

# エッジコンピューティングにおける時間的相関を考慮した近似的耐障害性保証 Approximate Fault Tolerance using the Time Correlation for Edge Stream Processing

高尾 大樹  
Daiki Takao

杉浦 健人  
Kento Sugiura

石川 佳治  
Yoshiharu Ishikawa

陸 可鏡  
Kejing Lu

## 1 はじめに

分散データストリーム処理システムにおいて耐障害性保証は重要な要件であり、これまで様々なアプローチが提案されている [1-4]. Flink [1] など多くのシステムでは複数タスクを介したパイプライン処理を採用しているが、このような処理形態ではあるタスクからの出力が途絶えると以降全てのタスクでの処理が停滞してしまう. システムの可用性や低遅延性を担保するためには、障害が発生しても処理を継続する仕組みが必要である. 既存システムでは一般的に待機系への処理移譲、及び損失データの再処理によって頑健な耐障害性を保証する. すなわち、あるノードに障害が発生した際には待機ノードがその処理内容を引き継ぎ、また障害によって損失したデータを再処理することで、誤差なく処理を継続する.

しかし、エッジコンピューティング環境においては、デバイス単位での物理的な多重化コストが課題となる. エッジコンピューティングとは、集約やフィルタリングなどの簡単なデータ処理をネットワークエッジで施すことで、データ通信量の削減や処理の負荷分散を可能とする処理形態である [5]. 例としてスマートシティでの環境センシングを考える. 各区画や建物内に広く設置された環境センサデバイスは計測データを無線通信によってそれぞれ近くのエッジデバイスへと送信し、各エッジは一定時間ごとに計測データを集約しクラウドサーバへ転送する. このようなシナリオにおいて、ある地点のエッジが故障したとき距離の離れた別のエッジが代わりに計測データを受信できるとは限らない. すなわち、従来のような誤差のない頑健な耐障害性保証のためにはデバイス単位での多重化が必要となるが、仮想マシンなどで代替ノードをすぐに立ち上げられるクラウド環境と異なり、そのコストは無視できない.

この問題に対して、エッジ単位での物理的多重化を必要としない近似的な耐障害性保証のアプローチを提案する. 本手法では計測データ間に存在する相関関係を活用し、各障害により損失したデータを他のデバイスの出力から近似的に復元することで、システム全体として処理を継続する. 本手法の概要図を図 1 に示す. 図 1 において、各エッジは窓幅  $w = 6$  の各センサの平均値を計算しサーバへと送信するが、各デバイスの障害により“-”で示した一部の計測値及び処理結果が損失した状況を表している. このとき、エッジ 2 ではセンサ D の計測値からセンサ C の集約結果を近似的に計算する. また、サーバでは損失したセンサ A・B の集約結果をエッジ 2 から得られるセ

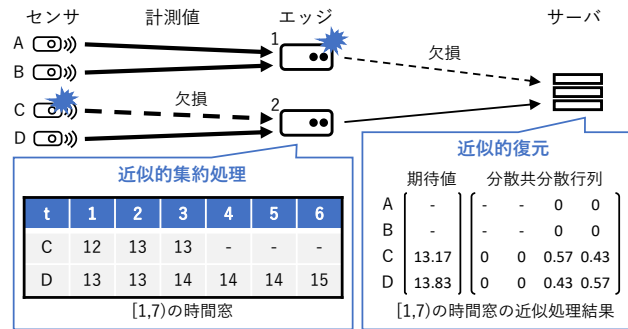


図 1 近似的耐障害性保証アプローチの概要図

ンサ C・D の集約結果を用いて近似的に復元する. この際、各エッジでの近似処理による不確実性の情報を確率分布として表現し、サーバはこの不確実性の情報を考慮して障害エッジ出力を復元することで、誤差上限を正しく推定できる.

本稿では特に、信頼性向上を目的とし、時間的な相関関係を考慮した近似処理手法について検討する. 著者らはこれまでエッジやサーバそれぞれでの近似処理手法について研究してきた [6,7] が、時間的な相関を考慮できていないため、実データに対して誤差上限を正しく推定できないという課題が残っている. そこで、時間的な相関関係を考慮できるようにエッジでの近似的集約処理手法を拡張し、誤差上限の推定精度向上を目指す. また、サーバでの近似的復元に必要な確率モデルの構築方法についても議論する.

本稿の構成は以下のとおりである. 2 章では、ストリーム処理システムにおける耐障害性保証に関する先行研究について概説する. 3 章では、近似的耐障害性保証アプローチについて説明する. また、4 章では時間的な相関を考慮したエッジでの近似的集約処理の拡張について、5 章ではサーバで保持する確率モデルの構築方法についてそれぞれ検討する. 最後に、本研究のまとめと今後の課題を 6 章で述べる.

## 2 関連研究

従来のデータストリーム処理システムの多くは耐障害性保証セマンティクスとして、at-most-once セマンティクスや at-least-once セマンティクスを保証する [3,8]. at-most-once セマンティクスを保証するシステムでは、内部状態やデータのバックアップは取らず待機ノードへの処理の移譲のみを行うため、データの損失の恐れがある. 一方、at-least-once セマンティクスを保証するシステムでは、障害復帰処理として入力データを再送するが、同一データを複数回処理する可能性がある. これらセマンティクスを保証するシステムでは、障害により生じる

誤差量を保証できないため信頼性に問題がある。

また、Flink [1] や Spark Streaming [2] などのストリーム処理システムでは、誤差のない頑健な耐障害性保証として exactly-once セマンティクスを保証できる。これらシステムでは、チェックポイントによる内部状態のバックアップ、及び入力データの再処理に基づく復旧により、障害の発生によらず入力されたデータを過不足なく 1 回処理した結果の出力を保証する。クラスタ環境に注目した場合、豊富なリソースが物理的に集中しており、またシステム前段のメッセージングキューに入力データの管理を一任できるため、この耐障害性保証のアプローチはコストと信頼性の両面において優れている。

しかし、エッジコンピューティング環境に注目した場合、物理的多重化コストの問題が生じる。従来の耐障害性保証アプローチでは、具体的な障害復帰方法に違いはあるものの、待機系への処理移譲を基本としている。しかし、エッジコンピューティング環境ではその物理的な制約からデバイスやエッジ単位での物理的な多重化が必要となり、システム規模が大きくなるほど多重化コストも増加する。また、障害復帰に入力データの再処理が必要な場合、再処理中にも次々と新しい入力生成されるため、データ生成時刻から実際の処理完了時刻までの遅延が増大する問題もある。クラスタ環境においてはこの遅延を抑制するために障害復帰処理を並列化するアプローチ [9] などが提案されているが、物理的多重化コストが無視できないエッジ環境においては異なるアプローチが求められる。

耐障害性保証におけるコストの問題に対して、結果の信頼性を保証した近似的耐障害性保証のアプローチ [10] がある。この手法は、未処理データ数と内部状態の変量に対する閾値を超した場合にのみバックアップを取ることで、バックアップデータ量及びその頻度を削減し、スループットを向上する。しかし、この手法は計算コストの問題にのみ着目しており待機系への処理移譲が必須であり、物理的多重化コストの問題は解決できない。

### 3 近似的耐障害性保証

本稿ではエッジコンピューティング環境における多重化コストの問題に対して、データ間に存在する相関関係を活用した近似的な耐障害性保証のアプローチによる解決を目指す。例えば図 1 においてセンサ C が故障したとき、損失した時刻 4 以降のセンサ C の計測値を、同時刻のセンサ D の値からエッジ 2 で推定する。またエッジ 1 の故障に対しては、損失したセンサ A、及び B の集約結果を、エッジ 2 から得られるセンサ C、及び D の集約結果からサーバ側で近似的に復元する。このように各障害により損失したデータを他のデータから近似的に復元することで、障害デバイスの復旧を待たずシステム全体として処理を継続する。しかし、各エッジで単純に推定値を集約するだけでは近似処理による不確実性の情報が失われてしまうため、サーバでの近似的復元に対する誤差保証が破綻するという信頼性の問題がある。そこで、本手法では近似集約結果を全て確率分布として計算することで、処理結果の誤差上限を理論的に保証する。以降では、エッジでの近似的集約処理、及びサーバでの障害エッジ出力推定処理についてそれぞれ述べる。

#### 3.1 エッジでの近似的集約処理

エッジでは、センサから受け取った計測値系列を時間窓により分割し、時間窓ごとに各センサの集約値を計算する。時刻  $t \in \mathbb{N}^+$  における  $n$  台のセンサ  $\mathbf{X} = \{X_1, X_2, \dots, X_n\}$  の計測値の真値を  $\mathbf{x}^t = \langle x_1^t, x_2^t, \dots, x_n^t \rangle$  と表す。エッジ  $i$  では、時間窓の始点  $t'$ 、及び窓幅  $w$  に対して、センサ  $\mathbf{X}_i \subseteq \mathbf{X}$  の計測値系列  $\mathbf{x}_i^{[t', t'+w)} = \langle x_i^{t'}, x_i^{t'+1}, \dots, x_i^{t'+w-1} \rangle$  から各センサ  $X \in \mathbf{X}_i$  の集約値  $\mathbf{Y}_X$  を計算する。ただし、本稿では集約問合せとして平均値及び合計値問合せを想定する。

センサ故障や通信障害などの問題に対して、センサ間に存在する相関関係を活用し、各時刻で得られた計測値から欠損値を確率分布の形で推定する。本手法では、センサ  $\mathbf{X}$  間の相関関係を多変量正規分布  $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$  によりモデル化する。ここで、 $\boldsymbol{\mu}$  及び  $\boldsymbol{\Sigma}$  はそれぞれ  $\mathbf{X}$  の平均ベクトルと分散共分散行列を表す。エッジ  $i$  では、時刻  $t$  で受け取ったセンサ  $\mathbf{O}_i^t \subseteq \mathbf{X}_i$  の計測値  $\mathbf{o}_i^t$  から、センサ  $\mathbf{Z}_i^t = \mathbf{X}_i \setminus \mathbf{O}_i^t$  の事後確率を表す正規分布  $\mathcal{N}(\boldsymbol{\mu}_{\mathbf{Z}_i^t | \mathbf{o}_i^t}, \boldsymbol{\Sigma}_{\mathbf{Z}_i^t | \mathbf{o}_i^t})$  を次のとおり求められる [11]。

$$\boldsymbol{\mu}_{\mathbf{Z}_i^t | \mathbf{o}_i^t} = \boldsymbol{\mu}_{\mathbf{Z}_i^t} + \boldsymbol{\Sigma}_{\mathbf{Z}_i^t \mathbf{O}_i^t} \boldsymbol{\Sigma}_{\mathbf{O}_i^t}^{-1} (\mathbf{o}_i^t - \boldsymbol{\mu}_{\mathbf{O}_i^t}) \quad (1)$$

$$\boldsymbol{\Sigma}_{\mathbf{Z}_i^t | \mathbf{o}_i^t} = \boldsymbol{\Sigma}_{\mathbf{Z}_i^t \mathbf{Z}_i^t} - \boldsymbol{\Sigma}_{\mathbf{Z}_i^t \mathbf{O}_i^t} \boldsymbol{\Sigma}_{\mathbf{O}_i^t}^{-1} \boldsymbol{\Sigma}_{\mathbf{O}_i^t \mathbf{Z}_i^t} \quad (2)$$

ここで、各記号の添字はベクトルないし行列から添字に対応する次元を抽出したことを示す。例えば、 $\boldsymbol{\mu}_{\mathbf{Z}_i^t}$  は  $\mathbf{Z}_i^t$  に対応する次元の値を平均ベクトル  $\boldsymbol{\mu}$  から抽出したもの（つまり  $\mathbf{Z}_i^t$  の平均ベクトル）であり、 $\boldsymbol{\Sigma}_{\mathbf{Z}_i^t \mathbf{O}_i^t}$  は分散共分散行列  $\boldsymbol{\Sigma}$  の  $\mathbf{Z}_i^t$  に対応する行から  $\mathbf{O}_i^t$  に対応する列を抽出したもの（つまり  $\mathbf{Z}_i^t$  と  $\mathbf{O}_i^t$  の共分散を表すブロック行列）である。

そして、各時刻で求めた確率分布の系列を時間窓で集約し、サーバへと出力する。例えば、センサ  $X \in \mathbf{X}_i$  の平均値  $Y_X = (\sum_{t \in [t', t'+w)} X^t) / w$  に対して、各時刻の平均ベクトル  $\boldsymbol{\mu}_{\mathbf{X}_i | \mathbf{o}_i^t}$  及び分散共分散行列  $\boldsymbol{\Sigma}_{\mathbf{X}_i | \mathbf{o}_i^t}$  から次の正規分布  $\mathcal{N}(\boldsymbol{\mu}_{\mathbf{Y}_{\mathbf{X}_i}}, \boldsymbol{\Sigma}_{\mathbf{Y}_{\mathbf{X}_i}})$  が得られる [6]。

$$\boldsymbol{\mu}_{\mathbf{Y}_{\mathbf{X}_i}} = \frac{1}{w} \left( \sum_{t \in [t', t'+w)} \boldsymbol{\mu}_{\mathbf{X}_i | \mathbf{o}_i^t} \right) \quad (3)$$

$$\boldsymbol{\Sigma}_{\mathbf{Y}_{\mathbf{X}_i}} = \frac{1}{w^2} \left( \sum_{t \in [t', t'+w)} \boldsymbol{\Sigma}_{\mathbf{X}_i | \mathbf{o}_i^t} \right) \quad (4)$$

ただし、 $\boldsymbol{\mu}_{\mathbf{X}_i | \mathbf{o}_i^t}$  及び  $\boldsymbol{\Sigma}_{\mathbf{X}_i | \mathbf{o}_i^t}$  について、 $X \in \mathbf{O}_i^t$  に対する期待値はその計測値  $\mathbf{o}_i^t \in \mathbf{O}_i^t$  であり、分散及び他センサとの共分散はすべて 0 とする。

集約結果を確率分布として計算することで、誤差上限を理論的に保証できる。センサ  $X \in \mathbf{X}_i$  の平均値  $Y_X$  が  $\boldsymbol{\mu}_{\mathbf{Y}_{\mathbf{X}_i}}$  である確率はある誤差上限  $e$  を用いて以下のように求められる。

$$P(Y_X \in [\boldsymbol{\mu}_{\mathbf{Y}_{\mathbf{X}_i}} - e, \boldsymbol{\mu}_{\mathbf{Y}_{\mathbf{X}_i}} + e]) = \int_{\boldsymbol{\mu}_{\mathbf{Y}_{\mathbf{X}_i}} - e}^{\boldsymbol{\mu}_{\mathbf{Y}_{\mathbf{X}_i}} + e} f(y | \boldsymbol{\mu}_{\mathbf{Y}_{\mathbf{X}_i}}, \boldsymbol{\Sigma}_{\mathbf{Y}_{\mathbf{X}_i}}) dy \quad (5)$$

ここで、 $f(y | \boldsymbol{\mu}_{\mathbf{Y}_{\mathbf{X}_i}}, \boldsymbol{\Sigma}_{\mathbf{Y}_{\mathbf{X}_i}})$  は正規分布  $\mathcal{N}(\boldsymbol{\mu}_{\mathbf{Y}_{\mathbf{X}_i}}, \boldsymbol{\Sigma}_{\mathbf{Y}_{\mathbf{X}_i}})$  に対する確率密度関数を表す。式 (5) より、ユーザ要求として要求信頼度  $\delta$  が与えられたとき、 $P(Y_X \in [\boldsymbol{\mu}_{\mathbf{Y}_{\mathbf{X}_i}} - e', \boldsymbol{\mu}_{\mathbf{Y}_{\mathbf{X}_i}} + e']) = \delta$  を解く

ここで、要求信頼度  $\delta$  を満たす最小の推定誤差  $e'$  を計算できる。このとき、 $\delta$  の確率で  $Y_X$  の真値が  $\mu_{Y_X} \pm e'$  以内に存在することを保証する。ただし、 $\Sigma_{Y_X} = 0$  のときは時間窓中の全時刻において計測値が得られているため、常に  $e' = 0$  であり集約結果に誤差は含まれない。

### 3.2 サーバでの障害エッジ出力推定手法

サーバでは、エッジ障害により損失した処理結果を他のエッジの処理結果から近似的に復元する。障害エッジに割り当てられたセンサ  $Z \subseteq X$  の集約結果は全て損失し、センサ  $O = X \setminus Z$  から集約結果として  $N(\mu_{Y_O}, \Sigma_{Y_O})$  が得られるとする。ただし、 $N(\mu_{Y_O}, \Sigma_{Y_O})$  は複数のエッジからの出力を合わせたものである点に注意する。エッジ  $i, j (i \neq j)$  はそれぞれが独立して近似集約処理を行っているため、本稿では  $Y_{X_i}, Y_{X_j}$  間の共分散を表す要素は全て 0 と扱う。

例として図 1 を考えると、サーバは [1, 7] の時間窓に対して、エッジ 1 の故障によりセンサ  $Z = \{A, B\}$  の集約結果を損失しており、センサ  $O = \{C, D\}$  の集約結果として正規分布  $N(\mu_{Y_O}, \Sigma_{Y_O})$  のみを得ている。なお、エッジ 1 と 2 はそれぞれ独立して処理しているため、センサ  $Z$  及び  $O$  間の共分散は全て 0 としている。このとき、 $N(\mu_{Y_O}, \Sigma_{Y_O})$  から  $N(\mu_{Y_Z}, \Sigma_{Y_Z})$  を近似的に復元できれば 3.1 節と同様の誤差保証が可能であるので、エッジ 1 の復旧を待たずサーバの処理を継続できる。

センサ  $Z$  の集約結果  $Y_Z$  に対する事後確率分布は次のとおり計算できる。

$$P(Y_Z) = \int P(Y_Z | Y_O = y) P(Y_O = y) dy \quad (6)$$

ここで、集約結果  $Y_X$  間の相関関係を正規分布  $N(\mu', \Sigma')$  により表現できるとき、式 (6) を計算することで  $Y_Z$  の事後確率として次の正規分布  $N(\mu_{Y_Z|Y_O}, \Sigma_{Y_Z|Y_O})$  が得られる [7]。

$$\mu_{Y_Z|Y_O} = \mu'_Z + \Sigma'_{ZO} (\Sigma'_{OO})^{-1} (\mu_{Y_O} - \mu'_O) \quad (7)$$

$$\Sigma_{Y_Z|Y_O} = \Sigma'_{ZZ} - \Sigma'_{ZO} (\Sigma'_{OO})^{-1} \Sigma'_{OZ} + \Sigma'_{ZO} (\Sigma'_{OO})^{-1} \Sigma_{Y_O} (\Sigma'_{OO})^{-1} \Sigma'_{OZ} \quad (8)$$

ここで、各エッジでの近似処理による不確実性は  $\Sigma_{Y_O}$  によって表現されており、式 (8) の第 3 項によってこれを考慮して分散の値を補正することで、誤差上限を正しく保証できる。

## 4 時間的相関の考慮による拡張

3 章で述べた手法では時間的相関を考慮できていないため、実データに対する誤差保証に課題がある。そこで本章では、時間的な相関を考慮したエッジでの近似的集約処理の拡張について検討する。ただし、本稿では議論の単純化のためマルコフ性を仮定する。すなわち、時刻  $t+1$  の各センサの計測値  $X^{t+1}$  は時刻  $t$  の計測値  $X^t$  のみに依存し、その相関関係は遷移モデル  $P(X^{t+1} | X^t)$  として表現できるとする。

エッジでは各時刻で得られた計測値を用いて欠損値に対する事後確率分布を計算するだけでなく、次の時刻の事前確率分布も推定することで、時間的相関を考慮した近似的集約処理が可能となる。まず、エッジ  $i$  が時刻  $t$  でセンサ  $O_i^t$  の計測値  $o_i^t$  を得たとき、 $X_i^t$  の事後確率分布  $P(X_i^t | o_i^1, \dots, o_i^t)$  を計算す

る。そして、 $X_i^t$  の事後確率分布及び遷移モデルから次の時刻  $t+1$  における事前確率分布  $P(X_i^{t+1} | o_i^1, \dots, o_i^t)$  を計算する。この処理を各時刻で反復的に繰り返し得られた事後確率分布の系列を時間窓で集約すればよい。以降では各処理の具体的な計算方法について述べる。

各時刻での事後確率分布の計算は 3.1 節と同様の議論ができる。時刻  $t$  での事前確率分布が正規分布として与えられたとき、センサ  $Z_i^t = X_i^t \setminus O_i^t$  の事後確率分布も正規分布となり、その期待値ベクトル、及び分散共分散行列はそれぞれ式 (1), (2) と同様に計算できる。このとき、 $X_i^t$  の事後確率分布を  $N(\mu_{X_i^t|o_i^t}, \Sigma_{X_i^t|o_i^t})$  として表す。ただし、センサ  $O_i^t$  に対応する期待値は  $o_i^t$  とし、分散及び他のセンサとの共分散は全て 0 とする。

時刻  $t+1$  における事前確率分布は時刻  $t$  での事後確率分布、及び遷移モデルを用いて次のとおり求められる。

$$P(X_i^{t+1} | o_i^1, \dots, o_i^t) = \int P(X_i^{t+1} | X_i^t = x) P(X_i^t = x | o_i^1, \dots, o_i^t) dx \quad (9)$$

また、式 (9) は式 (6) と同様の手順で計算できる。すなわち、時刻  $t+1$  及び  $t$  の計測値の同時確率分布  $P(X_i^{t+1}, X_i^t)$  が  $N(\mu'', \Sigma'')$  として表せるとき、時刻  $t+1$  の事前確率を表す正規分布  $N(\mu_{X_i^{t+1}|o_i^t}, \Sigma_{X_i^{t+1}|o_i^t})$  を次のとおり求められる。

$$\mu_{X_i^{t+1}|o_i^t} = \mu''_A + \Sigma''_{AB} (\Sigma''_{BB})^{-1} (\mu_{X_i^t|o_i^t} - \mu''_A) \quad (10)$$

$$\Sigma_{X_i^{t+1}|o_i^t} = \Sigma''_{AA} - \Sigma''_{AB} (\Sigma''_{BB})^{-1} \Sigma''_{BA} + \Sigma''_{AB} (\Sigma''_{BB})^{-1} \Sigma_{X_i^t|o_i^t} (\Sigma''_{BB})^{-1} \Sigma''_{BA} \quad (11)$$

ただし、 $\mu''$ 、及び  $\Sigma''$  の添字について  $A$  は先の時刻の計測値に対応する次元を、 $B$  は前の時刻の計測値に対応する次元をそれぞれ抽出することを示す。

例として図 1 を考えると、時刻 3 においてエッジ 2 はセンサ  $C, D$  の計測値を得ているため、 $\mu_{X_3^3|o_3^3} = o_3^3$  であり、 $\Sigma_{X_3^3|o_3^3}$  の全要素は 0 となる。この事後確率分布を式 (10), (11) に代入することで、時刻 4 における事前確率分布  $N(\mu_{X_4^4|o_4^3}, \Sigma_{X_4^4|o_4^3})$  が得られる。こうして得られた事前確率分布を用いて時刻 4 のセンサ  $C$  の欠損値を推定することで、推定精度の向上、及び実データに対する誤差保証の改善が期待できる。

以上の処理を反復的に繰り返すことで時間的な相関を考慮しながら欠損値を近似的に復元できるので、あとはこれを単純に集約すればよい。例えば、時間窓  $x_i^{[t', t'+w]}$  の平均値に対する正規分布は式 (3), (4) で求められる。

## 5 サーバ用のモデル構築方法の検討

サーバでも適切な遷移モデルを構築できれば、4 章と同様の手順で時間的な相関を考慮して障害エッジ出力を近似的に復元できると考えられる。しかしエッジとは異なり、サーバでは各エッジでの近似的な集約結果が入力として与えられるため、その窓幅  $w$  及び滑り幅  $l$  に応じた遷移モデルの構築が必要である。時間窓  $x^{[t, t+w]}$  に対する集約結果を  $Y_{X^{[t, t+w]}}$  とすると、次の時間窓に対する事前確率分布を計算するためには遷移モデル  $P(Y_{X^{[t+l, t+l+w]} | Y_{X^{[t, t+w]}})$  が必要となる。

しかし、この遷移モデルの効率的な構築は困難である。指定窓幅及び滑り幅で学習データを集約し得られた結果から構築するという単純な方法も考えられるが、問合せ毎に対応する遷移モデルを一から構築しなおさなければならず非効率的である。エッジ用に準備した遷移モデルからの構築も考えられるが、サーバで必要な遷移モデルとは時間の粒度が異なり、計測値間の相関から集約結果間の相関関係もモデル化する必要がある。

そこで本章では、サーバ用の確率モデルとして遷移モデルではなく、集約結果に対する事前確率分布  $P(\mathbf{Y}_{\mathbf{X}}^{[t,t+w]})$  の構築を検討する。このモデルではセンサ間の相関しか考慮できないが、十分な精度で障害エッジの出力を復元できるのであれば単純な方法よりも効率的であると考えられる。ただし、各エッジでは時間的相関を考慮して近似処理することを想定したモデルの構築が必要となる点に注意する。全時刻の事前確率分布が同じと仮定してセンサ  $\mathbf{X}$  間の相関関係を表した  $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$  を窓幅  $w$  で単純に集約してしまうような単純な方法は適切でない。

まず、時間窓中の任意時刻での事前確率分布について考える。時刻  $t$  の事前確率分布  $P(\mathbf{X}^t)$ 、及びエッジ用に準備した遷移モデル  $P(\mathbf{X}^{t+1} | \mathbf{X}^t)$  に対して、時刻  $t+i$  での事前確率分布は次式のとおり求められる。

$$P(\mathbf{X}^{t+i}) = \int \prod_{j=1}^i P(\mathbf{X}^{t+j} | \mathbf{X}^{t+j-1}) P(\mathbf{X}^t = \mathbf{x}) d\mathbf{x} \quad (12)$$

$\mathbf{V} = \boldsymbol{\Sigma}_{AB}'' (\boldsymbol{\Sigma}_{BB}'')^{-1} \boldsymbol{\Sigma}_{BA}'$ ,  $\mathbf{W} = \boldsymbol{\Sigma}_{AB}'' (\boldsymbol{\Sigma}_{BB}'')^{-1}$  とおくと、式 (12) を計算することで以下の正規分布  $\mathcal{N}(\boldsymbol{\mu}_{\mathbf{X}^{t+i}}, \boldsymbol{\Sigma}_{\mathbf{X}^{t+i}})$  を得る。

$$\boldsymbol{\mu}_{\mathbf{X}^{t+i}} = \sum_{j=0}^{i-1} \mathbf{W}^j \boldsymbol{\mu}_A'' + \mathbf{W}^i \left( \boldsymbol{\mu}_{\mathbf{X}^t} - \sum_{j=1}^i \mathbf{W}^{j-i} \boldsymbol{\mu}_B'' \right) \quad (13)$$

$$\boldsymbol{\Sigma}_{\mathbf{X}^{t+i}} = \sum_{j=0}^{i-1} \mathbf{W}^j (\boldsymbol{\Sigma}_{AA}' - \mathbf{V}) (\mathbf{W}^j)^T + \mathbf{W}^i \boldsymbol{\Sigma}_{\mathbf{X}^t} (\mathbf{W}^i)^T \quad (14)$$

時間幅  $[t, t+w)$  の各時刻の事前確率分布を集約することで、 $P(\mathbf{Y}_{\mathbf{X}}^{[t,t+w]})$  を構築できる。例えば時間窓  $\mathbf{x}^{[t',t'+w)}$  での平均値  $\mathbf{Y}_{\mathbf{X}}$  に対して次の正規分布  $\mathcal{N}(\boldsymbol{\mu}_{\mathbf{Y}_{\mathbf{X}}}, \boldsymbol{\Sigma}_{\mathbf{Y}_{\mathbf{X}}})$  が得られる。

$$\boldsymbol{\mu}_{\mathbf{Y}_{\mathbf{X}}} = \frac{1}{w} \left( \sum_{t \in [t', t'+w)} \boldsymbol{\mu}_{\mathbf{X}^t} \right) \quad (15)$$

$$\boldsymbol{\Sigma}_{\mathbf{Y}_{\mathbf{X}}} = \frac{1}{w^2} \left( \sum_{t \in [t', t'+w)} \boldsymbol{\Sigma}_{\mathbf{X}^t} \right) \quad (16)$$

例として図1を考えると、各エッジでは窓幅  $w=6$  での平均値を求めているため、サーバでは式 (13), (14) によって [1, 7) の各時刻の事前確率分布を求め、これを集約することで近似的復元に必要な確率モデルを構築できる。この事前確率分布は各エッジでの時間的相関を考慮した近似処理を想定したものであるため、3.2節と同様の手順で、エッジ2の時間的相関を考慮した近似的集約結果からエッジ1の出力を高精度に推定できると考えられる。また、一定の時間において遷移モデルが変わらないとき、以降の時間窓も同じモデルを用いて処理できる。

## 6 おわりに

本稿では、エッジ環境での耐障害性保証における物理的多重化コストの問題に対して、近似的な耐障害性保証のアプローチを提案した。本手法では計測データ間に存在する相関関係を活用し、障害により損失したデータを他のデータから近似的に復元することで、障害デバイスの復旧を待たずシステム全体として処理を継続する。また、本稿では実データに対する誤差保証の改善を目的とし、時間的な相関を考慮したエッジでの近似的集約処理の拡張、及びサーバでの近似的復元に必要な確率モデルの構築方法について検討した。

今後の課題として、実データを用いた提案手法の性能評価が挙げられる。時間的相関の考慮による誤差保証の改善の効果について検証するとともに、サーバでの近似的復元処理の誤差保証に対する定量的な評価が必要である。その他、より効率的な近似的耐障害性保証のアプローチ実現のためのエッジとサーバ間の協調の方針の検討も課題である。

## 謝辞

本研究は JSPS 科研費 (16H01722, 20K19804) の助成、及び国立研究開発法人新エネルギー・産業技術総合開発機構 (NEDO) の委託業務 (JPNP16007) から得られた結果による。

## 参考文献

- [1] “Apache Flink: Stateful Computations over Data Streams.” <https://flink.apache.org/> (Accessed: Apr. 20, 2021).
- [2] “Spark Streaming | Apache Spark.” <https://spark.apache.org/streaming/> (Accessed: Apr. 20, 2021).
- [3] “Apache Storm.” <https://storm.apache.org/> (Accessed: Apr. 20, 2021).
- [4] J.-H. Hwang, M. Balazinska, A. Rasin, U. Çetintemel, M. Stonebraker, and S. Zdonik, “High-availability algorithms for distributed stream processing,” in *Proc. ICDE*, pp. 779–790, April 2005.
- [5] W. Shi, J. Cao, Q. Zhang, Y. Li, and L. Xu, “Edge computing: Vision and challenges,” *IEEE Internet of Things Journal*, vol. 3, pp. 637–646, October 2016.
- [6] 高尾大樹, 杉浦健人, 石川佳治, “エッジコンピューティングにおける低遅延かつ高信頼度なデータストリームの近似的集約処理,” *電子情報通信学会論文誌 D*, vol. J104-D, pp. 463–475, May 2021.
- [7] D. Takao, K. Sugiura, and Y. Ishikawa, “Approximate fault tolerance for edge stream processing (in press),” in *Proc. of the 1st Int. Workshop on Time Ordered Data*, 2021.
- [8] L. Neumeyer, B. Robbins, A. Nair, and A. Kesari, “S4: Distributed stream computing platform,” in *2010 IEEE Int. Conf. on Data Mining Workshops*, pp. 170–177, January 2010.
- [9] M. Zaharia, T. Das, H. Li, T. Hunter, S. Shenker, and I. Stoica, “Discretized streams: A fault-tolerant model for scalable stream processing,” tech. rep., California Univ Berkeley Dept of Electrical Engineering and Computer Science, 2012.
- [10] Q. Huang and P. P. C. Lee, “Toward high-performance distributed stream processing via approximate fault tolerance,” *Proc. VLDB*, vol. 10, pp. 73–84, November 2016.
- [11] C. M. Bishop, *Pattern recognition and machine learning*. Springer, 2006.