

平成 22 年 4 月 11 日現在

研究種目：基盤研究 (C)
 研究期間：2007～2009
 課題番号：19500081
 研究課題名 (和文) 大規模ネットワーク内情報源の効率的なモニタリングに関する検討
 研究課題名 (英文) A Study on Efficient Information Sources Monitoring Method in Large-scale Network
 研究代表者
 菅原 真司 (SUGAWARA SHINJI)
 名古屋工業大学・大学院工学研究科・准教授
 研究者番号：90313410

研究成果の概要 (和文)：

本研究では、ネットワーク上の独立の情報源により提供されている情報を継続的に取得し、常に効率的に最新の状態に維持する手法を求めている。特にインターネット上の多数の情報源では、提供情報の更新頻度がそれぞれ異なる点に着目し、効率的な探索スケジュールを求め、計算機シミュレーションにより提案手法の有効性を確認した。さらに、提案手法を実装したシステムのプロトタイプを設計・製作し、現実のインターネット上での有効性も確認した。

研究成果の概要 (英文)：

In this study, we investigated an efficient method of information retrieval from the information sources scattered on a large-scale network like the Internet. Focusing on the information update time-cycle which varies with each information source, the method enables to make an effective schedule of information sources' monitoring so as to keep the retrieved information freshly updated. Furthermore, we confirmed the effectiveness of the method by using computer simulations and prototype system implementations.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2007 年度	1,300,000	390,000	1,690,000
2008 年度	1,100,000	330,000	1,430,000
2009 年度	1,100,000	330,000	1,430,000
年度			
年度			
総計	3,500,000	1,050,000	4,550,000

研究分野：総合領域

科研費の分科・細目：情報学・メディア情報学・データベース

キーワード：情報検索

1. 研究開始当初の背景

近年、インターネットに代表される大規模ネットワーク上に発信される情報は増加と多様化の一途を辿り、現在これらを利用する

ことは、ごくありふれた光景となっている。特殊な種類の情報でなければ、あるユーザの必要とする情報は他のユーザにより既にネットワーク上に提供されていることが多く、検索エンジンを用いることでユーザの利便

性は飛躍的に向上している。しかしその反面、稀少な情報や、高い精度で情報を発見することは、逆に膨大な情報量に紛れて困難となることがある。また、それらの情報内容は各情報源の管理者により、それぞれ独立かつ自由に更新、削除され、あるいは新しく出現する。そのため、ネットワーク上に提供される情報を纏めてひとつの巨大なデータベースとして捉えようとした場合、通常のデータベースとは異なる困難が生じており、通常の検索エンジン等を用いても完全には対処できない。このようなデータベースを申請者は、超分散データベースと呼び、その中に存在する情報の効率的な発見と取得の手法について研究を継続してきた。インターネットにおけるWWWは、情報発信と取得の機能を有し、システムとして明らかに超分散データベースに該当する。

超分散データベースではそれを構成する各情報源で情報の更新が頻繁に生じるが、その更新の発生を正確に把握するのは現在のWWWでは困難である。しかし、ネットワーク上に発信される情報を収集し、それらを常に最新の状態に維持したいと考えるユーザが多数存在することは、ニュース配信、オークション、株式情報、電子商取引などの現実のサービスが成立していることから明らかである。

2. 研究の目的

本研究では、ネットワーク上の独立の情報源により提供されている情報を継続的に取得し、常に効率的に最新の状態に維持する手法を求め、また、インターネット上での現実のサービスを探索対象として仮定し、効率的な探索スケジュールを求め、監視対象となる情報が複数ある場合も考慮すると一般にそれらの情報を提供するサーバの更新頻度は異なるため、サーバ側の特徴を考慮した問合せおよび更新情報の取得が必要となる。

3. 研究の方法

本研究ではこれらの事項について問題を定式化し解法の提案と評価を行う。評価では計算機シミュレーションを用い、提案方式の有効性を確認する。また、情報源監視システムのプロトタイプを設計・製作し、現実のインターネット上の情報源を一定の期間、提案方式に従って監視し、その効果を調査する。

4. 研究成果

4.1 ネットワーク内情報源モニタリングのモデルと問題の定式化

本研究では、ネットワーク内に存在する複数の情報源からユーザが情報（コンテンツ）を取得する際に、ある種の満足度を得ると考える。これは、そのユーザから見た当該コンテンツの重要度、公開されてからの時間経過による陳腐化の度合い等により異なるが、一般に時間経過とともに減少すると考えられる。一方、ユーザが多数の情報源を常に監視するためには、ネットワーク利用に要するコスト、監視の処理を行うための計算機への負荷など、監視回数に応じたコストが必要となる。

このようなモデル上で、効率的なモニタリングを行うことが問題となるが、本研究ではこれを以下のように考える。

問題：モニタリングの期間を通しての、ユーザの満足度の総和から監視のためのコストを差し引いたものをユーザの利得と定義し、これを最大化する手法を求め。

4.2 提案モニタリング手法

効率的なモニタリングを行うためには、コンテンツの更新がある際には監視を行い、ない場合には監視を行わないようにすることが理想であるが、情報源のコンテンツの更新は、各情報源のオーナーが自由に他と独立に行うため、一般に予め更新スケジュールを把握することは困難である。しかし、多くの場合、概ね一日に一度、三日に一度、週に一度、月に一度などのように、情報の更新に大まかな周期が存在するケースが多い。よって、更新周期の比較的短い情報源には頻繁に、長い情報源にはそれなりに監視周期をシフトすることが有効である。また、更新周期も変動することがあるため、これに追従する仕組みも必要である。

そこで本研究では、問題解決の鍵は、更新間隔のより正確な把握と、それを動的に監視間隔に反映させる監視スケジュールリングにあると考え、各情報源に対して、現在の監視間隔で得られたユーザ満足度と、監視間隔を延長または短縮する場合に得られるユーザ満足度を推定した値とを比較し、より高いユーザ満足度を得られる監視間隔を動的に選択してモニタリングを継続する手法を提案した。

このアルゴリズムを、本質的な部分を中心に簡潔に表記すると次のようになる。

<アルゴリズム>

- ①監視回数 M' を 0 に初期化し、初期監視間隔決定プロセス（詳細は省略）を用いて暫定的な監視間隔を決定する。
- ② M 回の監視を実行する。

- ③更新の検出回数より，監視間隔を決定する。
 (a) $M' > M$ の条件で M' 回の監視において，
 毎回更新を検出した場合，または1回も
 更新を検出なかった場合

①に戻る

- (b) 過去 M 回の監視で毎回更新を検出した
 場合

M 回の監視で得られたユーザ満足度と，
 同期間において監視間隔を短縮した場
 合に得られると推定されるユーザ満足
 度を比較し，監視間隔を短縮した方が得
 られるユーザ満足度が大きいならば監視
 間隔を短縮し，それ以外の場合は監視
 間隔を変更しない。

$M' \leftarrow M' + M$ として②に戻る。

- (c) それ以外の場合

M 回の監視で得られたユーザ満足度と，
 同期間において監視間隔を延長した場
 合に得られると推定されるユーザ満足
 度を比較し，監視間隔を延長した方が得
 られるユーザ満足度が大きいならば監視
 間隔を延長し，それ以外の場合は監視
 間隔を変更しない。

$M' \leftarrow M' + M$ として②に戻る。

<アルゴリズム 終>

4.3 評価結果

提案手法の有効性を明らかにするために，
 本研究では，計算機シミュレーション，およ
 び，プロトタイプ作成による性能データの
 収集を行った

4.3.1 計算機シミュレーションを用いた検 証例

下記の図1に示すように，提案方式は監視
 コストの重みを表す α の値が極端に大きい
 場合と小さい場合を除くと，監視時間間隔を
 ランダム，または一定にした監視と比較して，
 ユーザの利得が最も大きくできることが明
 らかになった。なお，最適解とは理論的なユ
 ーザ利得の最大値で，どのような監視スケ
 ジュールを用いてもこの値を越えることが
 できない限界値であるが，通常そのような理
 想的な監視は不可能である。詳細なシミュ
 レーションに関する条件の説明は，割愛した。

4.3.2 プロトタイプ作成による評価例

次に，図2の設計概要図に示すように構成
 され，提案アルゴリズムを実装した情報源モ
 ニタリングシステムのプロトタイプを製作
 し，現実のインターネット上にある複数の情
 報源を監視させた場合の結果の例を説明す
 る。

図3に示すように，現実のインターネット

上の情報源を監視する場合でも，計算機シ
 ミュレーションで評価を行った場合とほぼ同
 様の傾向の結果を得た。提案手法はやはり，
 監視コストの重みが極端に大きい場合と小
 さい場合を除き，一定時間間隔で監視する
 方法よりもユーザの利得を大きくすることが
 できた。

以上のことから，提案した監視方式は十分
 有効であることが明らかとなった。

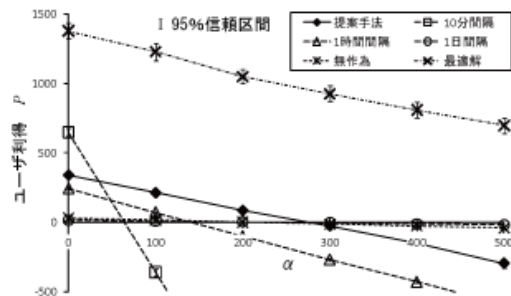


図1 計算機シミュレーション結果の例

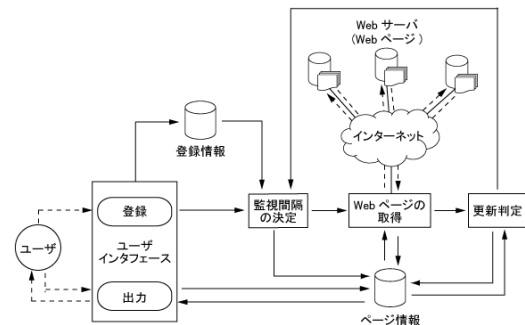


図2 情報源モニタリングシステム
 設計概要図

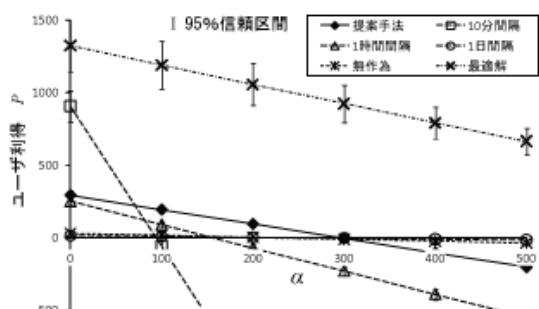


図3 プロトタイプを用いた評価例

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者に
 は下線)

[雑誌論文] (計 0 件)

[学会発表] (計 23 件)

[1] 世一 亮太, 菅原真司, 石橋豊, "ハイブ

- リッド型 Peer-to-Peer ネットワーク上のデータ共有における複製の一貫性維持に関する検討,” 電子情報通信学会総合大会, 2010年3月19日, 仙台・東北大学.
- [2] 白木夢人, 菅原真司, 石橋豊, “複数のモバイルエージェントを用いたネットワーク内情報探索に関する一検討,” 電子情報通信学会総合大会, 2010年3月17日, 仙台・東北大学.
- [3] 中村浩俊, 菅原真司, 石橋豊, “大規模ネットワーク内の遊休計算資源を用いた複数のターゲット情報の効率的探索,” 電子情報通信学会総合大会, 2010年3月17日, 仙台・東北大学.
- [4] 堀内拓真, 菅原真司, 石橋豊, “自律分散型データベースを用いた情報精練手法の検討,” 電子情報通信学会インターネットアーキテクチャ研究会, 2010年2月19日, 東京・機械振興会館.
- [5] 本多謙吾, 菅原真司, 石橋豊, “コンテンツ提示条件の違いに着目した効率的な大規模ネットワーク内情報源モニタリング手法の検討,” 電子情報通信学会ネットワークシステム研究会, 2010年1月28日, 福岡・福岡システム LSI 総合開発センター.
- [6] 井上友介, 菅原真司, 石橋豊, “ストレージ資源の抑制を考慮したハイブリッド型 P2P ネットワークにおけるコンテンツ共有のための効率的複製配置手法,” 電子情報通信学会ネットワークシステム研究会, 2010年1月28日, 福岡・福岡システム LSI 総合開発センター.
- [7] 井上友介, 菅原真司, 石橋豊, “ハイブリッド型 P2P ネットワークにおけるストレージ資源の抑制を考慮したデータ共有のための複製配置手法,” 情報学ワークショップ(WiNF2009), 2009年11月28日, 名古屋・名古屋工業大学.
- [8] 本多謙吾, 菅原真司, 石橋豊, “ユーザ効用を考慮した大規模ネットワーク内情報源モニタリング手法の提案とインターネット環境におけるその有効性の検証,” 情報学ワークショップ(WiNF2009), 2009年11月28日, 名古屋・名古屋工業大学.
- [9] 本多謙吾, 菅原真司, 石橋豊, “ユーザ効用を考慮した大規模ネットワーク内情報源のモニタリング手法の検討,” 電子情報通信学会情報ネットワーク研究会, 2009年1月23日, 名古屋・名古屋工業大学.
- [10] 井上友介, 菅原真司, 石橋豊, “ストレージ容量の制限を考慮したハイブリッド型 P2P ネットワークにおける複製配置手法,” 電子情報通信学会情報ネットワーク研究会, 2009年1月23日, 名古屋・名古屋工業大学.
- [11] 井上友介, 菅原真司, 石橋豊, “ハイブリッド型 P2P ネットワークにおけるストレージ資源の抑制を考慮した複製配置手法に関する検討,” 情報学ワークショップ(WiNF2008), 2008年9月26日, 名古屋・名古屋大学.
- [12] 本多謙吾, 菅原真司, 石橋豊, “大規模ネットワーク内情報源の効率的モニタリング手法の提案と考察,” 情報学ワークショップ(WiNF2008), 2008年9月25日, 名古屋・名古屋大学.
- [13] 井上友介, 菅原真司, 石橋豊, “ハイブリッド型 P2P ネットワークにおけるストレージ資源の抑制を考慮した複製配置手法,” 電子情報通信学会ソサイエティ大会, 2008年9月16日, 東京・明治大学.
- [14] 本多謙吾, 菅原真司, 石橋豊, “大規模ネットワーク内情報源の効率的モニタリング手法の検討,” 電子情報通信学会ソサイエティ大会, 2008年9月17日, 東京・明治大学.
- [15] S. Sugawara, H. Ohnishi, Y. Ishibashi, “Efficient Information Searching Methods Based on User Utility in Super Distributed Environment,” IEEE SAINT 2008, Aug. 1, 2008, Turku, Finland.
- [16] 矢田真輝, 菅原真司, 吉原貴仁, 山岡克式, “成功率管理型探索法における多様なユーザ要求に対する最適探索パラメータの導出,” 電子情報通信学会総合大会, 2008年3月20日, 北九州・北九州学術研究都市.
- [17] S. Kondo, S. Sugawara, Y. Ishibashi, “An Effective Scheduling Scheme for Information Searching with Computational Resources Scattered over a Large-Scale Network,” LKR2008, pp. 351-365, Mar. 5, 2008, Tokyo Institute of Technology, Tokyo.
- [18] 大西宏之, 菅原真司, 石橋豊, “ユーザ効用を考慮した情報探索手法における効率と準最適解の比較,” 電子情報通信学会情報ネットワーク研究会, 2008年1月18日, 名古屋・名古屋工業大学.
- [19] 近藤将太, 菅原真司, 石橋豊, “輻輳を考慮した情報探索のための効率的計算機資源配分スケジューリングに関する検討,” 電子情報通信学会情報ネットワーク研究会, 2008年1月17日, 名古屋・名古屋工業大学.
- [20] 大西宏之, 菅原真司, 石橋豊, “ユーザ効用を考慮した超分散環境における効率的情報探索手法,” 電子情報通信学会ソサイエティ大会, 2007年9月10日鳥

取・鳥取大学.

- [21]近藤将太, 菅原真司, 石橋豊, “分散環境に散在する計算資源を用いた効率的情報探索スケジューリングに関する検討,” 電子情報通信学会ソサイエティ大会, 2007年9月10日鳥取・鳥取大学.
- [22]近藤将太, 菅原真司, 石橋豊, “分散環境に散在する計算機資源を用いた効率的情報探索に関する検討,” 電子情報通信学会情報ネットワーク研究会, 2007年7月20日, 神戸・神戸大学.
- [23]大西宏之, 菅原真司, 石橋豊, “ユーザ効用を考慮した効率的情報探索に関する検討,” 電子情報通信学会情報ネットワーク研究会, 2007年7月20日, 神戸・神戸大学.

[その他]

ホームページ等

<http://nma.web.nitech.ac.jp/>

6. 研究組織

(1) 研究代表者

菅原 真司 (SUGAWARA SHINJI)

名古屋工業大学・大学院工学研究科・准教授

研究者番号：90313410

(2) 研究分担者 なし

(3) 連携研究者 なし

以上