

政府関係機関事業紹介

広域データ収集・解析プログラム開発を支援するソフトウェアパッケージ「SINETStream」

国立情報学研究所 学術基盤推進部 学術基盤課 クラウド支援室

国立情報学研究所（以下、NII）では、学術情報ネットワーク「SINET5」を介して広域に分散するデータを収集・解析する研究を支援するためのソフトウェアパッケージ「SINETStream」を開発し、2019年12月にオープンソース・ソフトウェアとして公開しました。NIIが実証実験を進めているSINETとモバイル通信環境を直結した「SINET広域データ収集基盤」と合わせて利用することで、IoT等に関わる研究者が必要とするプログラム開発が容易になり、Society 5.0の実現につながる広域に分散したデータを対象とする研究に取組みやすくなります。本稿では、広域ネットワークを介したデータの収集・解析の課題と、SINETStreamの概要について説明します。

1. 広域ネットワークを介したデータの収集・解析の課題

環境測定、生体観測、IoTなど、広域に分散したデータを活用する研究では、広域ネットワークを介して、センサー等から取得されるデータを欠損なく確実に収集し、解析に用いることが求められます。NIIでは、そのような研究を支援するため、学術情報ネットワークSINET5とモバイル通信環境を直結させた「SINET広域データ収集基盤」^[1]を構築し、実証実験を進めています。SINET広域データ収集基盤を利用することで、センサー等と解析を行うクラウドや大学等に設置されたサーバの間を隔離された安全なネットワークVPN（Virtual Private Network）で接続することができるようになりました。しかしながら、データの収集や解析を行いたい研究者にとって、広域ネットワークやVPNを介してデータを収集・解析するプログラムを作成することは、ネットワークに関する高度な知識やプログラミングスキルが必要とされ、容易ではありませんでした。

この問題に対応するために、広域データ収集・解析プログラム開発支援ソフトウェアパッケージ「SINETStream」^{[2][3]}を開発しました。SINETStreamは、広域ネットワークを介してデータを欠損なく確実に収集・解析するための機能を提供するソフトウェアパッケージです。具体的には、

- センサー等から収集されるデータをクラウドや大学等に設置されたサーバへ書き込む
- サーバに収集されたデータを解析プロ

ラムに読み出す

ための機能を提供します。研究者は、SINETStreamが提供するAPI（Application Programming Interface）を利用すると、広域ネットワークを介したデータの収集・解析を行うためのプログラムを容易に開発することができます。また、研究に用いられるデータには機微な情報が含まれる場合もありますが、SINETStreamにはセンサー等のデバイスの認証・認可と通信やデータの暗号化を行う機能も含んでいるため、安全な広域データ収集を実現できます。SINET広域データ収集基盤とSINETStreamを合わせて利用することにより、Society 5.0の実現につながる広域に分散したデータを対象とする研究環境を容易に実現できます。

2. SINETStreamの概要

図1に、SINET広域データ収集基盤上でのSINETStreamの利用イメージを示します。研究者は、SINETStreamのAPIを利用してセンサー用と解析用のプログラムを作成することにより、センサーからのデータを確実に収集し、解析できるように

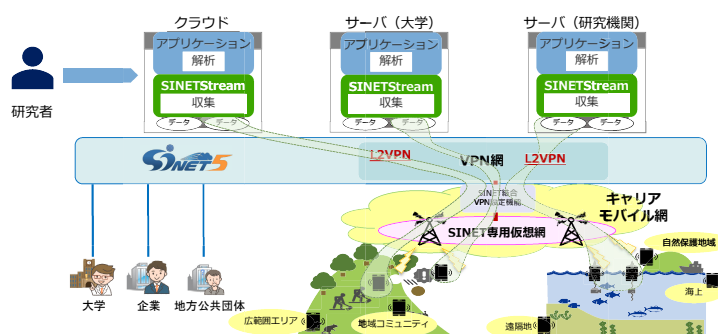


図1 SINET広域データ収集基盤上でのSINETStreamの利用イメージ

なります。また、SINETStreamのセキュリティ機能を利用することで、より安全な解析プログラム環境が構築できます。ここでは、SINETStreamのAPIおよびセキュリティ機能について説明します。

(1) SINETStreamのAPI

SINETStreamでは、「トピック」を用いた出版－購読型の非同期メッセージングモデルを採用しています。データの送信（出版）者と受信（購読）者との間にメッセージブローカーを仲介させることで、データの確実な収集を可能にします。トピックは、メッセージを送受信するときの識別子であり、各種センサーで生成される温度、湿度、画像などのデータストリームごとに定義します。APIを用いてトピック名を指定してデータの読み書きの手続きを行うことで、各データのセンサーからの書き込みと、解析プログラムでの読み出しが容易に行えます。

SINETStreamのバックエンド（SINETStreamサーバ）には、MQTT（Message Queuing Telemetry Transport）等の既存のメッセージブローカーを利用することができます。解析プログラムの可搬性を高めるため、SINETStreamでは各メッセージブローカーとの通信に必要なパラメータは別途設定ファイルで定義できるようにしています。

(2) SINETStreamのセキュリティ機能

より安全な解析プログラム環境を構築するために、認証・認可と、通信やデータの暗号化の機能を提供しています。これらの機能は、SINETStreamの設定ファイルでのパラメータの指定のみで利用できるようにしています。

SINETStreamでは、「認証」は書き込み、読み出しをするセンサーや解析プログラムが適切なものであるかを確認する手続き、「認可」はそれらに対してデータ等へのアクセス権限を制御する手続きを表します。認証・認可機能を利用するには、SINETStreamサーバで利用するメッセージブローカーでの設定も必要になります。主要なメッセージブローカーでの設定方法については、SINETStreamのウェブサイト^[2]での情報共有を行っています。

通信の暗号化はTLS（Transport Layer Security）、データの暗号化はSINETStream独自の機能として提供しています。通信の暗号化は、通信時のみ暗号化されるのに対し、データの暗号化ではセンサーまたは解析プログラム内部で暗号化されるた

め、SINETStreamサーバ内部でもデータが暗号化された状態で管理できます。暗号化アルゴリズム等は設定ファイルで指定できます。

(3) SINETStream利用事例

SINETStreamのウェブサイトでは、図2に示すような温度、湿度センサーデータを随時アップロードして可視化するライブデモを公開しています。SINETStreamのAPIを用いて取得された温度、湿度データを随時SINETStreamサーバに書き込み、解析プログラムで平均値を計算した結果をJavaScriptで随時読み出して可視化しています。

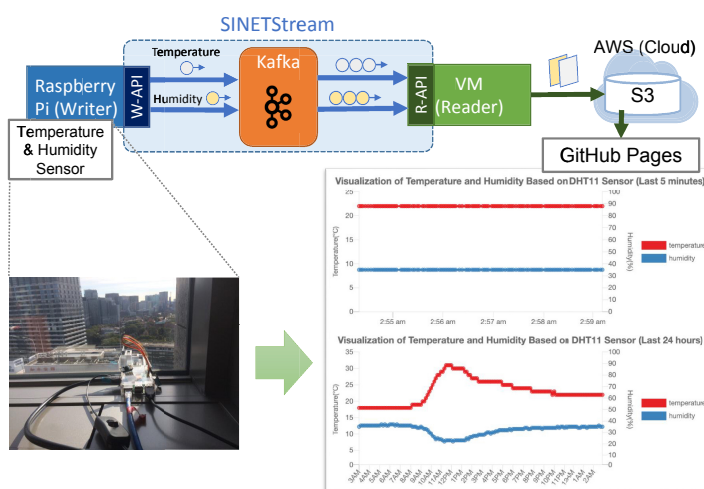


図2 SINETStreamのライブデモ

3. まとめ

本稿では、広域ネットワークを介したデータの収集・解析の課題と、SINETStreamの概要について説明しました。SINET広域データ収集基盤とSINETStreamを合わせて利用することにより、Society 5.0の実現につながる広域に分散したデータを対象とする研究環境を容易に実現できます。

NIIでは、SINETStream 1.0を2019年12月24日よりオープンソース・ソフトウェアとして公開しています。今後は、利用者の声をもとに機能強化を図り、広域に分散したデータを活用する研究を支援してまいります。

参考文献

- [1] SINET広域データ収集基盤、<https://www.sinet.ad.jp/wadci>
- [2] SINETStream <https://nii-gakunin-cloud.github.io/sinetstream/>を参照。Apache License 2.0に基づいて利用可能。
- [3] IoTストリームデータ処理のためのソフトウェアライブラリSINETStreamの開発、竹房、孫、藤原、吉田、合田、情報処理学会研究報告2020-IOT-48, No. 19, pp. 1-8、2020年3月。