

修士学位論文要約（平成29年3月）

# オンラインデータ解析のためのエージェント指向ストリーム処理に関する研究

山田 良介

指導教員: 木下 哲男, 学位論文指導教員: 笹井 一人

Research for agent-oriented stream processing for on-line data analysis

Ryosuke YAMADA

Supervisor: Tetsuo KINOSHITA, Research Advisor: Kazuto SASAI

As network devices increase in recent years, it becomes more important to extract useful information from large-scale data collected from these devices. For efficient data processing in distributed computers, stream processing is developed, which addresses data as data stream, and data processing as chain of small data processing component. But constructing stream processing is hard work. In order to solve this problem, we proposed agent-oriented stream processing framework, which helps constructing stream processing.

## 1. 序論

Web サービスの発達や、家電やセンサをネットワーク接続する Internet of Things(IoT) の取り組みなどにより、ネットワーク機器が増大、多様化し、それらの監視、管理が困難になってきている。この問題を解決するため、それらネットワーク機器から得られるデータを収集しデータ解析を行うことで管理者に対しより高度な情報を提供することが考えられている [1].

一方で従来のデータ解析システムは大規模なデータベースにデータを蓄積することを前提とし、また収集したデータに対し一括で処理を行うため計算負荷が大きく、解析に関わる機器負担が大きいものであった。例として、蓄積したネットワークデータに対する複数の統計手法の適用による攻撃検知手法 [2] が挙げられる。この問題を解決するものとして、近年ストリーム処理が注目されている。これはデータを連続的入力(ストリーム)として受け取り、オンラインで処理を行うことでストレージを必要とせず、また処理を分割、細分化し分散計算機環境上へ配置することで複数計算機を利用した効率的実行を可能とするものである。しかしストリーム処理を利用してデータ解析を行う上では、分割した処理間での入出力仕様の整合性をとる必要があり、設計の負担が大きい。また実装の上では各計算機の性能、特性を把握しそれぞれに適切な設定、処理分配、処理経路の構築を行う必要があり、実装の負担も大きいものであった。またデータ解析の上ではパラメータを変更しての試行を繰り返すことがあるが、ストリーム処理においては処理に関わる各機器への再設定が必要となり、複数回の試行にあたって負担を要する。

そこで本研究ではデータ解析のためのストリーム処理系を容易に設計、構築することを目的とした、エージェント指向ストリーム処理基盤を提案する。

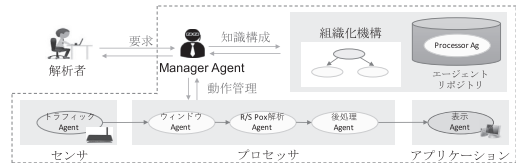


図1 提案基盤の概要

## 2. エージェント指向ストリーム処理基盤の提案

本稿では、エージェント間連携によりユーザー要求に応じてストリーム処理が能動的に構築されることによりユーザの設計、実装負担を軽減することを目指す。本稿における提案基盤の概要を図1に示す。本提案基盤は以下の2要素よりなる。

### (S1) ストリーム処理系における処理要素のエージェント化

ストリーム処理において、分散処理のために分割された処理要素それぞれをエージェントとし、プロセッサエージェントと呼称する。これは解析処理を呼び出す機能及びパラメータ設定機能を持ち、他エージェントからのデータ入力に応じて必要なパラメータを与えた解析処理を起動し、結果を次のエージェントへ出力するものである。エージェント間通信によりデータ入出力に必要な接続情報を取得し、設定を行うことで処理経路を確立する。また経路自身の行う処理及び必要パラメータについての知識を持ち、後述の組織化において使用する。

### (S2) 処理要求に応じたエージェントの組織化

ユーザの処理要求に応じて計算機環境上に能動的に処理を構築するために、契約ネットプロトコルによるエージェント組織化を導入し、そのためのエージェントとしてマネージャエージェントと呼称するエージェントを導入する。マネージャエージェントはユーザからの要求を受け、予めプロセッサエージェントが格納されているエージェントリポジトリに対しタ

スク通知を行う。プロセッサエージェントは自身の知識に照らし、タスク実行可能ならば応答する。マネージャエージェントは応答したエージェントから適切なエージェントを選択して契約を行い、環境上に生成する。

以上の2要素によりユーザ要求に応じた適切なエージェントが環境上に生成され、ストリーム処理系が構築される。

### 3. 設計

提案に基づき、マルチエージェントフレームワーク DASH/IDEA を用いて試作システムの設計と実装を行った。ストリーム処理におけるデータ入出力は ZeroMQ プロトコルに基づく TCP ソケット通信として実装され、通信に必要なホスト名とポート番号は組織化において契約終了直後に当該エージェント間で直接メッセージ交換を行うことで受け渡される。

### 4. 実験と評価

試作システム上で正常に解析が行えることを確認し、本提案基盤の有効性を示す実験を行った。時系列パケット量のデータに対する特性解析手法である R/S 解析 [3] を題材とし、これをユーザ要求に基づき提案基盤上で構築、実行されることを確認した。

本実験では R/S 解析における処理要素を、パケットキャプチャファイルを入力しパケット数のカウントを行う Traffic Count, ストリーム展開を行う Flat Map, 時間解像度を決定する Sum, ウィンドウ処理を行う Time Window, R/S 解析を行う RS Pox の5要素に分割し、それぞれをプロセッサエージェントとして構築しエージェントリポジトリに格納した。これらプロセッサエージェントをマネージャエージェントである Manager により組織化し、解析処理を構築した。実験の様子を図2に示す。ユーザ入力を代行する Make RS Topology エージェントを予め初期状態として配置し、このエージェントが持つユーザインターフェースを通してユーザ要求を受け、マネージャエージェントが生成された。そしてマネージャエージェントによる組織化により適切なエージェントが選択され、ストリーム処理系が構築されることが確認できた。

この処理系に対しパケットキャプチャファイルによる入力を与え、得られた出力をグラフ上にプロットすると、図3のような R/S Pox Diagram と呼ばれる R/S 解析特有の結果が得られた。この結果により、構築されたストリーム処理系が正しく動作することが確認された。

以上の結果により、ユーザ要求に応じて適切なストリーム処理形を能動的に構築することが可能であることがわかり、本提案基盤の有効性が示された。

### 5. 結論

本稿では、データ解析のためのストリーム処理系を容易に設計、構築することを目的とし、エージェント指向ストリーム処理基盤を提案した。そして提案に基づき試作システムを設計、実装し、R/S 解析

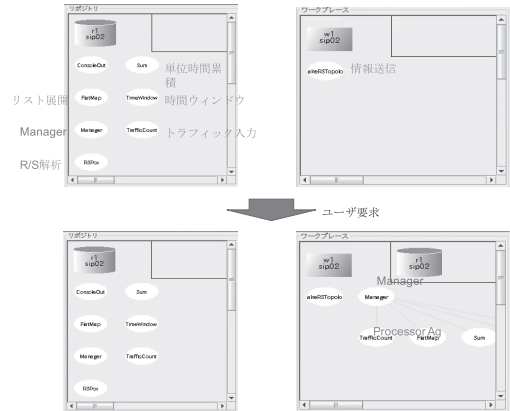


図2 実験結果

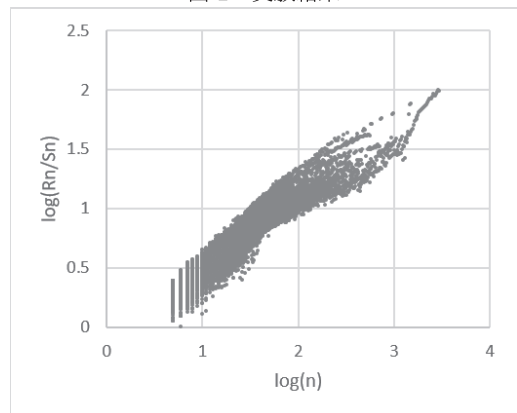


図3 実験より得られた出力

を題材としてその有効性を示す実験と評価を行った。

今後の課題として、ユーザにより多くの情報を与え、処理系構築を助ける対話インターフェースの構築、プロセッサエージェントの再利用性を高める設計知識の階層化などが必要となると考えられる。

### 参考文献

- 1) K. Kalegele, K. Sasai, H. Takahashi, G. Kitagata, T. Kinoshita, "Four Decades of Data Mining in Network and Systems Management," *IEEE Transactions on Knowledge and Data Engineering*, vol. 27, no. 10, pp. 2700–2716, 2015.
- 2) Lin Li, De-bao Xiao, "Research on the network security management based on data mining," *2010 3rd International Conference on Advanced Computer Theory and Engineering*, vol. 5, pp. 184–187, 2010.
- 3) 高橋秋典, 五十嵐隆治, 上田浩, 岩谷幸雄, 木下哲男, "R/S Pox レッグライン特性," 情報科学技術フォーラム講演論文集, vol. 11, no. 4, pp. 9–16, 2012.