

高速時空間データ管理技術「Axispot[®]」と 時空間データ高速検索技術

本稿では、コネクティッドカーどうしの車両間通信や拡張現実等の次世代サービスで求められる、大量のヒトや自動車等の動くオブジェクト（動的オブジェクト）が一齐に送信する情報を蓄積しながら、ある時刻に、ある特定の地域にいる動的オブジェクトだけをリアルタイムに検索・分析する高速時空間データ管理技術「Axispot[®]」の取り組みとして、そのコア機能となる「時空間データ高速検索技術」について述べます。

はなだて まさゆき きむら たつろう おき のぶひろ
花館 蔵之 /木村 達郎 /沖 宣宏
しげまつ なおこ うえの いそお ないとう いちべえ
重松 直子 /上野 磯生 /内藤 一兵衛
くぼ たかし みやはら かずひろ いそむら あつし
久保 貴司 /宮原 和夫 /磯村 淳

NTTソフトウェアイノベーションセンタ

はじめに

実空間上のヒトやモノ、自然環境等に関するさまざまな情報の収集とクラウド上での一元管理を可能にするIoT (Internet of Things) は、ヒトや自動車等の動く物体（動的オブジェクト）を管理する次のような次世代サービスが必要不可欠となってきています。

- ① 車両間通信サービス：インターネットとつながる自動車（コネクティッドカー）が収集した情報を、自動車の走行位置や情報送信時刻と合わせてクラウドに蓄え、事故や渋滞等の交通状況を、関係する道路を走行する車両にリアルタイムに通知します。
- ② 拡張現実：ヒトが装着したウェアラブルデバイスやスマートフォンが送信するさまざまな情報を、ヒトの歩行位置や情報送信時刻と合わせてクラウドに蓄え、店舗の混み具合やリコメンド情報等、ヒトの行動の意思決定にかかわる情報をリアルタイムに通知します。

NTTソフトウェアイノベーションセンタ (SIC) では、これらの次世代サービスの実現で必要不可欠となる、

大量の動的オブジェクトが一齐送信する情報の蓄積と、蓄積した動的オブジェクトの中から、ある時刻に、ある特定の地域にいる動的オブジェクトをリアルタイムに検索する高速時空間データ管理技術「Axispot[®]」に取り組んでいます。本稿では、Axispot[®]が用いているデータベース（時空間データベース）と、その技術課題を解決するために昨年度開発した「時空間データ高速検索技術」⁽¹⁾を紹介します。また、この時空間データ高速検索技術をベース技術として実現されるAxispot[®]のアーキテクチャ全体像と今後の展望を述べます。

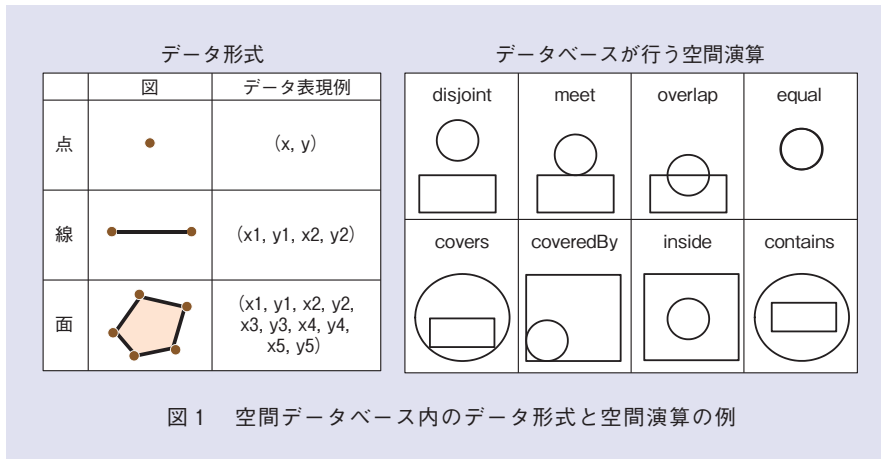
時空間データベースの現状

「時空間データベース」とは、緯度や経度といった空間上の位置（空間情報）と、時刻、期間といった時間情報の双方に関連付いたデータ群を効率的に蓄積・検索するデータベースです^{(2),(3)}。一方、時間情報や空間情報だけを用いるデータベースとして、それぞれ「時系列データベース」「空間データベース」が知られていますが、時空間データベースは、一般的に空間データベースで時間情報も管理できるように拡張することで実現されます⁽³⁾。そ

こで、時空間データベースを説明する前に、空間データベースの実現方法を簡単に説明します。

空間データベースは、土地や建物の範囲に関する空間情報を、「点」「線」「面」等の幾何的なデータ形式で格納し、同様に幾何的なデータ形式でクエリとして与えられる空間情報と空間演算を行い、該当する空間情報を返却するデータベースです。このデータ形式と空間演算の例を図1に示します⁽²⁾。例えば、ある地域に含まれる建物群を検索する場合、緯度と経度を用いて建物の範囲（面）を建物ごとに格納し、緯度と経度を用いて地域の範囲（面）をクエリとして与え、そのクエリで与えられた面と、建物の範囲を表す面の包含関係を計算します。

空間データベースは、既存の関係データベース (RDB: Relational Database) や分散キーバリューストア (KVS: Key Value Store) を用いて実現できます。RDBを用いる場合、空間情報をカラムとするテーブルを利用します。しかし、緯度や経度ごとにカラムを用意する場合、カラムごとに検索が発生するため、検索回数が2回となり検索効率が悪くなります。この解決方法として二次元情報向けの検索木



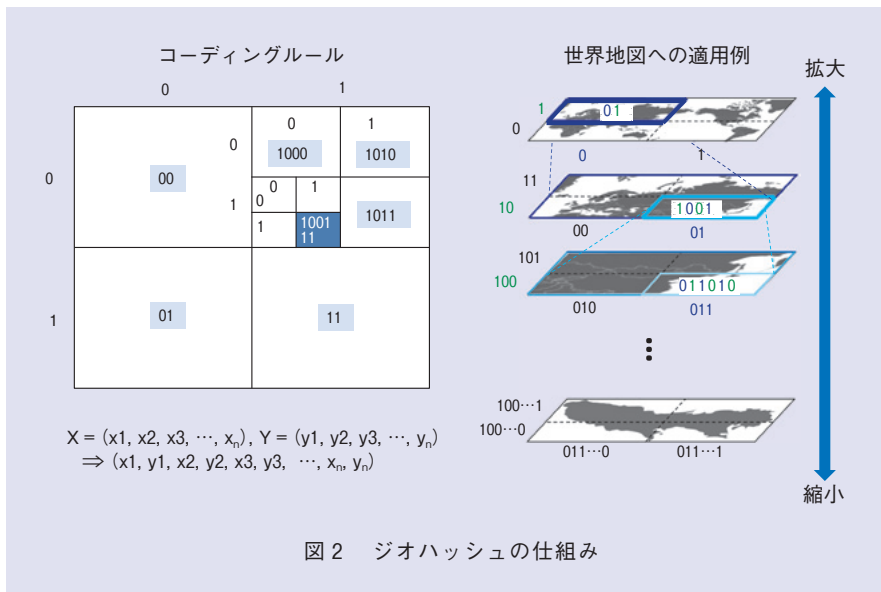
データ格納時に検索木の更新が頻繁に発生するため、データ格納の処理効率が低下します。そこで、膨大な数の動的オブジェクトの時空間データベースは、RDBを用いるよりも、このような検索木の更新が発生しない分散KVSを用いるほうが適していると考えられます。

次に、この分散KVSを用いてSICが開発した時空間データ高速検索技術を説明します。

時空間データ高速検索技術

時空間データ高速検索技術は、実空間上の大量の動的オブジェクトが収集したセンサ情報と、そのセンサ情報に紐付けられた時間情報と空間情報を格納しつつ、ある特定の時間情報と空間情報を矩形範囲で指定して、そこに含まれるセンサ情報を高速に検索する技術です。特に、本技術は、分散KVSとSICが提案する「時空間コード」と「限定的ノード選択アルゴリズム」を用いることにより、以下の要件を満たします。

- ① 多次元情報検索の効率化：「時空間コード」を分散KVSのキーとして用いることにより、時間、緯度、経度、高度等から構成される多次元情報の検索を1回で行います。
- ② 範囲指定：「時空間コード」の前方一致検索により、現在から1時間前、東経10度から20度、北緯30度から40度等、時間情報と空間情報の検索範囲を指定できます。
- ③ 時間的・空間的な偏りの平滑化：「限定的ノード選択アルゴリズム」を用いることにより、「午前は駅周辺が混雑し、夕方は郊外



(R-Tree⁽⁴⁾等)が提案されています。一方、分散KVSを用いる場合、Key-Value構造のテーブルのキーに格納できるデータ数は1つだけとなるため、空間情報のような多次元情報は、そのままでは格納できません。この解決方法として、「空間重点曲線」を用いて多次元情報を一次元情報に変換する方法が提案されています。特に、空間情報の一次元化では、空間重点曲線の1つであるZ曲線を用いて一次元化する「ジオハッシュ⁽⁵⁾」がもっとも用いられています。その理由として、一次元化された空間情報の長さを変えて、表

現できる空間の範囲を拡大・縮小できるズーム効果が挙げられます(図2)。本稿では、このジオハッシュを用いて一次元化された空間情報を「空間コード」と呼びます。

時空間データベースは、このRDBや分散KVSを用いた空間データベースで時間情報も管理できるようにすることで実現できます。しかし、数千万人のヒトや数千万台の接続デバイス等、膨大な数の動的オブジェクトを対象とする場合、移動によって頻繁に位置や時刻が変わります。この結果、RDBのような検索木を用いる場合、

が混雑する」等、実空間上の時間的、空間的な情報発生量の偏りの影響を減らすために、分散KVSを構成する各ノードに分散格納し

ます。

- ④ 全ノード探索の回避：「限定的ノード選択アルゴリズム」を用いることにより、分散KVSの全ノード

ドに分散格納されたデータを検索する際、全ノード検索を回避し、情報が格納されている可能性のあるノードだけから検索を行います。

■時空間コード

時空間コードは、前述した空間コードを時間領域に拡張した情報であり、Z曲線を用いた変換ルールに従って、時間情報と空間情報の各ビットを並べ直した次元情報です。時空間コードの例を図3に示します。この例では、時間36 bit、緯度30 bit、経度30 bitで構成しています。この場合、最小30 ms × 3 cm × 3 cm四方の矩形を表現可能となります。

この時空間コードを用いた「データ格納」と「データ検索」の処理の流れを説明します(図4)。

データ格納では、まず、格納したい時間情報と空間情報を用いて時空間

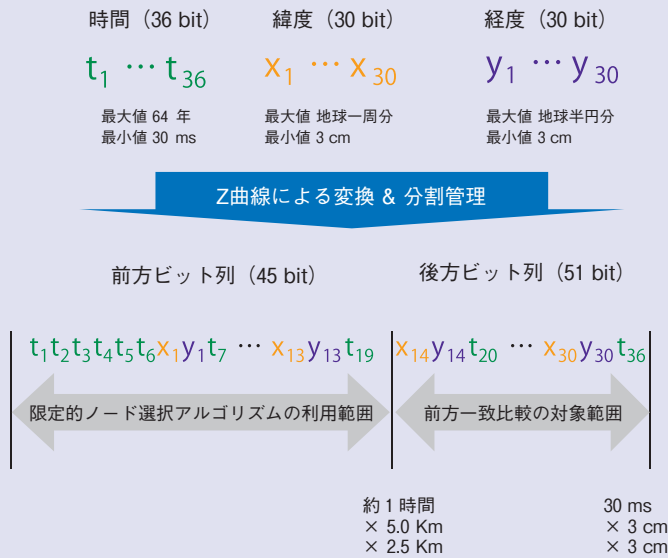


図3 時空間コードの例



※ 格納先ノードの候補が2, 4, 5となる場合の例

格納先ノードの候補だけ検索が実行される

図4 データ格納とデータ検索の処理フロー

コードを生成します。次に、この時空間コードの前方ビット列を用いてハッシュ計算を行います。これによって、全ノードから複数のノードの組み合わせを格納先ノードの候補として一意に求められます。さらに、この候補の中からランダムに1つのノードを選択しデータを格納します。このアルゴリズムを「限定的ノード選択アルゴリズム」と呼びます。

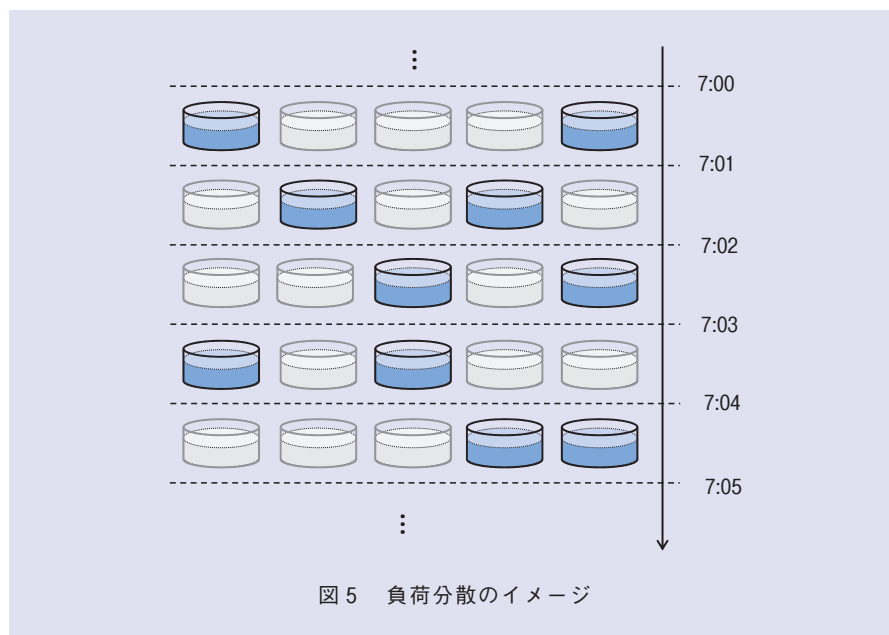
データ検索では、検索条件となる時間情報と空間情報を用いて時空間コードを生成します。次に、この時空間コードの前方ビット列を抽出し、この値を用いてハッシュ計算を行い、格納先ノードの候補を決定します。そして、候補となった格納先ノードは、そこに格納されている時空間コードと、検索条件の時空間コードの前方一致比較を行い、一致するデータを返却します。

■限定的ノード選択アルゴリズム

限定的ノード選択アルゴリズムは、時空間コードの前方ビット列に含まれる時間が時々刻々と変化するため、格納先ノードの候補の組み合わせも時々

刻々と変化します。仮に、ある地域で渋滞が発生した場合、時間の経過による格納先ノードの候補の組合せは変化します。したがって、データ格納時に、ある特定のノードに永続的に負荷が集中し続けることは回避できます。この変化を図5に示します。また、データ検索時は、同様に検索条件で与えられた時空間コードの前方ビット列を用いて、格納先の候補となるノードだけを検索するため、全ノードに対する検索負荷は発生しません。

データ検索時は、データベースに格納されている時空間コードと、検索条件で与えられた時空間コードの前方一致比較によって、時間と空間の範囲検索を一度で実行できます。また、アプリケーションは、検索条件として与える時空間コードの長さを変えることで、検索したい時間や空間の範囲を変更できます。例えば、広い範囲の時間や空間を検索したい場合、短いビット列を用います。狭い範囲の時間や空間を検索したい場合、長いビット列を用います。



本技術を実装した結果、従来方式⁽⁶⁾と比較し、スループットはデータ格納で約13倍、データ検索で約5倍、改善することを確認しました⁽¹⁾。

Axispot[®]のアーキテクチャ

今後は、時空間データ高速検索技術を土台として、道路などの矩形以外の範囲検索など、より高度な時空間データ管理機能を提供し、車両間通信や拡張現実等の次世代サービスへの適用をめざします。

Axispot[®]の全体アーキテクチャを図6に示します。Axispot[®]は、「データベース層」「データベース管理層」「ジオメッシュ層」「ジオメトリック検索層」「ジオメトリック分析層」の5層から構成されます。データベース層は、データを管理する複数のノードで構成される分散KVSから構成されます。データベース管理層では、前述の時空間データ高速検索技術を用いてデータ格納時やデータ検索時に、時間情報や空間情報を用いてデータ格納先ノードを計算します。ジオメッシュ層では、Z曲線を用いて空間情報と時間情報を一次元化した情報（時空間コード）を生成します。この層で行われる検索範囲は、常に100 km四方等の時空間コードで定められる固定長の矩形範囲となります。そこで、ジオメトリック検索層では、このジオメッシュ層の検索結果（矩形範囲）から、特定範囲に含まれる情報だけを絞り込みます。例えば、矩形範囲に含まれる動的オブジェクトから、多角形（ポリゴン）で表現される道路や公園、学区等の複雑な形状に含まれる動的オブジェクトのみを抽出します。ジオメトリック分析層では、ジオメッシュ層やジオメトリック検索層の検索結果を用いて集計や判定処理



図6 アーキテクチャ

等を行います。例えば、検索結果を「MapReduce技術⁽⁷⁾」と組み合わせて、動的オブジェクトを高速集計します。また、フェンスと呼ばれる特定領域に対する動的オブジェクトの出入りを検知します (GeoFencing)。

さらに、本技術をサイバー空間上に応用することで、サイバー空間上の仮想都市や自然環境等を再現した静的オブジェクトと、同じサイバー空間上で行動するヒトやロボット等の動的オブジェクトを、時間と位置情報を用いて合成可能となります。そこで、時間と空間を用いたデジタルツイン合成技術として発展させ、デジタルツインコンピューティング構想⁽⁸⁾の実現もめざします。

■参考文献

- (1) A. Isomura: "Real-time spatiotemporal data utilization for future mobility services," RedisConf19, Jun 2019.
- (2) T. Abraham and J. F. Roddick: "Survey of

- Spatio-Temporal Databases," GeoInformatica, Vol.3, No.1, pp.61-99, 1999.
- (3) N. Pant, M. Fouladgar, R. Elmasri, and K. Jitkajornwanich: "A Survey of Spatio-Temporal Database Research," LNCS, Vol. 10752, 2018.
- (4) M. Hadjieleftheriou, Y. Manolopoulos, Y. Theodoridis, and V. J. Tsotras: "R-Trees: A Dynamic Index Structure for Spatial Searching," 2017.
- (5) <https://blog.labix.org/2008/02/26/geohashorg-is-public>
- (6) <https://www.slideshare.net/DanielHochman/geospatial-indexing-at-scale-the-15-million-qps-redis-architecture-powering-lyft>
- (7) J. Dean and S. Ghemawat: "MapReduce: Simplified Data Processing on Large Clusters," OSDI 2004, San Francisco, U. S. A., Dec. 2004.
- (8) <https://www.ntt.co.jp/news2019/1906e/190610a.html>

(上段左から)花館 蔵之/ 木村 達郎/
 沖 宣宏/ 重松 直子/
 上野 磯生

(下段左から)内藤 一兵衛/ 久保 貴司/
 宮原 和夫/ 磯村 淳

時空間データ管理技術は、リアル空間上の物体の動きをサイバー空間上に再現するリアル・サイバー空間の融合のキー技術となります。今後は、その基本フレームワークとしてAxispot[®]を実現していく予定です。

◆問い合わせ先
 NTTソフトウェアイノベーションセンタ
 第三推進プロジェクト
 TEL 0422-59-3964
 FAX 0422-59-3145
 E-mail info-axispot-p-ml@hco.ntt.co.jp