

平成 21 年度

修士学位論文

サラウンディングキャンパスにおける
コンテキストに基づく
情報推薦システムの構築

Development of Information Recommendation
System base on the Context in Surrounding Campus

1125103 山崎雄大

指導教員 妻鳥貴彦

2010 年 3 月 1 日

高知工科大学大学院 工学研究科 基盤工学専攻
情報システム工学コース

要 旨

サラウンディングキャンパスにおける コンテキストに基づく 情報推薦システムの構築

山崎雄大

近年, ICT(Information and Communication Technology) の発展は, 大学の情報環境に影響を与えている. 一世代前までは紙媒体で管理されていた情報が電子化され, ネットワーク環境の整備が進むと, 簡単に情報サービスを取得することができるようになった. 近年では, 無線通信や携帯端末を始めとするユビキタスコンピューティング技術を大学に適用した情報環境が整えられつつある. これにより, 新たな学習支援や教育支援が行われると考えられる. しかし, ユーザからすれば, 多くの情報サービスから必要な情報サービスを取得するためには多くの手間がかかると考えられる. また, 多くの時間をかけて探し出した情報サービスが実質的に必要とする情報サービスではない場合が考えられる. そこで, 我々はこの問題を解決するサラウンディングコンピューティングに着目し, サラウンディングコンピューティング技術を大学に適用したサラウンディングキャンパスの提案を行ってきた.

本研究では, サラウンディングキャンパスのシステムの一部として, ユーザのおかれている状況を把握するために, コンテキストから状況の推測をおこなうモデルを提案し, 散在する情報サービスに対して情報フィルタリングを適用することで, ユーザにとって必要な情報サービスの推薦システムの構築を行う.

キーワード サラウンディングキャンパス, サラウンディングコンピューティング, ユビキタスコンピューティング, 情報フィルタリング

Abstract

Development of Information Recommendation System base on the Context in Surrounding Campus

Yudai YAMASAKI

In recent years, computer and network environments have been created in a lot of places of the world. It is affected the information environment of universities. For example, introduction of LMS(Learning Management System), electronic bulletin board system and library search system. Thus, universities digitalize paper based information service. Nowadays, universities have been create information environment which applied Ubiquitous Computing technology Teachers and students utilize new service of quality study support and education support. However, these universities have some problems. Teachers and students spend the time and effort necessary to get information service.

Therefore, we on focus Surrounding Computing environment. Users get necessary information service which correspond to situation in Surrounding Computing environment. We proposed Surrounding Campus which Surrounding Computing environment applies to university.

In this study, we develop information service recommendation system which relate information filtering.

key words Surrounding Campus, Surrounding Computing, Ubiquitous Computing, Information filtering

目次

第1章	はじめに	1
第2章	背景	2
2.1	大学の情報環境	2
2.1.1	大学の情報環境の変化	2
2.1.2	大学へのユビキタスコンピューティング技術の適用	2
2.1.3	大学の情報環境の問題点	3
2.2	サラウンディングコンピューティング	3
2.2.1	サラウンディングコンピューティング環境	3
2.2.2	サラウンディングコンピューティング技術	4
2.3	サラウンディングキャンパス	5
2.4	サラウンディングキャンパスの構成	6
2.5	過去の研究	7
2.6	本研究の目的	8
第3章	情報推薦	9
3.1	情報検索	9
3.2	情報フィルタリング	10
3.3	コンテンツ内容に基づくフィルタリング	10
3.3.1	ルールベース方式	11
3.3.2	メモリベース方式	11
3.3.3	モデルベース方式	11
3.4	協調フィルタリング	12
3.4.1	ユーザベース方式	12

目次

3.4.2	アイテムベース方式	12
3.4.3	モデルベース方式	12
3.5	内容型検索のモデルと全文検索のモデル	13
3.6	索引語の抽出と重み付け	13
3.7	ベクトル空間モデルに基づく情報サービスの検索	14
3.7.1	ベクトル空間モデル	14
3.8	文字列照合に基づく全文検索	15
3.8.1	KMP(Knuth-morris-Pratt) 法	15
3.8.2	BM(Boyer-Moore) 法	15
3.8.3	AC(Aho-Corasick) 法	16
3.9	索引を用いた全文検索	16
3.9.1	特徴ベクトル法	16
3.9.2	転置ファイル法	17
3.9.3	パトリシア・トライ法	18
第 4 章	情報推薦システムの検討	19
4.1	状況のモデル化	19
4.1.1	大学における状況	19
4.1.2	場所の持つ役割	20
4.1.3	時間の持つ役割	20
4.1.4	ユーザの持つ役割	21
4.1.5	周囲のユーザの持つ役割	21
4.1.6	状況を予測するためのデータモデルの作成	21
4.2	情報サービスの推薦	22
4.2.1	ユーザプロファイルとコンテンツモデル	22
4.2.2	コンテンツ内容に基づくフィルタリングと協調フィルタリング	22

目次

4.2.3	類似度の計算	23
4.3	情報推薦システムの提案	23
第 5 章	情報サービス推薦システムの構築	25
5.1	情報サービス推薦システムの設計	25
5.1.1	状況を把握する機能	26
5.1.2	情報サービスの推薦を行う機能	26
5.1.3	情報サービスの推薦と履歴の格納を行う機能	28
5.2	システムの概要	28
第 6 章	評価・考察	30
6.1	評価の概要	30
6.2	評価と考察	31
第 7 章	おわりに	32
	謝辞	33
	参考文献	34

図目次

2.1	サラウンディングキャンパスの概念図	6
2.2	SunSpot	8
3.1	情報検索の構成	9
5.1	情報サービス推薦システム全体の構成図	25
5.2	システムの概要図	29

表目次

4.1	情報サービスが利用される状況の一例	20
5.1	ユーザプロファイルの構成	27
5.2	コンテンツモデルの構成	27
6.1	テストデータの一部	31

第 1 章

はじめに

近年, ICT(Information and Communication Technology) の発展は, 大学の情報環境に影響を与えている. 一世代前までは, 学生情報や掲示板, シラバスなどは紙媒体であり, 情報の追加や削除といった管理に多くの手間がかかっていた. また, 学生や教職員はキャンパスライフに必要な情報を取得するために自ら事務室や掲示板に足を運ぶ必要があった. しかし, 情報の電子化が進み, ネットワーク環境の整備が進むと, ネットワーク環境さえ整っていればどこからでも情報サービスを取得することができるようになった. 近年では, 無線通信や携帯端末を始めとするユビキタスコンピューティング技術を大学に適用し, いつでも, どこでも情報サービスを取得することができる環境が整いつつある. また, ユビキタスコンピューティング環境の構築により, 新たな学習支援や授業支援などの情報サービスが提供が行われつつある. しかし, 今後多種多様な情報サービスの提供が始まるとユーザは多くの情報サービスの中から必要な情報サービスのみを取得するために多くの手間をかける必要がある. また, 多くの時間をかけて探し出した情報サービスが実質的に必要とするサービスではなかったという場合が考えられる. そこで, 我々はこの問題を解決するサラウンディングコンピューティングに着目し, サラウンディングコンピューティング技術を大学に適用したサラウンディングキャンパスの提案を行ってきた.

本研究では, ユーザのおかれている状況を把握するために, コンテキストから状況を推測するモデルを提案し, 状況の判断を行う. さらに, 散在する情報サービスに対して情報フィルタリングを適用することで, ユーザにとって必要な情報サービスの推薦を行う.

第 2 章

背景

2.1 大学の情報環境

2.1.1 大学の情報環境の変化

近年の ICT の発展は大学における学生や教員、職員のキャンパスライフに変化をもたらしている。一世代前、学生は手書きでレポートの作成を行う、教員は黒板を利用して講義を行う、職員は学生情報を紙媒体で管理する、という光景は一般的であった。しかし、コンピュータの導入により情報の追加、削除、更新などの作業効率の向上が図られると、これまで紙媒体で管理されていた多くの情報がデジタル化され、データとして扱われるようになった。さらに、ネットワーク技術の発展により、掲示連絡や履修登録、学生情報の管理、成績管理まで Web サービスとして管理、処理されることが一般的となった。現在ではネットワーク環境さえ整っていればいつでも、どこでも情報サービスを楽しむことができる。

2.1.2 大学へのユビキタスコンピューティング技術の適用

近年では大学にワイヤレス通信や携帯端末などのユビキタスコンピューティング技術を適用する取り組みが行われている。代表的な取り組みとして徳島大学の u-Campus(Ubiquitous Campus) 構想に基づいて大学にユビキタスコンピューティング環境の構築を行う取り組みがある [1][2]。

u-Campus 構想は徳島大学の教育理念の実現に向けて、キャンパスのデジタル化、効果的な教育実践、新しい教育の提案・開発の 3 つを軸に大学の情報化を行うことを目指して

2.2 サラウンディングコンピューティング

いる。u-Campus 構想を実現する具体的な取り組みの一つとして、u-Learning(Ubiquitous Learning) プロジェクトがある。u-Learning プロジェクトでは大学において、無線 LAN や携帯電話、PDA、RFID 等のユビキタスコンピューティングデバイスを活用することで真に「いつでも、どこでも、だれでも」講義を受けることができる環境の提供を目指している。

2.1.3 大学の情報環境の問題点

徳島大学の取り組みのように、大学ではユビキタスコンピューティング技術を適用し「いつでも、どこでも、だれでも」教育支援サービスや学習支援サービスを享受できる環境提供が行われている。ユーザからすれば、ユビキタスコンピューティング環境が整えられることで「いつでも、どこでも」情報サービスを取得することが可能となる。さらに、ユビキタスコンピューティング環境の特徴を生かした新しい情報サービスが期待できる。しかし、多くの情報サービスから必要な情報サービスを取得することは容易ではない。また、苦労して取得した情報サービスが実質的にユーザにとって必要ではない場合も考えられる。

そこで、我々はユビキタスコンピューティングに着目し、大学に適用することでユーザの状況に応じて必要な情報サービスの提供を行う、サラウンディングキャンパスを構築する取り組みを行ってきた。

2.2 サラウンディングコンピューティング

2.2.1 サラウンディングコンピューティング環境

サラウンディングコンピューティング環境とは、ユーザからすれば、身の回りに常に情報サービスがあるような環境であり、生活する中で起こる場面に応じて必要な情報サービスを取得することができる情報環境である [3][4]。

ユビキタスコンピューティング環境ではユーザは多くの情報サービスの中から自身が必要とする情報サービスを探し出す必要があるがあり、情報サービスを取得するまでに多くの時間と手間がかかる。また、探し出した情報サービス自体が実質的に必要とする情報サービス

2.2 サラウンディングコンピューティング

でない場合の問題がある。これに対して、サラウンディングコンピューティング環境ではユーザの状況に応じて必要な情報サービスの提供が行われるため、ユーザが散在する情報サービスの中から情報サービスを探し出さなければならないといった手間を省くことができる。さらにユーザの状況に応じた情報サービスを提供することで、実質的に必要でない情報サービスを取得することが軽減される。

2.2.2 サラウンディングコンピューティング技術

ユビキタスコンピューティング環境では散在している情報サービスを必要に応じて提供するために、常に膨大な情報処理を行う必要がある。ユーザの要求に対して情報サービスを収集してリアルタイム処理を行うこともできるが、一部のネットワークやコンピュータデバイスへの負担が増大する恐れがある。そこで、ネットワーク上の計算資源やデータベース資源を有効に活用して分散して処理することが求められる。そもそも、ユーザからすれば要求が満たされていれば良いため、データの流れを考えたとき、わざわざ情報サービスを集約する必然性はなく、分散処理することが自然である。このことから、ネットワークや情報機器の機能を意識せずに、ネットワーク上の計算資源やデータベース資源を自由に活用できる、進化するユビキタスコンピューティング環境であるサラウンディングコンピューティング環境の確立を目指している。

情報ネットワークを介して転送される情報サービスをより有益なものとするには、あらゆる情報サービスの有機的な連携や分散している情報の自動配信が求められる。そこで、サラウンディングコンピューティング環境で提供される情報サービスへの高度な価値の付与とその情報サービスを転送・提示するための方式についての研究を、超高速・高機能研究開発テストヘッドネットワーク JGN2(Japan Gigabit Network) における次世代高機能ネットワーク基盤技術、利活用技術に関する研究開発プロジェクトのテーマとして実施している [5]。更に、多様な価値を付与された情報サービスを転送するためには、送信するデータの種類や用いる通信媒体、その他外的要因によらずリアルタイムでの送受信や再現を可能にする事が求められる。そこで、高速、高効率符号処理、画像、音声情報などの多種多様な情報サービスを

2.3 サラウンディングキャンパス

統合的に再現するための信号処理システムの研究開発を行っている。

2.3 サラウンディングキャンパス

近年のユビキタスコンピューティング技術の発展に伴って、大学においてもいつでも、どこでも情報サービスを享受することができる環境が整いつつある。今後、情報環境の整備により、さらに多くの情報サービスが提供されるようになると予測できる。このため、ユーザからすれば多くの情報サービスの中から必要な情報サービスを探す必要があり、情報サービスを取得するまでに多くの手間がかかってしまう。

そこで、我々はユーザの状況に応じて必要な情報サービスの提供を行うサラウンディングコンピューティングに着目し、サラウンディングコンピューティング技術を大学に適用することでより快適な情報環境であるサラウンディングキャンパスの提案を行ってきた。サラウンディングキャンパスで提供される情報サービスは主に学習支援、授業支援、研究活動支援を想定しており、大学に関わる全ての学生、教員、職員を対象としている。図 2.1 にサラウンディングキャンパスの概念図を示す。

サラウンディングキャンパスにおける情報サービスの提供例としては、人工知能という科目を履修している学生が人工知能の講義が行われている教室へ向かうとその科目の講義資料が提供される。教員が教員室へ向かうと研究資料、担当している講義の資料、学生の成績管理情報などの情報サービスが提供される。このように、サラウンディングキャンパスではユーザの状況に応じて必要な情報サービスの提供が行われる。

2.4 サラウンディングキャンパスの構成

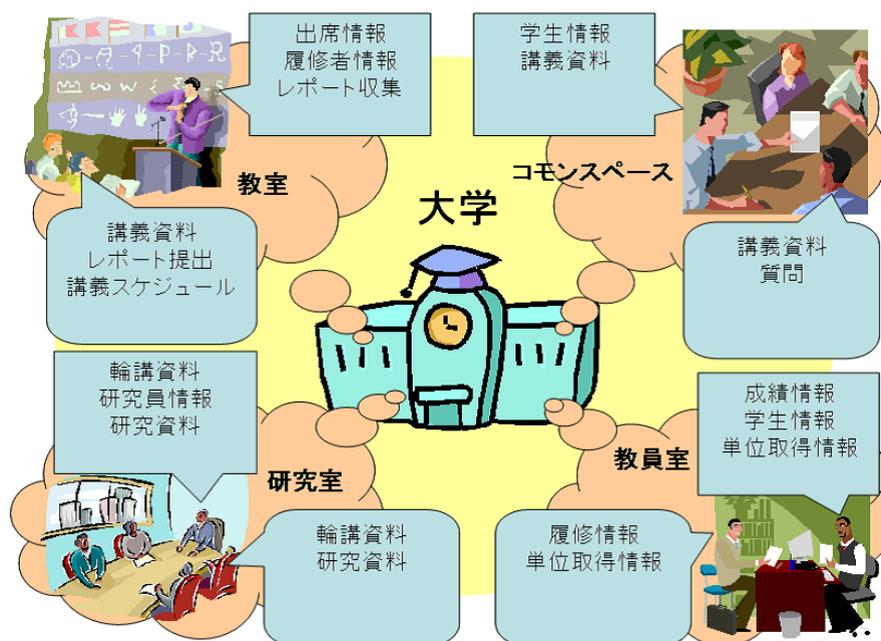


図 2.1 サラウンディングキャンパスの概念図

2.4 サラウンディングキャンパスの構成

サラウンディングキャンパスの全体構成は、情報の取得、取得した情報の解析とユーザの状況の把握、把握した状況に対する情報サービスの選択・提供、履修情報の更新から構成される。

- コンテキストの取得

情報サービスの提供を行うためにはユーザの状況を把握する必要がある。そのため、コンテキストに着目し、コンテキストからユーザの状況を推測する。

- 取得した情報の解析と状況の把握

コンテキストは状況を表す要素の集合であることからコンテキストを取得しただけでは状況を把握したことはない。コンテキストをサラウンディングキャンパスで扱うデータモデルに従って体系化することで状況の把握を行う。情報を整理するデータモデルはあらかじめ設定しておく必要がある。ユーザの状況に対してどのような情報サービスの提供を行うのかによってデータモデルの精巧さを調整する必要あがる。データモデ

2.5 過去の研究

ルが精密になるほど実装が複雑になるため、情報サービスの提供にどの程度の情報があれば良いかを考慮する必要がある。

例えば、教室で人工知能の科目の開講されており、学生に対して人工知能の科目の講義資料を提供する場合、教室名と時間の情報が取得できれば、教室に訪れた全ての学生に対して情報サービスの提供を行うことができる。しかし、科目の履修を行っている学生に対してのみの情報サービスの提供を行う場合を想定すると、教室名、時間、ユーザの履修情報が必要となる。

- 状況に対する情報サービスの選択

ユーザの状況をユーザからの要求と捉え、事前に用意しておいた情報サービスの中から要求に見合う情報サービスの選択を行う

- 履歴情報の蓄積

発生したユーザの状況に応じてどのような情報サービスが提供されたかを履歴情報として蓄積する。他のユーザに対する情報サービスの提供を行うとき、同様の状況が発生した場合に蓄積された情報履歴を基に情報サービスの提供を行う。

2.5 過去の研究

ユーザの状況に応じた情報サービスの提供を行うためには、ユーザのおかれている状況を把握する必要がある。これまで、先行研究として小型無線デバイスを用いたユーザのコンテキストの取得手法が提案されている。これによりユーザのおかれている状況のコンテキストの取得を行うことができる [6]。

コンテキストとは物体の状況を特徴づける全ての情報のことであり、サラウンディングキャンパスではユーザの状況を把握するため要素の取得を行った。コンテキストの取得には、SunMicrosystems 社の SunSpot センサデバイスを使用した。図 2.2 に利用した小型無線センサデバイスを示す。

2.6 本研究の目的



図 2.2 SunSpot

2.6 本研究の目的

サラウンディングキャンパスではユーザの状況に応じて必要な情報サービスの提供が行われる。サラウンディングキャンパスを実現するためには情報の取得、取得した情報の解析とユーザの状況の把握、把握した状況に対する情報サービスの選択、履歴情報の更新を行う機能が必要である。過去の研究として、ユーザのおかれている状況を把握するためにコンテキストの取得を行う方法を提案した。次の段階として、取得したコンテキストからユーザのおかれている状況を推測し、ユーザの状況において必要であると考えられる情報サービスの提供を行う必要がある。

そこで本研究では、ユーザのおかれている状況を把握するために、コンテキストから状況の推測をおこなうためのモデルを提案し、散在する情報サービスに対して情報フィルタリングを適用することで、ユーザにとって必要な情報サービスの推薦を行う情報サービス推薦システムの構築を行う。

第 3 章

情報推薦

3.1 情報検索

情報検索は、大量の情報の中からユーザの要求を満たす情報を見つけ出すことである。検索対象となる情報としては、文書や画像、動画があり、一般的に情報の検索には自然言語や索引語が用いられる。情報検索では情報対象となる複数の文書を文書集合、検索対象者の要求を検索質問と呼ぶ。検索質問は一般的にキーワードが用いられ、キーワードと文書集合との一致を計ることにより、検索結果を得ることができる。検索の処理の流れとしては検索質問と文書集合を中間的な内部表現にして照合を行い、ユーザの要求に適合している情報を結果として出力する [7][8][9]。

情報検索には大きく 2 つの検索モデルがある。検索質問は常に変化するが、文書集合は比較的固定されているような一般的な情報検索のモデルと検索質問は比較的固定されており、文書集合が常に変化する検索モデルがある。前記のような情報検索をアドホック検索、後記を情報フィルタリングと呼ぶ。図 3.1 に一般的な情報検索の構成について示す。

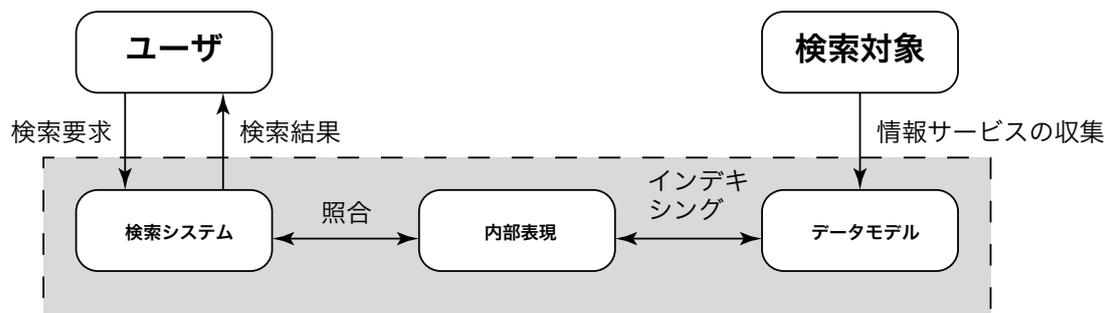


図 3.1 情報検索の構成

3.2 情報フィルタリング

情報フィルタリングは大量の情報の中から、ユーザにとって必要な情報を取捨選択する技術の総称である。一般的には膨大な情報資源の中からユーザの嗜好、関心、興味に合った情報を抽出・推薦するサービスに適用される。推薦システムを使用した代表的なサービスとしては Amazon.com、文房具販売のアスクル、CD・ビデオ・ゲームソフトのレンタルと販売の TSUTAYA online などがある。これらの情報推薦サービスはユーザの過去の購買履歴やアンケートなどからユーザの嗜好を推測し、商品の推薦が行われている。また、あるアイテムの推薦において、他のユーザが評価を行った情報を基にした推薦手法もある。情報フィルタリングでは、情報検索における検索質問をユーザプロフィール、文書集合をコンテンツモデルと呼ぶ。情報推薦にはユーザプロフィールとコンテンツモデルの類似度を求めることにより情報サービスが決定される [10][11][12][13][14]。

- ユーザプロフィール

ユーザプロフィールはユーザの過去の行動履歴やユーザの嗜好に関する情報を獲得しモデル化した情報である。嗜好情報を獲得しモデル化することを、嗜好抽出またはユーザプロフィールと呼ぶ。

- コンテンツモデル

コンテンツモデルは検索対象となるコンテンツの特徴量を抽出し、モデル化したものである。コンテンツの特徴量は、音楽データや映像データであれば、テンポや周波数成分、色情報、差分画像情報などから特徴量の抽出を行う。

3.3 コンテンツ内容に基づくフィルタリング

コンテンツ内容に基づくフィルタリングは、ユーザプロフィールとコンテンツモデルの類似度を計ることで情報を推薦する手法である。コンテンツ内容に基づくフィルタリングは情報サービスの内容に基づく検索が行われることからコンテンツが追加された場合でも安定した推薦が行われる。しかし、ユーザプロフィールとコンテンツモデル間で比較する要素が

3.3 コンテンツ内容に基づくフィルタリング

多数ある場合、嗜好予測の精度は向上するが、実装が難しくなるという問題点がある。また、ユーザの嗜好の変化に追従することが難しいため、ユーザプロファイルの更新が必要である。コンテンツ内容に基づくフィルタリングはさらに3つの方式に分類される。

3.3.1 ルールベース方式

システム開発者がユーザの評価したアイテムと推薦したいコンテンツとの対応ルールを設計する方式である。作成されたルールに従って、ユーザプロファイルとコンテンツモデルの嗜好の類似度を計ることで情報サービスの推薦を行う。実現したい情報サービスがあらかじめ決まっている場合、少ない推薦コストで実現することができる。

3.3.2 メモリベース方式

ユーザプロファイルとコンテンツモデルの両方をベクトルで表現し、ベクトル空間上にそれぞれを配置し、距離を比較することによって推薦する情報サービスを決定する方式である。推薦対象のコンテンツがテキストである場合は、tf・idf法などを用いてキーワードの出現頻度を求め、キーワードの重み付けを行う方法が用いられる。

3.3.3 モデルベース方式

過去の閲覧や購読を機械学習の教師信号と考え、コンテンツモデルとそれに対する正負の判断の組を機械学習のアルゴリズムに入力することで学習を行い、新たなユーザプロファイルとする方式である。コンテンツが文書の場合は、文書から文書ベクトルを生成する。ユーザは興味あり/興味なしという評価を文書に対して付けているとする。これらの組を機械学習アルゴリズムにかけることで、興味あり/興味なしを判断するモデルを生成する。学習には時間がかかるが、推薦実行時には高速な処理が可能である。

3.4 協調フィルタリング

3.4 協調フィルタリング

協調フィルタリングは興味・関心・嗜好が類似した人間は、同じ情報を求めているという仮定に基づいて検索が行われる手法である。ユーザと嗜好や関心が類似する他のユーザの利用・評価した情報を参考にして、ユーザが利用・評価していない情報サービスを推薦する。協調フィルタリングでは情報サービスの持つ内容については問わない。ユーザがどのコンテンツにどのような評価を行ったかのみに着目して情報サービスの推薦を行う手法である。

3.4.1 ユーザベース方式

ユーザと嗜好パターンが類似した標本ユーザを探索し、そのユーザが好むものを推薦する方式である。ユーザプロファイリングをもとに、ユーザの嗜好に近い他のユーザのユーザプロファイリングとの嗜好の類似度の計算を行う。類似度の計算からユーザは知らないが他のユーザは知っている情報サービスに対して、嗜好の予測値の計算を行う。ユーザベース方式はモデルの構築を行わないのでユーザやコンテンツの数が変化したとしても柔軟に対応ができる。しかし、推薦時に多くの項目の走査を行う必要があるため時間がかかる。

3.4.2 アイテムベース方式

コンテンツの評価値に着目した方式である。各ユーザがそれぞれのコンテンツに付けた評価値の類似度の計算を行い、類似度が高いコンテンツに対して評価予測値の計算を行うことでコンテンツの推薦を行う。ユーザベース方式と同様にユーザやコンテンツの数が変化しても柔軟に対応ができるが、推薦時の項目の走査に対して多くの時間を要する。

3.4.3 モデルベース方式

モデルベース方式は、ユーザやコンテンツ間の関係をあらかじめモデル化しておき、そのモデルを用いて推薦候補を予測する方法である。モデル化に用いられる代表的な手法として、クラスタリングがある。クラスタリングを用いる手法では、ある特徴を有するユーザ集合を

3.5 内容型検索のモデルと全文検索のモデル

事前にクラスタ化しておき、そのクラスタの特徴を表す代表的なベクトルを生成する。クラスタ化したグループとの類似度の計算を行うことで推薦候補を予測する。モデルベース方式は事前にモデルの予測を行うため、コンテンツの推薦時間が短くなる。しかし、ユーザやコンテンツの数が増えた場合、モデルの再構築が必要となる。

3.5 内容型検索のモデルと全文検索のモデル

コンテンツ内容に基づくフィルタリングと協調フィルタリングの情報検索を実現する方法として、内容型検索のモデルと全文検索のモデルを用いる方法がある。全文検索のモデルはユーザからの要求に対して、情報サービスの持つ情報が完全に一致する部分を探し出し、一致した情報サービスを検索結果とする。全文検索のモデルは文書内の文字を一文字ずつ照合する逐次検索と事前に文書から作成された索引を用いる索引検索の2種類に分類することができる。逐次検索の代表的な手法には KMP 法, BM 法, AC 法が、索引検索の代表的な手法には特徴ベクトル法, 転置ファイル法, パトリシア・トライ法がある。

一方、内容型の検索のモデルは全文検索のモデルのような完全な照合ではなく、ユーザの要求と意味的な類似を計ることに主眼をおいている。内容型検索ではユーザの要求と情報サービスのそれぞれのもつ情報を特徴付ける索引語の抽出と索引語の重みの抽出を行い、ユーザの要求と情報サービスの間にある類似度あるいは距離を定義することにより、ユーザの要求に適合する情報サービスの抽出を行う。内部型の検索モデルの代表的な手法にはベクトル空間モデルがある。

3.6 索引語の抽出と重み付け

情報検索において、情報サービスが持つ全ての文書や単語が必ずしも一律に情報サービスと関係しているわけではない。日本語では助詞や助動詞、英語では冠詞や前置詞などは、情報サービスと一般には無関係にある。このことから情報検索を行う前に情報サービスと関連のある重要な文書や単語を抽出するインデキシングを行うことで、検索時間を短縮することが

3.7 ベクトル空間モデルに基づく情報サービスの検索

できる。また、インデキシングを行う際に抽出する文書や単語の重み付けを行うことで、ユーザの要求を意味的に満たす情報サービスの抽出を行うことができる。情報検索ではインデキシングにより情報サービスと意味的な関連がある文章や単語を索引語と呼ぶ。

情報サービスの索引語の抽出には人の手によるインデキシングと計算機による自動インデキシングがある。人の手によるインデキシングは抽出される索引の一貫性を保つことが難しい。また、一般的に検索対象の情報サービスは大量にある場合が多いため、膨大な時間がかかってしまう。このため一般的には自動インデキシングが用いられる。インデキシングの検索対象が英語やフランス語の場合には単語間の空白に着目して、日本語や中国語の文章の場合には、形態素解析を用いて索引語の抽出が行われる。

索引語に対する重み付けの方法には一般的に、情報サービスに対する索引語の出現回数により重みを決める局所的重み付け、情報サービス内における索引語の分布を考慮した大域的重み付け、情報サービスが文章であった場合、文書の長さとその中に含まれる索引語の数に着目した文書正規化係数による重み付けの方法が用いられる。

3.7 ベクトル空間モデルに基づく情報サービスの検索

3.7.1 ベクトル空間モデル

ベクトル空間モデルではユーザの要求と情報サービスのそれぞれから索引語の重みを要素とするベクトルで情報サービスの表現を行う。抽出したベクトル間の類似度を計ることで情報サービスの推薦を行う方法である。検索対象の情報サービスが文書であった場合、検索対象の文書における索引語の重みを要素とする文書ベクトルを作成する。文書ベクトルの集合は $m \times n$ の検索対象全体の文書ベクトルの行列として表すことができる。この行列の各行は索引語ベクトルと呼ぶ。ユーザの要求も同様に索引語の重みを要素とするユーザ要求ベクトルとして表すことができる。

情報検索ではユーザの要求に類似する情報サービスを見つけ出す必要があるが、ベクトル空間モデルではユーザ要求ベクトルと文書ベクトルの類似度を計算することにより行う。類

3.8 文字列照合に基づく全文検索

似度の計算にはコサイン尺度や内積がよく用いられる。

3.8 文字列照合に基づく全文検索

3.8.1 KMP(Knuth-morris-Pratt) 法

KMP 法は単一のキーワードを高速に検索するアルゴリズムの一つである。単純なキーワード検索のアルゴリズムでは、キーワードの長さを n としたとき、検索対象のテキストストリングの 1 文字目から n 文字目まで、2 文字目から $n+1$ 文字目まで、順次キーワードと一致するかをチェックすればよい。しかし、このアルゴリズムではテキストストリング中の各文字がキーワードと複数回照合されることがあり、この点が冗長であるといえる。この欠点を解決する方法が KMP 法である。

KMP 法では、情報サービスのテキストストリングに対して先頭から文字の照合を行い、照合が失敗したときに次はテキストストリングのどの位置から照合を開始するべきかという情報を利用する。KMP 法は単純なキーワード検索アルゴリズムに対して、計算量が改善されている。しかし、一般的にキーワードはテキストストリングの長さに対して 10~20 文字程度と短い。また多くの場合、照合を行うときに 1 文字目で失敗する。さらに、KMP 法では失敗した際に次の照合開始位置を計算する時間が必要となる。これらのことから実際に KMP 法をキーワード検索に適用したとしても速度的には単純なキーワード検索アルゴリズムと大差はない。

3.8.2 BM(Boyer-Moore) 法

BM 法は単純なキーワード検索のアルゴリズムや KMP 法がキーワードの先頭からテキストストリングと文字列照合を行っていたのに対して、キーワードの末尾から文字列の照合を行って行く方法である。BM 法ではまず、キーワードとテキストストリングの文字列の 1 文字目を揃え、キーワードの末尾の位置の文字列の照合を行う。テキストストリング中の文字がキーワード中に存在する場合には、その位置が揃うようにキーワードをシフトし、キー

3.9 索引を用いた全文検索

ワード中に存在しない場合には、キーワードの長さだけ一挙にキーワードをシフトしながら照合処理を進める。

BM 法ではキーワードが長くなるとシフト量も平均的に大きくなることから、高速な文字列照合を行うことができる。またテキストストリング中の文字がキーワード中に現れる確率が小さいほど効率がよくなる。しかし、テキストストリング内の文字とキーワードの文字が同一の一文字で構成されている場合は文字照合一のシフト量は常に 1 文字なので速度的には単純なキーワード検索のアルゴリズムと大差がなくなってしまう。

現実的には、テキストストリング内の文字とキーワードの文字が同一の一文字で構成されることは稀であることから一般的には非常に高速な文字列照合アルゴリズムである。

3.8.3 AC(Aho-Corasick) 法

単純なキーワード検索のアルゴリズム、KMP 法、BM 法は単一のキーワードに対する文字列照合を行うアルゴリズムである。これに対して AC 法は複数のキーワードに対する文字列照合を行うアルゴリズムである。AC 法では、与えられたキーワード集合からマシン AC と呼ばれる一種の有限オートマトンを構成し、この有限オートマトンにテキストストリング入力として与えることにより文字列の照合を行う。

AC 法はテキストストリングをバックトラックすることなく 1 回走査するだけで、複数のキーワードを同時に検出することができる優れたアルゴリズムである。

3.9 索引を用いた全文検索

3.9.1 特徴ベクトル法

特徴ベクトル法は、特徴ベクトルと呼ばれる 2 次元情報を索引とし、検索処理を高速化する手法である。特徴ベクトルとは検索対象となる文書内にどのような索引語が出現したかという特徴を 0 と 1 からなるビット列で表現したものである。この索引語の特徴をビット列で表したものを単語シグネチャーと呼ぶ。また、検索対象の文書中から作成した複数の単語

3.9 索引を用いた全文検索

シグネチャーの論理和をとった、文書全体の特徴を表すビット列を文書シグネチャーと呼ぶ。

特徴ベクトル法を用いた情報検索では情報サービスの文書シグネチャーと検索単語の単語シグネチャーのビットベクトルを比較することでユーザの要求に合う情報サービスの検索を行う。

単語シグネチャーの作成には、事前処理として検索対象となる文書内から索引語となる単語の抽出を行う。次に索引語の文字情報をハッシュ関数により整数値に変換し、ビット長に応じて1,0のビット列からなる単語シグネチャーを生成する。ビット長は扱う文字の種類や文字数から決定する。文書シグネチャーの作成は単語シグネチャーの論理和から生成する。

特徴ベクトルはハッシュ関数を用いて、索引語をビット情報に変換することから他の手法に比べて索引のサイズが小さくなる。また一般的に索引語の構築には時間がかかってしまうが、特徴ベクトル法では各索引語のハッシュ値を計算し、順次、論理和演算を行うだけなので比較的短時間で索引を構築することができる。検索候補の絞り込みについてもビットごとの論理演算だけなので絞り込みは高速である。しかし、索引語から単語シグネチャーを生成することから特定の文字列の検索を行うことができない。また、特徴ベクトル法ではハッシュ関数を使うことから検索結果にユーザの要求を満たしているものと満たしていないものが含まれる場合がある。この場合は逐次検索を行う必要がある。

3.9.2 転置ファイル法

転置ファイル法は特徴ベクトル法の短所を補う全文検索手法である。転置ファイル法では各索引語が出現する文書内での出現位置情報を転置ファイルと呼ばれる表構造の索引に格納し、検索時にはこのファイルのみにアクセスして単語を検索する手法である。

転置ファイル法では索引内に全ての索引語の出現文書や出現位置情報を格納しているので、逐次処理による確認作業が必要なく、検索時間も高速となる。また、Nグラム索引を用いることにより、任意の文字列に対する検索が可能となる。しかし、索引ごとに位置情報を格納するので索引のサイズが大きくなる。また、転置ファイルでは索引語が追加、削除されると索引全体を作成し直す必要があるため情報サービスの更新が難しい。

3.9 索引を用いた全文検索

3.9.3 パトリシア・トライ法

特徴ベクトルや転置ファイル法が、単語や文字をベースに索引を形成する手法であったのに対して、パトリシア・トライ法では文書内の半無限文字列から索引を行う。パトリシア・トライでは構造的に2進木トライに属しており、2進木構造のノードの削除方法により完全2進木トライ、正規2進木トライ、パトリシア・トライの3種類の木構造に分類される。パトリシア・トライでは2進木トライの中で一本の枝しか持たないノードを全て削除する。このとき削除したノードの数を示すカウンタ、削除されたラベル値を蓄えたレコード情報へのポインタを各内部ノードが持つ必要がある。

パトリシア・トライは転置ファイル法と同様に登録文書内の全ての情報をパトリシア・トライに格納しているため、逐次検索による確認を行う必要がない。また登録文書内に登録文書内のすべての文字位置から始まる半無限文字列をパトリシア・トライに格納しているため、任意の文字列検索が可能である。さらにパトリシア・トライが木構造であることから情報サービスの追加や削除を容易に行うことが可能となる。しかし、パトリシア・トライでは特徴ベクトルの様なビット列ではなく半無限文字列を登録する必要があるため、索引の構築時間が遅くなる。

第 4 章

情報推薦システムの検討

4.1 状況のモデル化

4.1.1 大学における状況

ユーザの状況に応じて必要な情報サービスの提供を行うためには、ユーザの状況を把握し、その状況でどのような情報サービスが利用されているかを把握する必要がある。

大学において状況ができる場面は、大学生活の中で常に発生している。最適な情報サービスを提供するためには、大学内のユーザの状況を全ての把握する必要がある。しかし、刻一刻と変化する状況を全て把握し、的確に情報サービスの提供を行うことは難しい。また、大学には様々なユーザが存在することから状況は千差万別である。仮に全ての状況を把握できたとしてもサラウンディングキャンパス内で情報サービスが提供される機会は限られる。例えば講義資料であれば講義が行われているときは利用するが、休み時間などの移動中には必要ない。このことから情報サービスを利用する状況はある程度限られると考えられる。そこで、サラウンディングキャンパスにおける状況の把握は情報サービスが利用されるときに行う。情報サービスが必要な状況はコンテキスト内の要素が揃うことにより起きる。例えば、人工知能科目の講義であれば、“人工知能科目の履修生”，“AM9:00”，“A107 教室”というように要素が揃うことで、人工知能科目の履修生は AM9:00 から A107 教室で人工知能科目の講義を受講していることがわかり、人工知能科目の講義資料を提供する必要があると判断できる。このことから情報サービスが利用される状況を構成する要素を把握することで、ユーザの状況に応じた情報サービスの提供が可能になると考えられる。そこで、大学内で情報サービス

4.1 状況のモデル化

表 4.1 情報サービスが利用される状況の一例

場所	時間	ユーザ	周囲のユーザ	状況の予測
教室	9:00 ~ 18:00	教員		講義が開講されている
教室	9:00 ~ 18:00	教員	学生	講義が開講されている
教室	18:00 ~ 21:00	教員		研究室輪講が行われている
教員室		教員	学生	履修相談, 研究相談が行われている
図書館	8:00 ~ 19:00	学生		テスト勉強, 研究が行われている
研究室		学生	学生	研究室輪講, 研究, テスト勉強が行われている
研究室		学生		研究, テスト勉強が行われている

が利用される状況を構成する要素の一覧表を作成した。表 4.1 に一覧表を示す。

4.1.2 場所の持つ役割

作成した一覧表から場所の要素に着目し、場所が状況に与える影響について検討する。大学には教室、教員室、図書館、研究室、事務室などの場所が存在し、それぞれの場所は何らかの役割を持っている。教室であれば講義が行われる場所であり、研究室であれば、研究が行われる。場所の要素に時間やユーザが影響する場合もあるが、多くの場合、場所の持つ役割から状況を推測することができる。このことから、場所が決まれば場所の持つ役割から状況をある程度推測することができる。

4.1.3 時間の持つ役割

大学における時間の持つ役割は標準時刻と 1 限, 2 限, 3 限といった大学の時間がある。状況の推測において、時間の要素だけで状況を推測することは難しい。しかし、他の要素と組み合わせることによって状況を推測できる。例えば、教室という場所の要素だけでは、講義や輪講を行う

4.1 状況のモデル化

場としての役割も持っているため、状況を推測することが難しい。しかし、AM9:00 という時間が加わるだけで、1 限に教室で講義が行われているという状況を推測することができる。つまり、時間は場所やユーザの要素と組み合わせることによって状況を決定する役割を持っている。

4.1.4 ユーザの持つ役割

大学におけるユーザは教員、学生、職員がおり、それぞれのユーザは履修科目や担当科目、所属研究室などの役割を持っている。ユーザの要素から状況を推測することは難しいが、場所や時間、周囲のユーザの要素と組み合わせることによって状況を推測できる。教室で AM9:00 に講義が開講されることが予定されていた場合、ユーザがその時間に教室に向かうことによって状況を推測することができる。また、図書館ではユーザの履修科目や過去の貸し出し記録から本の推薦を行う。このことからユーザの要素を他の要素と組み合わせることによって状況の推測を行うことができる。

4.1.5 周囲のユーザの持つ役割

周囲のユーザ情報は自身の周辺にいるユーザが持つ役割である。周囲にいるユーザの状況への影響は、自身と周辺ユーザに共通の役割があり、共通の役割を持つ周囲のユーザに状況が起こる場合、自身も同様の状況が起こるという考えから状況の推測を行う。

4.1.6 状況を予測するためのデータモデルの作成

4.1.1 において、大学内で情報サービスが利用されるときに状況の要素一覧表から状況は各要素の組み合わせから成り立っていることがわかる。そこでユーザの状況に応じた情報サービスの推薦を行うために、ユーザの状況を把握するためのデータモデルの作成を行う。状況を決定するパラメータは以下ようになる。

- 場所
- 時間

4.2 情報サービスの推薦

- ユーザ
- 周囲のユーザ

4.2 情報サービスの推薦

4.2.1 ユーザプロフィールとコンテンツモデル

コンテキストにデータモデルを適用することで状況の把握を行い、把握した状況に対して情報サービスの推薦を行う。情報サービスの推薦は、データモデルにより体系化された要素をユーザプロフィールとし、検索対象となる情報サービスの持つ体系化された要素をコンテンツモデルとしたとき、それぞれを比較することにより情報サービスの推薦が行われる。対象となる情報サービスが持つ要素はあらかじめ体系化されたものが格納されている。

4.2.2 コンテンツ内容に基づくフィルタリングと協調フィルタリング

状況に応じて必要な情報サービスの提供を行うために、情報サービスはどのような状況で提供するか明確にしておき、その状況が起きると情報サービスの提供を行う。例えば人工知能科目講義の資料は人工知能科目講義に必要であることは明確であることから一意に情報サービスの提供を行う。しかし、人工知能科目のテスト勉強をしたい場合にも人工知能科目講義の資料は必要である。このとき、人工知能科目のテスト勉強を行う場所や時間は各ユーザが決めるため、情報サービスを提供する状況を一意に定めておくだけでは、状況に応じた情報サービスの提供をおこなうことは難しい。状況に対する情報サービスが一意に定まらない場合にも情報サービスの提供を行う必要がある。

そこで、状況に対して一意に情報サービスが定まる場合と定まらない場合の両方に対して情報サービスの推薦を行う。状況に対して情報サービスが一意に定まっている場合の情報サービスの提供にはコンテンツ内容に基づくフィルタリングのルールベース方式を適用する。データモデルにより体系化されたユーザの状況を表すユーザプロフィールと情報サービスのコンテンツモデルの一致を計ることで情報サービスの提供を行う。一方、状況に対して

4.3 情報推薦システムの提案

必要な情報サービスが一意に定まっていない場合には協調フィルタリングのユーザベース方式を適用する。協調フィルタリングでのユーザベース方式ではユーザの状況を表すユーザプロフィールと情報サービスのコンテンツモデルの類似度を計り、類似度が高い場合に情報サービスの提供を行う。

4.2.3 類似度の計算

情報サービスを推薦するためにはユーザプロフィールとコンテンツモデルの類似度を計り、情報サービスを推薦する必要がある。類似度は文字列の比較による類似度の計り方と文字列を数値やベクトルに置き換えて比較を行う方法がある。サラウンディングキャンパスにおける類似度の計算は、ユーザプロフィールに対してコンテンツモデルが要素を持っているか、持っている場合には類似度はどのくらいかを計る必要がある。ユーザプロフィールとコンテンツモデルは類似度計算の方法における索引に相当することから索引による検索を行う、特徴ベクトル法、転置ファイル法、パトリシア・トライ法の適用について検討する。転置ファイル法、パトリシア・トライ法は文書に対して索引語を用いて高速な検索を行う方法である。サラウンディングキャンパスにおける類似度の計算は検索対象のコンテンツモデルが文書でないため、高速な検索を行う必要はない。このことから特徴ベクトル法を用いて類似度の計算を行う。

4.3 情報推薦システムの提案

サラウンディングキャンパスで扱う情報サービスは状況に対して一意に情報サービスが定まる場合と決定まらない場合がある。サラウンディングキャンパスを実現するためにそれぞれに対して情報サービスの推薦を行う。

状況に応じて情報サービスを推薦するためには、ユーザの状況を把握する必要がある。そのため、コンテキストに対してデータモデルを適用することにより、ユーザプロフィールの作成を行う。また、あらかじめ用意されている情報サービスの持つ要素をコンテンツモデル

4.3 情報推薦システムの提案

とし、類似度計算を行うことでユーザの要求に対して情報サービスの提供を行う。類似度の計算は、ユーザの要求が持つ要素が情報サービスが持っているかを判断する必要があることから特徴ベクトル法を適用する

本研究ではサラウンディングキャンパスの機能の一部として、情報フィルタリングのコンテンツ内容に基づくフィルタリング技術と協調フィルタリング技術を用いることで、ユーザの状況に応じて必要な情報サービスの提供を行う、情報サービス推薦システムの構築を行う。

第 5 章

情報サービス推薦システムの構築

5.1 情報サービス推薦システムの設計

サラウンディングキャンパスを実現するために情報サービス推薦システムの構築を行う。本システムでは第 4 章で作成したデータモデルに基づいて作成した、ユーザプロフィールとコンテンツモデルに情報フィルタリングを適用することで情報サービスの推薦を行う。図 5.1 に情報サービス推薦システム全体の構成を示す。

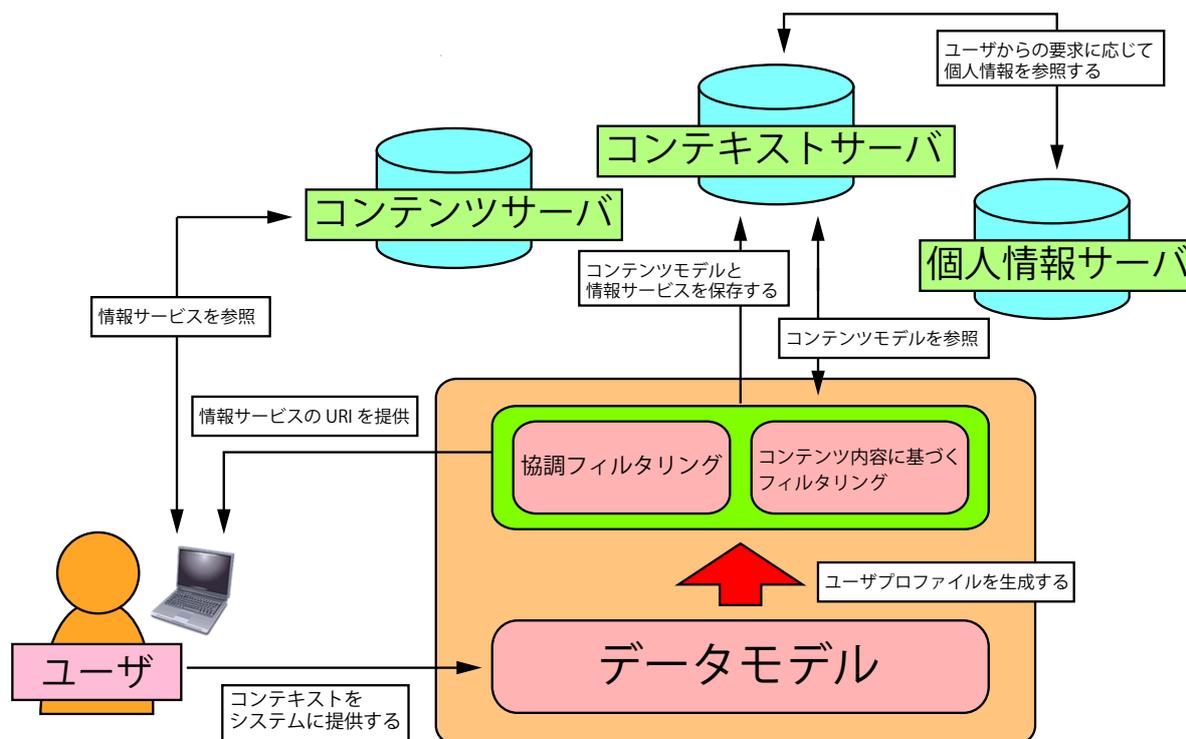


図 5.1 情報サービス推薦システム全体の構成図

5.1 情報サービス推薦システムの設計

本システムは大きく 3 つの機能を持つ。

- 状況を把握する機能
- 情報サービスの推薦を行う機能
- 情報サービスの推薦と履歴の格納を行う機能

5.1.1 状況を把握する機能

状況を把握する機能では、コンテキストから体系化されたユーザプロファイルの作成を行う。コンテキストの取得には過去の研究より、小型無線センサデバイスを用いる。取得したコンテキストに対してデータモデルを適用することによりユーザプロファイルを生成する。

5.1.2 情報サービスの推薦を行う機能

情報サービスの推薦を行う機能では、状況を把握する機能より取得したユーザプロファイルと情報サービスの持つ要素を体系化したコンテンツモデルの類似度を計ることにより、情報サービスの推薦を行う。本システムにおけるユーザプロファイルとコンテンツモデルの類似度は特徴ベクトル法を用いる。表 5.1, 5.2 にユーザプロファイルの構成とコンテンツモデルの構成について示す。

5.1 情報サービス推薦システムの設計

表 5.1 ユーザプロフィールの構成

ユーザプロフィールの持つ情報	説明
場所の要素	状況においてユーザがいる場所の情報
時間の要素	状況における時間
ユーザの要素	ユーザに関する情報
周囲のユーザの要素	状況においてユーザの周囲にいるユーザの情報

表 5.2 コンテンツモデルの構成

コンテンツモデルの持つ情報	説明
情報サービス	各要素によって表される状況に対する情報サービス
場所の要素	状況においてユーザがいる場所の情報
時間の要素	状況における時間
ユーザの要素	ユーザに関する情報
周囲のユーザの要素	状況においてユーザの周囲にいるユーザの情報

特徴ベクトル法は検索質問と検索対象の特徴を 1, 0 のビット列で表現し、類似度を計る方法である。ビット列化された単語を単語シグネチャーと呼び、単語シグネチャー同士の論理和より求められるものを文書シグネチャーと呼ぶ。本システムではユーザプロフィールの持つそれぞれの要素を単語シグネチャーとし、単語シグネチャー同士の論理和をとった文書シグネチャーを状況として扱う。情報サービスの推薦における単語シグネチャーの作成には、単語 w に対して、ハッシュ関数 $H(w)$ を用いる。式 (5.1) に本システムで用いたハッシュ関数の式を示す。ハッシュ関数では、ビット長を F 、その内 m 個のビットが 1、残りが 0 のビット列からなる単語シグネチャーを生成する。

$$H(x) = (\text{Sum}(w) \bmod F) + 1 \quad (5.1)$$

この関数により、得られたハッシュ値と同じ位置に 1 を立てたビット列を単語シグネチャー

5.2 システムの概要

とする。文書シグネチャーは単語シグネチャー同士の論理和をとることで求める。

ユーザプロファイルの文書シグネチャーとコンテンツモデルの文書シグネチャーの類似度を計り、推薦する情報サービスを決定する。

5.1.3 情報サービスの推薦と履歴の格納を行う機能

情報サービスの推薦を行う機能より推薦された情報サービスをユーザに提供すると共に、推薦した情報サービスと情報サービスの要素を履歴として履歴サーバに格納する。

5.2 システムの概要

情報サービス推薦システムは MySQL, Java, Apache を用いることにより構築される。本システムの動作について述べる。

サラウンディングキャンパス下で、状況が発生した場合、過去の研究より、小型無線センサデバイスを用いてコンテキストの取得を行う。取得したコンテキストの要素はそのままでは意味を持たないため、データモデルを適用し、体系化を行う。体系化されたコンテキストの要素をユーザプロファイルとし、情報サービスの推薦を行う。検索対象は、あらかじめ情報サービスと情報サービスの持つ要素を格納しておき、ユーザから要求があれば、コンテンツモデルとしてユーザプロファイルとの照合を行う。ユーザプロファイルとコンテンツモデルの照合にはコンテンツ内容に基づくフィルタリングと協調フィルタリングを適用することで、状況に対して一意に決まる情報サービスの提供と一意に決まらない場合の情報サービスの推薦を行う。推薦された情報サービスはユーザに提供されると共に履歴サーバに格納され、次の情報サービスの推薦時に参照される。図 5.2 にシステムの概要図を示す。

5.2 システムの概要

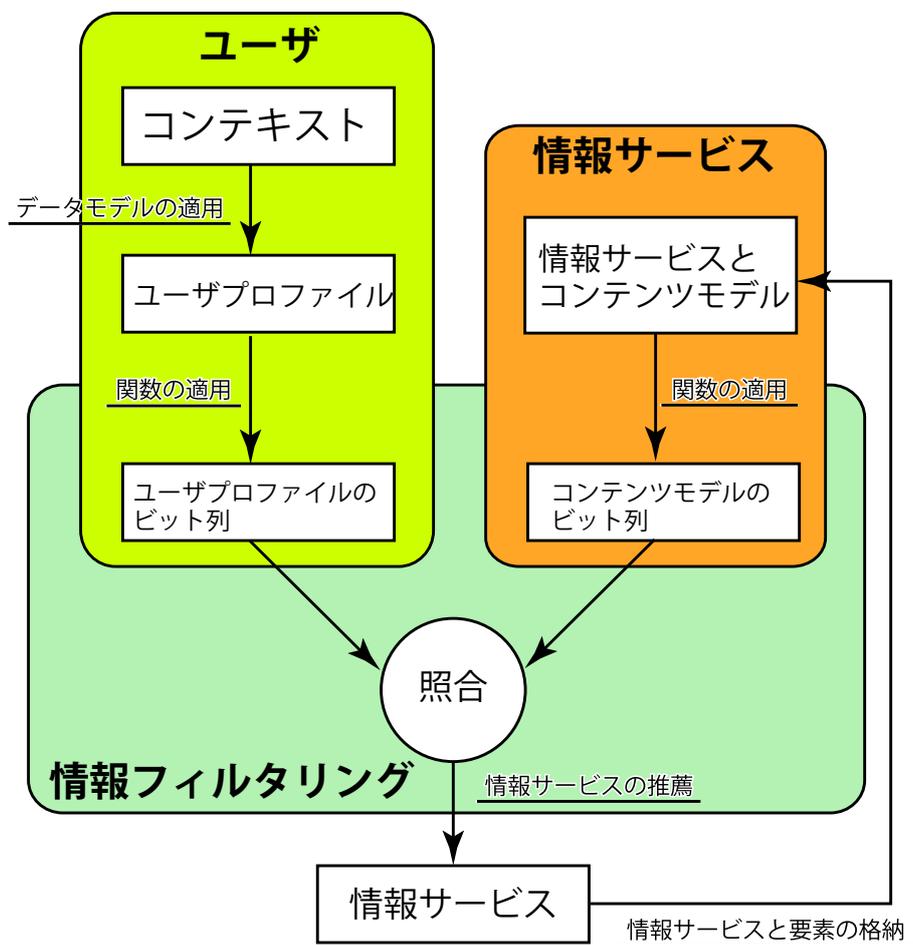


図 5.2 システムの概要図

第 6 章

評価・考察

6.1 評価の概要

サラウンディングキャンパスを実現するための情報サービス推薦システムの有用性を確かめるために、情報サービス推薦システムのコンテンツ内容に基づくフィルタリングと協調フィルタリングに対して、サラウンディングキャンパスで想定される状況をモデル化したテストデータの適用を行う。評価に用いるテストデータにはデータモデルを作成する際に用いた、情報サービスと状況の組み合わせの一覧表から状況の異なるものを適用する。表 6.1 にテストデータの一部を示す。これによりサラウンディングキャンパス下における全ての状況に対して、必要な情報サービスを提供できるか確認する。ユーザ側からすれば状況に対して、一意に決まる情報サービスの提供が行われるか、一意に決まらない場合にはユーザにとって必要と思われる情報サービスの推薦が行われるかについて評価を行う。

6.2 評価と考察

表 6.1 テストデータの一部

場所	時間	ユーザ	周囲のユーザ	予測される情報サービス
教室	9:00 ~ 18:00	教員		講義資料, 講義スライド
教室	9:00 ~ 18:00	教員	学生	講義資料, 講義スライド
教室	18:00 ~ 21:00	教員		輪講資料, 研究資料
教員室		教員	学生	履修資料, 研究資料
教員室		教員	教員	学科関連資料
図書館	8:00 ~ 19:00	学生		講義資料, 研究資料, オススメの本, 講義関連図書, 研究関連図書
研究室		学生	学生	輪講資料, 講義資料, 研究資料
研究室		学生	研究	資料, 輪講資料

6.2 評価と考察

システムに対して、テストデータを適用した。その結果、一意に決まる情報サービス、一意に決まらない情報サービスとも情報サービスの推薦が行われることが確認できた。しかし、一意に決まらない情報サービスの推薦では情報サービスの推薦を行う必要のない情報サービスが提供される場合があった。この結果に対して、システム側から考えた場合、一意に決まる情報サービスではユーザプロファイルとコンテンツモデルの単純な一致をとればよいことから、ひとつでも要素が違えば提供を行わないようにしたことからテストデータの予想通りの結果を得ることができた。一意に決まらない情報サービスの推薦では状況に対して必要と思われる情報サービスを推薦する必要がある。情報推薦が上手くいかなかった結果では本来情報サービスが必要のない場所で情報サービスの推薦が行われていた。これはユーザプロファイルとコンテンツモデルの要素がいくつ一致するかを計り、一致する数が多いものの推薦を行うため、要素同士の影響が考慮されなかったと考えられる。

第7章

おわりに

本研究ではサラウンディングキャンパスのシステムの一部として、状況を推測するためのデータモデルの作成を行い、モデル化したコンテキストに情報フィルタリングを用いることでユーザの状況における必要な情報サービスの提供を行うシステムを構築した。データモデルを作成する際に用いた、大学における情報サービスと情報サービスを利用するときの状況の組み合わせ一覧を基に状況を表すデータをシステムに適用し、情報サービスが推薦されるか確認した。出力結果から一意に決まらない情報サービスを出力する際に状況に適していない情報サービスが提供されることがわかった。考えられる要因としては本システムでは状況の要素が状況や他の要素にどの程度の影響力を持つかを考慮していないためだと考えられる。今後の展開として、状況に対する要素の影響力をシステムに反映させ、情報サービスの推薦精度を上げると共に、サラウンディングキャンパスのインターフェース部や情報サービスの格納方法について検討し、実際にサラウンディングキャンパスの構築を行った上で、必要な情報サービスが提供されるかを評価する必要がある。

謝辞

本研究を進めるにあたり、ご指導を賜りました主指導教員である情報システム工学科、妻鳥貴彦講師に心より深く感謝いたします。また、副査としてご助言をいただきました情報システム工学コース、福本昌弘教授及び吉田真一講師に心より感謝いたします。そして、本研究室修士 2 回生、畠山博和氏、藤原健太郎氏、同研究室修士 1 回生、清水雅也氏、森拓也氏には本研究において有益なご指摘、ご助言を頂いたことを心から感謝いたします。さらに、同研究室 4 回生、中澤大樹氏、濱野純平氏、細川恭平氏、前田晃宏氏、松井勇貴氏、松本直樹氏同研究室 3 回生、香西佑紀氏、池上裕貴氏、岡崎雄太氏、金子淳史氏、小松原健氏、中山陽介には本研究に対し、ご助言を賜り心から感謝いたします。

参考文献

- [1] 金西計英, 松浦健二, 光原弘幸, 緒方広明, 三好康夫, 森川富昭, 矢野米雄, “徳島大学における u-Learning の実践”, 北海道大学 現代 GP フォーラム, (2006).
- [2] 緒方広明, 矢野米雄, “CLUE:語学学習を対象としたユビキタスラーニング環境の試作と実験”, 情報処理学会論文誌, Vol.37, No.3, (1996).
- [3] 福本昌弘, 岩田誠, 濱村昌則, 島村和典, “サラウンディング?コンピューティング技術の研究開発”, 情報通信研究機構季報, Vol.51, pp.145-151, (2005).
- [4] 福本昌弘, 岩田誠, 酒居敬一, 吉田真一, 妻鳥貴彦, 濱村昌則, 島村和典, “サラウンディング・コンピューティング技術による情報転送?再現システム”, 電子情報通信学会技術研究報告, IA2007-52, Vol.107, No.449, pp.69-72(Jan.2008).
- [5] JGN2 四国リサーチセンター, “サラウンディング・コンピューティング技術の研究開発”, JGN2+AKARI シンポジウム 2008, (2008).
- [6] 清水雅也, 山崎雄大, 福田将行, “サラウンディングキャンパスの構築”, 教育システム情報学会, Vol.24, No.2, pp.88-94(July.2009).
- [7] 北研二, 津田和彦, 獅々堀正幹, “ 情報検索アルゴリズム ”, 共立出版株式会社. 2002.
- [8] 根岸裕也, 河口信夫, “ユビキタスコンピューティングにおけるコンテキストセンシングとデータ処理”, 人工知能学会誌, Vol.23, No.5, pp.597-603(2008).
- [9] 上野真臣, “e ラーニングにおけるデータマイニング”, 日本教育工学会論文誌, Vol.31, No.3, pp.271-283(2007).
- [10] 松辻智之, 王 軼群, 土方嘉徳, 西田正吾, “目的を考慮した情報フィルタリング手法に関する基礎調査”, 電子情報通信学会技術研究報告, DE2007-25, Vol.107, No.131, pp.365-372(2004).
- [11] 土方嘉徳, “ 情報推薦・情報フィルタリングのためのユーザプロファイリング技術 ”, 人工知能学会, Vol.19, No.3, pp.365-372(2004).

参考文献

- [12] 土方嘉徳, “嗜好抽出と情報推薦技術”, 情報処理学会, Vol.48, No.9, pp.957-965(Sep.2007).
- [13] 石川徹也, 宇田隆幸, “情報フィルタリングの利用システム”, 情報の科学と技術, Vol.56, No.10, pp.458-463(2006).
- [14] 奥健太, 中島伸介, 宮崎純, 植村俊亮, “状況依存型ユーザ嗜好モデリングに基づく Context-Aware 情報推薦システム”, 情報処理学会論文誌: データベース, Vol.48, No.SIG11(TOD 34), pp.162-176(July.2007).