

時間情報を持つ領域情報のキャッシュ方式

根岸 幸生[†] 渡辺 新[†] 大森 洋平[†] 大沢 裕[†]

[†] 埼玉大学工学部情報システム工学科 〒338-8570 さいたま市桜区下大久保 255

E-mail: † {negishi, shin, yohei, ohsawa}@mm.ics.saitama-u.ac.jp

あらまし 時間情報を持つ領域のキャッシュ方式について提案する。本稿で対象とする時空間情報管理システム STIMS では、空間データ管理方式として暗示的トポロジー方式を採用している。この方式は時空間データベース構築には適した方式である。一方、統計データの数値に応じて領域ごとに、塗りわけの処理や、地域区分の境界線の描画処理には、多くの市町村・町丁目の領域復元が必要であり、長い処理時間を要する。また、本稿で対象とするデータは、市町村・町丁目界の合併や分割により境界が変化する。本稿では、領域復元時に領域を構成するリンクの有効期間を求めることで、領域情報をキャッシュする手法を提案する。また、本キャッシュ方式を、STIMS を基盤とした自治体業務向け簡易型 GIS である文房具 GIS 上に実装し、大幅な速度向上が得られることを示す。

キーワード キャッシュ, 時空間 GIS, 領域情報

1. はじめに

筆者らは、時間情報を扱うことが可能な GIS (Geographic Information System) である時空間情報管理システム STIMS (Spatio-Temporal Information Management System) の研究・開発を行っている。

現在、STIMS をベースとして地方自治体業務向けの簡易型 GIS である文房具 GIS の開発を行っている。本システムの機能の 1 つに、自治体に存在する保健所管内、警察署管内などのような管内図を電子地図上で作成可能な管内図作成機能がある。本機能では、ユーザが Microsoft Excel を用いて管内図の定義を行う。管内図が定義された Excel ファイルをシステム側で読み込むことで、市町村や町丁目の領域を指定された色で塗りわけたり、地域を分割する線を引いた管内図が自動的に生成される。また、高齢化率や犯罪発生率など示した地図の生成など、応用も可能であり、自治体業務において非常に需要の高い機能である。

管内図を表示する際には市町村や町丁目界の領域情報が必要である。現在、多くの GIS では空間データの管理方式として、線の接続情報や面の形成情報などのトポロジー情報を空間データ中に記載するトポロジー暗示方式を用いている。そのため、領域情報を高速に取得することが可能である。一方、STIMS は空間データにトポロジー情報を記述しないトポロジー暗示方式[1]を採用している。そのため、時空間データベースの構築が容易である反面、トポロジー情報を必要に応じて演算により復元する必要がある。よって、トポロジー明示方式を採用している GIS と比べ、領域情報の取得に時間を要する。

1 つ 1 つの領域復元処理の時間は非常に短いものだが、管内図作成機能では一度に大量の領域を復元するため、処理に長い時間を要する。ユーザが文房具 GIS

を利用する上で、種類の異なる管内図や、時系列に沿った複数の管内図を何度も切り替えて表示する場合がある。その度に領域復元処理が行われ、同じ処理時間を要している。

そのため、本稿では時空間データを扱うシステムにおいて、一度復元された領域情報をキャッシュ化し、取り扱うための手法について述べる。それにより、以降同じ領域情報が必要になった際に、高速に取得可能にする。

2. 文房具 GIS と管内図作成機能

2.1. 文房具 GIS の概要

文房具 GIS エラー! 参照元が見つかりません。は地図の表示機能の加えて、以下の各機能を有している。

- 管内図作成機能
- グラフ表示機能
- 作図機能
- 経路略図作成機能
- デジタルアルバム機能

これらの機能の中から、本稿に密接に関連している管内図作成機能について詳細を述べる。

2.2. 管内図作成機能

地方自治体では、各業務毎に、その自治体を複数のエリアに分割している。県単位では、保健所管内、警察所管内、税務所管内などが存在し、市町村単位では、連合町会地区、学区などが存在する。自治体業務の中で作成される資料には、このようなエリアを色塗り分けや、境界線による分割などをして表示した地図を利用するものが多い。昨年度共同研究をした埼玉県庁はもちろん、今年度共同研究をしている川口市においても非常に需要の高い機能である。

この管内図を簡単に作成できるのが管内図作成機能である。表示例を図 1 に示す。

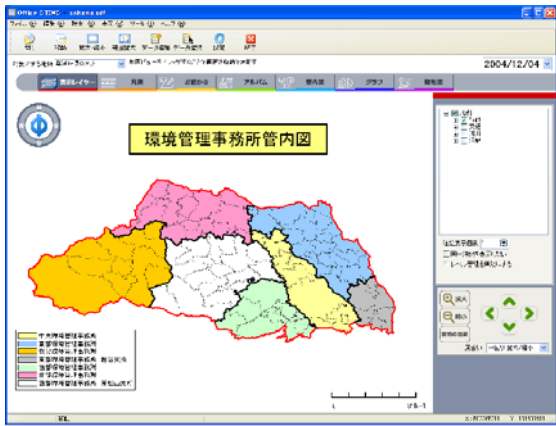


図 1 管内図表示例

このように、市町村や町丁目領域を、任意の色による塗りわけや、太線の描画（以下、動的境界生成）により、区分することが可能である。この様な図を文房具GISで表示するためには、図 2 に示す Excel 表で、ユーザが管内図データの定義を行う必要がある。

凡例		色		
1	名称			
2	中央環境管理事務所			
3	東部環境管理事務所			
4	東部環境管理事務所			
5	狭小環境管理事務所			
6	東部環境管理事務所	越谷支所		
7	西部環境管理事務所			
8	北部環境管理事務所			
9	西部環境管理事務所	東松山支所		
データ				
11	x	y	市町村名 地域区分	
12	502690000	129180000	さいたま市(多野)	6
13	502140000	129390000	川越市	2
14	501780000	130140000	越谷市	3
15	503040000	129000000	川口市	6
16	502140000	130140000	行田市	4
17	500700000	129480000	秩父市	7
18	502079000	129890000	所沢市	2
19	501300000	129180000	蕨田市	7
20	502618000	130080000	加須市	4
21	501120000	130440000	本庄市	3
22	501840000	129720000	東松山市	7
23	502879000	129360000	岩槻市	4
24	503158000	129480000	春日部市	4
25	501899000	129060000	狭山市	2
26	502380000	130000000	羽生市	4
27	502258000	128840000	鴻巣市	6
28	501420000	130320000	深谷市	3
29	502500000	129480000	上尾市	6
30	503280000	129000000	草加市	5

図 2 管内図定義ファイル例

まず、表の 1 行目は、それ以下に凡例が記述されていることを示している。第 2 行は第 1 カラムが凡例に表示する文字列を、第 2 カラムがその文字列に対応する色を定義している。第 3 行目から 9 行目までが、凡例の内容である。

第 10 行目では、以降が管内図定義のデータであることを示している。第 11 行目は、1 レコードを構成するデータの属性が列挙されている。この行以降は、1 つの市町村の情報が 4 つのコラムから構成されており、

第 1 カラムと第 2 カラムが、その市町村の代表点座標の x 値と y 値を表している。第 3 カラムは市町村名を示している。管内図や面塗りのための Excel 表では、1 つの市町村に対してこの 3 つの情報が必須である。市町村を塗り分ける際に、(x,y) の点を種として、領域の復元処理が行われる。第 4 カラムは、先に凡例で定義した実際に表現したい情報の数値が示されている。この Excel 表ではマクロが定義されており、この第 4 カラムに数値を記入すると、その数値に従って第 4 カラムが凡例で定義した色で自動的に塗られる。12 行目からは、市町村の管内図データである。ここで定義された色により市町村領域の塗りわけが行われ、定義された数字に従って、動的境界の生成 (3.2 節) が行われる。

汎用時空間 GIS としての STIMS は、主として地物の位置形状を時間軸を含めて管理するシステムである。地物の属性情報は通常のリレーショナルデータベース管理システムで管理することを前提としている。

一方、文房具 GIS は情報システムには素人である一般の職員を対象とするものであるため、DBMS を SQL で操作することには抵抗が強い。そこで、文房具 GIS においては、各種属性項目を日常業務でなじみがある Excel で管理する方式を採用した。

3. 領域復元と動的境界生成の仕組み

3.1. 領域形状の復元

市町村や町丁目の領域を面塗りする際には、以下で挙げる領域復元処理が行われる。

- ① 領域復元のための代表点座標を取得する。
- ② 代表点を元に、水平線アルゴリズムにより、検索対象となる主題に含まれる最も近い水平線と交差する右側の線オブジェクトを検索する。
- ③ 交差する線オブジェクトが発見されたら、その線の両端点の内、Y 座標値が大きい方の端点を選択する。
- ④ 選択された端点に接続するリンクの一覧を、空間データから取得する。
- ⑤ 取得したリンクのうち、時計回りに見て最も角度の小さいリンクを選択する。
- ⑥ 選択されたリンクの両端点のうち、④ で選択された点ではない端点を選択する。
- ⑦ ④ - ⑥ の処理を繰り返し、⑥ での選択点と、③ で選択された端点とは反対側の端点座標と一致するまで続ける。

② で述べた主題とは、1 つ以上の地図オブジェクトを 1 つの概念としてまとめたものである。例として、県境界や市町村境界をまとめた「行政界」、一般道、高速道路、有料道路などをまとめた「道路」などの

主題がある。空間データ検索や空間演算を効率よく行うために導入された概念である。

必要なトポロジーの取得を効率よく行い、地図オブジェクトの検索処理を柔軟に設定するために必要である。

以上の処理により、領域情報が復元される。復元された領域を、図 2 に示した Excel ファイルで指定された色で塗りわけをする。

3.2. 動的境界生成

動的境界生成の際には、最初に前節で述べた面復元処理が行われた後、次のような処理が行われる。図 3 において P1~P3 は同一管内に、また P4 は別の管内に属するものとする。その時、同一管内の境界を得るために以下の処理を実行する。

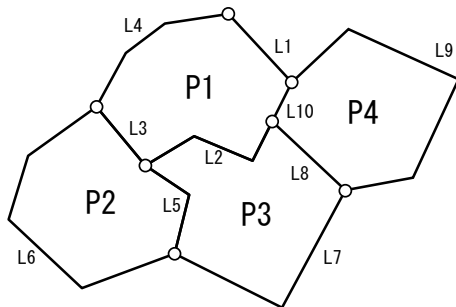


図 3 管内境界線検出の例

- ① P1 に対して領域復元を行い、P1 を構成するリンクを得る。得られたリンクをカウントしていく。
- ② 同様に P2, P3 に対してもそれぞれ領域復元処理を行い、得られたリンクをカウントしていく。
- ③ その結果、各々のリンクのカウントは図 7 の左側のリストのようになり、このうちカウントが 2 以上のリンクを破棄することで、異なる地域区分間の境界を得ることが出来る。
- ④ 上記の処理手順を他の全ての地域区分に対しても行う。

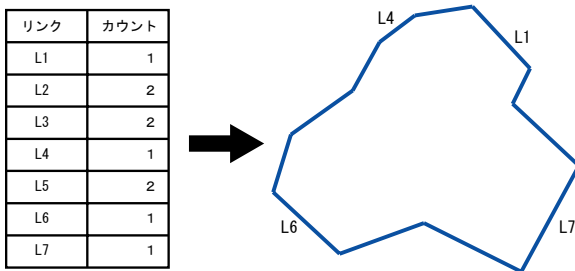


図 4 管内外枠の抽出

3.3. 管内図作成機能の問題点

管内図作成機能には、処理時間に関する問題が 2 つ存在する。1 つは、管内図データを定義した Excel ファイルの読み込み時間が長いこと。他方は、多数の領域復元処理を行うために要する処理時間が長くなることである。

図 1 に示した管内図生成時間を測定した。測定環境と測定結果をそれぞれ表 1 と表 2 に示す。測定環境は、自治体で使用される職員用パソコンの平均的な性能とほぼ同じである。

表 1 測定環境

CPU	PentiumM 1.73GHz
メモリ	512MB
地図	JSGI 数値地図 25000 埼玉県全域 (市町村数：約 90)

表 2 管内図生成時間

定義ファイル読み込み時間	6.50 秒
全領域復元時間	0.90 秒

表 2 に示した通り、管内図作成にはおよそ 7 秒~8 秒の時間を要する。以下で、その原因と解決法について簡潔に述べる。

Microsoft Excel は巨大なアプリケーションであるため、ファイルの読み込みにも長大な時間を要する。そのため、筆者らは、一度読み込んだ Excel の定義ファイルをデータベース化し、次回以降は SQL を利用して、データベース化された定義ファイルを読み込むことで、高速化する手法を提案し、実装した。図 5 に処理の流れを示す。

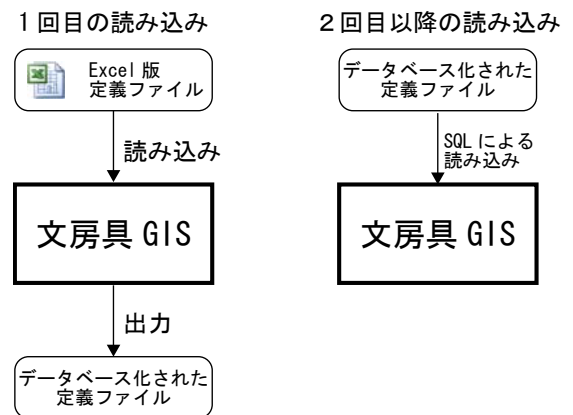


図 5 管内図定義ファイルのコンバート

領域復元処理は、3.1 節で述べた通り、様々な空間演算や検証を行う。1 つ 1 つの領域復元処理時間は、非

常に短いものであるが、大量の領域復元を行う際には長い時間を要する。そのため、筆者らは1度領域復元処理で復元されたポリゴンをキャッシュ化することで、2度目以降の領域復元処理の短縮化を図った。次章にて、キャッシュ方式の詳細を述べる。

4. 領域情報のキャッシュ化

本章では、時間情報を持つ領域情報が復元された後、復元された領域情報をキャッシュ登録し、登録後、同じ領域情報が必要になった際に、高速に取得するための手法について詳細を述べる。

4.1. キャッシュの登録

STIMSで復元された領域情報は、次に示す手順でキャッシュとして登録される。

- ① 領域復元のための代表点を得る。
- ② 領域復元処理を行う。(3.1節)
- ③ 復元された領域情報をキャッシュとしてメモリ上に登録する。
- ④ ③と同時に、キャッシュ情報をハードディスクに書き出し保存する。

キャッシュ登録された領域の情報は配列で管理される。

4.2. キャッシュデータの内容

キャッシュ化された領域情報には、復元されたポリゴンの情報に加え、ポリゴンを構成するリンクの情報が含まれる。ポリゴンとリンクの関係を図6に示す。

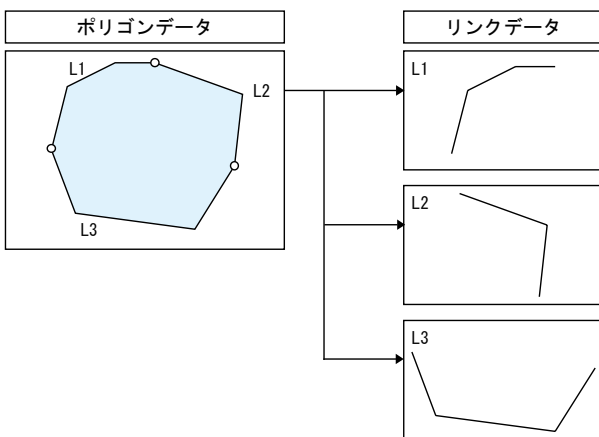


図6 ポリゴンとリンクの関係

ポリゴン情報には、空間演算により復元されたポリゴンの頂点列などの情報が含まれる。この頂点列は、領域復元時に取得したリンクの頂点列を結合して復元されたものである。

リンク情報には、各々が独自にカテゴリ番号や時間情報などの情報を持っている。カテゴリ番号とは、市町村境界、丁町目界など、線の種類別に割り当てられる番号である。時間情報とは、リンクに付与されている生成時間と消滅時間であり、これらの情報によりリンクの存在時間が決定される。これらの情報は、次節で述べるキャッシュの検索時に利用されるため、キャッシュとして登録しておく必要がある。

ポリゴンとリンクがそれぞれ保持する情報を、表3と表4にまとめる。ポリゴン情報は、ポリゴンの頂点列数と頂点列座標、ポリゴンのMBR、ポリゴンを構成するリンクの配列で構成されている。ポリゴンを構成するリンクは、カテゴリ番号、リンクの方向、時間情報、頂点数、頂点列座標で構成されている。時間情報には、リンクの生成時間と消滅時間が含まれている。

表3 ポリゴンキャッシュデータ

変数名	備考
VertexNum	ポリゴンの頂点数
Vertex[0]...[n-1]	ポリゴンの頂点列の座標値
MBR	ポリゴンの外接長方形
LinkList	ポリゴンを構成するリンクへのポインタ

表4 リンクキャッシュデータ

変数名	備考
Category	リンクのカテゴリ番号
Direction	リンクの方向(現在未使用)
TimeInfo	リンクの存在時間情報
VertexNum	リンクの頂点数
Vertex[0]...[n-1]	リンクの頂点列の座標値

4.3. キャッシュの検索

本節では、領域情報のキャッシュ登録後、領域情報が必要になった際の、取得方法について述べる。領域情報のキャッシュが登録されている場合は、領域復元のための代表点を取得後、登録されたキャッシュ中に、取得した代表点に対応する領域情報が存在するか、検索を行う。キャッシュ探索は、図7手順で行われる。

領域復元のための代表点を取得後、取得した代表点座標を元に、後述するキャッシュの検索処理を行う。検索処理でヒットするキャッシュが存在した場合は、キャッシュ登録されている領域情報をシステム側で取得する。キャッシュがヒットしなかった場合は、3.1節で述べた領域復元処理とキャッシュ登録を行う。

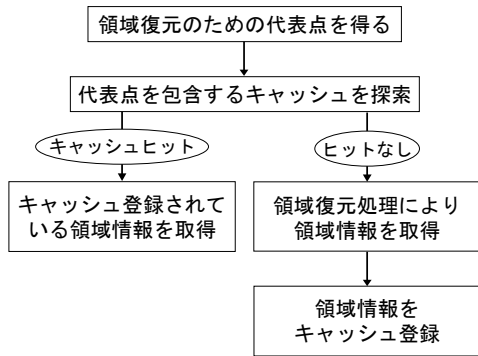


図 7 キャッシュ登録後の面復元処理

次に、キャッシュの探索の方法について詳細を述べる。取得した代表点を元に、配列管理されているキャッシュを先頭から順に図 8 に示す方法で検索していく。

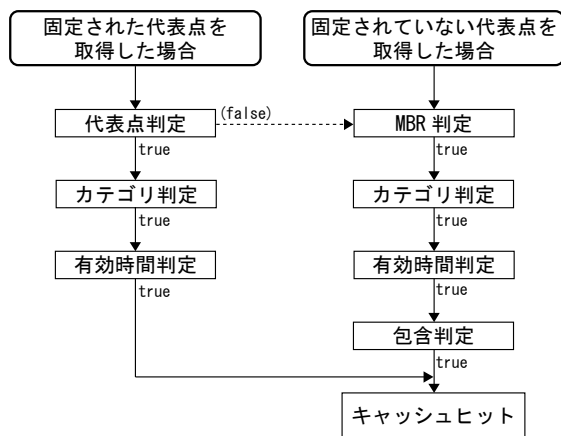


図 8 ポリゴンキャッシュ探索

キャッシュの検索方法は、代表点の取得方法によって方法が異なる。代表点の取得方法は、固定された代表点を取得する場合と、固定されていない代表点を取得する場合の 2 つに分別される。

固定された代表点を取得する場合とは、図 2 に示した管内図定義ファイルから代表点を得る場合である。この場合、1 つの市町村領域に対応する代表点の座標値が確定している。よって、代表点が確定していない場合と比較して、簡略化された処理が可能である。

固定されていない代表点を取得する場合とは、GUI に表示された地図上で、ユーザが取得したい領域内のポイントを指定する場合である。

代表点の取得法ごとに、図 8 に示した流れで判定が行われる。各々の判定処理の詳細を以下で述べる。

- **代表点判定**…管内図作成時は、図 2 に示した Excel ファイルから代表点を取得するため、1 つの領域に対応する代表点が常に一定である。

そのため、取得した代表点とキャッシュ登録されている代表点が一致すれば、対応する領域が確定する。包含判定などを行う必要がないため、キャッシュを探索する時間を大幅に短縮できる。本判定で、代表点一致しなかった場合は、固定されていない代表点を取得した場合と同様の手順で判定処理を行う。

- **MBR 判定**…本判定は、固定されていない代表点を取得した場合に最初に行われる。取得した代表点が、キャッシュ登録されたポリゴンの MBR 内に含まれるかを判定する。
- **カテゴリ判定**…ヒットしたキャッシュを構成するリンクに割り当てられたカテゴリ番号が、現在システムで指定されている主題に含まれる否かを判定する。
- **有効時間判定**…ヒットしたキャッシュの有効時間が、システムで指定されている興味対象時間内に含まれるか否かを判定する。詳細は 4.5 節で述べる。
- **包含判定**…代表点がポリゴンに含まれるかを判定する処理である。厳密な判定を行うため、他の判定に比べて処理時間が長い。そのため、他の判定処理により、本判定は最後に行われる。

以上の判定処理を経て、キャッシュがヒットした場合は、リンクの頂点列座標をシステム側で取得する。取得したリンクを結合して領域のポリゴンの描画や、動的境界の生成を行う。

4.4. 動的境界生成への対応

本節では、キャッシュ化された領域情報を元に、3.2 節の動的境界を生成するための手法について述べる。3.2 節で述べた通り、動的境界の生成には、領域のポリゴンを構成するリンク情報が必要となる。4.2 節で述べたように、領域情報のキャッシュには、ポリゴンを構成するリンク情報も含まれている。キャッシュがヒットした場合には、リンク情報を取得するため、この情報を元に動的境界の生成が可能である。

4.5. STIMS の時間管理機能との連動

STIMS では、興味対象時間の変化とともに地物も変化する。特に近年、市町村合併が盛んに進められているため、市町村境界の変化が激しい。管内図作成機能を利用する際も、市町村領域が変化する地図を利用する場合が当然ある。よって、時間情報を持つ領域情報のキャッシュ方式を実装する上では、システムで設定されている興味対象時間において存在する領域のキャッシュは有効とし、興味対象時間において存在しない

領域のキャッシュは向こうとする必要がある。

4.2節で述べた通り、キャッシュにはリンクの開始時間と終了時間が記述されている。これにより各々のキャッシュの有効時間が決定するので、キャッシュを検索する際に有効時間判定をすることでSTIMSの時間管理機能と連動することが可能である。

キャッシュの有効時間判定の例を図9に示す。図9は、A市とB市が、2005年4月1日に合併し、D市が誕生した様子を表すものである。隣り合うC市の領域に変化はない。リンクの時間情報により、キャッシュの有効時間が図9に示すように決定している。開始から終了までの時間が、システムの興味対象時間内に含まれている際は、そのキャッシュは有効となる。逆に、含まれていない場合、無効なキャッシュとなる。そのため、システムの興味対象時間がT1(2005年3月31日以前)の場合は、A市、B市、C市のキャッシュが有効となる。T2(2005年4月1日以降)の場合は、C市と新たに誕生したD市のキャッシュが有効である。

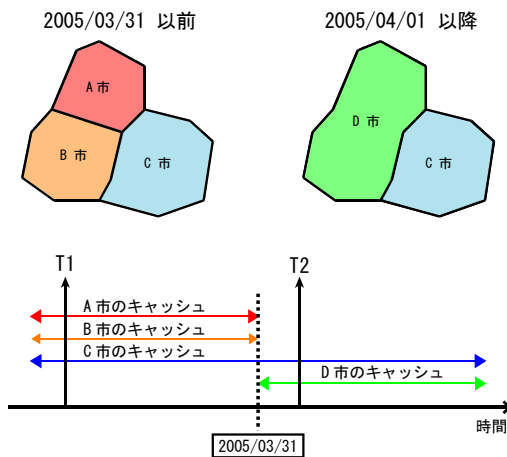


図9 ポリゴンキャッシュの時系列対応

4.6. 従来の方式との比較

1章で述べた通り、STIMSでは、トポロジー情報を演算することで復元する。そのため、復元時間短縮のために、トポロジーキャッシュが利用されている。

ここで、3.1節の領域復元処理を例に挙げ説明する。トポロジーキャッシュとは、領域復元処理④における端点と接続するリンクの一覧を、キャッシュ化したものである。オブジェクトの探索時間を短縮化するためのものであり、このトポロジー情報を元に、領域復元処理などの空間演算が行われる。

本稿で提案した領域情報のキャッシュ方式では、空間演算結果をキャッシュ化するため、さらなる時間短縮が可能である。次章では、本キャッシュ方式をSTIMSを基盤とした文房具GIS[3]上に実装し、管内図作成時間がどの程度向上されたか、測定結果を示す。

5. 管内図生成時間測定

埼玉県の管内図生成時間を図10に示す。測定環境は表1と同様である。

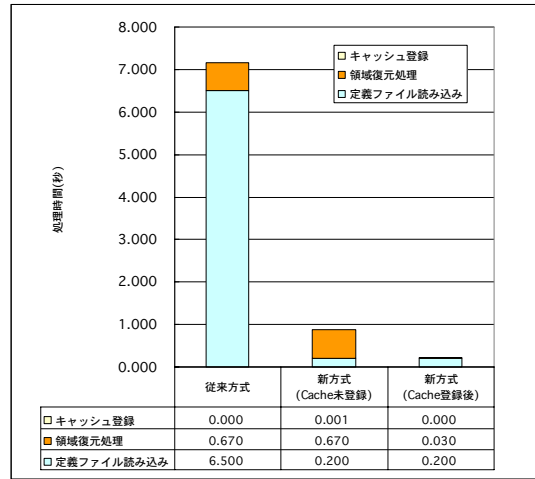


図10 管内図生成時間測定結果

図10の従来方式とは、Excel版定義ファイル読み込みとキャッシュを用いない領域復元処理による管内図生成方式である。新方式は、キャッシュ未登録時と、キャッシュ登録時の時間を示した。

新方式による管内図生成時間は、キャッシュが未登録時には1秒強の時間を要している。しかし、キャッシュ登録後は、0.2~0.3秒程で管内図生成が可能となった。これは、自治体職員が管内図作成機能を利用していく上で、実用上問題がない速度が実現できたとと言える。

6. まとめ

本稿では、STIMSにおける領域復元処理の高速化のための手法として、領域ポリゴンをキャッシュ化する手法について述べた。これにより、一度に大量の領域復元処理を行う場合でも、高速に領域情報を取得することが可能となった。

文献

- [1] 大沢裕, 長島敦, “トポロジー暗示型時空間情報管理システム: STIMS,” 第12回機能図形情報システムシンポジウム講演論文集, pp.12-36,2000.
- [2] 野中秀樹, 大沢裕, “位相構造を持たない地理情報システムに関する考察,” 第8回機能図形情報システムシンポジウム講演論文集, pp.27-32,1997.
- [3] 根岸幸生, 渡辺新, 鄭泉, 大森洋平, 大沢裕, “行政のための簡易型時空間情報管理システムの構築,” 情報処理学会研究報告 2005,No.68, pp.685-690,2005.
- [4] 天笠俊之, 有次正義, 金森吉成, “時間的に変化するデータに対する索引技術,” 情報処理 2001.10 pp.972-979, 2000.