

分散型 GIS 「GLOBALBASE」 の実装

森 洋久

国際日本文化研究センター
〒610-1192 京都府京都市西京区御陵大枝山町 3-2
Tel: 075-335-2165 Fax: 075-335-2090
E-mail: joshua@nichibun.ac.jp

あらまし

本研究では、空間情報の蓄積の目的や性質の違いで別々に選択された座標系同士を、分散環境で一つにつながり合わせ、あたかも一続きの空間としてブラウジング可能な GLOBALBASE アーキテクチャについて説明する。本研究では、まず座標系同士の座標変換を記述したマッピングと呼ばれる記述を新たに導入した。また、このマッピングと座標系のブラウジングを可能とするプロトコルの実装について説明する。

以上の説明の後、マッピングを導入したときの効果について、国際日本文化研究センターで行っている地理情報の蓄積の例より評価する。

キーワード GIS、座標変換、マッピング、GLOBALBASE

An Implementation of GLOBALBASE: the distributed GIS

Hirohisa Mori

International Research Center for Japanese Studies (IRCJS)
3-2 Oeyama-cho, Goryo, Nishikyo-ku, Kyoto 610-1192, JAPAN
Tel: 075-335-2165 Fax: 075-335-2090
E-mail: joshua@nichibun.ac.jp

Abstract

This study investigates and evaluates GLOBALBASE architecture that ties together coordinate systems that are selected individually due to different accumulation purposes and characteristics of the spatial information, so that they can be browsed through as if they were one continuous space. In this study, a new mapping describing the coordinate transfer between such coordinate systems is first introduced. Secondly, the protocol layers to browse the coordinate and mapping data structure are explained.

At last, this study evaluates the effect of mapping technology for heterogeneous spatial information such as old maps and illustrated maps using some examples of IRCJS spatial information database.

Key words GIS, Coordinate Transfer, Mapping, GLOBALBASE

1. はじめに

我々の住んでいる土地がどのように変化してきたのかという興味は古今東西変わることはない。地理情報を重ねたり、つなぎ合わせたいという要求はかなり古くから存在する。

日本においては、我々の調べでは、18世紀の中頃、森幸安という人物がおり、日本全国の地理情報を集積し、膨大な地図を残している。「緯度経度」という用語を日本人に紹介したのは、伊能忠敬よりも古い。幸安は、収集した様々な国絵図などの情報を、さらに長久保積水らが作成したと思われる、測量地図に近い日本図により、位置精度を高めた地図を作っている。各国に分割されているが、実際に国の境界線で接続できるように構成されており、おそらく、全部つなぎ合わせると、巨大な一つの日本全図になると思われる。さらには、韓国や中国といった近隣諸国の地図、アジア地図、はてはマテオリッチ世界図なども参照しており、森幸安の頭の中には、先の日本地図の作製の延長線上に、地球が構成されていたと思われる。

同じような興味は、ヨーロッパなどでも存在し、パスカルなどが良い例である。現代でも、Powers of Ten [1]といった作品があり、地球全体と、自分自身の生活空間レベル、あるいはそれ以下のミクロな地域とを連続に結ぶスケラブルな空間感覚は、人間の基本的な欲求のようである。

現代における地理研究古地図研究においても、地図を重ねたり、つなぎ合わせる技術は重要な技術であり、コンピュータを利用した試みが行われている[2][3]。

この著しい成果の一方で、これらの試みの限界は、一人、あるいは一グループで地理情報を集めているということである。この限界を超えて不特定多数、あるいは、それに近いグループで地理情報を集積しようという試みは、ECAI[4]に見られる。ECAIは文系研究情報を地理情報に結びつけ、ある一定の基準を満たしていれば、だれでも、サーバに蓄積することの出来るシステムを提供している。また、WMT[5]は別々に作られた地図を一つのクリアリングサーバを介して、共有、重ねあわせする実験であった。

しかしこれらのシステムは中心に一つのサ

ーバが存在するために、そのサーバのサポートする以上のことは出来ないという機能上のボトルネックが存在する。その限界の中で最も重要な問題の一つは座標系の問題であろう。これらのシステムでは、情報の重ね合わせを統一的な基準で行おうとするために、座標系を統一し、共有することを前提としている。しかし、世界中に散らばっている様々な地理情報は、それぞれ最適な座標系があり、それらはすべて異なっていると言っても過言ではない。

この矛盾は古地図を集積する場合を考えれば顕著であろう。古地図は以上の共有サーバには載せることの出来ない地理情報の一つである。竹原好兵衛の京図という、江戸時代に非常に広く流布した木版摺りの京都の地図を例に取ってみよう。誰かがある竹原好兵衛の京都図を多大な労力をかけて現代図に重ね合わせ公開したとする。それに続いて、内容は違いますが、同じ系統の竹原好兵衛の地図を同様に現代図に重ねた人がいたとする。しかし、最初のひとが重ね合わせた結果の地理情報だけ公開したのでは、次のひとは、おそらく同じ労力を最初から払わなければならない。しかし、もし、重ね合わせ方も同時に公開していれば、同じ系統の版木であれば、次のひとはほぼ自動的に現代図に重ねることが出来る。

現代の測量図であっても、同じことが言える。サイバーマップジャパンによれば、マピオンによって公開されている日本全国図は、各公共座標系の境界部分にあたる地域で、座標変換が理論通り行かず、微妙にずれが生じているという。それを補正する作業は結局人間のアドホックな判断によるという。実際われわれも、国土地理院の数値地図2500を日本全国つなぎ合わせるときに、既存のGISを使い、公共座標系を緯度経度に変換しつなぎ合わせると、微妙なずれを生じることを観測している。現代図でもこのような理論に寄らない部分があり、別々の系統の地図で集積された地理情報を正確に重ねようとした場合、人間の経験と勘によるしかない。しかし、そのような労力をもし共有する事が出来れば、同じことを二度行う必要は無くなる。これが、世界図や隣接する隣の国同士の地理情報となると測量図であっても、どのような誤差が生じるかはほとんど予想がつかない。しかし、現在の座標系を統一、共有するシステムでは、このような労力を二度三度行うこともやむを

得ない。結果的にお互いに接続することのないいくつものサーバが立ち上がる結果になり、また、載せられる地理情報にも制限が付いてくることになる。

この問題に対する我々の解決方法はすでにおわかりであろうが「座標系を統一し共有するのではなく、座標系間の重なり情報を共有する」ということである。二つの地理情報の異なる座標系の同一地物や座標点同士の対応表を公開すれば、二つの地理情報を重ね合わせる事が出来る。この情報を我々はマッピング情報と呼んでいる。さらに座標系 A と B、B と C の間にマッピングが存在すれば、その二つのマッピングをたどることにより、A と C もさほど苦労せず重ねることが出来るだろう。マッピング情報によって相対化された座標系とマッピングのネットワークの情報の詳しい構造については、論文[6]で詳しく述べているので参照されたい。

本論文では、複数のマッピングをたどり、任意の二つの座標系を重ね合わせることを、ブラウジングをしながらダイナミックに行うためのプロトコルとアーキテクチャである、GLOBALBASE の実装について説明し、その定性的な評価を行う (図1)

本論文 2 章では、GLOBALBASE のクライアントの動作がユーザにどのように見えるのかを示す。そのあとで、その動作を実現するためのプロトコルを示す。3 章では、国際日本文化研究センターにおいて、当システムを利用して地理情報を集積している実例を示し、GLOBALBASE のマッピングを基礎としたデータ構造の効果と問題点を定性的に明らかにしていく。最後に、実装されたプロトコルの問題点を示す。

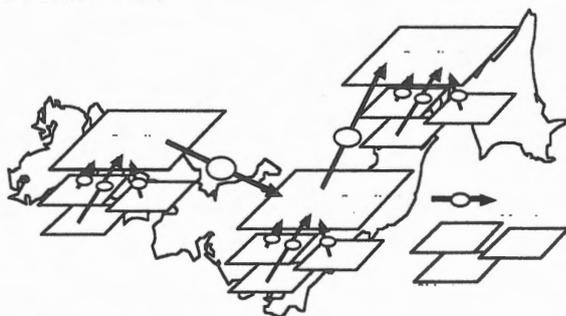


図1. マッピングを基本としたデータ構造

2. クライアントの動作の定義

本章では、ユーザが GLOBALBASE 上の地理情報を検索する方法を示す。プロトコルは

この動作を実現するために存在すると考えられる。

ユーザは、まず、クライアントを立ち上げると、地理情報における座標系の検索条件を設定する。地理情報の持つ内容年代は条件の中で必須である。これを受けプロトコルは、現在ユーザが画面上で見ている地理的範囲と見ている拡大率 (ディスプレイに対する縮尺) を AND 条件でユーザの提示した検索条件へ付加し、GLOBALBASE サーバへ問い合わせに行く。どのサーバへ問い合わせにいくかは後述する。

得られた検索結果、座標系のリストのうち、クライアントは基準となる座標系を一つ選ぶ。その後、リストの中にある他の座標系をこの基準座標系へ変換するための、重ね合わせ情報をサーバに取得しに行く。これに対するサーバの返答は、複数のマッピングとなることもある。つまり、基準座標系と、リストの他の座標系の間直接張られたマッピングが存在するとは限らない。さらに他の座標系が挟まっており、複数のマッピングをたどらないといけない場合がある。このマッピングの列を二つの座標系間のマッピング・パスと呼ぶことにする。

マッピング・パスがリストの各座標系に対し、検索終了したならば、あとは、このマッピングに従い座標変換し、結果をディスプレイに表示する。

ユーザが見る位置を移動し、見えなくなった座標系があると、その座標系を表示対象から除く。その一方で、位置が移動したために新しい座標系が表示可能となっているかも知れない。そのため、見ている位置や拡大率が変化するたびに、クライアントは定期的にサーバへ問い合わせを行う。得られた結果を座標系のリストへ加え、基準座標系からのマッピング・パスを計算する。

もし、基準座標系がユーザの見ている範囲からはずれた場合は、現在あるリストの中から別の座標系を新たな基準座標系として選び直す。

ブラウジングの動作は以上の繰り返しである。当然、ユーザは途中で検索条件を変更することも出来る。

3. プロトコル

2 章の考察から、サーバは分散環境上で、

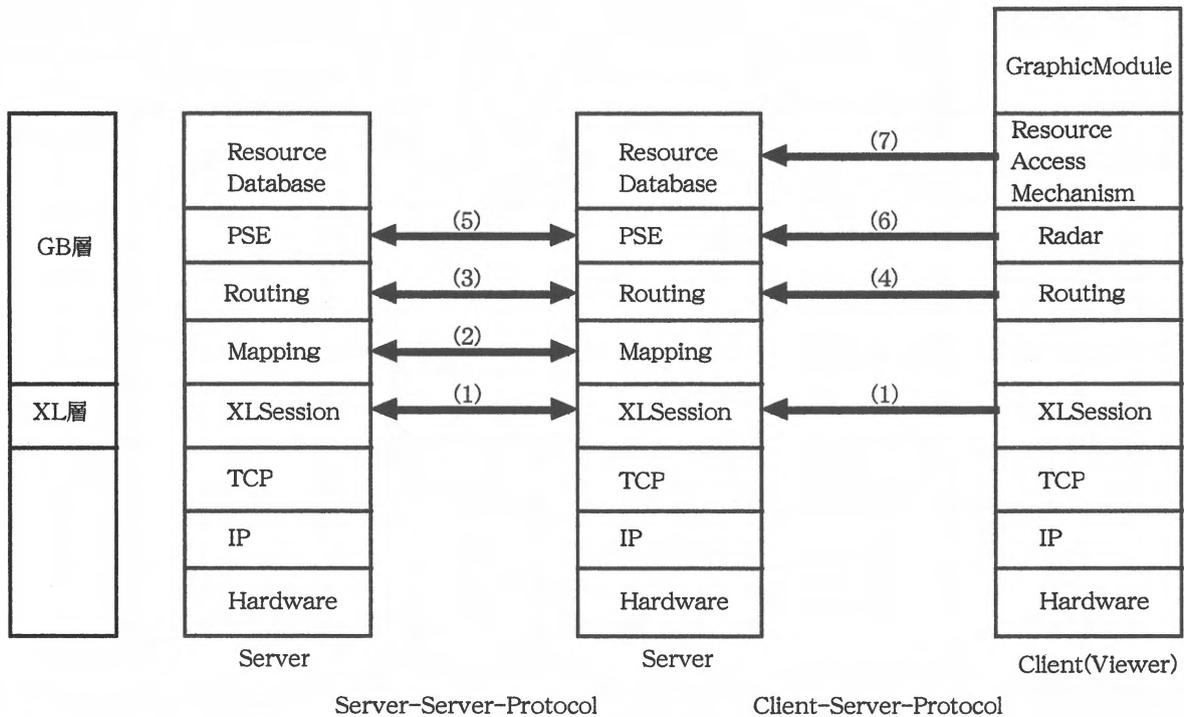


図2.GLOBALBASE プロトコルスタック

- i. どのようにクライアントからの座標系の検索要求をどのように検索するのか
- ii. マッピング・パスを検索する方法はどのようにしたらよいか。

という二つの問題を解決しなければならぬ。そのため、GLOBALBASE プロトコルでは、図 2 に示すプロトコルスタックを提供する。

図 2

まず、この二つの問題を解決する前提として、座標系間、時にはサーバ間をもまたがるマッピングのネットワークをマッピングの正方向からも、また逆方向からもたどれるように情報を管理しなければならない。つまり、各座標系はどの座標系からマッピングが張られているのかという情報をすべて保持している。もし、サーバや途中のネットワークがダウンして、マッピングをたどることが出来なくなった場合は、その状態を検知し、一時的にマッピングを無効にする。再びネットワークが復活すると、マッピングを有効に戻す、といった処理をしなければならない。この処理を行うネットワーク層をマッピング層と呼んでいる。このメカニズムについては論文[6]を参照されたい。

3.1. 分散型検索エンジン(Partial Search Engine)

その上で、まず一番上層の PSE (Partial Search Engine) について説明する。これは、クライアントからの座標系の検索要求を処理する層であり、i. の問題を解決する。検索エンジンであっても、完全に分散している必要がある。これは、クライアントが定期的に発行するクエリを分散化し、サーバに対する負荷を軽減するためである。

各サーバには小さな検索エンジンを持たせる。この検索エンジンは日頃からこのサーバの周りの座標系の情報を、マッピングをたどって収集しており、インデックス化している。この小さな検索エンジンを PSE と呼んでいる。

検索エンジンは座標系の位置、および縮尺や解像度、内容年代は必須のキーとして持っている。その他に様々な座標系に付加されたメタデータを検索することが可能になっている。

この検索エンジンはサーバ上の一つの座標系に張り付いており、集められた座標系の位置情報は、この張り付いている座標系の値に変換されエンジンに登録される。この座標系と検索エンジンのセットをランプ (固まり) と呼んでいる。その座標変換方法は、張り付いている座標系と、対象の座標系間のマッピ

ング・パスが検索できれば良いので、ii. の問題に帰着できる。

次に各 PSE どの範囲の座標系を検索対象とすれば良いかということであるが、これはかなりアドホックに決めている。各座標系は、自分の近くのランプがマッピングを何ステップたどって到達できるかという情報を保持しており、一番近いものから n 番目を超えると、その情報は破棄する。n の値を最適に決定する方法が今のところ内での、現在は 5 としている。この上で、各座標系はその n 個のランプの PSE に対して、定期的に登録クエリを発行する。このクエリにより、各ランプにはその周辺の座標系の情報が収集されることになる。

ネットワーク上のランプの数については、各サーバに一つと考えている。この数は、各サーバが完全に周囲のネットワークから切断されたとしても、自立的に動作できる最小の数である。

以上のようにして各サーバに集められた検索情報を、クライアントはたぐることによって必要な座標系の情報を集めることができる。まず、クライアントは立ち上がった瞬間にどのサーバのどの座標系に接続するかは明示的に指定しておく必要がある。その上で、接続した座標系の周囲のランプの情報を取得し、発見された PSE に対してクエリを発行し、座標系のリストを得る。得られた座標系のリストの各座標系の周囲にある PSE も、同様に問い合わせた対象へ入れる。このようにして PSE の数を増やしていくが、ユーザのしている範囲の移動や拡大縮小により、見ている範囲と、PSE の対象座標系の範囲が重ならなくなった PSE に関しては、問い合わせ対象から外していく。

以上のように PSE のアーキテクチャは、一つのサーチエンジンが全世界、あるいはそれに近い地理情報を包含することが無いので、小さく軽い実装が可能な反面、必要な地理情報が検索できないのではないかという懸念が生じる。実際その通りなのであるが、おそらく、これで検索出来ない地理情報は重ねても無意味であると考えられる。なぜならば、一つの PSE、あるいは近隣の PSE に入っていない二つの座標系は、相当に長いマッピング・パスをたどらないと重ね合わせ出来ないということである。短いパスがないということは、それだけ重ね合わせの需要が無かったということであるし、また、実際にマッピング・パ

スを基に重ね合わせても誤差が大きくなってしまふであろう。

もし、この二つの座標系を有効に重ね合わせたいと思ったユーザが現れた場合は、そのユーザがあらたにマッピングを定義するであろう。そうしたならば、そのマッピングの存在により、二つの座標系は同時に検索可能な範囲の PSE に収まることになる。

3.2. マッピング・パス検索

マッピング・パスの検索は、インターネットにおける経路選択アルゴリズムと等価である。各座標系に経路選択可能アドレスを割り振る。インターネットで言えばコンピュータに振られた IP アドレスに相当する。各サーバは、インターネットにおけるルータと同じように、日頃から経路選択テーブルを準備しており、その上で、二つの座標系間のマッピング・パスは、座標系間でトレースルートをすれば検索出来る。

しかし、ここで問題になるのは、経路選択可能アドレスを各座標系に振ることを、ほとんど不特定多数となり得る地理情報発信者に任せるわけにはいかないということである。サーバの管理者などに任せても、マッピングは上記不特定のユーザにより縦横無尽に張られるわけで、管理不能となってしまふ。これは、インターネットで言えば、ルータや国境を越すネットワークを不特定多数のユーザが日夜増設し続けている現象と似ている。そこで、経路選択アドレスを自動割り当てするメカニズムが必要である。

このためのアルゴリズムとして、我々が以前にネットワークのルータをも含めたノードの経路選択アドレスを自動割り当てを行うアルゴリズム ACRP (Auto - Configured Routing Protocol) が有効である[7]。このプロトコルをインターネットで実現しようと考えると、一つのマシンの IP アドレスが時々変換するので DNS の仕組みが大きな問題となった。しかし幸いなことに GLOBALBASE では、座標系は固定された URL を持っているので、この URL をもとにその座標系自身にアドレスを問い合わせれば良いわけである。

4. 座標系の重ね合わせと、マッピング導入の効果

4.1. 日文研におけるベースマップ作成

日文研では、現在 GLOBALBASE を利用した地理情報データベースのベースマップとして、世界地図および日本列島全域の地図の入力を行っている。

まず日本列島であるが、国土地理院より公開されている数値地図25000と、数値地図2500を利用して、前者は25000分の1の地図より入力したビットマップの地図である。緯度経度座標系に載っているデータである。後者は2500分の1の地図より入力されたベクタデータであり、公共座標系1~19にそれぞれ載っている。

以上の二種類の地図をそれぞれマッピングによりつなぎ合わせ、数値地図25000のシームレスな日本列島、および、数値地図2500によるシームレスな日本列島を作製した。

最後に数値地図25000と数値地図2500同士をマッピングで接続し、あらい解像度の時は、数値地図25000による地図、細かい解像度の時には、数値地図2500による地図が表示されるようにする予定である。

次に世界地図であるが、これは実験的にメルカトル地図をテスト用にスキャンし、緯度経度座標系に張り付けた。

最後に世界地図、日本地図、および、以前から日文研において蓄積してきた京都の地理情報群をすべてマッピングにより接続する予定である。

4.2. 日本列島接続時のマッピングの導入の効果

マッピングの導入における一つ目の効果は、数値地図2500における、公共座標系間の接続の問題がこれによって単純に解決出来たということである。通常 GIS を利用する場合、隣接する公共座標系上のデータを同時につなぎあわせて利用する場合には、いったん両者のすべての情報を緯度経度座標系（この場合ベッセル楕円体座標系）へ置き換えて、つなぎ合わせる。しかし、実際にはつなぎ合わせ時には誤差が生じる。しかし、今回、出来るだけ離れた地点の二つの地物を目標に、回転と、若干の拡大縮小ですべての隣接する公共座標系をずれなくつなぎ合わせる事が出来た（図3）。



図 3-1 地球



図 3-2 25000 分の 1 数値地図（国土地理院）近畿地方

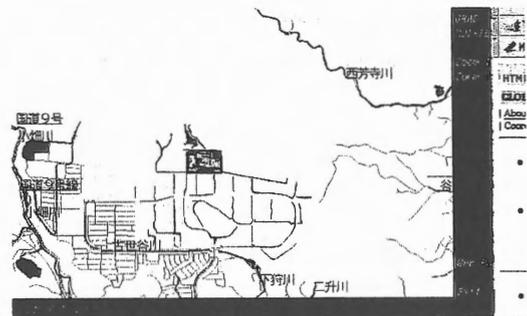


図 3-3 2500 分の 1 数値地図（国土地理院）国際日本文化研究センター周辺

当初の計画では、3点あるいはそれ以上の地物を目標に重ねる必要があると思われたが、実際は2点で十分重ね合わせることが出来た背景には、数値地図2500の基のデータは、座標変換を自由に施すことの出来ない紙の地図から作られており、実用性を考えると、もともと公共座標系間で単純に接続出来るように作図されていたと考えることが出来る。こう考えると、理論的な座標変換をして、緯度経度座標系へ持っていっても重ならないのは当然である。

ここで、理論に寄らず、地図作製における人間の手作業を忠実に保存する、マッピングの効果の一つ示されたことになる。

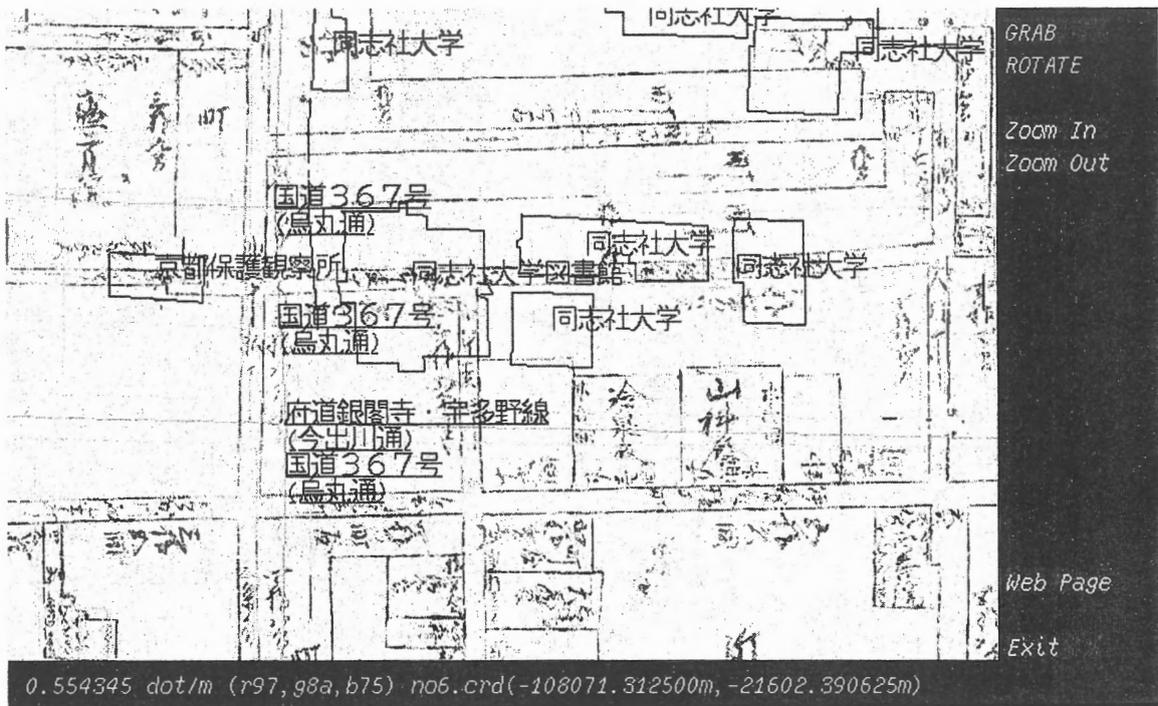


図4. 古地図との重ね合わせ

4.3. 古地図との重ね合わせ

京都市内にあたる部分の1638年の様子を描いた、宮内庁書陵部所蔵中井洛中絵図と現代図の重ね合わせは、道の辻を目標点にして重ねた(図4)。現在は全部で300点くらいの目標点である。目標点の間はTIN(三角パッチ)を生成し、線形補完により補完している。もし、重なり具合が不十分な場合は目標点を増やしていけばよい。同様の地図が京都大学にもあるが、現在重ね合わせは進行中である。

5. プロトコルの評価と諸問題

現在プロトコルに関しては最初の実装が終了したばかりで十分な評価が行われたとは言えないが、現時点で発生している諸問題を観察し、評価に変えたい。

まず、あらゆるネットワークを超えた処理に共通して発生している問題は、ネットワークレイテンシーの問題である。問い合わせを発生させ、それに返答が来るまでに、TCP/IPのフロー制御などの低レイヤの問題をも含めて様々なネットワーク処理が発生するためかなり返答が遅くなる現象が発生している。この問題を解決するためには、二つの方法が考えられる。一つは、問い合わせは出来るだけ多くの命令をまとめて送信すること。こうすることにより、命令に対するパケットの数

が減り、低レイヤの処理時間の問題も解決する。

次に、ネットワークがダウンした場合のプロトコルの動作である。これに関しては、まだ様々な問題が存在する。ネットワークがダウンしても長時間たてば、サーバは、アクセス出来ない部分は切り離し、残りの部分で正常動作するように、各状態の修正が行われることは確認できた。しかし、ネットワークの異常にサーバが気づき、修正に要するまでの時間が現在1週間程度である。

この時間の原因は、マッピングの両側の座標系に保持されている、マッピングおよび対向側の座標系の情報が、タイムアウトによって削除されるまでの、タイムアウト時間がかなり長くなるためである。タイムアウトによってマッピングが切り離されないと、マッピング・パス検索のための経路選択テーブルの書き換えや、近隣ランプの情報の書き換えが発生しないのである。

しかしタイムアウトを短くすると、マッピングは短い間隔でポーリングせねばならず、サーバに負担がかかってしまう。そこで現在検討しているのが、クライアントからアクセスがあり、異常が発生した場合に、その発生座標系の周辺で部分的にマッピングの接続チェックを行う仕組みを導入することである。これにより、アクセス頻度の多い場所は異常

6. 結論

本論文では、地理情報を分散環境上で有効に交換するためのアーキテクチャ、GLOBALBASE について説明した。得に座標系を共有せず、座標系間の重ね合わせ、つなぎ合わせのための情報をネットワーク上で共有することが重要であり、そのためのマッピングと相対化された座標系について論じた。このデータ構造を保持し、自動的な座標系の検索、重ね合わせ、つなぎ合わせを行う GLOBALBASE プロトコルについて説明した。これらはまだ最初の実装が完成したばかりであり、まだ評価は不十分であるが、マッピングの効果、プロトコルの諸問題を、定性的に観察した。

その結果、マッピングによるデータ構造は重ね合わせ誤差の解消、古地図などの標準的な座標系に重ならない図面の重ね合わせなどに効果を発揮した。

参考文献

- [1] Philip Morrison and Phylis Morrison, Powers of Ten Scientific American Library Paperback
- [2] エーピーピーカンパニー「江戸東京重ね地図」丸善株式会社
- [3] 原 正一郎, 安永尚志「メタデータを利用した分散データベースの統合」人文科学とコンピュータシンポジウム、情報処理学会 2000.
- [4] <http://www.ecai.org/>
- [5] 嶋田 茂, 山浦 晃裕, Lance McKee 「OpenGIS が拓く地理空間情報処理の世界」情報処理学会, 情報処理, Vol.41, No.6, pp. 661-665, 2000.
- [6] 森 洋久「歴史研究における GIS: GLOBALBASE のための座標変換メカニズムの検討」情報処理学会研究会報告 Vol.2000, No.49, pp.63-66, 2000
- [7] Hirohisa Mori, Ken Sakamura *A New Network Layer Protocol with Routing Addresses and Tables Auto-Configuration*