M-004

能動的情報資源による異種コンテンツの自律的連携手法
Autonomous Cooperation Method of Heterogeneous Contents based on Active Information Resource

浅沼 亜紀[†] 魏 文鵬[†] 伊藤 大視[†] 高橋 秀幸[†,‡] 笹井 一人[†,‡] 北形 元[†,‡] 木下 哲男[†,‡]
Aki Asanuma Wenpeng Wei Taishi Ito Hideyuki Takahashi Kazuto Sasai
Gen Kitagata Tetsuo Kinoshita

1. はじめに

インターネット，スマートフォンの普及に伴い，コンテンツのデジタル化，ネットワークのブロードバンド化が進み，映像や音楽などのマルチメディアコンテンツがネット上に数多く公開・共有されるようになった．今後，膨大な数のコンテンツから目的とする情報を素早く，容易に取得する事が重要となる．例えば，災害などの緊急時には，リアルタイムな情報や現在地に特化した情報など，利用者の要求に沿った情報を迅速に抽出する事が求められ，これに適した機能やツールの開発が課題となっている．

本研究では能動的情報資源(Active Information Resource: AIR)[1]の枠組みに基づき，多様な異種コンテンツを能動的に提供するための情報資源の自律的連携手法を提案する．具体的には，分散環境上で生成される表現形式の異なる異種コンテンツの構造化，AIR 間の協調・連携による情報の関連付けを行い，状況や要求の変化に応じた異種コンテンツの提供を実現する．本稿では，AIR 化された異種コンテンツの集積・管理及び情報資源の関連付けや情報提供を行う際の処理などを含む連携手法について議論する．

2. 関連研究と技術的課題

コンテンツの提供システムに関する様々な研究が行われている．例えばコンテンツを自動推薦するテレビ[2]がある．これは視聴中のコンテンツに関連したコンテンツを，メタデータを利用して自動で検索し，利用者に提示するシステムである．選択したコンテンツに付与されたメタデータを検索キーとして関連性を求め，利用者の検索操作を軽減する事が可能である．一方，学術情報(論文・史料など)を能動的情報資源(AIR)として構成し，検索，統合，分析などの作業を AIR に一部代行させ，ユーザによる学術情報の活用を包括的に支援することを目指した知識強化型の学術情報検索支援システム[3]がある．

従来のシステムでは内容に関するメタデータを全てデータベースで一元管理している．しかし要求の変化，例えばコンテンツそのものの追加や削除，メタデータの更新などといった変化が発生した場合にも柔軟に対応できる事が求められる．また，実験対象のコンテンツは全て番組（映像）やテキストコンテンツであり，異種のコンテンツを扱う場合について検討する必要がある．

異種コンテンツを対象にした場合，以下の 2 点が課題として考えられる．

（１）表現形式の異なる異種コンテンツの対応付け

テキストだけでなく，動画・音声・画像など多種なメディア情報を柔軟に取り扱う事が要求される．これらの情報に含まれるメタデータは，各メディアによって表現形式が異なるため，そのままでは比較する事ができない．そこでメディアタイプの異なる情報資源にも対応できる仕組みが必要となる．

（２）状況変化やユーザの要求変化への対応

例えば災害時における情報は偶発的であり，状況や要求が頻繁に変化する事が予測される．状況とはコンテンツ自身の追加・削除・更新，またユーザの使用状況を指す．要求とは利用者毎に異なるその時々の条件付きの要求である．こうした変化にも柔軟に対処できるシステムアーキテクチャや動作環境の仕組みが必要である．

上記 2 点の課題を解決するため，本研究では能動的情報資源(AIR)の枠組みに基づき，情報資源を自律的に連携・協調させる事で，情報資源自身が能動的に相互の関連性の判断や，追加・更新などの変動に対応する事が可能な手法を提案する．具体的には表現形式の異なるメタデータへの対応と，動的な変化にも対応可能な情報資源同士の連携・関連付けについて述べる．

3. 異種コンテンツの自律的連携手法

3.1 提案手法の概要

能動的情報資源(AIR)とは，分散環境上にある情報資源の構造を強化・拡張する事で，利用者の要求に対し，情報資源自身が能動的に反応し，より高度な活用を図る機構である．AIR による異種コンテンツの提供例を図 1 に示す．

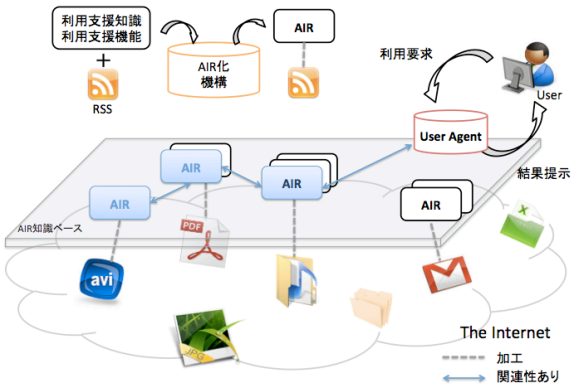

図 1 AIR による異種コンテンツ提供の例

利用支援知識はメタ知識を属性と属性値の組の集合として表現する．各属性は種々の目的・用途に応じて定義されるものとする．2 章で述べた(1)の技術的課題を解決するため，知識の一部として組み込まれる情報資源のメタ情報の属性を Dublin Core[4]等のメタデータセットを基に異種コンテンツのメタ情報を構造化する．具体的には共通の属性項目を抽出し，メタデータの再編を行う．またコンテンツの種類や性質に応じて異なる所定の属性セットを導入する．

利用支援機能は[3]を参考に，AIR 相互間で交換される基本的なメッセージの処理機能，AIR の協調連携処理の他，

[†]東北大学大学院情報科学研究科
[‡]東北大学電気通信研究所

Copyright © 2012 by
The Instiute of Electronics, Information and Communication Engineers and
Information Processing Society of Japan All rights reserved.

関連度の判定機能を付与する．動的な変化にも対応可能な情報資源同士の連携・関連付けについては次節(ステージ 2)で述べる．

3.2 能動的情報資源による異種コンテンツの連携

連携処理のプロセスは次の 2 つのステージから成る．

（ステージ１）異種コンテンツの AIR 化

表現形式に対応するメタデータセットを考慮した利用支援知識形式に基づき，該当コンテンツへのリンク情報を含む利用支援知識を記述する．また AIR の基本動作を行うための利用支援知識と利用支援機能を付与する．これにより AIR 相互間で連携・協調処理を行い，利用者の要求を AIR 側が自律的に実行し，利用者への負担を軽減する事が可能になる．

（ステージ２）AIR の協調動作による連携処理

情報資源の相互関係と関連付けを行う．具体的な処理として AIR が新規に追加，削除された場合，及び利用者の要求が変化した場合について述べる．

新規に AIR が追加されると(図 2(1))，新規 AIR は既存の AIR と自律的にメッセージの交換を行い，相互の類似性や関連性を判定する．この際，所定の閾値を超える関連度を有する AIR が発見されれば，その情報を双方の AIR の利用支援知識として獲得し，関連付けが行われる(図 2(3)～(4))．このように新規 AIR が追加される度に関連付けを行う事で，漸次関連性が強い AIR 群が形成される．

AIR を削除する場合，その削除対象 AIR と関連性のある全 AIR に対し，自身が削除対象である事を通知する．通知を受信した AIR は，自身の利用支援知識からその削除対象 AIR の情報を削除する．

利用者の要求が変化した場合，例えば，リアルタイムな情報に関する要求の場合は，リアルタイム性重視の関連度算出式を付加したメッセージを User Ag から該当 AIR に送信する．そのメッセージを受け取った該当 AIR は既に関連付けられている全 AIR に対し，再度関連性を求めるメッセージを送信する．そこで所定の閾値以上の関連度を有する AIR が発見されれば，その情報を双方の AIR の利用支援知識として獲得し，関連付けを行う．関連付け終了後，関連性の高い AIR が全て User Ag に回収され，利用者に提示される．

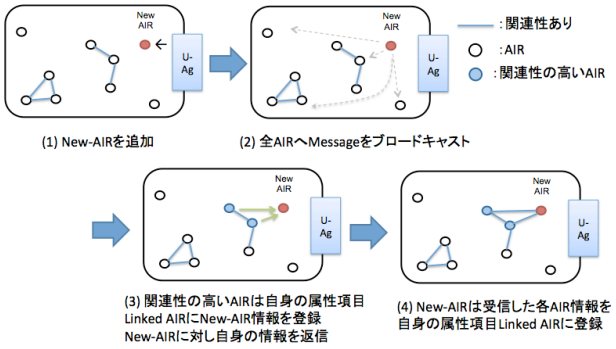

図 2 新規 AIR の連携プロセス

二つのコンテンツ間の関連度算出方法について，分類した属性項目毎に各要素を比較し類似度 S を求め，その属性項目における重み要因 α(重要度)を掛けあわせたものを加算する事で関連度 R を求める．重みは要求に応じて変化するものとする．

コンテンツ C_i と C_j の類似度 S の計算は各属性項目における集合の類似度を用いる．集合の類似度の計算方法は様々存在するが，今回は閾値付き Simpson 係数を採用した．

4. 設計と実装

提案における AIR をエージェントとして設計し予備実験を行った．エージェントとして ADIPS/DASH フレームワーク[5]を用いた．今回の対象コンテンツはニュースサイト記事 10 個とし，メタ情報の属性項目は RSS(RDF site summary)を参考に Dublin Core 規定の 15 項目で定義している．また，今回関連度の算出は単純に共通属性項目の数の和としている．

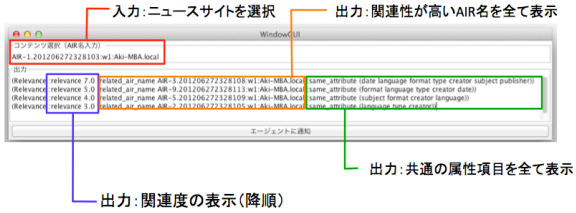

図 3 プロトタイプシステムの動作画面

実験のシナリオは，ユーザがあるニュースサイトをひとつ選択した時，そのニュースサイトと共通した属性項目を持つ他サイトで，かつ関連度が一定値より高いもの(今回は Relevance 3.0 以上)を提示するというものである．

入力画面にニュースサイト名（対応する AIR 名）を入力すると，そのニュースサイトと関連性の高い他のニュースサイト(AIR)の情報が全て出力される事を確認した(図 3)．この際，出力される情報とはニュースサイトに対応する AIR 名，共通した属性項目名，関連度の値であり，関連度の降順に提示される．

5. おわりに

本稿では，多種多様なコンテンツの探索や利活用の負担を軽減するため，能動的情報資源 AIR による異種コンテンツの自律的連携手法を提案した．今後，関連度の計算方法の変更やコンテンツの数を増やし，より複雑な処理への対応を検討する．

謝辞

本研究の一部は，総務省平成 23 年度第 3 次補正予算「情報通信技術の研究開発」委託課題「災害情報を迅速に伝達するための放送・通信連携基盤技術の研究開発」の援助を受けて実施した．

参考文献

1) 木下哲男：分散情報資源活用の一手法：能動的 情報資源の設計，信学技報，A199-45, pp. 13-19 (1999)
2) NHK 技研 R&D/No.121/2010.5
http://www.nhk.or.jp/strl/publica/rd/rd121/pdf/P16-P25.pdf
3) Baoning Li, Tetsuo Kinoshita, “Active Support for Using Academic Information Resource in Distributed Environment”, Int. J. Computer Science and Network Security, Vol.7, No.6, pp.69-73 (2007.6)
4) Dublin Core
http://dublincore.org/
5) ADIPS/DASH フレームワーク
http://www.agent-town.com/dash/index.html

Copyright © 2012 by
The Instiute of Electronics, Information and Communication Engineers and
Information Processing Society of Japan All rights reserved.