

意味フィルタにおける転送空間データのルール生成

*根路銘もえ子 NEROME Moeko 屋比久友秀 YABIKU Tomohide 松田善臣 MATSUDA Yoshitaka
姜東植 KANG Dongshik 宮城隼夫 MIYAGI Hayao 翁長健治 ONAGA Kenji

通信・放送機構 沖縄リサーチセンター

Okinawa Research Center, Telecommunications Advancement Organization

1. はじめに

本研究では、データ転送におけるネットワークの伝送効率の向上を目的として、ユーザにとって不必要なデータをフィルタリングし、意味のあるデータのみを抽出するシステム「意味フィルタ」を提案する。意味フィルタの有効性を検証する対象として、地理情報システム (Geographic Information System: GIS) [1] を採用する。GIS に意味フィルタを適用する場合には、提示する空間データである建物の重要度をユーザに応じて決定する必要がある。これまでのシステムにおいて、建物の重要度を決定するルールは、予め設計者の意思により生成され、それに基づき建物の重要度が決定されていた。本稿では、よりユーザの目的や好みに応じた建物を提示するために、アンケートをもとに C4.5[2] を用いて重要度決定ルール生成を行う。また、建物の幾何および位相情報を基にした重要度の決定をファジィ推論により行う。2手法で得られた重要度を合成することにより、建物の重要度を決定する。本稿では、これら建物重要度の決定方法について説明する。

2. 意味フィルタ

本研究が提案する「意味フィルタ」を用いたデータのフィルタリングの流れを以下に示す (図 1)。

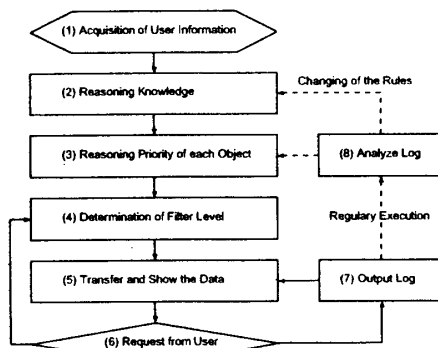


図 1: フィルタリングの流れ

[step1:] ユーザ情報の入手

各ユーザの認証を行い、予め入力されたユーザ情報を入力する。

[step2:] 利用情報に対する知識度の推論

一般的なルールを用いたファジィ推論によって利用情報に対するユーザの知識度を推測する。

[step3:] データの重要度の決定

ユーザに応じたデータの重要度合いをユーザ情報を用いて決定する。

[step4:] フィルタレベルの決定

得られた知識度と、ユーザの通信・端末環境の情報をもとに、データをフィルタする度合いを決定する。

[step5:] データのフィルタリング

フィルタレベルに応じてデータをフィルタリングする。

[step6:] ユーザ操作

転送されたデータの結果に対して、ユーザ指示を受け、フィルタレベルの調整と差分情報の転送等を行う。

[step7:] ログ出力

得られたユーザ指示や、転送されたデータの情報などを適宜アクセスログとして保存しておく。

[step8:] ログ解析

収集された全てのアクセスログに対してリコメンデーション方式による解析を行う。これにより、同グループに分類できるユーザに対して、推論を行わずにデータ転送を行うことが可能となる。

本研究では、意味フィルタの有効性を検証するために、GIS における意味フィルタの設計を行っている。これまでの意味フィルタを用いた GIS については、文献 [5, 6] にて報告している。

3. 転送空間データの決定ルール生成

本節では、2. 節で提案した意味フィルタを GIS へ適用する場合における step3 の手続きについて詳しく述べる。

意味フィルタを GIS へ適用する場合、図 1 の step3 は、転送する空間データ、つまり建物データの重要度を決定する手続きに相当する。

建物重要度の推測方法は大きく分けて以下の 2 通りがある。

- 建物の持つ属性値とユーザの年齢・職業等から推測
- 建物の持つ幾何情報や位相情報から推測

よりユーザの目的や好みに応じた建物を提示するために、前者に関して、アンケートをもとに、C4.5[2]を用いたルール生成を行う。また、後者における建物の幾何および位相情報を基にした重要度の決定をファジィ推論により行う。これら2手法で得られた重要度の和をとることによって、建物の重要度を決定する。

3.1 ユーザ情報による重要度決定ルール生成

本研究では、ユーザ情報から各種建物の重要度を決定するルールを生成するために、J.Quinlanによって開発された帰納学習システム C4.5 を用いる。C4.5 は、与えられた教師データに対するパターン解析を行うことにより、帰納的に決定木を作るシステムである。

重要度決定ルールを生成するための教師データを収集するために、各種建物について、ユーザの目的に応じてどれだけ重要であるかをアンケート調査する。そのアンケートに基づき、各種建物の訓練事例を生成する。訓練事例は、以下の通り、3種の属性と5つのクラスが存在することとする。

[3種の属性:] 目的 (買う, 食べる, 楽しむ), 性別 (男, 女), 年齢

[5つのクラス:] 提示すべき, 提示してもよい, どちらでもない, 提示しなくてもよい, 提示しない

これらで構成される訓練事例の1例は以下の通りである。

例: デパートに関する1訓練事例

[買う, 女, 25, 提示すべき]

3.2 幾何および位相情報による重要度決定ルール生成

建物の幾何および位相情報を用いてファジィ推論により、重要度を決定する。前件部としては、建物の面積および交差点からの距離を用いる。推論ルールの例を表1に示す。

表1: 建物重要度の推論に用いるルールの例

ルール1:	交差点から	近く	面積が	広ければ	高い
ルール2:	交差点から	やや近く	面積が	やや広ければ	やや高い

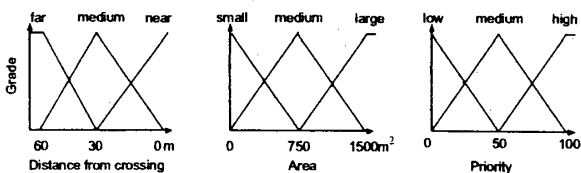


図2: 三角型メンバーシップ関数の例

表1では、「交差点からの距離」と「面積」から、「建物の重要度」を推論している。このとき、各ファジィ変数は図2のよう

な三角型メンバーシップ関数で定義しておく。ここで、三角型メンバーシップ関数を用いるのは計算の簡単化のためである。

これらのルールとメンバーシップ関数をもとに、Min-Max-重心法[3]を用いてデファジィ化を行うことによって、建物の重要度を求める。

3.3 重要度の合成

ユーザ情報をもとに決定された重要度と幾何情報をもとに決定された重要度の和を取り、建物の重要度とする。これにより、ユーザ情報だけでなく、建物自身の情報を考慮した重要度の決定が可能となる。

このように決定された建物の重要度をもとに、図1のstep4で求めたユーザのフィルタレベルに応じて、提示する建物を決定する。

4. おわりに

本研究では、ユーザが望むデータを抽出可能な意味フィルタをGISへ適用する場合における建物重要度の決定方法を提案した。これまでのシステムにおいて、建物の重要度を決定するルールは、予め設計者の意思により生成され、それに基づき建物の重要度が決定されていた。それに対し本提案手法は、アンケート結果に基づき、ユーザ情報を用いた建物重要度の決定ルールを帰納学習アルゴリズム C4.5 により生成している。これにより、よりユーザに適したルールの生成が可能となっている。また、建物の幾何および位相情報を用いた建物重要度の決定をファジィ推論により行っている。これらの2手法で得られた重要度の和をとることによって、ユーザ情報だけでなく建物情報を考慮した重要度の決定が可能であるといえる。今後は、ユーザのログを解析することによって、重要度決定ルールの再構築について検討を行う。また、オブジェクト指向GIS[4]において、G-XMLデータを用いたフィルタリングについても検討する。

参考文献

- [1] Jeffrey Star, John Estes, 訳者 岡部篤行, 貞弘幸雄, 今井修: 入門地理情報システム, 共立出版 (1998).
- [2] J. R. Quinlan, 監訳 古川康一: AIによるデータ解析, トップラン (1995). (1998).
- [3] 中島信之, 竹田英二, 石井博昭: 社会科学の数理 ファジィ理論入門, 裳華房 (1994).
- [4] 賢良則, 名嘉村盛和, 宮城隼夫, 翁長健治: オブジェクト指向技術を用いた空間データ転送システム「龍潭」の設計と実装, 地理情報システム学会, GIS-理論と応用, Vol.8 (2000).
- [5] 根路銘もえ子, 屋比久友秀, 松田善臣, 姜東植, 宮城隼夫, 翁長健治: ユーザ情報に基づく地理空間データの適応フィルタリング, 電子情報通信学会総合大会講演論文集, p.408 (2002).
- [6] 松田善臣, 宮城隼夫, 姜東植, 上間淳也, 翁長健治: 意味フィルタを用いた空間データの抽出, 第10回インテリジェント・システム・シンポジウム講演論文集, pp.417-420 (2000).