

2005-01198A

厚生労働科学研究研究費補助金

健康科学総合研究事業

地理及び社会状況を加味した地域分析法の開発に関する研究

平成17年度 総括・分担研究報告書

主任研究者 浅見 泰司

平成18(2006)年 4月

目 次

I. 総括研究報告

地理及び社会状況を加味した地域分析方法の開発に関する研究	1 - 1
------------------------------------	-------

浅見泰司

II. 分担研究報告

1. 空間ドキュメント管理システムの設計と開発に関する研究	2 - 1
-------------------------------------	-------

有川正俊

2. 集積地域同定のための新たな検定手法の開発に関する研究	3 - 1
-------------------------------------	-------

丹後俊郎

3. 健康危機情報における空間分析手法の開発に関する研究	4 - 1
------------------------------------	-------

浅見泰司

4. 北九州市の小学校欠席率からみた 2004 年～2005 年	5 - 1
--	-------

冬季インフルエンザ流行の時空間解析

郡山一明

III. 研究成果の刊行に関する一覧表	6 - 1
---------------------------	-------

IV. 研究成果の刊行物・別刷	7 - 1
-----------------------	-------

厚生労働科学研究費補助金（健康科学総合研究事業）
総括研究報告書

地理及び社会状況を加味した地域分析方法の開発に関する研究

主任研究者 浅見 泰司 東京大学空間情報科学研究センター教授

研究要旨

(1)さまざまなデジタルドキュメントをドラッグ&ドロップするだけで自動的に住所や地名を抽出し、緯度経度を導出して、地図上で空間分布を閲覧できるソフトウェアシステム「空間ドキュメント管理システム (SDMS)」の基本設計および開発の研究を行った。(2)地域同定のための検定法の評価法として新たな指標を提案し、flexible scan 法と circular scan 法の比較を行い、flexible scan 法が circular scan 法より精確に地域を同定できることを確かめた。また日々のサーベイランスに向けて、この flexible scan 法を時間変化も考慮した空間-時間的解析への拡張を試みた。(3)心臓疾患による救急車搬送発生地点を用いた AED (自動体外式除細動器) の設置地点に関する分析をおこなった。地域における装置の需要量および生存退院率にもとづく救命確率を考慮した、AED の最適配置地点を求める方法を開発した。(4)北九州市における 2004 年~2005 年冬季インフルエンザ流行を小学校欠席率に基づく時空間解析を行い検討した。小学校欠席率はインフルエンザ流行を把握するのに極めて有効であり、地域健康危機管理の日常的なサーベイランスとして地図とリンクさせることが必要であることを導き出した。

分担研究者

丹後 俊郎

国立保健医療科学院
技術評価部部長

郡山 一明

救急救命九州研修所教授

有川 正俊

東京大学空間情報科学
研究センター助教授

現状実施されている統計調査はリアルタイムの報告ではない上に、市町村等空間的に大きな集計単位の統計数値となっているため、社会の複雑化、高速化に対応して迅速に社会の状態を捉えた現状把握や Community 単位や Family 単位での詳細な状況分析、問題の早期発見をも困難としている。本研究では、この状況を打開するため、迅速かつ効率的に現状把握等を実施するために、因子間の相互関係の解明、知的情報処理技術による時空間データを含む社会情報の収集・入力技術の確立、それに合った空間解析・統計モデルの開発をおこなう。

A. 研究目的

現在の地域保健行政において、現状把握、問題抽出、原因分析等は、社会に存在する情報の一部の情報である保健・医療・福祉に関する統計調査等を画一的側面から分析しているため、正確な現状把握及び本質的原因の解明を行うことは困難である。また、

B. 研究方法

健康危機などに際して、リアルタイムで精密な行政対応を支援する総合的な地域診断システムの開発を目的として、以下の通

り研究を実施した。

- 1)さまざまなデジタルドキュメント（テキスト，HTML，EXCEL など）をドラッグ&ドロップするだけで自動的に住所や地名を抽出し，緯度経度を導出して，地図上で空間分布を閲覧できるソフトウェアシステム「空間ドキュメント管理システム（SDMS: Spatial Document Management System）」の基本設計および開発の研究を行った。昨年度は試作の段階であったが，今年度は一般ユーザに利用可能にするときに生じる問題点の整理と解決および SDMS の枠組みの再検討を行った。
- 2)空間的データからある症候の集積性を早期発見しその地域の同定を行う際に，疾病集積性の検定という方法が用いられるが，その手法の精度を測るための新しい評価指標を提案した。この指標によって，従来よく用いられる SaTScan 法よりも flexible scan 法のほうが精確に地域を同定される場合があることが観察された。さらにその flexible scan 法を時間変化も考慮した space-time 解析に適用できるよう拡張を試みた。
- 3)現在の社会的状況から，AED（自動体外式除細動器）の配置地点に対する検討が急務であるとの認識にたち，実際の心停止発生地点から推定された確率密度分布に，生存退院率にもとづく救命確率を加味した AED の最適配置問題を提案した。さらに，この方法を弘前市で過去に生じた心停止発生地点の場合に適用し，装置の最適配置地点を示すとともに，各地点に設置されたときの救命確率，および需要に対する設置効果の分布を表す地域地図を作成することで，本研究で開発された方法の有効性を実証した。特に，心停止発生地点の取得に際しては，本研究班で開発された SDMS の中核となるアドレスマッチングシステムを使用し，SDMS の医療分野への応用可能性について検証を行った。
- 4)インフルエンザは毎年，その抗原を変化

させて流行する。近年，新型インフルエンザの流行が懸念されており，保健所と中心として地域健康危機管理を実施することが求められている。現状のサーベイランスは，医療機関を受診した患者数を把握して週ごとに集積されたデータを感染症法に基づいて県単位で広報されているが，時間的空間的乖離があり，感染初期を把握するには有効とは言えない。そこでインフルエンザの感受性が高い小学校の欠席率に着目して解析を行った。

（倫理面への配慮）

本研究においては，原則として公開されたデータを用いて技術開発を行なう。また，プライバシーにかかわる個人情報扱う場合は，個人情報の保持・公開には十分留意すると共に，疫学倫理指針に基づき研究を行なうこととする。

C. 研究成果

本研究によって得られた研究成果は主に以下の通りである。（項目の数字は「B. 研究方法」のものにしたがう）

- 1) 空間ドキュメント管理システム（SDMS）の実用化を中心に研究を進めた。その結果，試験的に配布可能な品質のレベルまで持ってくることができた。また，Windows 対応にしたために，一般ユーザでも気軽に使えるようになった。
- 2) 集積性の検定の新しい指標を用いることで，いくつか提案されている手法の特徴を表現することができ，実際の解析の場合に，どの手法を用いればよいかを選択する際の参考になるものと考えられる。
- 3) 心臓疾患による救急車搬送発生地点の点分布パターンに着目した，AED（自動体外式除細動器）の最適設置地点を数理的に得る手法の開発を行った。特に，実際の地域社会に反映される結果とともに，健康危機情報における空間的な要素の重要性が再認識された。
- 4) 日常の学校欠席率は $1.5 \pm 0.5\%$ であった

こと、学校欠席率と定点観測値は高い相関が見られたこと、小学校の欠席率を把握することは地域の感染拡大状況を把握するのに有効であること、学校欠席率を定量的に解析して比較することで地域の流行状況を初期から把握できることが判明した。

D. 考察

- 1) 今年度は、空間ドキュメント管理システム (SDMS) の実用化を中心に研究を進めた。その結果、試験的に配布可能な品質のレベルまで持ってくる事ができた。また、Windows 対応にしたために、一般ユーザでも気軽に使えるようになった。来年度は、引き続き、SDMS の品質を向上させ、ユーザビリティの観点から、ユーザに負担をかけない理想的なツールに可能な限り近づける努力を続ける予定である。われわれは、SDMS の開発を通して、SDMS は単なる1つのソフトウェア以上に、GIS のあり方を根本的に変える枠組みの提案につながるのではないかという洞察を持つようになってきた。その意味で、場所情報管理の大衆化を招き、医療情報をはじめ関係分野において、来るべきユビキタス情報化社会にふさわしい空間情報コミュニケーションツールの画期的な枠組みを示す事ができたと考えている。
- 2) 集積性の検定の新しい指標を用いることで、いくつか提案されている手法の特徴を表現することができ、実際の解析の場合に、どの手法を用いればよいかを選択する際の参考になるものと考えられる。
- 3) 本年度に開発された AED の配置地点に関する方法については、今後、「コンビニエンスストア」や「交番」などの身近で利便性の高い施設への併設や、地域内で既に装置が設置されているなどの条件を加味した応用が図られ、地域社会の実情により即した検討が期待される。また、本年度の成果については、今後、報道機関への公表や、東京などの他地域への適用を予定している。
- 4) 2004 年～2005 年のインフルエンザは小

倉北区から流行が始まったものと考えられた。その拡大状況は市内各区の学校欠席率を定量的に比較することでも把握できる。学校欠席率を毎日調査することで、地域の感染状況を把握することが即時に把握することが可能である。来年度は、既に得られたこれらのデータをさらに解析を進める予定である。

E. 結論

- 1) 本研究で開発されたシステムは、従来の地理情報システムでは取り扱えなかった、人間活動で日常的に利用する位置を表現するテキスト情報を、知的言語処理と住所マッチング技術を用いて自動的に空間データ化するシステムを構築しようというユニークなものである。さらに品質を上げるための努力を行い、ソフトウェアツールの品質を向上させることにより、医療従事者が日常的に気軽に利用できる空間情報コミュニケーションツールへと発展させるための改良の努力を行っていきたい。
- 2) flexible scan 法のサーベイランスへの適用について、その評価も踏まえ詳しく検討し、実際の日本でのサーベイランスに向けた研究を行っていきたい。
- 3) 他の空間オブジェクトにも対応可能な、健康危機情報に関わる有意義な分析方法の開発を行っていきたい。
- 4) 学校欠席率を毎日調査することで、地域の感染状況を把握することが即時に把握することが可能であることが判明したことから、今後、既に得られたこれらのデータをさらに解析を進める予定である。

以上を統合して、地域診断システムとして、健康で安全な社会の構築に貢献できるシステム、技術開発を目指す。

F. 研究発表

1. 論文発表
[1] 浅見泰司(2005)「環境分析のための GIS の現状と展望」『環境管理』41(8), 1-6.

[2]浅見泰司, 有川正俊, 白石 陽, 片岡裕介, 相良 毅, 「空間ドキュメント管理システムの設計と開発に関する研究」, 第 8 回東京大学空間情報科学研究センターシンポジウム(CSIS DAYS 2005), 全国共同利用研究発表大会, セッション E:空間 IT と要素技術開発, 2005 年 9 月.

[3]浅見泰司(2005)「情報化の進展で都市は消滅するか?」都市住宅学会(編)『データで読みとく都市居住の未来』学芸出版社, pp.50-56.

[4]Tango T and Takahashi K. A flexibly shaped spatial scan statistic for detecting clusters. *International Journal of Health Geographics* 2005, 4:11.

[5]Takahashi K and Tango T. An extended power of cluster detection tests. *Statistics in Medicine* 2006, 25:841 - 852.

2. 学会発表

[1]相尚寿, 浅見泰司, 貞広幸雄 (2005)「寄り道支援路線図の自動作成システム」『地理情報システム学会講演論文集』14, 273-276.

[2]Masatoshi Arikawa, Toru Hayashi, Kaoru Sezaki: Weblog based mapping with track backs on spatio-temporal relations, the 22nd International Cartographic Conference, the International Cartographic Association, 2005, CD-ROM Proceedings, A Coruna, Spain, July 2005.

[3]浅見 泰司, 有川 正俊, 白石陽, 片岡 裕介, 相良 毅: 空間ドキュメント管理システムの設計と開発に関する研究, 第 8 回東京大学空間情報科学研究センターシンポジウム(CSISDAYS 2005), 全国共同利用研究発表大会, セッション E:空間 IT と要素技術開発, 2005 年 9 月.

[4]Yoh Shiraishi, Masatoshi Arikawa: A Framework for Interactive Searching and Mapping of Personal Spatial Information, the 22nd International Cartographic Conference, the International Cartographic Association, 2005, CD-ROM Proceedings, A Coruna,

Spain, July 2005.

[5]白石 陽, 有川 正俊, 浅見 泰司: 実空間と Web 情報空間から収集したパーソナル空間情報の記述, 発信, 閲覧のためのフレームワークの提案, 地理情報システム学会, 第 9 回空間 IT ワークショップ, 2005 年 7 月.

[6]白石 陽: 空間ドキュメント管理システムによる POI の生成と表示, 都市の OR ウィンターセミナー in つくば, 2006 年 1 月 10 日

(<http://infoshako.sk.tsukuba.ac.jp/~toshiw3/Labo/koshizuka/uor/index.html>)

[7] Tango T and Takahashi K. FleXScan: A flexible scan statistic for detecting clusters. *Fifth International Interdisciplinary Conference on Geomedical Systems*, University of Cambridge, 16-17 September 2005.

[8]高橋邦彦, 横山徹爾, 丹後俊郎, Martin Kulldorff. サーベイランスのための Flexible scan 法. 第 16 回日本疫学会学術総会, 名古屋市, 2006 年 1 月 23 日・24 日.

G. 知的財産権の出願・登録状況

- | | |
|-----------|------|
| 1. 特許取得 | 該当なし |
| 2. 実用新案登録 | 該当なし |
| 3. その他 | 該当なし |

厚生労働科学研究費補助金（健康科学総合研究事業）
分担研究報告書

空間ドキュメント管理システムの設計と開発に関する研究

分担研究者 有川 正俊 東京大学空間情報科学研究センター助教授
研究協力者 白石 陽 東京大学空間情報科学研究センター助手

研究要旨

さまざまなデジタルドキュメント（テキスト，HTML，EXCEL など）をドラッグ&ドロップするだけで自動的に住所や地名を抽出し，緯度経度を導出して，地図上で空間分布を閲覧できるソフトウェアシステム「空間ドキュメント管理システム（SDMS: Spatial Document Management System）」の基本設計および開発の研究を行った。昨年度は試作の段階であったが，今年度は一般ユーザに利用可能にするときに生じる問題点の整理と解決およびSDMSの枠組みの再検討を行った。

A. 研究目的

医療情報管理において，疾病・事故・事件が発生した場所の情報は極めて重要であり，またその情報無しでは現場での処理や判断を行うことは困難である。しかしながら，場所の情報のIT化はなかなか進んでいない。場所の情報を扱う従来のシステムとして地理情報システム（GIS）があるが，それを使える技能を得るためにはかなりの時間と努力を必要とする。本来の業務に加えて，GISの技能を各人に習得してもらうことは現実的には難しい。理想的には，地図技術者のようなGISを活用する専門家を各部署に配置して，場所情報のIT化を図る枠組みが良いが，現実的にはGIS利用だけのために新たに人を雇用するという枠組みは理解されにくい。また，GISで扱える空間データは，基本的に緯度経度などのような座標値(x,y)にして，データベースとして形式的に整ったものである必要があり，これを作成・整備するにはコストがかかるために，空間データ自身を作ることも蓄積することも難しい。

一方，現場での処理や判断では，場所の情報は重要であり，人間が分かるアナログ形式で，つまり住所や地名や経路記述などの自然言語で必ず記録されていると言って良い。しかしながら，それらはアナログ場

所情報であり，貴重な情報ではあるが再利用が難しく，死蔵してしまう場合がほとんどである。そこで，IT化への期待が大きくなっている。それらの自然言語の場所記述を座標値に変換して，GISで管理するという枠組みが理想的だが，前段でも説明したように人的リソース割り当ての観点から現実的ではない。一方，インターネットの世界では，人間が人間に読んでもらうために作成した一般ドキュメントが大量に作成され流通しており，これらはデータベースとして形式的に整っていないが，サーチエンジンで検索でき，多くの一般市民が日常的に利用している。

ここで，デジタル情報を以下の2つに分類して考えることにする。

- (a) デジタルドキュメント（デジタルコンテンツ）：人間が人間に情報伝達（コミュニケーション）するために作成したデジタル情報
- (b) デジタルデータ：機械による処理に向く形式化されたデジタル情報

ウェブが普及する以前は，情報検索の対象は(b)デジタルデータであった。人々は，コストをかけて機械で処理しやすい，品質の高いデジタルデータを作成していたが，

あくまでも、それらのシステムもデータも、作成と管理のためには専門家が必要であった。それらを使うためには、人々は努力して専門家になる必要があった。現在の情報検索の主流であるサーチエンジンは、人間が日常的に使っている言葉で作成した(a)デジタルドキュメントそのものが検索の対象となり、検索方法も単に連想するキーワードを入れるだけという簡単な枠組みになっている。つまり、以前は、人が機械の環境に歩み寄っていたが、現在は、機械が人の環境に歩み寄る枠組みにするのが主流である。この枠組みは人間中心設計 (human

centered design) と呼ばれる。人間中心、個人中心になった他の例としては、大型計算機がパーソナルコンピュータへ、ラインコマンドがデスクトップへ、マスメディアがウェブへなど、社会構造自体が階層的なものからフラットなものへと民主化が進む変革の流れと同じように、IT 環境も人間化・個人化への変革が進んでいる。

場所情報を扱う現在の主流の情報システムである GIS を考えてみると、人間中心のシステムとは言いがたい。本研究では、自然言語の場所記述を含むデジタルドキュメントを対象として、ドラッグ&ドロップと

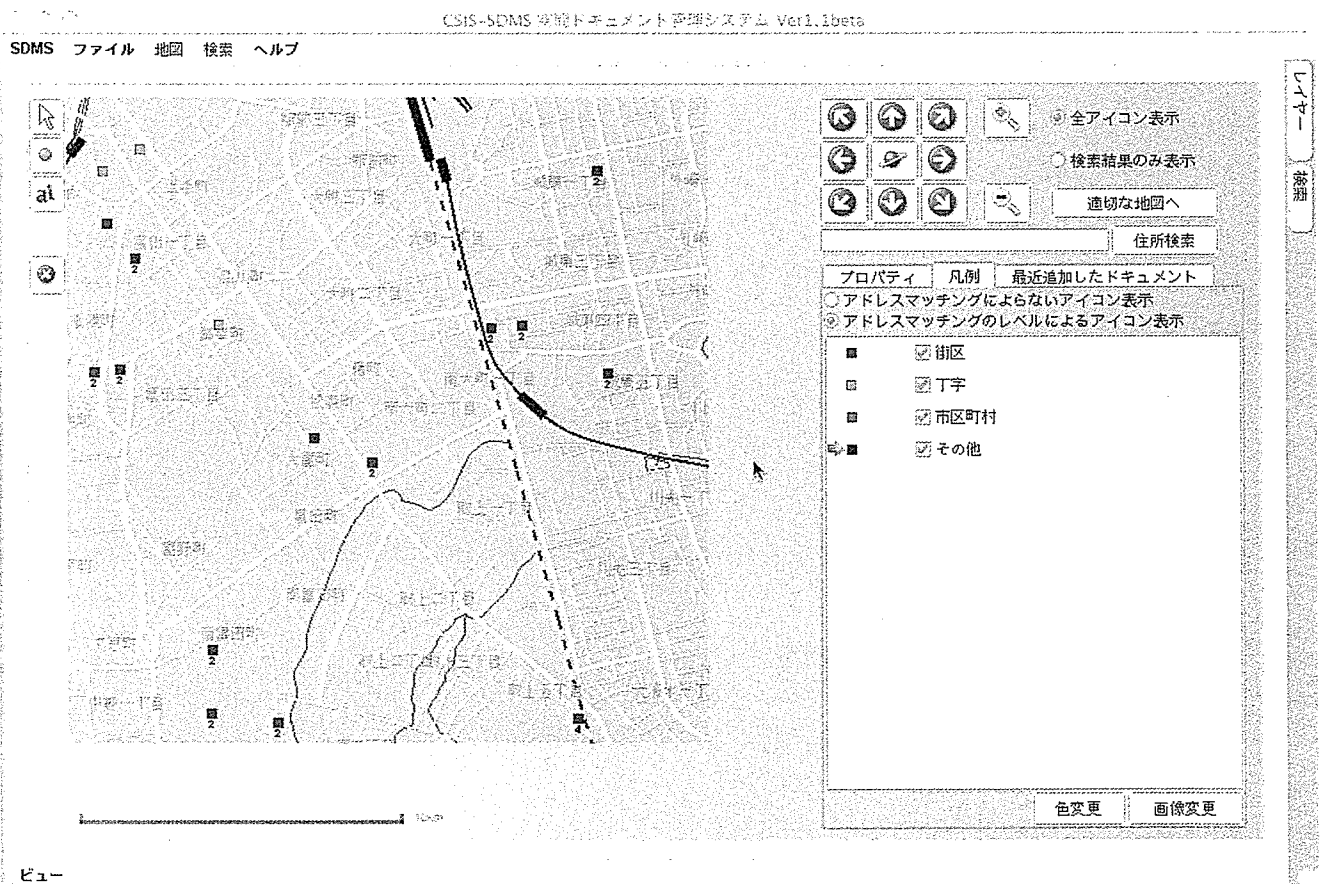


図 1: 空間ドキュメント管理システムの画面の例

あるテキストファイルの中に住所のテーブルが含まれていた場合、そのテキストファイルを空間ドキュメント管理システムにドラッグ&ドロップするだけで、含まれる住所記述を地図上の点として抽出し、表示できる。図では、住所がマッチングできたレベルにより、点の色を違えて表示している。

いう簡単で自然な操作だけで、デジタル地図を作成・操作できる人間中心の場所情報管理システムを検討し、実際に試作システムを構築して、医療現場での場所情報利活用の促進を実現する現実的な枠組みの体系化に関して研究を行った。本研究は、GISを人間中心の観点から再設計する試みと見なせ、従来の空間データ中心システムから、空間コミュニケーション中心システムへと変革させるものと位置付けている。

B. 研究方法

本研究は、空間情報工学、ユーザインタフェース、自然言語処理、データベースの観点から、人間中心場所情報管理システムとして、空間ドキュメント管理システム (SDMS: Spatial Document Management System) の開発研究を行った。本研究の最終目的は、SDMSを多くの医療従事者が日常的に簡単に利用できる環境を実現することである。これにより、医療従事者の場所に関する語彙が増加し、空間的洞察力と空間的判断能力を強化できる。そのために、SDMSの試作システムの完成度を向上させ、機能変更・追加するための、SDMSの基本設計・実装・試験を行った。また、一般ユーザが利用するための、操作マニュアルやインストールマニュアルなどのドキュメンテーションの作業も行った。

以下、SDMSの特徴・データ処理の主な流れ・改善点に関して説明を行う。

(B-1) SDMSの特徴

- ・一般ドキュメント (プレインテキスト, HTML, EXCEL, PDF など) を POI (Point of Interest) に変換する。ドキュメントから住所抽出を行い、アドレスマッチングにより、緯度経度に変換して POI を生成する。ドラッグ&ドロップ操作で上記の処理を実行可能。(ただし、MS-Word, Unicode, PDF 6.0 以上, など対応しない形式もある。その場合

は、プレインテキストに一度変換してから、SDMS を利用することを仮定している。)

- ・住所表記のゆれや省略をある程度許容する。
- ・地図縮尺・移動自動機能。
- ・ドキュメントを単位にしたレイヤ管理。POI の編集機能。表示や凡例の変更機能。
- ・テキスト検索と空間検索の組み合わせ利用。
- ・全国のアドレスマッチング (街区レベル) と全国の 25,000 分の 1 の背景地図を提供。アドレスマッチング用テーブルデータと背景地図のデータは、どちらも国土交通省から無料で利用可能なものを利用。
- ・POI を Shape, CSV, G-XML 2.0 の形式で出力可能。これにより、ほぼ全ての専門家向けの GIS との連携が可能。
- ・ユーザ空間辞書。ユーザが独自に地名と座標を登録できる機能。
- ・無料で配布可能な構成。
- ・ポータビリティを重視して、Java アプリケーションとして実現。

(B-2) SDMS のデータ処理の主な流れ

- (a) 一般ドキュメントあるいはそれらが入ったフォルダをユーザが SDMS にドラッグ&ドロップする。
- (b) 一般ドキュメントをプレインテキストに変換する。
- (c) プレインテキストを自然言語解析 (形態素解析) して、名詞句を抽出する。
- (d) 地名であるかどうかを、アドレスマッチング・サーバに問い合わせる。地名である場合は、正規化された地名と座標値とマッチングレベルが返される。マッチングレベルとは、県・市・区・丁目・番地などのどのレベルまでマッチしたかを示す値。

(このアドレスマッチング・サーバ spat は、東京大学生産技術研究所の相良毅助手が開発したものであり、サーバ自体は、東京大学空間情報科学研究センターで運用されている。)

- (e) プレインテキスト中で、地名と判定された部分に、`<spa>`タグが挿入され、地名として識別された部分の情報を残す。この場合、`<spa>`タグの属性として、`address`, `level`, `lat(itude)`, `lon(gitude)`, `url` なども同時に設定される。`url` は、もとのドキュメントの場所を指す。
(例) 「発生場所は`<spa id="5" address="東京都 北区 十番台 一丁目" level="5" lat="35.756924" lon="139.723083" label="北区十番台" url="http://juban/" alias="">`北区十番台1丁目`</spa>`です。」
- (f) `<spa>`タグの属性値の部分が、POI (Point of Interest) として、空間データベースに登録される。
- (g) 日本地図上に抽出された POI 集合が表示される。
- (h) ドラッグ&ドロップの単位である、ドキュメントかフォルダごとに、レイヤが設定され、POI 集合がレイヤごとに管理される。
- (i) POI に対して、キーワード検索も可能であり、キーワード検索結果は、新しいレイヤの POI 集合として表示される。たとえば、「インフルエンザ」や「食中毒」のキーワードが含まれるドキュメントの POI だけを生成することも可能である。
- (j) POI は、他の GIS などで利用するために、Shape, CSV, G-XML などの形式で export できる。この export 機能により、高度な空間解析や視覚化を GIS 上で行える。

(B-3) 今年度の SDMS の主な改良点

ソフトウェアを実際に利用できるようにするためには、細かい点をたくさん改善する必要がある。それらの細かい内容の中心はバグフィックスであり、ソフトウェアが当初の仕様とは異なる振る舞いをする点を改善した。バグフィックス以外で、本年度の大きな改良点は以下のとおりである。

(改良点 1 : プラットフォームを Macintosh から Windows へ拡大させた)

当初、緊急医療で利用する場合、セキュリティの面から、スタンドアロンで利用することが重要であると判断し、アドレスマッチング・サーバ spat が Unix 上でしか動作しないという理由から、SDMS のターゲットマシンとして Macintosh (Mac OS X) を選んでいた。しかし、SDMS を Windows 上で利用したいという要望が多く、spat は東大のサーバを利用して、ネットワーク経由で、SDMS を Windows で動作できるようにした。SDMS 自体は、Java で開発しているために、spat の部分を除けば、Windows 上で動作する環境ではあるが、ネットワーク対応にするためにいくつかの修正作業が生じた。結果として、現在の SDMS は、インターネットへ接続できる環境では Windows と Macintosh の両方で利用可能であり、スタンドアロンでの利用の場合は Macintosh のみとなっている。

(改良点 2 : アドレスマッチング・サーバ spat の高速化・XML 対応)

spat のプログラムは高速化されるとともに、AJAX に対応するために問い合わせ結果が XML 形式で返るようになり、これらの変更に対応する改良を施した。

(改良点 3 : 地図表示の適正化)

昨年度のバージョンでは、地図がメッシュ単位でしか表示できなかったが、今年度のバージョンではメッシュをまたがって表示できるように改良した。

(改良点4：CSVでの export 機能)

今年度、新たに、CSV形式で、POIを保存できる機能を追加した。

(改良点5：ユーザ空間辞書機能)

昨年度までは、spatに登録されていた住所と地名のみのアドレスマッチングが可能であったが、ユーザごとに、独自に、地名と座標値をCSV形式で登録できる機能を実現し、アドレスマッチングの拡張性を実現した。

C. 研究成果

SDMSの試作システムを改善し、配布可能なレベルまで品質の向上を図る努力を行った。この班の共同研究者や研究関係者にはすでにSDMS試作版を渡して、試験的利用をしてもらっている。実装に関しては、確実な部分を行ったが、同時に、次の研究展開を見据えて、高度なアドレスジオコーディングならびにBlogの地図利用への拡張に関する検討も行った。

D. 考察

医療情報管理において、場所に関する情報は重要であるにもかかわらず、現在のGISが人間中心システムでないがために、一般ユーザにとって使いにくいものになっている。この問題点を補うために、空間ドキュメント管理システム(SDMS)を設計し、実現を行い、現在、配布して試験的に利用してもらっている。現在配布しているソフトウェアの品質は、まだまだ気の利いたものになってはいないが、基本機能は確実に実現されており、便利なツールとして使えるようになってきている。今後、さらに品

質を上げるための努力を行い、ソフトウェアの品質を向上させることにより、医療従事者が日常的に気軽に利用できる空間情報コミュニケーションツールへと発展させるための改良の努力を行っていきたい。

E. 結論

今年度は、空間ドキュメント管理システム(SDMS)の実用化を中心に研究を進めた。その結果、試験的に配布可能な品質のレベルまで持ってくる事ができた。また、Windows対応にしたために、一般ユーザでも気軽に使えるようになった。来年度は、引き続きSDMSの品質を向上させ、ユーザビリティの観点から、ユーザに負担をかけない理想的なツールに可能な限り近づける努力を続ける予定である。われわれは、SDMSの開発を通して、SDMSは単なる1つのソフトウェア以上に、GISのあり方を根本的に変える枠組みの提案につながるのではないかという洞察を持つようになってきた。その意味で、場所情報管理の大衆化を招き、医療情報をはじめ関係分野において、来るべきユビキタス情報化社会にふさわしい空間情報コミュニケーションツールの画期的な枠組みを示すことができたと考えている。

本研究を遂行するにあたり、アドレスマッチングのエンジン部分を本システム向けに改良を加えて利用させていただきました東京大学生産技術研究所の相良毅助手に感謝いたします。また、相良毅助手には、空間ドキュメント管理システムの設計の際に多くの有意義なアドバイスとコメントをいただきました。

F. 研究発表

1. 論文発表
なし。
2. 学会発表

[1] Masatoshi Arikawa, Toru Hayashi, Kaoru Sezaki: Weblog based mapping with track backs on spatio-temporal relations, the 22nd

International Cartographic Conference, the International Cartographic Association, 2005, CD-ROM Proceedings, A Coruna, Spain, July 2005.

[2] Yoh Shiraishi, Masatoshi Arikawa: A Framework for Interactive Searching and Mapping of Personal Spatial Information, the 22nd International Cartographic Conference, the International Cartographic Association, 2005, CD-ROM Proceedings, A Coruna, Spain, July 2005.

[3] 白石 陽, 有川 正俊, 浅見 泰司: 実空間と Web 情報空間から収集したパーソナル空間情報の記述, 発信, 閲覧のためのフレームワークの提案, 地理情報システム学会, 第 9 回空間 IT ワークショップ, 2005 年 7 月.

[4] 浅見 泰司, 有川 正俊, 白石陽, 片岡裕介, 相良 毅: 空間ドキュメント管理システムの設計と開発に関する研究, 第 8 回東京大学空間情報科学研究センターシンポジウム(CSISDAYS 2005), 全国共同利用研究発表大会, セッション E:空間 IT と要素技術開発, 2005 年 9 月.

[5] 白石 陽: 空間ドキュメント管理システムによる POI の生成と表示, 都市の OR ウィンターセミナー in つくば, 2006 年 1 月 10 日
(<http://infoshako.sk.tsukuba.ac.jp/~toshiw3/Labo/koshizuka/uor/index.html>)

G. 知的財産権の出願・登録状況
なし.

厚生労働科学研究費補助金（健康科学総合研究事業）
研究報告書

集積地域同定のための新たな検定手法の開発に関する研究

分担研究者 丹後 俊郎 国立保健医療科学院技術評価部
研究協力者 高橋 邦彦 国立保健医療科学院技術評価部

研究要旨

空間的データからある症候の集積性を早期発見し、その地域の同定を行う際に利用される疾病集積性の検定といわれる方法がいくつか提案されているが、それらの精度の比較について従来の評価指標では不十分である。本研究ではこの地域同定のための検定法の評価法として新たな指標を提案し、本研究者が提案する flexible scan 法と、よく用いられる circular scan 法の比較を行い、集積地域の形状によって flexible scan 法が circular scan 法より精確に地域を同定できることを確かめた。また日々のサーベイランスに向けて、この flexible scan 法を時間変化も考慮した空間-時間的解析への拡張を試みた。

A. 研究目的

地域診断システムにおけるバイオ・サーベイランスの方法として、ある疾病や症状などの集積性の検出およびその集積地域を同定するために、集積性の検定法が用いられる。実際に米国でのバイオテロリズムの影響もあり、ニューヨークやワシントンDCをはじめ欧米のいくつかの地域で日々監視を目的としたサーベイランス解析が行われている。そのシステムで実際に利用されている方法がハーバード大学のKulldorff助教授によるscan統計量による方法とそれを組み込んだソフトウェアSaTScanであるが、この方法ではうまく同定できない地域があることが指摘されており、最近その点を改良したいくつかの方法が提案されている。本研究分担者らもflexible scan法とソフトウェアFleXScanを提案している。このようにいくつかの手法が提案された場合、その手法のよしあしを比較検討することは重要であり、その比較は検出力によって行われることが一般的であった。しかしながら、ここで扱う集積性の同定の検定法は一般的な検定法とは違い、地域を同定することも目的の一つであり、従来の一般的な検出力

による比較では不十分である。そこで本研究ではこのような同定を含む検定法の比較のために、新たな評価指標を提案する。

また、サーベイランス解析を目的とした場合、平面上の地域同定だけではなく、時間変化も考慮したspace-time解析が必要となる。平面上の同定手法をはじめ、集積性の検定手法はいくつか提案されているが、実際にこのspace-time解析を実用化の段階まで拡張できているものはSaTScan法をはじめほんのわずかしが公開されていない。そこで本研究では平面上におけるSaTScan法を改良する形で研究分担者らが提案しているflexible scan法をspace-time解析に適用できる形に改良を試みる。

B. 研究方法

平面上の地域同定を含む検定手法の比較として、KulldorffによるSaTScanで用いられるcircular scan法と、Tango and TakahashiによるFleXScanで用いられるflexible scan法の2つを検討する。まず、東京・神奈川の113地区の地図を用いて、真のホットスポットモデルを想定したシミュレーションデータを用いて、実際に2つ

の手法で検定・同定を行う。このシミュレーションを 1000 回行い、その比較を行った。この比較に際し、一般的にもちいられている集積地域があるかないかの決定のみをはかった usual power に対し、同定された地域の“地域数”と“設定した真の集積地域をいくつ同定しているか”を考慮した方法を提案する。これにより同定に対して、false positive、false negative を調べ、それを考慮した検定手法の評価を行い、従来の usual power による評価とも比較を行う。

次に、flexible scan 法の space-time への拡張については、定期的なサーベイランスを目的とした prospective space-time 解析への拡張を、ハーバード大学の Kulldorff 助教授らとともに試み、実際にマサチューセッツ州のデータに適用し、SaTScan 法とその同定の違いを観察する。

C. 研究成果

評価指標の提案として、以下の指標を提案した。

● Bivariate power distribution (表 1)

本研究で提案する方法のひとつで、同定されたクラスターに含まれる地域の数と、実際に想定した真の集積地域をいくつ同定しているかを 2 次元分布の表として表す。

● Extended power

本研究で提案する方法のひとつで、bivariate power distribution にもとづいて真のクラスターの同定の様子を考慮しながら、一般的な検定で用いられる検出力と同様に $[0,1]$ 上の実数として検定の精度を表す指標となっている。その際、同定の様子を false positive と false negative それぞれに重みをつけ、同定の見逃しにペナルティーを付け、幾何平均の形で表す。また特別の場合として usual power を含んだ形となっており、一般の検出力を拡張した指標となっている。

● Extended power の profile (図 1)

上記の extended power は false positive、

false negative のペナルティーを決める際の重みのパラメータに依存して値が変化する関数となっている。そこで、その重みを様々に動かすことで、その関数の plot をし、それによって検定手法の評価を行った。

上記の評価指標を用いて、circular scan 法と flexible scan 法の比較を行った。これらの指標を用いることによって、従来の usual power でははかれなかった、地域同定を含む検定指標の性能をより詳細に表すことができるようになった。

次に、flexible scan 法を space-time 解析に拡張し、実際にマサチューセッツ州の Respiratory のデータで解析を行い、米国で実際に用いられている SaTScan による結果と比較を行った。集積が認められた時期は 2 つの手法とも同じであったが、その集積地域は、flexible scan 法がより細かく同定を行っている様子が観察された (図 2)

D. 考察

同定を含む検定の評価としては、従来の usual power ではその違いがよくわからず、同定の様子を踏まえた場合の精度の評価と逆の結論を導き出す場合もあった。本研究で提案した bivariate power distribution によりその同定の様子を詳細に観察することができるようになった。しかしこの指標は同定の様子は詳細にわかるが、表の形で表されるため 1 つの値としての評価はできない。そこで従来の検出力と同様に $[0,1]$ 上の 1 つの実数値として表される評価指標として、特別の場合として usual power を含んだ形となる extended power を提案した。この指標により、false positive と false negative に重みを付けて、目的に応じた検定手法の選定を考えることが可能になった。しかし extended power はその重みのパラメータの設定は各研究者が予め決めなければならない、そのため自由度は大きい、その値を決めかねる研究者には扱いづらい面もある。そこで本研究ではそのパラメータ

の決め方を提案するとともに、パラメータが変動した際の power の plot を用いることで、視覚的にも評価が出来る方法として profile を提案した。この profile で 2 つの検定方法を比較することで、一様に片方の方法がよい場合や、パラメータ (false positive と false negative の重みの比) がどのように設定した場合はどちらがよいかなどを視覚的に判断できるようになった。特にこの評価指標により flexible scan 法が非円形の集積地域がある場合では SaTScan による方法よりも強力に精度よくその集積地域を同定していることが観察できた。

E. 結論

SaTScan 法を中心とした疾病集積性の検定は、集積性の有無の判定と同時に集積地域の同定も目的とした検定法であるが、同定の精度を踏まえた手法の評価は行われていなかった。本研究ではその同定の様子を詳細に表す bivariate power distribution、従来の検出力と同様に $[0,1]$ 上の実数値で評価する extended power とその視覚的表示 profile によって、これらの手法を評価することを提案している。これらの評価指標を用いることによって、従来の検出力による評価ではわからなかった手法の良い点、欠点などが表現でき、よりの確に評価が行われ、解析者の手法の選択に役立つものと考えられる。米国でのサーベイランスに実際に利用されている SaTScan 法と本研究者らの提案する FleXScan 法を比較した場合、usual power はほぼ同じ、もしくは SaTScan 法のほうが高い場合においても、実際にこれらの評価指標を用いることで FleXScan 法がより精度よく地域同定をしていることも観察された。

また、実際のサーベイランスに向け、FleXScan 法の space-time 解析への拡張については、試験的な運用として実現可能の段階にあり、平面状と同様に SaTScan 法よ

りも精度よく地域を同定できることが期待される。今後、実際場面でよりうまく利用できる形も踏まえ、検討を続けていくことが重要である。

F. 研究発表

1. 論文発表

- [1] Tango T and Takahashi K. A flexibly shaped spatial scan statistic for detecting clusters. *International Journal of Health Geographics* 2005, 4:11.
- [2] Takahashi K and Tango T. An extended power of cluster detection tests. *Statistics in Medicine* 2006, 25:841 - 852.

2. 学会発表

- [1] Tango T and Takahashi K. FleXScan: A flexible scan statistic for detecting clusters. *Fifth International Interdisciplinary Conference on Geomedical Systems*, University of Cambridge, 16-17 September 2005.
- [2] 高橋邦彦, 横山徹爾, 丹後俊郎, Martin Kullendorff. サーベイランスのための Flexible scan 法. 第 16 回日本疫学会学術総会, 名古屋市, 2006 年 1 月 23 日-24 日.

G. 知的財産権の出願・登録状況

なし

表 1 : bivariate power distribution

Kulldorff							Tango and Takahashi						
Length l	Include s hot-spot regions					Total	Length l	Include s hot-spot regions					Total
	0	1	2	3	4			0	1	2	3	4	
1	1	0				1	1	0	0				0
2	0	0	351			351	2	0	0	0			0
3	2	0	1	0		6	3	0	0	0	0		0
4	0	0	3	0	0	3	4	0	0	0	0	138	138
5	2	0	2	0	0	4	5	0	0	0	3	147	150
6	1	0	0	0	0	1	6	1	0	0	2	200	203
7	0	0	0	81	0	81	7	0	1	0	1	147	152
8	0	0	10	18	38	66	8	0	0	2	9	107	118
9	0	0	2	0	26	28	9	0	0	0	10	71	81
10	0	0	0	29	3	32	10	1	0	2	5	28	36
11	0	0	1	13	1	15	11	0	0	0	0	10	10
12	0	0	2	4	60	66	12	0	0	0	0	2	2
13	0	0	0	5	62	67	13	0	0	0	0	0	0
14	0	0	0	10	27	37	14	0	0	0	0	0	0
15	0	0	0	6	37	43	15	0	0	0	0	0	0
Total	6	0	375	166	251	801	Total	2	1	1	33	850	890

usual power= $\alpha=0.801$ usual power= $\alpha=0.890$

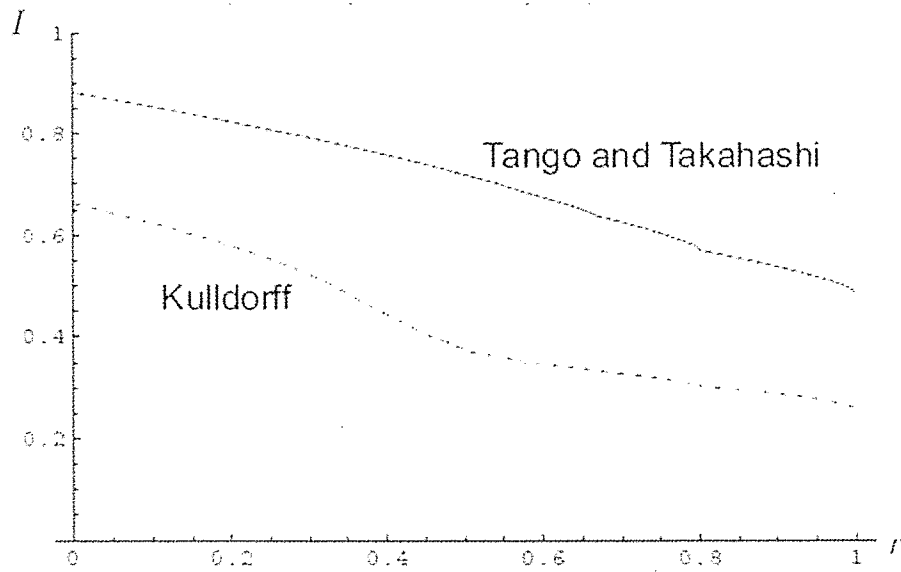
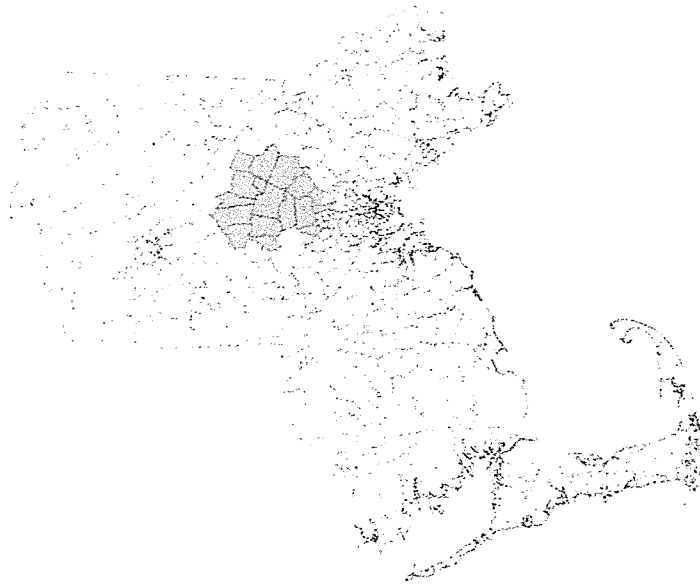
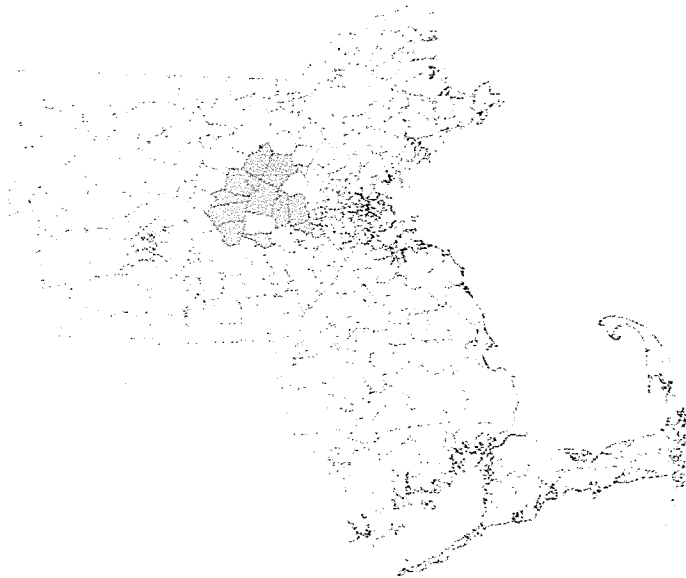


図 1 : extended power の profile



SaTScan 法による同定地域



FlexScan 法による同定地域

図 2 : マサチューセッツ州の Respiratory の集積地域—SaTScan(上)と FlexScan(下)

厚生労働科学研究費補助金（健康科学総合研究事業）
分担研究報告書

健康危機情報における空間分析手法の開発に関する研究

主任研究者 浅見 泰司 東京大学空間情報科学研究センター教授
研究協力者 片岡 裕介 東京大学大学院新領域創成科学研究科博士課程

研究要旨

本研究の目的は、地域保健行政を支援するための、健康危機情報に関わる空間分析手法を開発することである。本年度は、心臓疾患による救急車搬送発生地点を用いた、AED（自動体外式除細動器）の設置地点に関する分析をおこなった。研究成果として、地域における装置の需要量および生存退院率にもとづく救命確率を考慮した、AEDの最適配置地点を求める方法を開発した。さらに、この方法を弘前市で過去に生じた心停止発生地点の場合に適用した結果として、装置の最適配置地点を示すとともに、各地点に設置されたときの救命確率、および需要に対する設置効果の分布を表す地域地図を作成することで、本研究で開発された方法の有効性を実証した。

A. 研究目的

AED（自動体外式除細動器）は、突発的な心停止状態への救急措置に有効とされる医療機器である。取り扱いが簡便で、既に一般市民による使用が認められている状況のなか、全国の自治体には設置に向けた早急な対処が求められている。現状のAEDの設置場所としてその多くは駅、空港やイベント会場などの大規模な集客施設、もしくは医療機関が主となる傾向がみられる。人命に直接的に関わる事態であるAEDの設置に際しては、地域の社会的そして経済的な事情も加味されなければならない。そこで殊に強調されるべきは、より多くの装置を用意する努力もさることながら、装置の配置地点に関する慎重な検討であろう。

AED設置に関する最近の研究をみると、Malcom et al. (2004) は、地域の人口密度と心停止が発生した場所との関連についての調査を行った。その結果としては、人口密度が増加するに従って、病院外で心停止が発生する割合が減少する一方で、自宅外で発生する割合については増加する傾向があったことが報じられている。また、Crocco et al. (2004) では、集客施設で必要とされる、

AEDの設置数に関する数理モデルの提案がなされた。AEDの効果が注目されつつある現在の社会的状況を鑑みても、その設置地点に関する検討が今後求められていくであろうし、この要求に応える最も直接的で有効な方法は、理論的に得られた設置地点が示されることであろう。

そこで本研究では、AEDの最適配置地点に関する考察をおこなうにあたり、主に以下の点を目的としている。まず、心停止発生地点（需要点）の空間パターンにもとづいた地域全体の装置の需要量を表す密度分布に対して、装置との位置関係で決定される救命確率をもって各配置地点（供給点）周辺で重み付けした「供給効果」を最大化する配置問題を提案する。さらに、実際の心停止発生地点をデータとして用いた場合の最適配置地点を示すとともに、地域の救命確率および「供給効果」の地理的分布についても表すことで、本研究で提案される方法の有効性を実証するものとする*。

B. 研究方法

1) 装置の需要量の把握と供給効果の定式化
まず地域における需要量を把握する。さら

に、装置の設置に伴って装置周辺に生じると考えられる救命確率を定義し、需要量に救命確率を加味した「供給効果」を定式化する。

1.1) 需要点にもとづく密度分布の推定

実際に装置が必要とされた地点、つまり過去に心停止が発生した地点を AED の「需要点 (各点で均質)」とし、需要点の分布パターンにより地域の需要量が決定されるものと考えた。とはいえ、心停止が発生した地点のみに需要があるとするのは、同一地点で再発する可能性を考えてみても妥当性を欠く。よって、過去に数多く発生した地域を潜在的に需要が高い地域であると見なすこととする。従って各需要点周辺の需要量は、需要点で最大であり、周囲に拡がるにつれて減衰する確率密度分布で表現することができよう。ただし、ごく短期間のうちでは、心停止の発生が偶然であることを免れない。そのため需要点数については、需要を代表するのに十分な数である必要がある。

需要量の地域全体の確率密度分布についてはカーネル法を用いて推定する (Silverman, 1986)。カーネル法とは、点分布を形成する各点の中心に、カーネルと呼ばれる密度分布の山を置き、それら全てのカーネルの合計をもって全体の確率密度関数を求める密度推定法である (図 1 参照)。なお、本研究ではカーネル関数のパラメータとなるバンド幅選択に際して、Sheather and Jones (1991) の方程式解によるプラグイン法を採用する¹⁾。

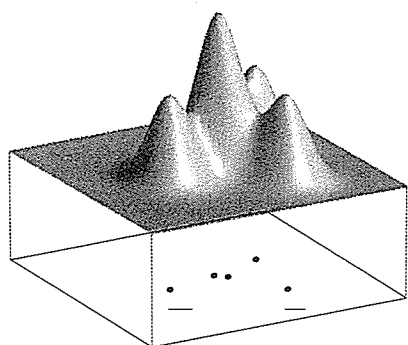


図 1 点分布と密度分布

カーネル法によると、平面上の点 (x,y) における確率密度関数 $f(x,y)$ は、点分布の各点 i の座標 (x_i,y_i) を用いて以下のように表せる (Bowman and Azzelini, 1997)。

$$f(x,y) = \frac{1}{n} \sum_i K\left(\frac{x-x_i}{h_x}\right) K\left(\frac{y-y_i}{h_y}\right) \quad (1)$$

($K(\cdot)$: カーネル関数, n : 点の総数,
 h_x, h_y : 各座標値に関するバンド幅)

また(1)式で、カーネル関数として正規分布を用いて、各需要点の座標が $(x_{1,i}, y_{1,i})$ とすると、点分布にもとづく地域全体の確率密度関数は、以下の(2)式のようにできる。

$$f(x,y) = \frac{1}{2\pi h_x h_y n_1} \sum_{i=1}^{n_1} \exp\left(-\frac{(x-x_{1,i})^2}{2h_x^2} - \frac{(y-y_{1,i})^2}{2h_y^2}\right) \quad (2)$$

(n_1 : 需要点数)

(2)式は、需要点から推定された確率密度関数であると同時に、需要量の地理的分布を表す空間パターンとして捉えられる。

1.2) 救命確率と供給効果の定式化

次に、設置される装置周辺の需要量に対して、装置による効果の程度を表す「供給効果」について考える。何よりも装置が有効に使用されるためには、心停止発生時点から限られた時間内に処置が施される必要がある。その際、各施設立地点から一定距離内にある領域に含まれる需要量を最大化する配置問題として知られる、最大被覆問題 (Maximal Covering Location Problem) があげられる。最大被覆問題は、Church and ReVelle (1974) により提案された後、救急施設などを対象とした離散平面における配置問題として発展するなかで、様々な解法がこれまでに提案されてきた (Brotcorne et al., 2003)。ただ先述のとおり、地域の需要量は点分布であるよりも、各地点の潜在的な需要を表す確率密度分布で表現される方が適切であろう。さらに AED の場合では、心停止に始まり処置に至るまでの時間に、救命確率が大きく依存するとされるため (American Heart Association, 2000)、被覆する領域内においても場所によ

っては効果の程度は著しく異なると考えられる。同様に Alsalloum et al. (2006) は、救急車の配置モデルにおいて、通常用いられる一様に被覆される領域を、定めた制限時間内に被覆される確率に置き換えた方法を提案しているが、AED の場合においても類似した状況にあると考えてよい。

そこで、装置から距離 d の地点における救命確率は、経過時間に依存する救命確率を、距離的な問題に置き換えた以下の関数型で近似できるとする。

$$g(d) = \exp(-ad^2) \quad (3)$$

(a : パラメータ)

本研究では、心停止が発生した現場に居合わせた「救助者」が装置を取りに行くことを前提としている。救助者の置かれた状況を考えると、設置場所がわかっている装置については、装置に向かって取りに行くことが可能である。ところが、当然ながら救助者は装置の全設置点を把握しているわけではない。事前知識も不確実であることに加え、このように一刻を争う事態では行動が直感に大きく左右され、状況として通常の施設選択の場合とは異なるともいえる。つまり、救助者は必ずしも現場から最も近くにある装置を常に把握しているわけではなく、ある地点において装置が心停止にもたらす効果は、心停止発生地点から最も近くにある装置との関係性のみに影響するとは限らない²⁾。すると装置の「認知度」の拡がり、装置からの距離に伴う救命確率の減衰に代替されているとしたときに、近くに設置されている装置が多いほど、設置に伴う効果は上昇するものとも考えられる。以上より、ある地点における AED 設置によって与えられる影響の程度は、各装置の救命確率の和として説明できるものと仮定する。

よって(3)式を用いると、地点 (x,y) での救命確率の和は、各装置の設置地点の座標 $(x_{2,j}, y_{2,j})$ を用いて以下のようにできる。

$$g(x,y) = \sum_{j=1}^{n_2} \exp\left(-a\left((x-x_{2,j})^2 + (y-y_{2,j})^2\right)\right) \quad (4)$$

(n_2 : 装置の設置数)

以上をふまえ、装置の効果が期待される装置周辺において、救命確率を需要量に加味することで得られる、需要と装置の効果に依存する量を本研究では「供給効果」と定義する。なお、理解を容易にするために一次元で考えると、需要量と「供給効果」の関係は図2のようになる。

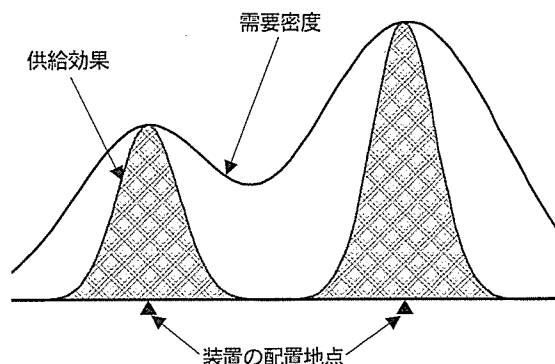


図2 需要密度に対する供給効果

これより、ある地点における需要に対して救命可能な量となる「供給効果」は、その地点の需要量に救命確率の和を乗じたものの地域全体の総和 (図2の網掛け部分) と考えることができ、以下の(5)式のように定式化される。

$$\begin{aligned} L_1(x,y) &= \iint \mathcal{D}(x,y) \cdot g(x,y) dx dy \\ &= \iint \left\{ \frac{1}{2\pi h_x h_y n_1} \sum_{i=1}^{n_1} \exp\left(-\frac{(x-x_{1,i})^2}{2h_x^2} - \frac{(y-y_{1,i})^2}{2h_y^2}\right) \right. \\ &\quad \left. \times \sum_{j=1}^{n_2} \exp\left(-a\left((x-x_{2,j})^2 + (y-y_{2,j})^2\right)\right) \right\} dx dy \\ &= \frac{1}{4n_1 \sqrt{1+2ah_x^2} \sqrt{1+2ah_y^2}} \\ &\quad \times \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \exp\left(-\frac{a(x_{1,i}-x_{2,j})^2}{1+2ah_x^2} - \frac{a(y_{1,i}-y_{2,j})^2}{1+2ah_y^2}\right) \\ &\quad \times \operatorname{erf}\left(\frac{x-x_{1,i}+2ah_x^2(x-x_{2,j})}{h_x \sqrt{2+4ah_x^2}}\right) \\ &\quad \times \operatorname{erf}\left(\frac{y-y_{1,i}+2ah_y^2(y-y_{2,j})}{h_y \sqrt{2+4ah_y^2}}\right) \quad (5) \end{aligned}$$

(5)式において、各項における積分定数については省略した。また、 $\operatorname{erf}(\cdot)$ は誤差関数をあ