - 劣線形計算量 $\sim O(\log n)$ を目指す (超軽量 & 超高速!)
 - リアルタイム解析に適したオンラインアルゴリズム

S	気温 a_1	降水量 a_2	風速 a_3	風向 a_4	日照 a_5
e_1	14	5	3	0	10
e_2	11	10	1	0	5
e_3	19	5	2	1	14
e_4	13	3	2	1	14

気温 12℃ 以上 & 降水量 5 mm 以下だった過去の日数?

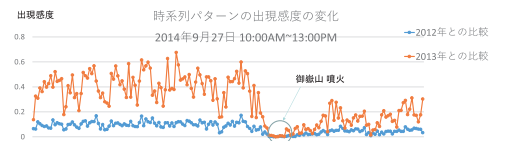
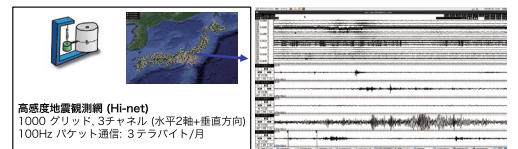


- 自然科学データ (地震観測網データ、天文測光データ) の突発現象をリアルタイム検知する共同研究を推進しています。
- 極地や閉システムのリアルタイムデータ解析**、例えば予兆検出や異常検知に興味のある方はお声がけください。
- スマートファクトリーやバイオDXといった融合領域研究も進めています。

参考文献

P. Thanapol, Y. Yamamoto and S. Sako:
 Detection for transient patterns with unpredictable duration using Chebyshev Inequality and dynamic binning, Proc. of CANDAR WANC, 2021

Y. Yamamoto, Y. Tabei and K. Iwanuma:
 Parasol: PARASOL: a hybrid approximation approach for scalable frequent itemset mining in streaming data, Journal of Int. Inf. Systems, 2020



■ その他の社会連携活動

- 静岡大学-ヤマハ発動機共同研究講座スマートファクトリー領域担当
- 理化学研究所AIP客員研究員
- 講演「姿勢推定を用いた組立作業の工程分解」(精密工学会第4回画像応用技術専門委員会, 2021)

■ 相談に応じられる関連分野

- 自然科学データのリアルタイム分析
- スマートファクトリー、バイオDX
- ストリームデータ基盤開発



山本 泰生
 学術院情報学領域
 情報科学系列
 准教授