

適合性フィードバックの効率化について

岩山真

(株)日立製作所中央研究所

iwayama@charl.hitachi.co.jp

本論文では、フィードバックする適合性判定数が少ないという状況での適合性フィードバックにおいて、増進的適合性フィードバックと検索結果自動分類の効果を比較した。TREC コレクションを用いた実験結果から、増進的適合性フィードバックは総合的な精度向上には寄与しないことがわかった。クラスタリングによる検索結果の自動分類は有効であることが実証されたが、検索要求によっては自動分類の弊害も見られた。この問題を解決するために、検索要求のバイアスを考慮したクラスタリングアルゴリズムを提案し、その有効性を示した。

Interactive Relevance Feedback with a Small Number of Relevance Judgements

Makoto Iwayama

Central Research Laboratory, Hitachi Ltd.

iwayama@charl.hitachi.co.jp

The use of incremental relevance feedback and document clustering were investigated in an interactive relevance feedback environment in which the number of relevance judgements was quite small. Through experiments on the TREC collection, the incremental relevance feedback approach was found not to improve the overall search effectiveness. The clustering approach was found to be promising, although it sometimes over-focuses on a particular topic in a query and ignores the others. To overcome this problem, a query-biased clustering algorithm was developed and shown to be effective.

1 はじめに

対話的な文書検索では、ユーザが検索システムに検索要求を与えると、検索システムは与えられた検索要求と検索対象文書との適合性(relevance)を計算し、適合性が高いと判断した文書をユーザに提示する。ユーザは次に、それら幾つかの文書に対して実際に適合性を判定する。ユーザが適合/不適合と判定した文書群は検索システムにフィードバックされ、検索結果の改善に用いられる。この適合性フィードバック(relevance feedback)は、ユーザが検索結果に満足するまで繰り返される。

適合性フィードバックでは、ユーザから検索システムへのフィードバック量が多いほど、検索システムは検索結果を改善することができる。Buckley等は、検索精度(recall/precision)はユーザが判定した適合文書数の対数にほぼ比例することを実験的に確かめた[5]。よって、できるだけ多くの適合性判定を行なうことがユーザに求められるが、これはインターネットを利用した検索など即時的な検索では過度の期待であることが多い。一方、文書フィルタリングでは、ユーザの検索要求が長く継続するため、各々の検索要求に対して比較的多くの適合性判定を集めることができる。従って、近年の適合性フィードバックの研究では、文書フィルタリングを対象にして、十分な数の適合性判定を仮定したものが多い[4, 15, 12]。

これに対し本研究では、対話的文書検索を対象にし、適合性判定数が10ないし20以下と非常に少ない状況を仮定する。インターネットにおける対話的文書検索では、ユーザに100以上の適合性判定を強いるのは事実上不可能である。また、文書フィルタリングにおいても、10ないし30の適合性判定をフィードバックするのみで、1000以上の適合性判定をフィードバックした結果より高々10%劣る検索精度に達するという研究結果もある[2]。つまり、検索精度が本質的に改善するのは、フィードバックする適合性判定の数が非常に少ない領域においてである。本研究では、このような状況下でいかに効率良く検索結果を改善していくかを目的に、増進的適合性フィードバック(incremental relevance feedback)と検索結果の自動分類の二つの手法を比較検討する。

通常の適合性フィードバックが幾つかの適合性判定をまとめてフィードバックするのに対し、増進的適合性フィードバックは、ユーザの適合性判定を増進的にフィードバックする手法であり、ユーザがある文書の適合性を判定した時点で即座に検索結果を更

新する。増進的適合性フィードバックはAalbersbergにより提案され[1]、後に、Allanにより文書フィルタリングにおいてその効果が調べられた[2]。本研究では、対話的な文書検索に増進的適合性フィードバックを適用し、少ない適合性判定で効率良く検索精度が改善できるかどうかを実験により確かめる。

一方、検索結果の自動分類は、文書クラスタリングにより検索結果を自動分類してユーザに提示する方法で、近年では検索絞り込みの支援に用いられることが多い[6, 9]。通常の順位付けされた文書表示に比べ、検索結果を自動分類することで、ユーザの情報要求(information need)に適合する文書を効率良く見つけることができる[8]。また、検索結果のクラスタリングは、自動適合性フィードバック(pseudo relevance feedback)の前処理としても有用である[3]。本研究では、フィードバックする適合性判定の数とそれによる検索精度改善率との関係に注目して、適合性フィードバックにおけるクラスタリングの効果を調べる。また、検索要求によるバイアスを考慮できるように従来のクラスタリングアルゴリズムを改良し、その有効性を調べる。

2 比較基準モデル：ベースラインモデルと上限モデル

実験にはTRECコレクションを用いた。ディスク1とディスク2に含まれる計742,709個の文書を検索対象とし、トピック101から150までの50のトピックそれぞれから“title”フィールド、“desc(description)”フィールドを抜き出し50の検索要求を作成した。適合性判定はTRECコレクションで用意されているものをそのまま用いた。

検索モデルはベクトル空間モデルを用いた。タームの重み付け方法としては、Singhal等が提案したLt.Lnc方式を用いた[14]。

適合性フィードバックによるタームの重み修正法は、以下の改良Rocchio法を用いた。

$$Q_i^{\text{新}} = \alpha Q_i^{\text{旧}} + \beta \frac{1}{|\text{適合文書}|} \sum_{\text{適合文書}} w_{ti} - \gamma \frac{1}{|\text{不適合文書}|} \sum_{\text{不適合文書}} w_{ti}$$

ここで $Q_i^{\text{旧}}$ はターム i の修正前の重み、 $Q_i^{\text{新}}$ は修正後の重みである。 w_{ti} はそれぞれの判定文書におけるターム i の重みである。パラメータ α, β, γ はそれぞ

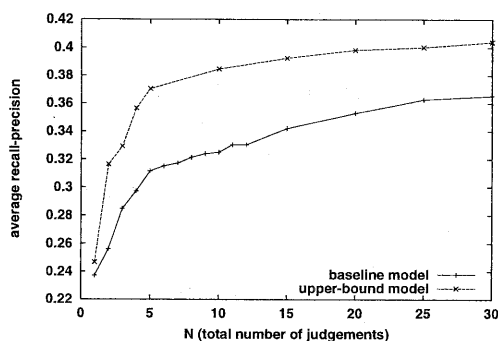


図 1: 比較基準モデルにおけるフィードバックの効果

れ 8, 16, 4 に設定した。これらは、使用した TREC コレクションにおいて最も良く使われている値である。予備実験を行った結果、不適合文書をフィードバックしても検索精度がほとんど向上しなかったため¹、本実験では適合文書のみフィードバックした。

比較基準モデルとして、ベースラインモデル、上限モデルを仮定した。それぞれのモデルではまず、初期検索要求から上述のベクトル空間モデルにより文書を検索し、順位付き文書リストを得る。ベースラインモデルでは、上位 N 位までの文書を調べ、適合する文書のみをフィードバックして、改良 Rocchio 法により新たな検索要求を作成する。上限モデルでは、同じ文書リストにおいて、最上位から下位方向に適合文書のみを N 個抽出し、この N 個の適合文書から新たな検索要求を作成する。つまり、上限モデルでは完璧な文書ランキングを仮定していることになる。ここでは、ユーザは適合文書のみ遭遇するを仮定している。一方、ベースラインモデルでは、実際の検索状況と同じく、初期検索精度に応じた数の不適合文書にも遭遇する。更新した検索要求からそれぞれ新たな順位付き文書リストを得て、この文書リストに対して平均 recall-precision を計算する²。

比較モデルにおける適合性判定数 N と平均 recall-

¹ 不適合文書のフィードバックも検索精度を改善したが、適合文書のフィードバックに比べるとその量は無視できるほど小さかった。我々の実験では、フィードバックに用いる適合性判定の数が少ないことに注意されたい。文書フィルタリングなど、十分な数の適合判定が利用できる環境では、不適合文書のフィードバックも効果的であることが知られている。

² つまり、単一のデータセットに対して検索、フィードバック、および評価を行う。文書フィルタリングなどのタスクでは試験用、統制用の二つのデータセットを用意し、評価は試験用データセットに対して行うのが一般的であり、これは実際の運用状況にも適合している。一方、実際の対話的文書検索では、単一のデータセットに対して検索、フィードバックが行われるため、今回は単一データベース内で評価を行った。

precision の関係を図 1 に示す。本研究では、適合性判定の数が少ない状況での精度改善に注目するため、 N は 30 以下に限った。

3 増進的適合性フィードバック

3.1 手法

Aalbersberg が提案した増進的適合性フィードバック [1] では、検索システムは現在の検索要求に対し最も適合性が高いと思われる文書をユーザに提示し、ユーザはその文書に対する適合性を判定する。ユーザの適合性判定は即座に検索システムにフィードバックされ、検索システムは検索要求を更新し、それに応じて全文書のスコアも再計算する。従って、検索結果の文書ランキングは常に、ユーザがこれまで行なった適合性判定に基づいた最新のものに保たれている。この貪欲 (greedy) な戦略により、ユーザはより多くの適合文書を見つけることができる。

例えば、ユーザがある順位付き文書リストから一つの適合文書を見つけたとする。ここで、次の適合文書を見つけるために二つの選択肢がある。最初の選択肢は、同じ順位付き文書リストから次の適合文書を見つける方法 (ベースラインモデル) であり、もう一つは、たった今見つけた適合文書をシステムにフィードバックして順位付き文書リストを更新し、更新したリストから次の適合文書を見つける方法 (増進的適合性フィードバック) である。図 1 より、フィードバックする適合性判定の数が増すにつれ検索精度も向上することがわかるため、更新した文書リストから探す後者の方が、より簡単に次の適合文書を見つけることができると期待できる。

Aalbersberg は、この仮説を実験的に確かめなかった [1]³。Allan は後に、増進的フィードバックに関する多くの実験を行ったが [2]、それら是对話的検索においてではなく、バッチ的な文書フィルタリングにおいてであった。Allan の実験では、適合性判定は増進的にフィードバックされるが、それらはあらかじめ用意した静的な判定集合の中からランダムに選ばれている。本研究で興味あるのは、フィードバックにより動的に更新される文書集合から、次のフィードバックに用いる文書を選ぶという状況である。

³ Aalbersberg の実験では、適合性フィードバックの回数は一回に限られていた。

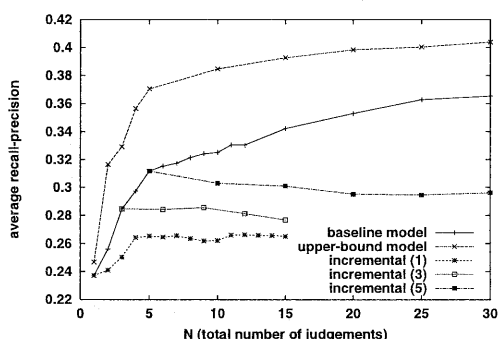


図 2: 増進的適合性フィードバックにおけるフィードバックの効果

3.2 結果と考察

図 2 に、実験結果を示す。増進的適合性フィードバックにおける N とは、増進的に与えられた適合性判定の総数のことである。実験では、元々の増進的適合性フィードバックに加え、判定を一時的にためておく (バッファリング)、まとめてフィードバックする方法も比較した。一時的にためておく判定数 (バッファの大きさ) は 3 および 5 を試みた。元々の増進的適合性フィードバックはバッファの大きさが 1 の場合に相当する。

図からわかるように、残念ながら増進的適合性フィードバックはベースラインモデルを上回ることができなかった。バッファの大きさが 1 の場合、最初の数サイクルのフィードバックでしか精度が改善していない。バッファの大きさが 3 および 5 の場合、全てのフィードバックサイクルにおいて精度の改善が見られない。これは、フィードバックした判定に含まれる適合文書数が少ないためではない。詳細ははぶくが、増進的適合性フィードバックは期待通りベースラインモデルより多くの適合文書を見つけている。

それではなぜ、より多くの適合文書をフィードバックするにも関わらず、増進的適合性フィードバックはベースラインモデルに劣るのだろうか。各フィードバックサイクルで新たに判定する文書は、以前に判定した文書とほとんど同じものであるか、現在の検索要求のサブトピックに関するものである可能性が高い。なぜなら、新たな判定は、現在の検索要求と深く関連するトップランクの文書に対してのみ行われるからである。よって、それらの文書を使って更新した検索要求は、ほとんど変化しないか、または

特定化されるにすぎず、結果、検索精度も向上しないか、または総合的には悪くなってしまう。バッファの大きさが 3 および 5 の場合、フィードバックを続けても初期検索結果を改善できないのは、上記の影響だと考えられる。これに対しベースラインモデルでは、判定文書は常に初期検索結果から選ばれるため、 N が大きくなるにつれ、ユーザの情報要求に関連する様々なトピックを含む文書がフィードバックされやすくなる。

この現象を詳しく調べるために、50 の検索要求を「高品質/低品質」の 2 つのグループに分け、それぞれに対して実験結果を集計した。実際には、上位 30 位の検索結果に 15 個以上の適合文書を含むような検索要求を高品質とし、それ以外を低品質とした。高品質検索要求は、情報要求に関連する様々なトピックを含む文書群を上位に検索できる可能性が高く、逆に低品質な検索要求は、情報要求の表現としてあまり有用ではない。上位 30 位における平均適合率は、高品質セットで 0.7500、低品質セットで 0.2345 であった。

図 3 にそれぞれの結果を示す。低品質検索要求では、フィードバックが進むにつれ、検索結果も改善されていくが、それでも絶対的な精度はベースラインモデルに劣る。一方、高品質検索要求では、バッファの大きさが 1 の場合を除き、初期検索結果にフィードバックを加えるにつれ、検索精度は急激に落ちていく。これは、初期検索結果に含まれる様々なトピックの一面にのみ検索が集中していくことを示唆している。

以上の結果から、増進的適合性フィードバックは検索精度の総合的な向上には不向きであることがわかった。反面、検索要求が特定のトピックに集中していくという意味で、検索結果を絞り込んでいく過程で有用かもしれない。また、適合性フィードバックにおいては、フィードバックする文書数を単に増やすことが総合的な精度向上にはつながらないこともわかった。総合的な精度を上げるには、むしろフィードバックする文書の多様性が重要となる。これは、文書分類で訓練例をいかに効率良くサンプリングするかという問題にも深く関係する [11]。

4 検索結果の自動分類

4.1 手法

「ある情報要求に適合する文書はお互いに類似している」というクラスタ仮説 [16] が正しければ、検

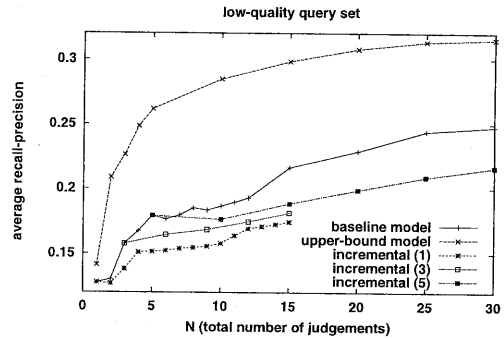
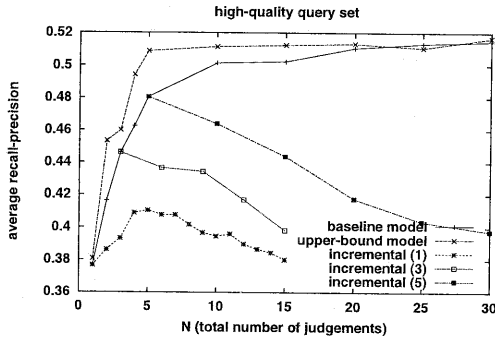


図 3: 増進的適合性フィードバックにおけるフィードバックの効果 (高品質/低品質検索要求)

検索結果をクラスタリングすることで、適合文書のみ含むクラスと不適合文書のみ含むクラスに分割することができる。更に、ユーザが適合クラスを選ぶことができれば、その中に含まれる多くの適合文書を効率良くフィードバックすることができる。ここでは、不適合クラスに集められた多くの不適合文書を調べる必要がない。

クラスタリングによる検索結果の自動分類は、多くの研究者により様々な検索環境で用いられている。Scatter/Gather [7] は、クラスタリングとクラス選択を繰り返すことで検索結果の絞り込みを支援する手法である。Buckley 等は、検索結果のクラスタリングは自動適合性フィードバックの前処理としても有用であることを示した [3]。Evans 等は、検索結果をクラスタリングして表示することで、ユーザが実際に効率良く適合性フィードバックを行えることを示した [8]。本研究では、フィードバックする適合性判断の数とそれによる検索精度改善率との関係に着目して、検索結果をクラスタリングすることの有用性を調べる。前節の実験結果からわかったように、単に多くの適合文書をフィードバックすれば検索精度もそれに応じて改善するわけではないため、クラス仮説により適合文書を効率良く集めることができて、それらがフィードバック情報として有用であるとは限らない。

実験では、各検索要求から 150 の文書を検索し、これら 150 の文書を 5 個のクラスに分割した。クラスタリングアルゴリズムとしては確率的クラスタリング [10] を用いた。5 個のクラスから最良のクラスを選ぶために、[13] で使われている “DENSITY” 法を用いた。“DENSITY” 法では、適合する文書 (正

解) を含む割合が大きいクラスから順にクラスを選ぶ。各クラス内の文書は、検索要求との関連度によりソートする。実験では、まず 1 位のクラスを選び、その中から N 個の文書を選びフィードバックに用いる。1 位のクラスに N 個以下の文書しか存在しない場合、2 位のクラスから残りの文書を選ぶ。

“DENSITY” 法を用いるということは、実際のユーザも適合文書を多く含むクラスを比較的容易に選択できると仮定しているのだが、これは少々強すぎる仮定かもしれない。ユーザはほとんどの場合 “DENSITY” 法による 1 位のクラスを選ぶことができるという実験結果 [9] もあるが、そのコストも含め更に詳しく調べる必要がある。

4.2 結果と考察

図 4 に実験結果を示す。増進的適合性フィードバックとは異なり、検索結果の自動分類法では、フィードバックにより常に検索精度が向上している。平均 recall-precision の値も常にベースラインモデルを上回る。また、詳細ははぶくが、自動分類した検索結果から文書を選ぶほうが、オリジナルの検索結果から文書を選ぶ (ベースラインモデル) よりも多くの適合文書を見つけている。以上はクラス仮説の正当性、およびその有効性を実証している。

ここで、増進的適合性フィードバックの場合と同様に、50 の検索要求を高品質/低品質の二つのグループに分け、それぞれに対して結果を集計し図 5 に示す。図より、高品質検索要求においてクラスタリングの問題点が見えてとれる。高品質検索要求は多くの

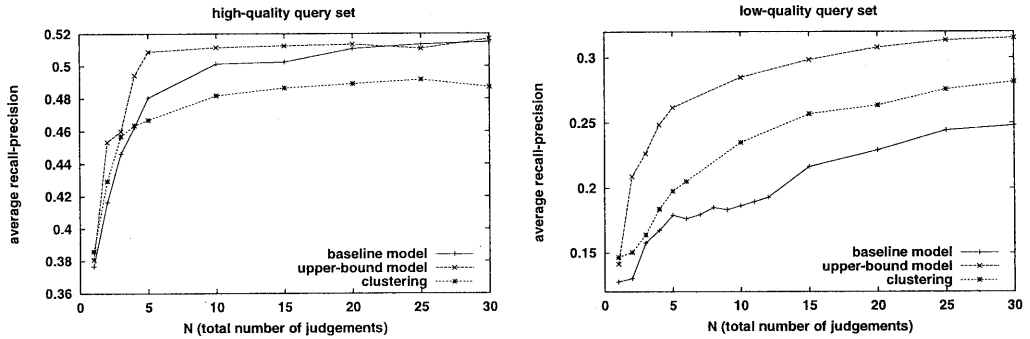


図 5: 検索結果自動分類法におけるフィードバックの効果 (高品質/低品質検索要求)

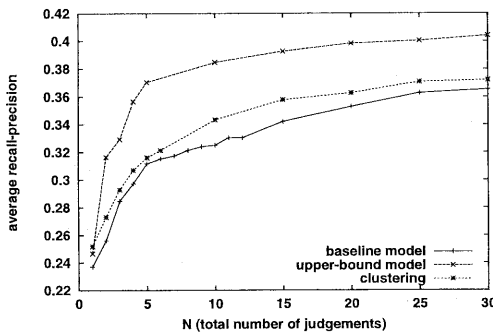


図 4: 検索結果自動分類法におけるフィードバックの効果

適合文書を検索するため、クラスタリングの対象文書もほとんどが適合文書である。今回用いたクラスタリングアルゴリズムは重複するクラスタを許さないため⁴，これら適合文書をクラスタリングし，その中の1つか2つを選ぶことは，適合文書がカバーする様々な適合トピックの一部分しか考慮しないことに相当し，これは総合的な検索精度の低下を招く。実際，クラスタリングによる検索結果自動分類法はベースラインモデルを上回ることができない。一方，低品質検索要求は比較的小数の適合文書しか検索しないため，適合文書はクラスタリングにより1つないしは2つのクラスタに集まりやすく，これらのクラスタを選ぶことで必要かつ十分なフィードバックが可能になる。実際，検索結果の自動分類法はベースラインモデルを上回る。高品質検索要求におけるク

ラスタリングの問題点を解決するためには，例えば，複数のクラスタから上位の文書のみを集めてフィードバックする方法もあるが，次節では別の本質的な解決法を試みる。

5 検索要求によるバイアスを考慮したクラスタリング

前節の結果は，適合性フィードバックにおける検索要求の重要性を示唆している。特に，検索要求が多く，多くの適合文書を検索できるほど，当然ながらその重要性も高い。実際，そのような高品質検索要求に関しては，検索要求との関連度が強い順に検索結果をフィードバックすること（ベースラインモデル）で総合的に良い結果が得られ，検索結果のクラスタリングは必ずしも有効ではない。本節では，検索要求によるバイアスを考慮したクラスタリングアルゴリズムを提案し，その有効性を調べる。

前節の検索結果自動分類法では，検索要求はクラスタリングの対象文書を決めるのみで，クラスタリングアルゴリズム自体は検索要求とは無関係である。本節では，検索要求など任意のバイアスを考慮できるようクラスタリングアルゴリズムを改良する。前節で用いた確率的クラスタリングは，文書 d が与えられたという条件でのクラスタ C の確率 $P(C|d)$ を計算し，クラスタ集合の確率 $\prod_C \prod_{d \in C} P(C|d)$ が最大になるようボトムアップにクラスタを併合していく [10]。ここで，確率 $P(C|d)$ の条件部にバイアス（検索要求） q の情報を加えた確率 $P(C|d, q)$ で，全ての $P(C|d)$ を置き換えることにより，バイアスの影響を考慮したクラスタリングアルゴリズムを得る。これ

⁴ ほとんどのクラスタリングアルゴリズムも同じである。

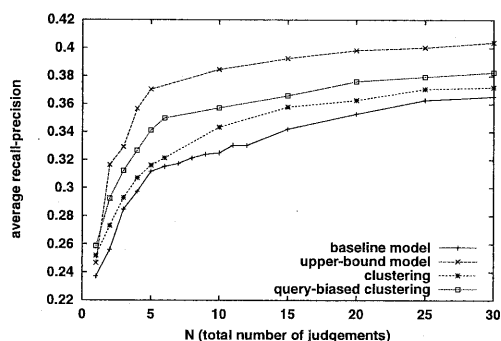


図 6: 検索要求でバイアスされたクラスタリングを用いた検索結果自動分類法におけるフィードバックの効果

は、検索要求 q に含まれる各ターム (単語) の重みを、文字通りバイアスとして、文書 d におけるそれらタームの重みに加えることに相当する。

前節の実験法において、クラスタリングアルゴリズムのみを上述のアルゴリズムで置き換えて実験した結果を図 6 に示す。図から、検索要求でバイアスされたクラスタリングは検索結果の分類法として効果的であることがわかる。オリジナルのクラスタリングを用いた検索結果分類法およびベースラインモデルを上回り、かつその差は有為である。これらの差は、高品質検索要求での差によるところが大きい。図 7 に高品質/低品質検索要求での結果を示す。高品質検索要求における結果を見ると、検索要求でバイアスされたクラスタリングは、ベースラインモデルと同じかやや上回る結果を出す。つまり、ベースラインモデルと同じく、検索要求の有効性を十分反映した手法になっていることがわかる。一方、低品質検索要求では、オリジナルのクラスタリングと同じかやや上回る結果を出す。つまり、検索結果をクラスタリングすることのそもそもの動機となったクラスタ仮説の有効性も十分反映した手法になっていることがわかる。また、詳細ははぶくが、二つのクラスタリング法はほぼ同じ数の適合文書をフィードバックしていることがわかる。即ち、検索要求でバイアスされたクラスタリングが優れている理由は、フィードバックする適合文書の数が多いからではなく、検索要求に強く関連した文書をより多くフィードバックするからである。加えて、クラスタ仮説による優位性も保たれている。

6 おわりに

本論文では、フィードバックする適合性判定数が少ないという状況での適合性フィードバックにおいて、増進的適合性フィードバックと検索結果自動分類の効果を比較した。TREC コレクションを用いた実験結果から、増進的適合性フィードバックは総合的な精度向上には寄与しないことがわかった。クラスタリングによる検索結果の自動分類は有効であることが実証されたが、検索要求によっては自動分類の弊害も見られた。この問題を解決するために、検索要求のバイアスを考慮したクラスタリングアルゴリズムを提案し、その有効性を示した。

増進的適合性フィードバックの結果は否定的であったが、情報要求のサブピックに効率良くフォーカスしていくためには有用かもしれない。本論文ではこの点に関して詳しく検討しなかったが、今後の課題として必要な事項である。なぜなら、適合性フィードバックは繰り返して適用するものという常識があるが、少なくとも、今回のようにフィードバックを小刻みに適用し、かつトップランクの文書しか見ないという方法では、総合的に良い結果を出すことができないからである。

検索結果の自動分類に関しては、実際のユーザが適合クラスタを容易に選ぶことができるかどうかを確かめる必要がある。Hearst と Pedersen による実験結果 [9] はこの仮説を部分的に支持しているが、本論文で提案したバイアス付きクラスタリングに関しては定かでない。著者の個人的な観察によれば、検索要求でバイアスされたクラスタリングは、通常のクラスタリングに比べ、適合文書と非適合文書をより明確に分けることができ、ユーザによるクラスタ選択も容易になると思われるが、この点についても実験を加える必要がある。

最後に、バイアス付きのクラスタリングは様々な検索環境に適用できる。例えば、Scatter/Gather のような検索絞り込みや、検索結果の自動要約において有用である。通常のクラスタリングとは異なり、同じ文書集合からでもバイアスに応じて異なる分類が可能になるため、ユーザの情報要求をバイアスとすることで、よりユーザに適応した自動分類法が実現できる。

参考文献

- [1] I. J. Aalbersberg. Incremental relevance feedback. In *Proceedings of the Annual International ACM SIGIR*

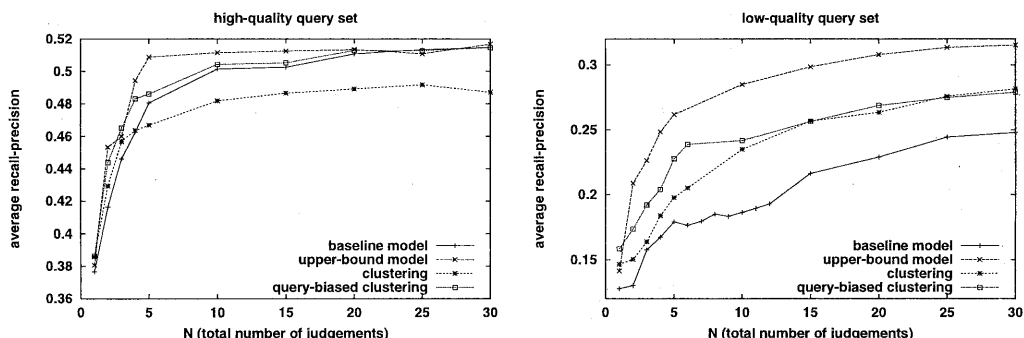


図 7: 検索要求でバイアスされたクラスタリングを用いた検索結果自動分類法におけるフィードバックの効果 (高品質/低品質検索要求)

Conference on Research and Development in Information Retrieval, pp. 11–22, 1992.

- [2] J. Allan. Incremental relevance feedback for information filtering. In *Proceedings of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 270–278, 1996.
- [3] C. Buckley, M. Mitra, J. Walz, and C. Cardie. Using clustering and SuperConcepts within SMART: TREC 6. In *Proceedings of the Sixth Text REtrieval Conference (TREC-6)*, 1998.
- [4] C. Buckley and G. Salton. Optimization of relevance feedback weights. In *Proceedings of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 351–357, 1995.
- [5] C. Buckley, G. Salton, and J. Allan. The effect of adding relevance information in a relevance feedback environment. In *Proceedings of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 292–300, 1994.
- [6] D. R. Cutting, D. R. Karger, J. O. Pedersen, and J. W. Tukey. Scatter/Gather: A cluster-based approach to browsing large document collections. In *Proceedings of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 318–329, 1992.
- [7] D. Cutting, J. Kupiec, J. Pedersen, and P. Sibun. A practical part-of-speech tagger. In *Proc. of the Third Conference on Applied Natural Language Processing*, 1992.
- [8] D. A. Evans, A. Huettner, Tong X., P. Jansen, and J. Bennett. Effectiveness of clustering in ad-hoc retrieval. In *Proceedings of the Seventh Text REtrieval Conference (TREC-7)*, 1999.
- [9] M. A. Hearst and J. O. Pedersen. Reexamining the cluster hypothesis: Scatter/Gather on retrieval results. In *Proceedings of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 1996.
- [10] M. Iwayama and T. Tokunaga. Cluster-based text categorization: A comparison of category search strategies. In *Proceedings of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 273–280, 1995.
- [11] D. D. Lewis and W. A. Gale. A sequential algorithm for training text classifiers. In *Proceedings of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 3–12, 1994.
- [12] R. E. Schapire, Y. Singer, and A. Singhal. Boosting and rocchio applied to text filtering. In *Proceedings of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 215–223, 1998.
- [13] H. Schütze and C. Silverstein. Projections for efficient document clustering. In *Proceedings of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 1997.
- [14] A. Singhal, C. Buckley, and M. Mitra. Pivoted document length normalization. In *Proceedings of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 21–29, 1996.
- [15] A. Singhal, M. Mitra, and C. Buckley. Learning routing queries in a query zone. In *Proceedings of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 25–32, 1997.
- [16] C. J. van Rijsbergen. *Information Retrieval*. Butterworths, London, 1979.