Vol.2019-DBS-169 No.12
Vol.2019-IFAT-136 No.12
2019/9/10

センサストリーム処理のための近似的耐障害性保証

高尾 大樹[1,a)]　石川 佳治[1]　杉浦 健人[1]

概要：大規模な分散並列ストリーム処理に注目が集まる一方，耐障害性保証のためのチェックポイント取得がシステムの性能に悪影響を与える例が報告されている．これに対し既存研究として，入出力キューや内部状態のチェックポイント取得を一部省略し，耐障害性を近似的に保証することで性能を改善する手法が提案された．しかし，1) パラメータ指定が直感的でなく適切なパラメータが実験的にしか決定できない，2) チェックポイント取得の必要性判断ための関数は適切なものをユーザが定義しなければならない，という 2 つの問題がある．そこで，本稿では処理対象としてセンサデータストリームを前提にし，統計モデルに基づき近似的に耐障害性を保証する手法を提案する．本手法では，ユーザ指定のパラメータとして許容誤差及び許容誤り率の 2 つを想定し，直感的な誤差保証の指定を可能とする．また，対象問合せとして最大・最小・合計・平均の集約問合せを想定し，各問合せにおける近似的なチェックポイントの取得方法を検討する．

キーワード：ストリーム処理，耐障害性保証，近似処理，統計モデル

1. はじめに

データの大規模化と意思決定の高速化を受け，分散データストリーム処理システムが多くの応用で利用されている．データが全てストレージにあることを想定するバッチ処理と異なり，ストリーム処理ではデータが断続的にシステムに入力されるため，Hadoop [1] のようなバッチ処理システムとは異なるアーキテクチャが必要となる．そのため，Storm [2] や Flink [3]，Samza [4] など多くの OSS（Open Source Software）ミドルウェアが活発に開発されており，企業・学術研究両方において大きな注目を集めている．

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の一つであるが，耐障害性保証のための処理が目的の解析処理の性能を悪化させてしまうという問題がある．多くのシステムではチェックポインティング，つまり処理途中におけるシステムの内部状態をローカルないしリモートのストレージに書き出すことで耐障害性保証を実現している．しかし，Huang らはこのチェックポインティングによって Spark Streaming のスループットが最大で 50%ほど低下してしまうことを示した [5]．このような著しい処理性能の低下は望ましくなく，より効果的な耐障害性保証の仕組みが必要である．

この問題に対し，既存研究として Huang らは耐障害性を近似的に保証する手法を提案した [5]．この手法では内部状態（入出力キュー，処理の途中結果）の変量に対し閾値を設定し，現在の内部状態とチェックポイントとして取得した過去の内部状態との変量が閾値を超えた場合にのみ新たなチェックポイントを取得する．つまり，バックアップするデータ量及び頻度を削減することでスループットを向上している．

しかし Huang らの手法には，ユーザが求める SLA（Service Level Agreement）を満たすための適切な閾値の設定が難しいという問題がある．Huang らは耐障害性の近似を内部状態の変量の閾値によって定めている．しかし，各閾値が処理の最終結果に与える影響の予測は難しい．例えば，入力キューの閾値は障害発生時の消失を許容する未処理のタプル数を示すが，もし消失したタプルの中に処理結果に大きな影響を与えるものが含まれる場合，最終的な処理結果はユーザが許容できない程に異なる可能性がある．具体的な例を挙げれば，時間窓中で最大値を計算するような処理の場合，本来最大値となるタプルが消失することで最終結果が大きく異なってしまう可能性がある．

そこで本稿では，センサストリーム処理を対象とした，確率モデルに基づく近似的な耐障害性保証を提案する．Huang らの手法において耐障害性保証のための閾値設定が困難となった本質的な原因は，任意のストリーム処理を対象としたことだと考えられる．そのため，本手法では対象となるデータをセンサデータに限定し，かつ Spark におけ

[1] 名古屋大学情報学研究科
Graduate School of Informatics, Nagoya University
[a)] takao@db.is.i.nagoya-u.ac.jp

sensor	time step					
	1	2	3	4	5	...
X_1	21	22	22	23	23	
X_2	21	22	23	23	23	...
X_3	19	20	19	19	18	

図 1　センサストリーム

る Structured Streaming [6] のように対象となる問合せを単純な集約問合せのみとする．これにより，センサデータ間の相関を用いた近似的な復元 [7] が可能となり，チェックポイントとして保持しなければならないデータ量を削減できる．更に，耐障害性の近似度を最終的な処理結果の許容誤差・許容誤り率という 2 つのパラメータで設定可能となる．

本手法の概要を以下に示す．本稿では，チェックポイントはシステムに対する入力キューにおいてのみ取得する．つまり，処理の途中結果のチェックポイントは取得せず，障害発生時にはバックアップした入力データを再送することで内部状態を復元する．ただし，入力された全てのデータを保持するのではなく，事前に計算した確率モデルで復元不可能なデータのみをチェックポイントとして保持する．例えば，図 1 のようなセンサストリームを考える．センサストリームは 3 つのセンサを持つが，このうち X_1 と X_2 には相関が見られ，多次元ガウス分布によるモデル化が可能である．つまり，ユーザが指定する許容誤差・許容誤り率を満たす範囲であれば，チェックポイントとして保持すべきデータは $\{X_1, X_3\}$ ないし $\{X_2, X_3\}$ のいずれかの組合せのみでよい．本稿では合計・平均・最大・最小の 4 つの問合せを対象に，各集約処理を近似的な復元可能なセンサの組合せを導出し，チェックポインティングを行う．

本稿の構成は以下の通りである．まず初めに，近似的な耐障害性保証の議論に必要な概念を 2 章で定義し，本研究の問題定義を 3 章で示す．次に，4 章では各問合せに対する誤差評価や障害復帰について検討し，バックアップデータの効率的な選択方法を 5 章に示す．最後に，本研究のまとめと今後の課題について 6 章で述べる．

2. 準備

本章では，近似的な耐障害性保証について議論するために必要な要素を述べる．まず，入力となるデータストリームと，滑り窓 (sliding window) による有限ストリームへの分割について定義する．次に，確率モデルに基づくセンサ値の推定について概略を説明し，最後に処理対象となる集約問合せについて述べる．

2.1 データストリーム

まず，対象となるデータストリームを定義する．

定義 1. 全 n 個のセンサの集合 $\boldsymbol{X} = \{X_1, X_2, \ldots, X_n\}$ で表す．また，全てのセンサは定期的にセンシングデータを測定し，同じ時刻を持つとする．つまり，データストリーム DS を各時刻 t におけるセンサ集合 $\boldsymbol{X}^t = \{X_1^t, X_2^t, \ldots, X_n^t\}$ の無限の系列とする．

$$DS = \langle \boldsymbol{X}^1, \boldsymbol{X}^2, \ldots, \boldsymbol{X}^t, \ldots \rangle \tag{1}$$

■

本稿では滑り窓の使用を前提とし，入力されたデータストリームを時間窓へと分割する．

定義 2. データストリーム DS，窓幅 w，及び滑り幅 l が与えられたとき，データストリーム DS を時間窓に基づき有限のストリームの集合 W_{DS} に分割する．

$$W_{DS} = \left\{ \boldsymbol{X}^{[t,t+w)} \mid t = 1 + i \cdot l, i \in \mathcal{N}^0 \right\} \tag{2}$$

ただし，有限のストリーム $\boldsymbol{X}^{[t,t+w)}$ は以下の式で表す．

$$\boldsymbol{X}^{[t,t+w)} = \langle \boldsymbol{X}^t, \boldsymbol{X}^{t+1}, \ldots, \boldsymbol{X}^{t+w-1} \rangle \tag{3}$$

■

つまり，添字の開区間ないし閉区間で表される範囲の有限のストリームである．

2.2 確率モデルに基づく推定

本研究ではセンサ $\boldsymbol{X}$ 間の相関関係を表現した多次元ガウス分布 $\mathcal{N}(\boldsymbol{x} \mid \boldsymbol{\mu}, \Sigma)$ を活用しセンサ値の推定を行う．なお，$\boldsymbol{\mu}$ 及び Σ はそれぞれ $\boldsymbol{X}$ の平均ベクトルと分散共分散行列である．

センサ集合 $\boldsymbol{X}$ が多次元ガウス分布に従うとき，範囲積分によってセンサ値を確率的に推定できる [7]．例えば，センサ $Y \in \boldsymbol{X}$ の値がある値 v である確率は，推定における許容誤差 ϵ を与えることで以下のように求められる．

$$P(Y \in [v-\epsilon, v+\epsilon]) = \int_{v-\epsilon}^{v+\epsilon} \mathcal{N}(y \mid \mu_y, \sigma_y) dy \tag{4}$$

つまり，ユーザから許容誤差 ϵ と許容誤り率 δ を受け取ることで，推定した値 $Y = v$ が信頼に足るか（確率が $1-\delta$ より大きくなるか）判断できる．

更に，事後確率分布を用いることで推定の精度を向上できる．いくつかのセンサ $\boldsymbol{O} \subseteq \boldsymbol{X}$ に対してその測定値 $\boldsymbol{o}$ が得られたとき，センサ $Y \in \boldsymbol{X} \setminus \boldsymbol{O}$ の事後確率を表すガウス分布 $\mathcal{N}(y \mid \mu_{Y|\boldsymbol{o}}, \sigma_{Y|\boldsymbol{o}})$ は以下の式から求められる [7]．

$$\mu_{Y|\boldsymbol{o}} = \mu_Y + \Sigma_{Y\boldsymbol{O}} \Sigma_{\boldsymbol{OO}}^{-1} (\boldsymbol{o} - \boldsymbol{\mu}_{\boldsymbol{O}}) \tag{5}$$

$$\sigma_{Y|\boldsymbol{o}} = \Sigma_{YY} - \Sigma_{Y\boldsymbol{O}} \Sigma_{\boldsymbol{OO}}^{-1} \Sigma_{\boldsymbol{O}Y} \tag{6}$$

ここで，各記号の添字はベクトルないし行列から添え字に対応する次元を抽出したことを示す．例えば，μ_Y は Y 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（つまりセンサ Y の期待値）であり，$\Sigma_{Y\boldsymbol{O}}$ は 分散共分散行列 Σ の Y 行 から $\boldsymbol{O}$ に対応する列を抽出したもの（つまりセ

ンサ Y に対する参照センサ集合 $\boldsymbol{O}$ の共分散ベクトル）である．事後確率分布の分散は事前分布のそれよりも減少するため，より高精度な推定が可能となる．

また，式 (6) に示す事後確率分布の分散は参照センサ集合 $\boldsymbol{O}$ の実測値 $\boldsymbol{o}$ に依存しない [7] ことから，以後 $\sigma_{Y|\boldsymbol{O}}$ と表記する．

2.3 対象集約問合せ

本稿では，合計・平均・最大・最小の 4 つの集約問合せを処理対象とする．

更に，障害発生時における各集約問合せ agg $\in$ $\{\text{sum}, \text{avg}, \max, \min\}$ の結果については，ユーザが指定する信頼度要求を満たす処理結果が得られればよいとする．本稿では，ユーザが信頼度要求として許容誤差 ϵ と許容誤り率 δ を指定するものとし，各時間窓 $\boldsymbol{X}^{[t,t+w)} \in W_{DS}$ において

$$P\left(Y \in [v-\epsilon, v+\epsilon]\right) > 1-\delta \tag{7}$$

を満たす各属性 $X \in \boldsymbol{X}$ の集約結果 v を出力する．ただし，$\pi_X(DS)$ をデータストリームからのある属性 $X \in \boldsymbol{X}$ の射影を表すとし，$Y = \text{agg}(\pi_X(\boldsymbol{X}^{[t,t+w)}))$ とする．

3. 問題定義

本章では，本研究で扱う近似的な耐障害性保証のためのバックアップの取得について問題を定義する．まず，近似的な耐障害性保証を行う際のバックアップ取得のコストと耐障害性の信頼度を定義する．次に，近似的耐障害性を満たす適切なバックアップの取得を，バックアップのコスト最小化問題として定義する．

3.1 バックアップ取得のコスト

本研究では，耐障害性保証のためのバックアップとして入力ストリームのみを保持する．つまり，処理の内部状態（e.g., 部分的な集約結果）などは保持せず，障害が発生した場合は処理対象となる全ストリームの再送により状態を復元する．

コストについて，本研究ではバックアップするセンサデータの数に基づいて定める．

定義 3. ある時刻 t までにバックアップするセンサデータを $\boldsymbol{O}^{[1,t]} \subseteq DS$ とするとき，バックアップに必要なコストを以下の式で表す．

$$C(\boldsymbol{O}^{[1,t]}) = \sum_{i=1}^{t} \left|\boldsymbol{O}^i\right| \tag{8}$$

■

ただし，データストリーム間の包含関係 $\subseteq$ は以下のように定める．

定義 4. ある無限のストリーム $DS = \langle \boldsymbol{X}^1, \boldsymbol{X}^2, \ldots \rangle$ とある有限なストリーム $\boldsymbol{O}^{[1,t]} = \langle \boldsymbol{O}^1, \boldsymbol{O}^2, \ldots, \boldsymbol{O}^t \rangle$ があるとき，それらの包含関係を以下の式で表す．

$$\boldsymbol{O}^{[1,t]} \subseteq DS \Leftrightarrow \bigwedge_{i=1}^{t} \boldsymbol{O}^i \subseteq \boldsymbol{X}^i \tag{9}$$

■

近似的処理及び近似的な耐障害性の保証を行う場合，バックアップするセンサデータを減らすことでコストを削減できる．一方，入力されたデータストリーム DS に対し近似的な耐障害性保証を行わない場合，全センサデータのバックアップが必要である．つまり，n 個のセンサを対象とする場合，ある時刻 t までに nt 個のセンサデータを保持しなければならない．しかし，前述したように本稿では近似問合せを対象としており，ユーザの要求信頼度を満たす範囲で耐障害性も緩和できる．例えば，図 1 のセンサストリームにおいて，$\{X_1, X_3\}$ ないし $\{X_2, X_3\}$ からユーザの要求信頼度を満たした推定が可能であれば，いずれかのみをバックアップすればよい．これにより，一つのセンサについてはバックアップを省略できるため，耐障害性コストを削減できる．そこで，本稿ではバックアップに必要なセンサデータの最小化を目指す．

3.2 耐障害性の信頼度

本節では，時間窓 $\boldsymbol{X}^{[t',t'+w)}$ に対して，部分ストリーム $\boldsymbol{O}^{[t',t'+w)}$ をバックアップする際の耐障害性の信頼度を定義する．本手法では，$\boldsymbol{X}^{[t',t'+w)}$ の処理結果が障害により失われたとき，バックアップデータ $\boldsymbol{o}^{[t',t'+w)}$ から確率モデルに基づき各属性 $X \in \boldsymbol{X}$ の集約値を復元する．このとき，復元における確率（推定の信頼度）がユーザ要求を上回っていることを保証できればよい．本稿では，各属性の信頼度の最小値を用いて時間窓全体の復元の信頼度を定める．

定義 5. 集約関数 agg $\in \{\text{sum}, \text{avg}, \max, \min\}$，許容誤差 ϵ，及び時間窓 $\boldsymbol{X}^{[t',t'+w)}$ に対するバックアップ $\boldsymbol{o}^{[t',t'+w)}$ が与えられたとする．ある属性 $X \in \boldsymbol{X}$ に対する集約結果を確率変数 $Y = \text{agg}(\pi_X(\boldsymbol{X}^{[t',t'+w)}))$ で，システムによる復元値を v_X で表すとき，時間窓 $\boldsymbol{X}^{[t',t'+w)}$ に対応するバックアップデータ $\boldsymbol{o}^{[t',t'+w)}$ による耐障害性保証の信頼度を以下の式で求める．

$$R(\boldsymbol{o}^{[t',t'+w)}) = \min_{X \in \boldsymbol{X}} P(Y \in [v_X - \epsilon, v_X + \epsilon] \mid \boldsymbol{o}^{[t',t'+w)}) \tag{10}$$

■

例として図 1 において，幅 $w = 3$ の時間窓 $\boldsymbol{X}^{[1,4)}$ の平均処理結果を $\boldsymbol{o}^{[1,4)} = \langle \boldsymbol{o}^1, \ldots, \boldsymbol{o}^3 \rangle$ から推定することを考える．ただし，$\boldsymbol{o}^i = \{x_1^i, x_3^i\}$ とする．このとき，X_1 および X_3 については全入力データをバックアップしていたため，処理結果復元の際に誤差は生じず信頼度は 100% となる．一方，X_2 についてはバックアップデータから処理結果を推定することになるが，X_1 との強い相関関係があることから精度よく推定できる．このときの信頼度を仮に 98% とする

と，$R(\boldsymbol{o}^{[1,4)})$ も最小値を取って 98% となる．

以上の定義より，$R(\boldsymbol{o}^{[t',t'+w)}) > 1-\delta$ を満たすとき，時間窓 $\boldsymbol{X}^{[t',t'+w)}$ に対する全問合せにおいて，ユーザ指定の信頼度要求を満たした出力が再処理によって得られる．

3.3 問題定義

時刻 t で障害が発生したとき，本手法ではバックアップデータ $\boldsymbol{o}^{[1,t]}$ を再送することで障害復帰を図ることから，本研究で取り組む問題をバックアップするデータの最小化問題として定義する．.

定義 6. データストリーム DS，窓幅 w，滑り幅 l，集約関数 $\mathrm{agg} \in \{\mathrm{sum}, \mathrm{avg}, \mathrm{max}, \mathrm{min}\}$，許容誤差 ϵ，許容誤り率 δ が与えられたとき，時刻 t において以下の式で表されるバックアップ $\boldsymbol{o}^{[1,t]}$ を保持する．

$$\begin{aligned} &\text{minimize} && C(\boldsymbol{o}^{[1,t]}) \\ &\text{subject to} && \boldsymbol{o}^{[1,t]} \subseteq DS \\ & && \forall \boldsymbol{X}^{[t',t'+w)} \in W_{DS}, \\ & && \quad t'+w-1 \leq t \rightarrow R(\boldsymbol{o}^{[t',t'+w)}) > 1-\delta \end{aligned} \tag{11}$$

■

4. 各問合せを考慮した耐障害性保証

本章では，最大・最小・合計・平均問合せを対象としたストリーム処理システムの近似的な耐障害性保証について述べる．各問合せに対して 式 (11) の厳密解を考えると，全探索や入力データごとの信頼度評価が必要となり，問合せの処理性能を著しく損なう恐れがある．そこで，2.2 節で述べた確率モデルによる推定の性質を考慮することで，効率的な耐障害性保証を検討する．具体的にはユーザからの信頼度要求を満たした推定に必要なバックアップデータを事前に計算することで，耐障害性保証のコストの削減を目指す．

4.1 節では合計及び平均値問合せに対して，4.2 節では最大及び最小値問合せに対して，耐障害性の信頼度及び適切なバックアップデータの選択をそれぞれ述べる．ただし，合計・平均値問合せと最大・最小値問合せはそれぞれ同様の議論が行えるため，以下では平均及び最大問合せを中心に述べる．

4.1 合計及び平均値問合せ

各入力データがガウス分布に従うとき合計及び平均値問合せの出力値もガウス分布に従うことから，出力値に対するガウス分布を用いて耐障害性の信頼度を評価する．以降では，ある時間窓 $\boldsymbol{X}^{[t',t'+w)}$ での平均値問合せを想定して議論を進めるが，合計値問合せに対しても同様の議論が可能である．

時間的な相関を考慮しない場合，あるセンサ $X_i \in \boldsymbol{X}$ の平均値 Y_i 及びそのガウス分布の期待値・分散はそれぞれ次の通りである．

$$Y_i = \frac{\sum_{t\in[t',t'+w)} X_i^t}{w} \tag{12}$$

$$\mu_{Y_i} = \frac{\sum_{t\in[t',t'+w)} \mu_{X_i^t}}{w} \tag{13}$$

$$\sigma_{Y_i} = \frac{\sum_{t\in[t',t'+w)} \sigma_{X_i^t}}{w^2} \tag{14}$$

つまり，ある時間窓におけるセンサ X_i の平均値 Y_i の期待値と分散は，各時刻におけるセンサ X_i の期待値及び分散の線形和でそれぞれ計算可能である．

ここで重要となるのは，式 (6) にあるように，事後確率分布の分散 $\sigma_{Y|\boldsymbol{O}}$ が計測値に依存しない点である．つまり，各時刻においてチェックポイントを取得するセンサの組合せ $\boldsymbol{O}^t \subseteq \boldsymbol{X}$ を決定したとき，各時刻の分散 $\sigma_{X_i^t|\boldsymbol{O}^t}$ の値は事前計算できる．これは，処理を開始する前に組合せ $\boldsymbol{O}$ で $R(\boldsymbol{o}^{[t',t'+w)}) > 1-\delta$ を満たせるかが確認可能であることを示す．ただし，コストを最小とする $\boldsymbol{o}^{[t',t'+w)}$ を愚直に求める場合 2^{wn} 通りのバックアップデータそれぞれに対して信頼度 $R(\boldsymbol{o}^{[t',t'+w)})$ を計算する必要があり，計算コストの観点から現実的ではない．

そこで，上述の耐障害性の信頼度の計算における性質を活用した，効率的な $\boldsymbol{o}^{[t',t'+w)}$ の決定方法を検討する．まず，あるセンサ $X_i \in \boldsymbol{X}$ の平均値 Y_i に着目する．近似的に復元した平均値として Y_i の期待値を用いる場合，センサ X_i に関する信頼度は以下の式で計算できる．

$$\int_{\mu_{Y_i}-\epsilon}^{\mu_{Y_i}+\epsilon} \mathcal{N}(y_i \mid \mu_{Y_i}, \sigma_{Y_i}) dy_i \tag{15}$$

更に， $Y_i' = Y_i - \mu_{Y_i}$ と置くとき，この確率変数に対する平均は 0，分散は σ_{Y_i} となるため，式 (15) は以下の式により計算できる．

$$\int_{-\epsilon}^{\epsilon} \mathcal{N}(y_i' \mid 0, \sigma_{Y_i}) dy_i' > 1-\delta \tag{16}$$

ここで，この式を満たす最小の σ_{Y_i} を $\sigma_{Y_i}^*$ としたとき，式 (14) から以下の式が導ける．

$$\begin{aligned} &\forall t \in [t', t'+w), \sigma_{X_i^t|\boldsymbol{O}^t} \leq w\sigma_{Y_i}^* \\ &\rightarrow \sigma_{Y_i|\boldsymbol{O}^{[t',t'+w)}} \leq \sigma_{Y_i}^* \end{aligned} \tag{17}$$

つまり，各 $\sigma_{X_i^t|\boldsymbol{O}^t}$ が少なくとも $w\sigma_{Y_i}^*$ 以下であれば，復元した平均値はユーザが指定した信頼度要求を満たす．時間窓 $\boldsymbol{X}^{[t',t'+w)}$ で確率モデルが不変であるとき， $\sigma_{X_i^t|\boldsymbol{O}^t} \leq w\sigma_{Y_i}^*$ を満たす $\boldsymbol{O}^t$ は時刻 t によらず窓内で共通であるため，あるセンサ X_i の平均値を近似的に復元したい場合 $\sigma_{X_i|\boldsymbol{O}} \leq w\sigma_{Y_i}^*$ を満たす組合せ $\boldsymbol{O} \subseteq \boldsymbol{X}$ を見つければよい．更に，信頼度要求を満たすための平均値の分散 $\sigma_{Y_i}^*$ はセンサ X_i に依存せず，$\forall X_i, X_j \in \boldsymbol{X}, \sigma_{Y_i}^* = \sigma_{Y_j}^*$ である．したがって，以下の式を満たす $\boldsymbol{O}$ を得られれば，時間窓中の全てのセンサにおける平均値を近似的に復元できる．

$$\forall X_i \in \boldsymbol{X}, \sigma_{X_i|\boldsymbol{O}} \leq w\sigma_Y^* \tag{18}$$

ただし，σ_Y^* は以下の式を満たす最小の値であるとする．

$$\int_{-\epsilon}^{\epsilon} \mathcal{N}(y \mid 0, \sigma_Y)dy > 1-\delta \tag{19}$$

4.2 最大及び最小値問合せ

最大及び最小値問合せでは出力値がガウス分布に従わないため，耐障害性の信頼度を各センサ値の分布から計算する必要がある．以降では，ある時間窓 $\boldsymbol{X}^{[t',t'+w)}$ での最大値問合せを想定して議論を進めるが，最小値問合せでも同様の議論が可能である．

センサ X_i の最大値を表す確率変数を $Y_i = \max(X_i^{[t',t'+w)})$ で表すとき，最大値の確率密度関数は以下の式で表せる．

$$P(Y_i = y_i) = \tag{20}$$
$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} P(x_i^{[t',t'+w)}) \mathbb{1}[\max(x_i^{[t',t'+w)}) = y_i] dx_i^{[t',t'+w)}$$

なお，$\mathbb{1}[\phi]$ は ϕ が真となるときに 1 となる関数である．しかし，この計算には重積分が必要になり計算コストの観点から現実的でない．

そこで, 本手法では最大値問合せに対して時間窓 $\boldsymbol{X}^{[t',t'+w)}$ 内で最大の推定値 $\mu_{X_i^{t^*}}$ を予測値として出力する．このとき，以下の式を満たすようなチェックポインティングの実行が行えればよい．

$$P(Y_i \in [\mu_{X_i^{t^*}|\boldsymbol{o}^{t^*}} - \epsilon, \mu_{X_i^{t^*}|\boldsymbol{o}^{t^*}} + \epsilon] \mid \boldsymbol{o}^{[t',t'+w)}) > 1-\delta \tag{21}$$

この式を，モデルとして保持する多次元ガウス分布によって計算可能な形へと変形するために，まず最大値を取る時刻 t^* を固定して考える．このとき，時刻 t^* における値を誤差 ϵ 以内で推定でき，かつそれ以外の時刻で最大値を超える値が出現しないことを保証できれば，少なくとも時刻 t^* が最大値となる際は信頼度要求を満たせる．ただし，他の時刻の値が最大値となる可能性もあるためこの確率は明らかに上述のものよりも小さく，センサ間の時間的な相関を考慮しない場合，このときの信頼度は以下のように表せる．

$$\begin{aligned}
&P(Y_i \in [\mu_{X_i^{t^*}|\boldsymbol{o}^{t^*}} - \epsilon, \mu_{X_i^{t^*}|\boldsymbol{o}^{t^*}} + \epsilon] \mid \boldsymbol{o}^{[t',t'+w)}) \\
&\geq P(X_i^{t^*} \in [\mu_{X_i^{t^*}|\boldsymbol{o}^{t^*}} - \epsilon, \mu_{X_i^{t^*}|\boldsymbol{o}^{t^*}} + \epsilon] \mid \boldsymbol{o}^{[t',t'+w)}) \\
&\prod_{t \in [t',t'+w) \wedge t \neq t^*} P(X_i^t < \mu_{X_i^{t^*}|\boldsymbol{o}^{t^*}} + \epsilon \mid \boldsymbol{o}^{[t',t'+w)})
\end{aligned} \tag{22}$$

例として，図 1 において，時間窓 $\boldsymbol{X}^{[1,4)}$ 内での X_2 に対する最大値問合せ $Y_2 = \max(\boldsymbol{X}_2^{[1,4)})$ の復元を考える．各時刻 t に対する X_2 のモデルが現在図 2 のようになったとする．このとき，信頼度計算のための積分範囲は青色及び橙色で示した部分全体となる．

この式は確率モデルから計算可能であるが，事後確率分布の期待値は取得値 $\boldsymbol{o}^{[t',t'+w)}$ に依存するため，事前に

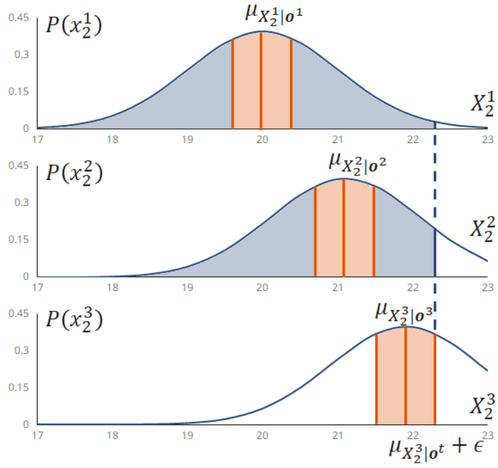

図 2 最大値問合せにおける信頼度

チェックポイントとして保持する属性は決定できない．この式を用いて $\boldsymbol{O}^{[t',t'+w)}$ を決定する場合，計測値が到着するたびに信頼度を評価し，ユーザ指定の信頼度を上回るまで全ての計測値をバックアップするなどの方法を取るしかなく，非常に非効率的である．

そこで，積分範囲を狭めたより厳しい式を考えることでこの問題の解決を考える．今，$\mu_{X_i^{t^*}|\boldsymbol{o}^{t^*}}$ は最大の期待値であるため，t^* 以外の時刻 t において $\mu_{X_i^t|\boldsymbol{o}^t} \leq \mu_{X_i^{t^*}|\boldsymbol{o}^{t^*}}$ である．つまり，t^* 以外の時刻 t において以下の関係が成り立つ．

$$[\mu_{X_i^t|\boldsymbol{o}^t} - \epsilon, \mu_{X_i^t|\boldsymbol{o}^t} + \epsilon] \subseteq [-\infty, \mu_{X_i^{t^*}|\boldsymbol{o}^{t^*}} + \epsilon] \tag{23}$$

これは，図 2 中の X_2^1 及び X_2^2 のグラフにおいて，橙色で示した範囲 $[\mu_{X_i^t|\boldsymbol{o}^t} - \epsilon, \mu_{X_i^t|\boldsymbol{o}^t} + \epsilon]$ が，青色で示した範囲分 $[-\infty, \mu_{X_i^{t^*}|\boldsymbol{o}^{t^*}} + \epsilon]$ より小さいことからも明らかである．したがって，各時刻における期待値を中心とした積分を行ったとき，以下の式が成り立つ．

$$\begin{aligned}
&P(X_i^{t^*} \in [\mu_{X_i^{t^*}|\boldsymbol{o}^{t^*}} - \epsilon, \mu_{X_i^{t^*}|\boldsymbol{o}^{t^*}} + \epsilon] \mid \boldsymbol{o}^{[t',t'+w)}) \\
&\prod_{t \in [t',t'+w) \wedge t \neq t^*} P(X_i^t < \mu_{X_i^{t^*}|\boldsymbol{o}^{t^*}} + \epsilon \mid \boldsymbol{o}^{[t',t'+w)}) \\
&\geq \prod_{t \in [t',t'+w)} P(X_i^t \in [\mu_{X_i^t|\boldsymbol{o}^t} - \epsilon, \mu_{X_i^t|\boldsymbol{o}^t} + \epsilon] \mid \boldsymbol{o}^{[t',t'+w)})
\end{aligned} \tag{24}$$

前述のとおり期待値を中心とした積分は任意の値を中心にでき，かつ本稿では時間窓内で共通の確率モデルを用いるため，測定値によらず事前計算可能な以下の式により推定の信頼度を評価できる．

$$\begin{aligned}
&\prod_{t \in [t',t'+w)} \int_{-\epsilon}^{\epsilon} \mathcal{N}(x_i^t \mid 0, \sigma_{X_i^t|\boldsymbol{O}^t}) dx_i^t = \\
&\left(\int_{-\epsilon}^{\epsilon} \mathcal{N}(x_i \mid 0, \sigma_{X_i|\boldsymbol{O}}) dx_i \right)^w > 1-\delta \\
&\therefore \int_{-\epsilon}^{\epsilon} \mathcal{N}(x_i \mid 0, \sigma_{X_i|\boldsymbol{O}}) dx_i > (1-\delta)^{\frac{1}{w}}
\end{aligned} \tag{25}$$

したがって，この式を満たす $\boldsymbol{O} \subseteq \boldsymbol{X}$ を確率モデルから事前計算すれば，ユーザの信頼度要求を満たした推定が行

Vol.2019-DBS-169 No.12
Vol.2019-IFAT-136 No.12
2019/9/10

Input: センサ集合 $\boldsymbol{X}$, 目標分散値 σ^*
Output: バックアップ対象のセンサ集合 $\boldsymbol{O}$

$\boldsymbol{O} \leftarrow \emptyset$
while $Var(\boldsymbol{O}) > \sigma^*$ **do**
  $X_{new} \leftarrow \arg\max_{X_i \in \boldsymbol{X} \setminus \boldsymbol{O}} Var(\boldsymbol{O}) - Var(\boldsymbol{O} \cup \{X_i\})$
  $\boldsymbol{O} \leftarrow \boldsymbol{O} \cup \{X_{new}\}$

図 3 貪欲法に基づくバックアップ対象センサ決定アルゴリズム

える．

5. バックアップデータの効率的な選択

4 章では各問合せに対する耐障害性保証について検討したが，どの場合も各センサのモデルが一定の分散以下となるようにセンサデータをバックアップすればよいことが分かった．あとはこれを満たす最適なバックアップ対象のセンサ集合を求めればよい．愚直に計算するならば，2^n 通りの組合せに対して全探索すれば厳密な最適解を求められるが，これは非効率である．そこで本章では，学習済みのモデルと目標の分散値が与えられたときの，バックアップ対象のセンサ集合の効率的な決定方法を提案する．

ここで，バックアップするセンサの組合せ $\boldsymbol{O} \subseteq \boldsymbol{X}$ に対して新しい要素 $X_i \in \boldsymbol{X} \setminus \boldsymbol{O}$ を加えることにより，各センサのモデルの分散は減少する．よって，$\boldsymbol{O}$ の選択の基本方針として，なるべく分散が減少するようなセンサを順に選んでいけばよい．

本稿では，各時刻においてバックアップ対象のセンサ集合 $\boldsymbol{O}$ に対する分散の評価関数 $Var(\boldsymbol{O})$ を次式の通り定義する．

$$Var(\boldsymbol{O}) = \min_{X_i \in \boldsymbol{X}} \sigma_{X_i | \boldsymbol{O}} \tag{26}$$

このとき，$\boldsymbol{O}$ に追加したとき分散の減少が一番大きくなるような $X_{new} \in \boldsymbol{X} \setminus \boldsymbol{O}$ は次式から求められる．

$$X_{new} = \arg\max_{X_i \in \boldsymbol{X} \setminus \boldsymbol{O}} Var(\boldsymbol{O}) - Var(\boldsymbol{O} \cup \{X_i\}) \tag{27}$$

上述の議論を，貪欲的なバックアップ対象のセンサ集合の決定アルゴリズムとして図 3 にまとめる．

6. おわりに

本稿では，センサストリームを対象とした，確率モデルに基づく近似的な耐障害性保証を提案した．本手法では対象となる問合せを簡単な集約処理に限定することで，出力結果に対する許容誤差及び許容誤り率という二つのパラメータにより耐障害性の近似度を設定可能にした．これにより，既存研究では難しかった，SLA に応じた直感的な誤差保証を実現できる．また，近似的な耐障害性保証をバックアップデータの最小化問題として定式化し，各問合せに対してこの問題の効率的なアプローチを提案した．厳密な最適解を求めるのは計算コストの観点で難しいことから，代わりに貪欲法に基づくアルゴリズムやヒューリスティックスに基づくアプローチを提案した．

今後の課題としては，実験による提案手法の妥当性及び性能の評価がある．特にヒューリスティックスに基づくアプローチでは妥当性の確認が必要である．また，センサストリームではデータの欠損や遅延の問題が考えられるため，これらを考慮した包括的な耐障害性についても現在検討している．そのほか，本手法では入力データのバックアップのみを検討したが，内部状態のバックアップについても取り組みたい．

謝辞 本研究は JSPS 科研費（JP16H01722, JP19K21530）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）の委託業務による．

参考文献

[1] "Apache Hadoop," https://hadoop.apache.org/ (Accessed: August 3, 2019).
[2] "Apache Storm," https://storm.apache.org/ (Accessed: August 3, 2019).
[3] "Apache Flink: Stateful Computations over Data Streams," https://flink.apache.org/ (Accessed: August 3, 2019).
[4] "Samza," http://samza.apache.org/ (Accessed: August 3, 2019).
[5] Q. Huang and P. P. C. Lee, "Toward high-performance distributed stream processing via approximate fault tolerance," *PVLDB*, vol. 10, no. 3, pp. 73–84, 2016.
[6] M. Armbrust, T. Das, J. Torres, B. Yavuz, S. Zhu, R. Xin, A. Ghodsi, I. Stoica, and M. Zaharia, "Structured Streaming: A declarative API for real-time applications in Apache Spark," in *Proc. SIGMOD*, 2018, pp. 601–613.
[7] A. Deshpande, C. Guestrin, S. R. Madden, J. M. Hellerstein, and W. Hong, "Model-based approximate querying in sensor networks," *VLDBJ*, vol. 14, no. 4, pp. 417–443, 2005.